JMIR Public Health and Surveillance

Impact Factor (2023): 3.5
Volume 11 (2025) ISSN 2369-2960 Editor in Chief: Travis Sanchez, PhD, MPH

Contents

Original Papers

Viewpoints

Reviews

Research Letter

COVID-19 Testing Equity in New York City During the First 2 Years of the Pandemic: Demographic Analysis of Free Testing Data

Daniel Rosenfeld[1,2], MSci; Sean Brennan[1,3], MPH; Andrew Wallach[1,4], MD; Theodore Long[1,5], MHS, MD; Chris Keeley[1], MUP; Sarah Joseph Kurien[1], MPA

[1]NYC Test & Treat Corps, The New York City Health + Hospitals Corporation, 50 Water St., New York, NY, United States
[2]Department of Housing Stabilization, The Massachusetts Executive Office of Housing and Livable Communities, Boston, MA, United States
[3]Memorial Sloan Kettering Cancer Center, New York, NY, United States
[4]Division of General Internal Medicine and Clinical Innovation, New York University Grossman School of Medicine, New York, NY, United States
[5]Department of Population Health, New York University, New York, NY, United States

Corresponding Author:
Daniel Rosenfeld, MSci
NYC Test & Treat Corps, The New York City Health + Hospitals Corporation, 50 Water St., New York, NY, United States

Abstract

Background: COVID-19 has caused over 46,000 deaths in New York City, with a disproportional impact on certain communities. As part of the COVID-19 response, the city has directly administered over 6 million COVID-19 tests (in addition to millions of indirectly administered tests not covered in this analysis) at no cost to individuals, resulting in nearly half a million positive results. Given that the prevalence of testing, throughout the pandemic, has tended to be higher in more affluent areas, these tests were targeted to areas with fewer resources.

Objective: This study aimed to evaluate the impact of New York City's COVID-19 testing program; specifically, we aimed to review its ability to provide equitable testing in economically, geographically, and demographically diverse populations. Of note, in addition to the brick-and-mortar testing sites evaluated herein, this program conducted 2.1 million tests through mobile units to further address testing inequity.

Methods: Testing data were collected from the in-house Microsoft SQL Server Management Studio 18 Clarity database, representing 6,347,533 total tests and 449,721 positive test results. These tests were conducted at 48 hospital system locations. Per capita testing rates by zip code tabulation area (ZCTA) and COVID-19 positivity rates by ZCTA were used as dependent variables in separate regressions. Median income, median age, the percentage of English-speaking individuals, and the percentage of people of color were used as independent demographic variables to analyze testing patterns across several intersecting identities. Negative binomial regressions were run in a Jupyter Notebook using Python.

Results: Per capita testing inversely correlated with median income geographically. The overall pseudo r^2 value was 0.1101 when comparing hospital system tests by ZCTA against the selected variables. The number of tests significantly increased as median income fell (SE 1.00000155; $P<.001$). No other variables correlated at a significant level with the number of tests (all P values were >.05). When considering positive test results by ZCTA, the number of positive test results also significantly increased as median income fell (SE $1.57e^{-6}$; $P<.001$) and as the percentage of female residents fell (SE 0.957; $P=.001$). The number of positive test results by ZCTA rose at a significant level alongside the percentage of English-only speakers (SE 0.271; $P=.03$).

Conclusions: New York City's COVID-19 testing program was able to improve equity through the provision of no-cost testing, which focused on areas of the city that were disproportionately impacted by COVID-19 and had fewer resources. By detecting higher numbers of positive test results in resource-poor neighborhoods, New York City was able to deploy additional resources, such as those for contact tracing and isolation and quarantine support (eg, free food delivery and free hotel stays), early during the COVID-19 pandemic. Equitable deployment of testing is feasible and should be considered early in future epidemics or pandemics.

(JMIR Public Health Surveill 2025;11:e52972) doi:10.2196/52972

KEYWORDS

COVID-19 testing; health disparities; equity in testing; New York City; socioeconomic factors; testing accessibility; health care inequalities; demographic analysis; COVID-19 mortality; coronavirus; SARS-CoV-2; pandemic; equitable testing; cost; poor neighborhood; resources

Introduction

In March 2020, New York City (NYC) emerged as one of the global epicenters of SARS-CoV-2 infections and COVID-19 deaths. This was roughly 2 months after the first detected case in the United States and 3 months following the first case in Wuhan, China [1]. Early in the pandemic, within the United States and especially within NYC, testing availability was extremely limited. In fact, there were no positive test results recorded in NYC until the beginning of March 2020 [2]. Over the ensuing 3 years, over 45,000 deaths and over 3.2 million infections were documented in NYC—a wave of mortality that contributed to a decline of 4.6 years in life expectancy from 2019 to 2020 alone [3].

It should be noted that COVID-19 mortality and SARS-CoV-2 infection rates have varied widely based on several spatial and socioeconomic factors [4]. US counties with higher poverty rates had higher COVID-19 case numbers and related death rates, and racial and ethnic minority individuals are at increased vulnerability for COVID-19 when considering both infection rates and mortality rates [5,6]. While life expectancy declined by 4.6 years citywide, it declined by 6 years for Hispanic and Latinx New Yorkers (to 77.3 years), by 5.5 years for Black New Yorkers (to 73 years), and by 3 years for White New Yorkers (to 80.1 years) [3]. These mortality disparities were not limited to race. Individuals with chronic conditions, such as obesity and diabetes, faced higher mortality rates than those faced by the general population as a whole, while socioeconomic factors, such as poverty, housing overcrowding, the effects of historical residential racial segregation, and an inability to perform some jobs remotely, have intersected with the aforementioned chronic health issues to create a perfect storm of increased COVID-19 mortality in particular neighborhoods [7].

Following the initial lack of testing at the onset of the pandemic, the federal and state governments sought to develop and implement a mass testing strategy. Once tests were made available, significant barriers to accessing testing remained in terms of costs and the geographic locations of testing sites. The NYC Health + Hospitals Corporation (H+H)—the largest public health system in the United States, with 11 acute care facilities, 6 diagnostic and treatment centers, and more than 50 neighborhood health centers (mostly, but not entirely, in economically depressed neighborhoods)—became the largest free testing provider in NYC. The NYC Department of Health and Mental Hygiene has always tracked health and socioeconomic disparities. This background provided the basis for establishing the TRIE (Taskforce on Racial Inclusion & Equity) neighborhoods [8]. These were neighborhoods that had economic disadvantages and were disproportionately impacted by COVID-19. We leveraged the neighborhoods, which essentially involved combining former disparity data related to the impact of COVID-19, to help set our strategy for allocating testing sites. By providing access to no-cost COVID-19 testing, we sought to understand whether our equitably distributed testing program had a relationship with health outcomes and whether the no-cost testing option can be used to increase testing in underserved areas.

Methods

Data Sources

Data on COVID-19 testing that was directly administered by the NYC H+H were collected from the in-house NYC H+H Microsoft SQL Server Management Studio 18 Clarity database, representing 6,347,533 total tests and 449,721 positive test results. These tests were sourced from the in-house NYC H+H instance of Epic Systems; these did not include tests conducted by external testing vendors contracted by the NYC H+H or tests conducted by the school testing program that mandated surveillance testing in NYC public schools from 2020 to 2022. Demographic data were taken from 2020 US Census Bureau data [9]. Citywide testing data came from the NYC coronavirus data repository [10].

Experimental Design

We initially sought to run a Poisson regression analysis comparing counts of tests and positive test results to demographic variables while using population as an offset; however, we found the data to be overdispersed and chose to use negative binomial regressions. In terms of the variables we chose to analyze, we were influenced by two studies from King County, Washington, and NYC that, among other methods, analyzed testing rates against demographic variables. The King County, Washington, study used a variable called “people of color (POC),” into which POC racial data were collapsed due to collinearity among POC groups [11]. Given similar patterns of collinearity in our data, we chose to do this in our study. The NYC study used a number of different demographic variables, such as income, education, and housing overcrowding, to calculate a single socioeconomic index score [2]. Although this is one way to handle collinearity between socioeconomic variables, such as income and poverty, we chose to use median income as our main predictor of socioeconomic status. As variables related to COVID-19 mortality and testing rates, we included age and sex breakdowns, as being older and being male positively correlate with higher mortality. As a measure of presumed difficulty in accessing health care, the percentage of English speakers was included.

Ethical Considerations

We used deidentified patient data. The Biomedical Research Alliance of New York (BRANY) Institutional Review Board (IRB) determined that our research in this paper did not constitute research involving human subjects that is regulated by US Department of Health and Human Services or US Food and Drug Administration regulations, and it was therefore not subject to further BRANY IRB review. The BRANY study ID is 23-15-564, the sponsor ID is RH2023-814, and our category was determined as “15-Not Human Subject Research.” We were advised by BRANY that we did not have to create a National Clinical Trial number with the US National Library of Medicine ClinicalTrials.gov website because our data-only study used preexisting data.

Analysis

Regressions were run in a Jupyter Notebook (Project Jupyter) using Python. Data visualizations were created by using Tableau

(Tableau Software LLC). A selection of data visualizations can be found in the *Results* section. As this was a spatial analysis, we took care to map each of our demographic variables, testing data points, and background health outcome data. We mapped out "Percent POC," "Percent Female," "Percent Speaks Only English," "Median Age," "Median Income," and "COVID-19 death rate" as demographic and health outcome variables by zip code tabulation area (ZCTA). To spatially display our testing data, we mapped the variables "Total Citywide Tests per 100,000," "NYC H+H COVID-19 Tests," "NYC H+H COVID-19 tests per 100,000," "NYC H+H positive tests per 100,000," and "NYC H+H COVID-19 test positivity rate" also by ZCTA. We also created scatter plots to illustrate the relationships between these variables.

Results

The overall pseudo r^2 value was 0.1101 when comparing NYC H+H tests by ZCTA against the selected variables. The number of NYC H+H tests significantly increased as median income fell (exponentiated β=.99998, SE 1.00000155; *P*<.001). No other variables correlated at a significant level with the number of tests (all *P* values were >.05). When considering positive test results by ZCTA, the number of positive test results also significantly increased as median income fell (exponentiated β=.99998, SE $1.57e^{-6}$; *P*<.001) and as the percentage of female residents fell (exponentiated β=.04014, SE 0.957; *P*=.001). Positive test results by ZCTA rose at a less significant level alongside the percentage of English-only speakers (β=1.81630, SE 0.271; *P*=.03). No other predictor variables correlated with positive test results by ZCTA (all *P* values were >.05). *P*<.05 was chosen as a standard biomedical level of significance for geographic area–level research. A selection of data visualizations depicting our results are shown in Figures 1-6.

Figure 1. Selected demographic, testing, and mortality variables by zip code tabulation areas (clockwise from upper left): "Percent POC," "Percent Female," "Percent Speaks Only English," "Total Tests per 100k," "Median Income" (US $), and "Median Age" (years). POC: people of color.

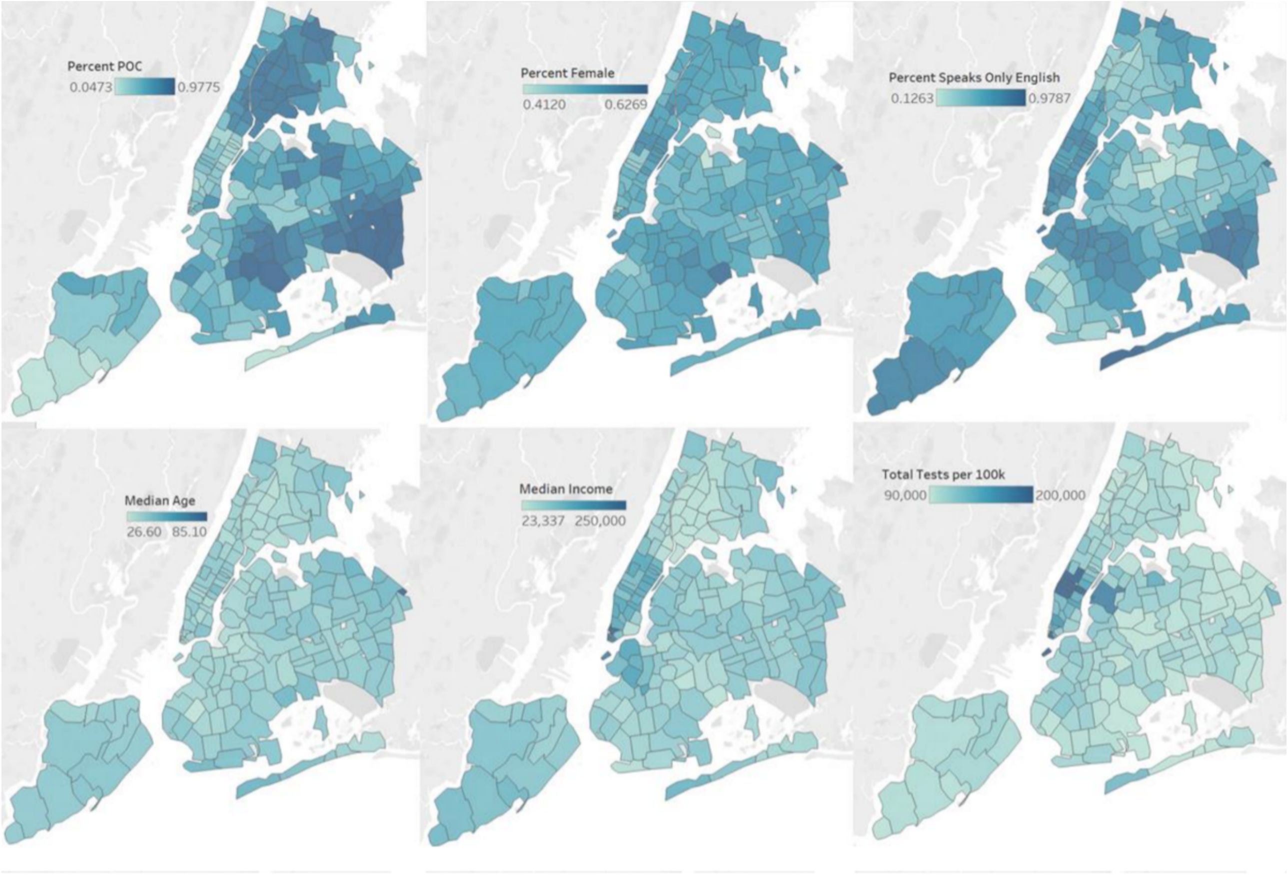


Figure 2. Selected demographic, testing, and mortality variables by zip code tabulation areas continued (clockwise from upper left): “NYC H+H COVID-19 Tests,” “NYC H+H Positive Tests,” “NYC H+H COVID-19 test positivity rate,” “COVID-19 death rate,” “NYC H+H positive tests per 100,000,” and “NYC H+H COVID-19 tests per 100,000.” NYC H+H: New York City Health + Hospitals Corporation.

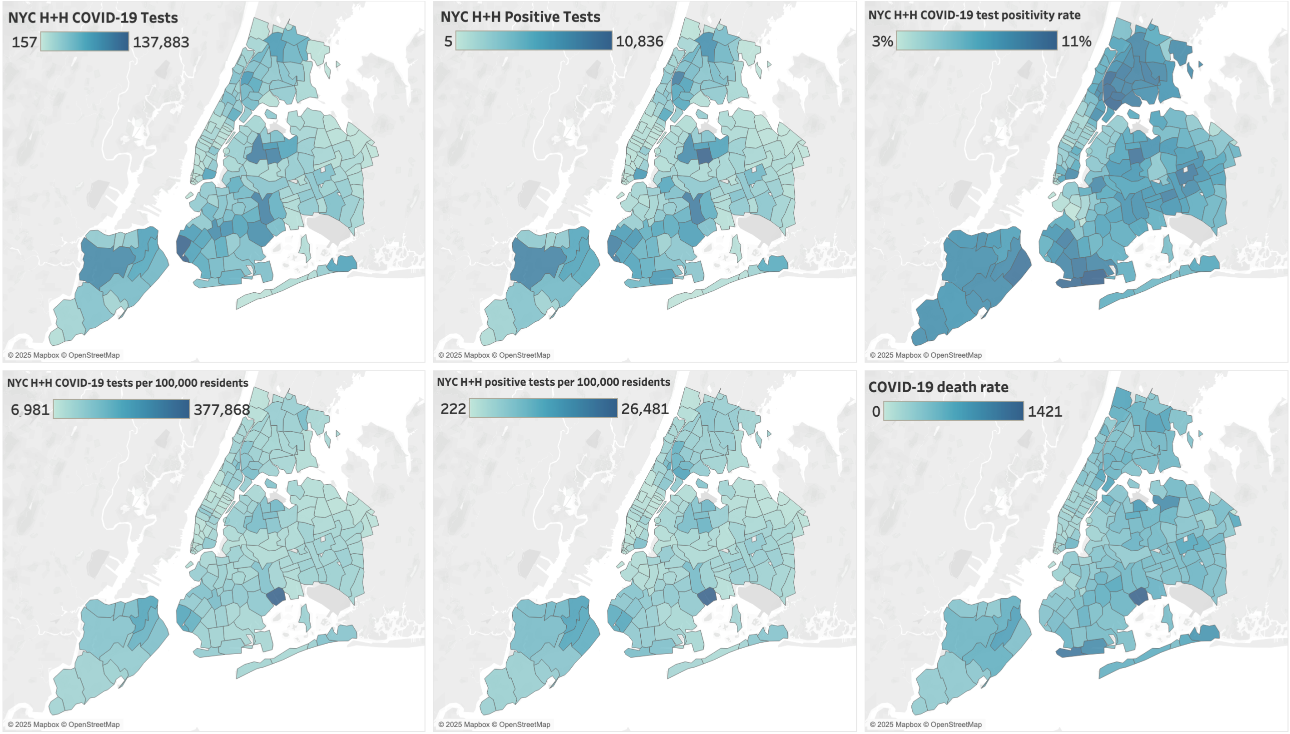


Figure 3. Selected demographics compared to NYC H+H COVID-19 tests per 100,000 residents by zip code tabulation area (from top to bottom): “Median Age” (years), “Median Income” (US $), “Percent POC,” and “Percent Speaks Only English” by NYC H+H COVID-19 tests per 100,000 residents. NYC H+H: New York City Health + Hospitals Corporation; POC: people of color.

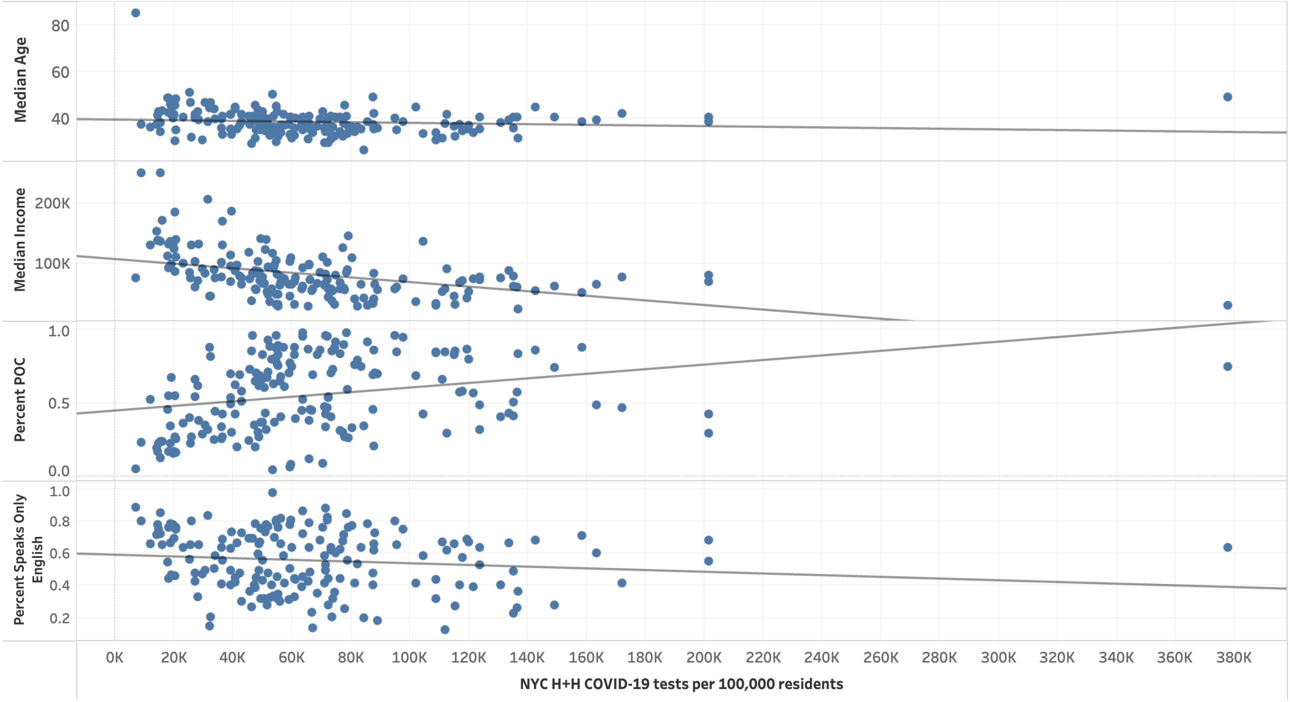


Figure 4. Selected demographics compared to NYC H+H COVID-19 test positivity rate by zip code tabulation area (from top to bottom): “Median Age” (years), “Median Income” (US $), “Percent POC,” and “Percent Speaks Only English” by NYC H+H test positivity rate. NYC H+H: New York City Health + Hospitals Corporation; POC: people of color.

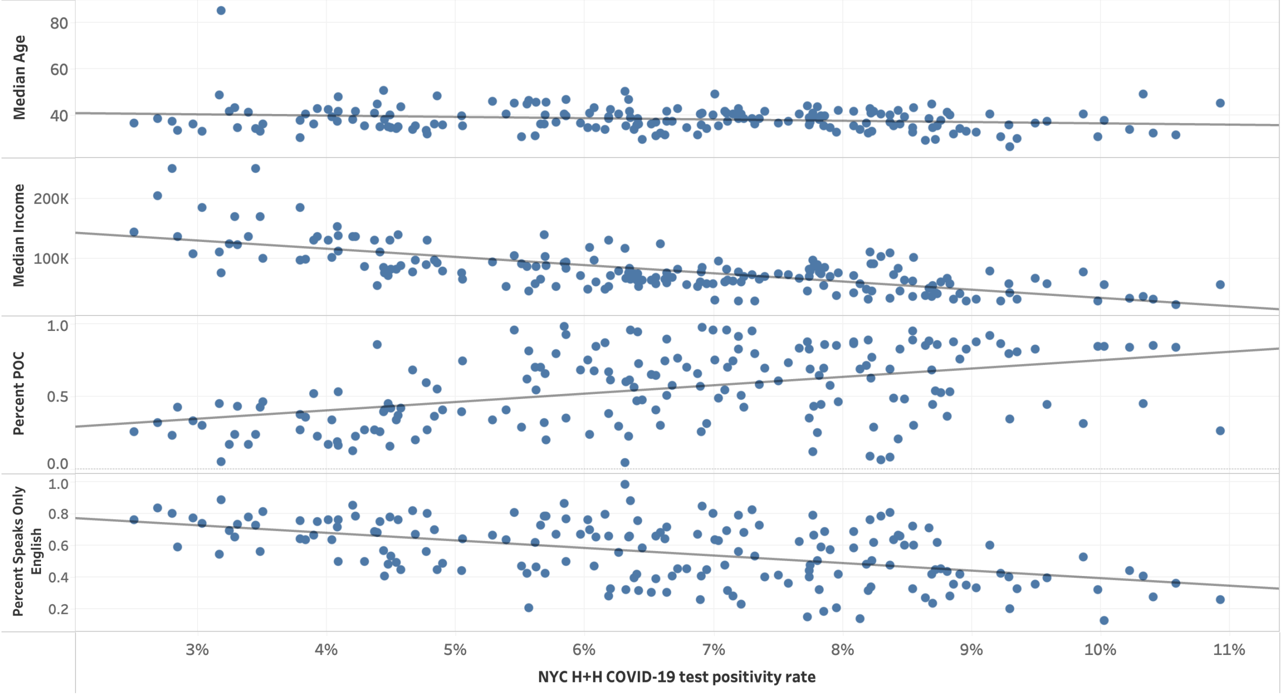


Figure 5. Selected test and death rates compared to NYC H+H COVID-19 tests per 100,000 residents by zip code tabulation area (from top to bottom): NYC H+H COVID-19 test positivity rate and COVID-19 death rate by NYC H+H COVID-19 tests per 100,000 residents. NYC H+H: New York City Health + Hospitals Corporation.

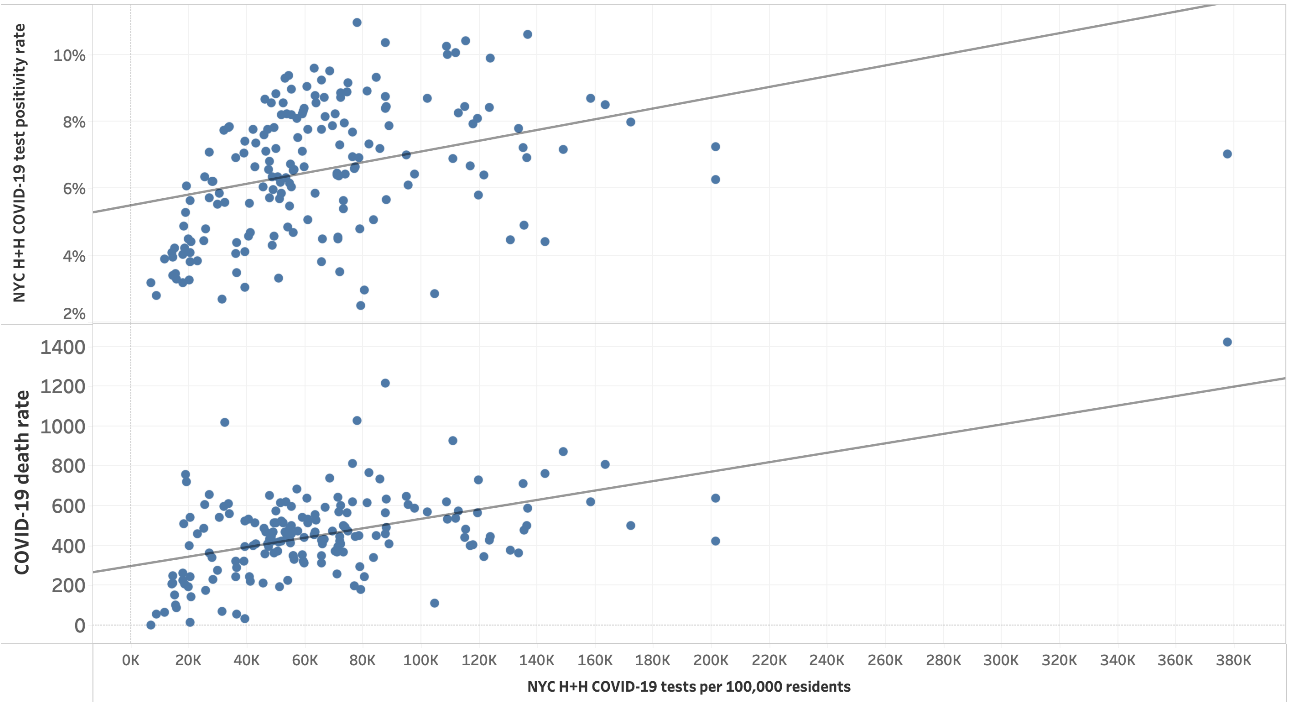
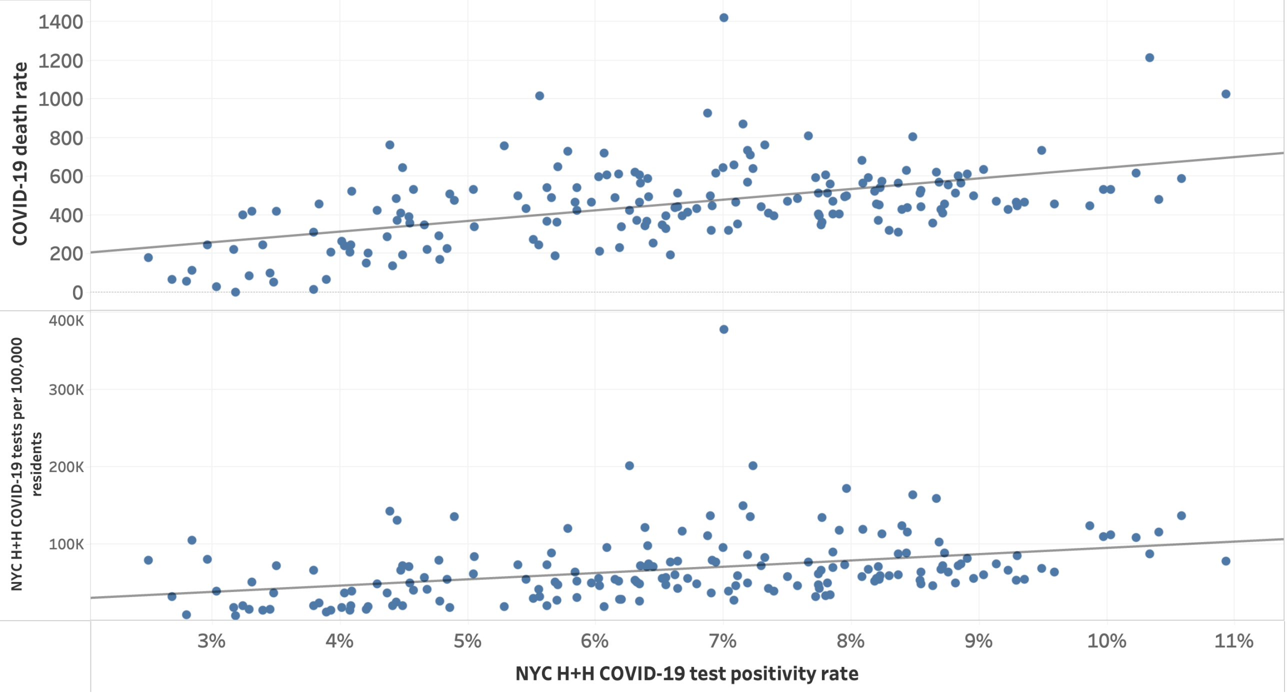


Figure 6. Selected test and death rates compared to NYC H+H COVID-19 positivity rate by zip code tabulation area (from top to bottom): COVID-19 death rate and NYC H+H COVID-19 tests per 100,000 residents by NYC H+H COVID-19 positivity rate. NYC H+H: New York City Health + Hospitals Corporation.



Discussion

The devastatingly high mortality rates seen in NYC during the height of the first COVID-19 wave in April 2020 were likely linked to undetected cases as a result of limited testing capabilities at the time. Roughly half of all deceased NYC patients with COVID-19 died during the first months of the first wave, making comparative analysis somewhat difficult in NYC when compared to other regions. Approximately 6 million COVID-19 tests were completed by the NYC H+H between early 2020 and late 2022, though most testing was done following the first wave in 2020 due to the time and resources required to set up testing sites. In our analysis, the number of NYC H+H no-cost COVID-19 tests increased as the median income of a discrete neighborhood fell—a trend that stood in opposition to overall citywide testing trends, wherein the highest testing rates were observed in wealthier areas of the city. The number of positive test results at the NYC H+H also rose as median income fell in a discrete neighborhood, even when controlling for other variables. This suggests that NYC H+H no-cost COVID-19 tests were equitably distributed in targeted increased-risk neighborhoods of NYC. This was important, considering the disproportionate negative health impacts of the pandemic on these neighborhoods, which already had a preexisting vulnerability to a range of adverse health outcomes.

Consistent and easily accessible large-scale testing may be a crucial element of successful strategies for avoiding additional mortality. Our research echoes that of some other studies that found an inverse relationship between COVID-19 positivity and test availability [11]. With regard to the racial and income equity issues that have been seen in COVID-19 testing within NYC and the United States, our findings on tests administered by the NYC H+H were more mixed for demographic variables, such as “Percent POC,” that otherwise might be related to COVID-19 mortality risk. Our study had limitations with regard to interpreting disparities in COVID-19 testing rates across space and disparities in time to COVID-19 mortality. COVID-19 testing was extremely limited early in the pandemic due to the lack of readily available testing, whereas from 2022 to 2023, at-home rapid testing became ubiquitous but came with the caveat that results were rarely reported. In addition, as we compared COVID-19 testing rates and results between geographic and demographic variables, we recognized the pitfalls that may arise when equating jurisdiction-level data with individuals who live within those jurisdictions. NYC is an extremely diverse city with substantial income inequality and variability even within the smallest geographic neighborhoods. The relationship between COVID-19 test positivity and the percentage of female residents could suggest that ZCTAs with a higher percentage of male residents had higher positivity rates, and this information could be used to inform future testing campaign strategies (eg, targeting areas with a higher percentage of male residents).

We take the inverse relationship between NYC H+H testing rates and median income as a positive sign that the no-cost COVID-19 testing program was able to penetrate areas of the city that were likely neglected by other testing providers. Despite the devastating consequences from the COVID-19 pandemic in NYC, this targeted, equitably distributed testing plan could have broader public health implications if similar outreach strategies are used in future pandemics.

Acknowledgments

We thank Kym Neck, Anna Gilbert, Cornelia Mills, Taehan Lee, Sharhzad Divsalar, Chi Gao, Cara Feldkamp, George Zhuo, Alexander Dobranic, Christina Pili, and George Cowan from the New York City Health + Hospitals Corporation (NYC H+H) for their help in completing this paper. We also thank Raffaella Hart, Laura Donohue, Linda Reuter, and Svetlana Abramova from Biomedical Research Alliance of New York (BRANY) for their assistance in the institutional review board review process, as well as Dr Bob Nolan, who provided guidance during the submission process.

Data Availability

All underlying data by zip code tabulation area can be found on GitHub (Multimedia Appendix 1) [12,13].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Python code.
[DOCX File, 14 KB - publichealth_v11i1e52972_app1.docx]

References

1. Thompson CN, Baumgartner J, Pichardo C, et al. COVID-19 outbreak - New York City, February 29-June 1, 2020. MMWR Morb Mortal Wkly Rep 2020 Nov 20;69(46):1725-1729. [doi: 10.15585/mmwr.mm6946a2] [Medline: 33211680]
2. Lieberman-Cribbin W, Tuminello S, Flores RM, Taioli E. Disparities in COVID-19 testing and positivity in New York City. Am J Prev Med 2020 Sep;59(3):326-332. [doi: 10.1016/j.amepre.2020.06.005] [Medline: 32703702]
3. 2020 vital statistics report reveals toll of COVID-19 on New Yorkers and life expectancy. City of New York. 2023 Apr 7. URL: https://www.nyc.gov/site/doh/about/press/pr2023/2020-vital-statistics-report-reveals-toll-of-covid-19.page [accessed 2025-06-13]
4. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020 May;20(5):533-534. [doi: 10.1016/S1473-3099(20)30120-1] [Medline: 32087114]
5. Liao TF, De Maio F. Association of social and economic inequality with coronavirus disease 2019 incidence and mortality across US counties. JAMA Netw Open 2021 Jan 4;4(1):e2034578. [doi: 10.1001/jamanetworkopen.2020.34578] [Medline: 33471120]
6. Parolin Z, Lee EK. The role of poverty and racial discrimination in exacerbating the health consequences of COVID-19. Lancet Reg Health Am 2022 Mar;7:100178. [doi: 10.1016/j.lana.2021.100178] [Medline: 35018358]
7. Bilal U, Tabb LP, Barber S, Diez Roux AV. Spatial inequities in COVID-19 testing, positivity, confirmed cases, and mortality in 3 U.S. cities : an ecological study. Ann Intern Med 2021 Jul;174(7):936-944. [doi: 10.7326/M20-3936] [Medline: 33780289]
8. About Taskforce on Racial Inclusion & Equity. City of New York. URL: https://www.nyc.gov/site/trie/about/about.page [accessed 2024-08-07]
9. S2502: demographic characteristics for occupied housing units. American Community Survey. 2022: ACS 5-year estimates subject tables. United States Census Bureau. 2022. URL: https://data.census.gov/table/ACSST5Y2022.S2502?q=S2502 [accessed 2025-03-02]
10. NYC Department of Health and Mental Hygiene. coronavirus-data/data-by-modzcta.csv at master · nychealth/coronavirus-data. GitHub. URL: https://github.com/nychealth/coronavirus-data/blob/master/totals/data-by-modzcta.csv [accessed 2022-09-29]
11. Seto E, Min E, Ingram C, Cummings BJ, Farquhar SA. Community-level factors associated with COVID-19 cases and testing equity in King County, Washington. Int J Environ Res Public Health 2020 Dec 18;17(24):9516. [doi: 10.3390/ijerph17249516] [Medline: 33353095]
12. testingequitydata / 2022-12-02 updated testing equity python code.ipynb. GitHub. URL: https://github.com/rosenfed/testingequitydata/blob/main/2022-12-02%20updated%20testing%20equity%20python%20code.ipynb [accessed 2025-02-27]
13. testingequitydata / 2022-10-12_acsimport.xlsx. GitHub. URL: https://github.com/rosenfed/testingequitydata/blob/main/2022-10-12_acsimport.xlsx [accessed 2025-02-27]

Abbreviations

BRANY: Biomedical Research Alliance of New York
H+H: Health + Hospitals Corporation
IRB: Institutional Review Board
NYC: New York City

POC: people of color
TRIE: Taskforce on Racial Inclusion & Equity
ZCTA: zip code tabulation area

Edited by A Mavragani; submitted 20.09.23; peer-reviewed by F Kaiksow, S Mathew; revised version received 08.09.24; accepted 10.09.24; published 13.03.25.

Please cite as:
Rosenfeld D, Brennan S, Wallach A, Long T, Keeley C, Kurien SJ
COVID-19 Testing Equity in New York City During the First 2 Years of the Pandemic: Demographic Analysis of Free Testing Data
JMIR Public Health Surveill 2025;11:e52972
URL: https://publichealth.jmir.org/2025/1/e52972
doi:10.2196/52972

Development and Validation of Prediction Models for Perceived and Unmet Mental Health Needs in the Canadian General Population: Model-Based Synthetic Estimation Study

Jianli Wang[1,2], PhD; Heather Orpana[3,4], PhD; André Carrington[5], PhD; George Kephart[1,6], PhD; Helen-Maria Vasiliadis[7], PhD; Benjamin Leikin[8], BSc

[1]Department of Community Health and Epidemiology, Faculty of Medicine, Dalhousie University, 5790 University Avenue, Halifax, NS, Canada
[2]Department of Psychiatry, Faculty of Medicine, Dalhousie University, Halifax, NS, Canada
[3]Centre for Surveillance and Applied Research, Public Health Agency of Canada, Ottawa, ON, Canada
[4]School of Psychology, University of Ottawa, Ottawa, ON, Canada
[5]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[6]School of Health Administration, Faculty of Health, Dalhousie University, Halifax, NS, Canada
[7]Department of Community Health Sciences, Faculty of Medicine and Health, Université de Sherbrooke, Sherbrooke, QC, Canada
[8]Community Health and Wellness Branch, Ottawa Public Health, Ottawa, ON, Canada

Corresponding Author:
Jianli Wang, PhD
Department of Community Health and Epidemiology, Faculty of Medicine, Dalhousie University, 5790 University Avenue, Halifax, NS, Canada

Abstract

Background: Research has shown that perceptions of a mental health need are closely associated with service demands and are an important dimension in needs assessment. Perceived and unmet mental health needs are important factors in the decision-making process regarding mental health services planning and resources allocation. However, few prediction tools are available to be used by policy and decision makers to forecast perceived and unmet mental health needs at the population level.

Objective: We aim to develop prediction models to forecast perceived and unmet mental health needs at the provincial and health regional levels in Canada.

Methods: Data from 2018, 2019, and 2020 Canadian Community Health Survey and Canadian Urban Environment were used (n=65,000 each year). Perceived and unmet mental health needs were measured by the Perceived Needs for Care Questionnaire. Using the 2018 dataset, we developed the prediction models through the application of regression synthetic estimation for the Atlantic, Central, and Western regions. The models were validated in the 2019 and 2020 datasets at the provincial level and in 10 randomly selected health regions by comparing the observed and predicted proportions of the outcomes.

Results: In 2018, a total of 17.82% of the participants reported perceived mental health need and 3.81% reported unmet mental health need. The proportions were similar in 2019 (18.04% and 3.91%) and in 2020 (18.1% and 3.92%). Sex, age, self-reported mental health, physician diagnosed mood and anxiety disorders, self-reported life stress and life satisfaction were the predictors in the 3 regional models. The individual based models had good discriminative power with C statistics over 0.83 and good calibration. Applying the synthetic models in 2019 and 2020 data, the models had the best performance in Ontario, Quebec, and British Columbia; the absolute differences between observed and predicted proportions were less than 1%. The absolute differences between the predicted and observed proportion of perceived mental health needs in Newfoundland and Labrador (−4.16% in 2020) and Prince Edward Island (4.58% in 2019) were larger than those in other provinces. When applying the models in the 10 selected health regions, the models calibrated well in the health regions in Ontario and in Quebec; the absolute differences in perceived mental health needs ranged from 0.23% to 2.34%.

Conclusions: Predicting perceived and unmet mental health at the population level is feasible. There are common factors that contribute to perceived and unmet mental health needs across regions, at different magnitudes, due to different population characteristics. Therefore, predicting perceived and unmet mental health needs should be region specific. The performance of the models at the provincial and health regional levels may be affected by population size.

(JMIR Public Health Surveill 2025;11:e66056) doi:10.2196/66056

KEYWORDS

population risk prediction; development; validation; perceived mental health need; unmet mental health need

Introduction

Mental disorders are prevalent and have a considerable impact on the people who experience them, as well as families, communities and society as a whole. To enhance the provision of timely and appropriate mental health services for people who are in need, effective health services planning should be adaptable and responsive to the changing needs and emerging opportunities. Research has shown that many people living with a mental disorder do not receive the services due to the barriers of availability, accessibility, and acceptability [1]. On the other hand, many people with symptoms of mental disorders that do not meet clinical criteria may also benefit from intervention and these people may actively seek treatment to prevent the symptoms from escalating to the level of a disorder [2]. These phenomena reflect different types of mental health care needs and the extent to which the mental health needs are being met.

Mental health need is difficult to define. Bradshaw [3] proposed six need types: (1) normative (presence of a mental disorder diagnosis), (2) felt (subjective perception of a mental health problem), (3) expressed (demand for mental health service), (4) comparative (population inequities in mental health), (5) medical (treatable disease), and (6) social (restoring quality of life). There is evidence to suggest that perceptions of a mental health need (felt need) are closely associated with service demands [4] and are an important dimension in needs assessment [5]. Data-driven evidence about perceived and unmet mental health needs at provincial, state and regional, or county levels can be very helpful in building the business case of demand [6]. However, population health data about perceived and unmet mental health needs have not been well adopted for this purpose [7].

Unmet health need is “the absence of sufficient or appropriate care and services.” [8] These needs are commonly assessed as being for information, medication, counseling or therapy, or another type of help. A need is considered to be fully met when a person receives help meeting all their expectations. In other cases, the help may only partly fulfill the need, or no support is provided at all and the need is unmet [9]. Based on 2012 Canadian Community Health Survey (CCHS)–mental health, 17.3% of Canadians aged 15 years and older reported having mental health needs in the past year, and about one-third of those with a mental health care need reported that it was unmet [9,10]. Addressing unmet needs is important, as people with untreated anxiety or depression are at higher risk of experiencing poor outcomes including persistence of symptoms, delayed recovery, poor personal and occupational functioning, and recurrence of these problems [11]. Unmet health care need has been identified to be an important indicator of health care access within various health care systems [12,13], and is a form of inefficiency that has economic implications, according to Mental Health Commission of Canada [14]. Both perceived and unmet mental health needs constitute the central part of demand which is a critical element in the process of health resources allocation.

Decision makers and health service planners have detailed information about existing mental health resources including available infrastructure, services and workforce at the provincial and regional levels. To allocate and use these resources efficiently, the magnitude of mental health needs and unmet needs in the community is critical information in the decision-making process. The ability to closely monitor and forecast the trends of perceived and unmet needs in populations can greatly facilitate this decision-making process. The objective of this study was to develop prediction models for estimating and forecasting perceived and unmet mental health needs at the provincial and health regional levels in Canada.

Methods

Study Design

For the objective of this study, data from the 2018, 2019, and 2020 CCHS and Canadian Urban Environment (CANUE) were used. CCHS is an annual national population health survey conducted by Statistics Canada to gather data on health status, health care use, and health determinants at the health region levels of geography. The CCHS covers the household population from 12 years of age and older living in the 10 provinces and 3 territories. The CCHS sample is selected using different frames according to the age group. For the adult population (18 y and older), the sample of households is selected from an area frame. For the youth population (12 to 17 y old) a list frame is used to select persons. The area frame used by the Labour Force Survey is used as a sampling frame for the adult population. The Labour Force Survey uses a 2-stage sample design. In the first stage, a sample of primary sampling units, corresponding to geographical regions called clusters, is selected. In each selected primary sampling unit, a sample of dwellings is drawn at the second stage. To sample persons for the youth population between the ages of 12 and 17 years, the CCHS uses a list frame created from the Canadian Child Benefit files. The sample for the youth population is selected from a list of individuals. For adult participants, 1 person is selected per household using varying probabilities taking into account the age and the household composition. Each year the CCHS collects information on approximately 65,000 Canadians (60,000 aged 18 y and older and 5000 aged 12 to 17 y) [15]. The CCHS uses 2 separate computer-assisted interviewing applications to collect data, 1 for telephone interviews (computer-assisted telephone interviewing) and 1 for personal interviews (computer-assisted personal interviewing). Approximately 25% of these completed cases were conducted in person using computer-assisted personal interviewing, and the other 75% were conducted over the phone using computer-assisted telephone interviewing. Starting in 2015, the CCHS undertook significant redesign in sampling and the content [15]. Data about the outcome variables of the proposed study (perceived and unmet mental heath needs) were not collected in 2015, 2016, and 2017 CCHS, therefore, we will analyze the 2018, 2019, and 2020 CCHS data.

CANUE is a Canadian Institutes of Health Research–funded collaboration that focuses on developing robust methods for producing measures of urban form that capture a wide range of characteristics for every postal code in Canada [16]. It has produced a unique repository of standardized metrics of urban, suburban, and rural characteristics. CANUE shares a range of data that can be directly linked with population health surveys through postal codes, including neighborhood-level material and social deprivation [17], marginalization [18], and walkability. CANUE is updated every 5 years. The most recent CANUE update was in 2016. The CANUE data can be obtained from CANUE consortium at no costs for research, and is linkable to the CCHS by postal codes.

Perceived and Unmet Mental Health Needs

In the CCHS, perceived and unmet mental health needs were measured by the Meadows et al [19] Perceived Needs for Care Questionnaire (PNCQ). The PNCQ was designed and field tested for the Australian National Survey for Mental Health and Wellbeing which was commissioned by the Federal Government of Australia [19]. The PNCQ assessed four types of help for problems with emotions, mental health, or the use of alcohol or drugs: (1) information about problems, treatments, or services; (2) medication; (3) counseling or therapy; and (4) other mental health services. Respondents were asked which types of help they had received in the past 12 months. For each type received, they were asked if they felt they had received enough. For each type of help not received, they were asked if they felt it was needed. Based on the PNCQ, the following binary outcome variables can be derived: [9,20]

- *Perceived mental health needs:* having reported "yes" to a perceived need of any type of help.
- *Unmet mental health needs:* For any type of help for which a respondent had a perceived need, having reported not receiving any or enough of that type of help. Interrater reliability of the PNCQ (κ value) was 0.62 [19]. Validity testing using homo-method agreements and hetero-method agreement demonstrated that the PNCQ has good discriminative validity [19]. The PNCQ has been used to assess perceived and unmet mental health needs in a number of population health surveys conducted in Australia [19], Canada [21], the United States [22], and the Netherlands [20].

Candidate Predictors

Overview

Guided by the Andersen Health Behavior model [23], we selected potential predictors from the core component of the CCHS which are consistently administered over time in all provinces and territories. This ensured that health regions can readily ascertain the predictor profile of their regions to make an estimation and prediction when PNCQ is not administered in the CCHS.

From the CCHS core component, we selected the predisposing factors (sex, gender, age, marital status, race or ethnicity, rural or urban residence, first official language spoken, immigration, household food security, self-reported life stress, work stress, life satisfaction, and sense of belonging), enabling factors (educational level, household income, employment status, and insurance coverage), and need factors (general health, comorbid physical conditions, self-reported mental health, mood, and anxiety, and primary care service use). Mental health problems such as depression, anxiety, and psychological distress are not part of the core component; data about these variables are not available in every CCHS and for every province and territory. Therefore, these variables were not selected. A number of studies have shown that health behaviors (smoking [24], problematic substance use [25,26], and physical inactivity [27,28]) are associated with depression, anxiety, and psychological distress. We selected these health behaviors from the core component as candidate predictors. From the CANUE, neighborhood-level social and material deprivation [17], marginalization [18], and walkability will be selected as enabling factors.

Model Development

The 2018 CCHS data were used to develop the prediction models. Given the vast geographic area of Canada, we developed the models by regions: Atlantic region (Newfoundland and Labrador, Prince Edward Island, Nova Scotia, and New Brunswick), Central region (Ontario and Quebec), Western region (Manitoba, Saskatchewan, Alberta, and British Columbia). For each region, we developed prediction models for perceived and unmet mental health needs through the application of regression synthetic estimation developed by the UK's Office of National Statistics [29], which is a type of small area estimation methodology. The same methodology has been used to estimate state and area-level prevalence of severe mental illness in the United States [30]. The regression synthetic estimation method involves several steps: (1) The development of a prediction model of the outcome at the individual level. (2) The coefficients derived from the foregoing model are used with a parallel set of predictors on the area level, using data obtained from CCHS, and coded using the same categories as used to estimate the individual-level model and to compute national and area-level estimates. (3) Internal and external validation.

To develop the predictive models for the outcomes at the health region level, multilevel, random intercept binary logistic regression models were used to analyze the perceived and unmet need for mental health services (level 1) nested within health regions (level 2). A backward selection method was used to identify the model with the best calibration and discrimination. The decisions of model selection were initially based on the changes in the values of Akaike information criterion and Bayesian information criterion [31].

Because the goal of models was to assist decision makers and service planners to estimate perceived and unmet needs at the population level (not to be used by clinicians to identify high risk individuals), we focused more on the calibration of the model, instead of discrimination, when it came to model performance. Discrimination is the ability of a prediction model to separate those who experienced the outcome events from those who did not, by predicting higher versus lower probabilities, respectively. We quantified this by calculating the area under the receiver operating characteristics curve, which is equal to the C statistic. Calibration measures how closely predicted outcomes agree with actual outcomes. For this, we used the calibration slope by which a value of 1 indicates perfect model calibration with data; less than 1 is overconfident and greater than 1 is underconfident [32]. Additionally, we used calibration plots to visually compare the mean predicted risk of the outcomes versus the observed risk (the cumulative fraction of events) by decile risk groups so that the overall calibration and the areas with over or under prediction can be identified.

The second step was synthetic estimation of the proportions at the population level with perceived and unmet mental health needs, which consisted of 2 stages. First, population proportions for predictor variables that were used in the initial modeling were estimated. For instance, if 5 household income categories are used in the initial modeling (<CAD $30,000, CAD $30,000 - $49,000, CAD $49,000 - $60,000, CAD $60,000 - $80,000, and CAD $80,000+; <US $21,175.80, US $21,175.80-$34,587.14, US $34,587.14-$42,351.60, US $42,351.60-56,468.80, and US $56,468.80+, respectively), the population proportion of participants in each of the same 5 income ranges were estimated from the CCHS. The regression coefficients were then applied to the corresponding proportions in the dataset, and to calculate the logit estimates for each province, which were then converted into probabilities, giving the predicted proportions of perceived and unmet mental health needs in the province. These procedures were repeated with data aggregated at the provincial and health regional levels. All analyses were conducted using bootstrap weights provided by Statistics Canada to account for the unequal sampling probabilities and design effects.

Validation

As internal validation, we used CCHS-2018 data to compare the predicted proportions (step 2 above) and the observed proportions. To ensure the validity of the predictive models, it is also important that the developed models are externally validated in related but independent populations. As external validation, we applied the developed synthetic models in CCHS-2019 and CCHS-2020, to examine the performance of the models, to investigate how the pandemic may affect the performance of the models by comparing the model predicted (synthetic) proportions and the observed proportions. Additionally, we randomly selected 10 subprovincial health regions from over 100 across the country. We applied the regional models directly to the health regions and compared the observed and predicted proportions of perceived and unmet mental health needs. All analyses were conducted using the bootstrap weights provided by Statistics Canada to account for the sampling and design effects of the CCHS.

Ethical Considerations

This is a secondary data analysis using the deidentified data collected by Statistics Canada. The data analysis was carried out at one of the Research Data Centers across the country, located on the campus of Dalhousie University. Researchers who conducted data analysis had to go through security clearances and become “deemed employees” of Statistics Canada. The results were vetted by data analysts of the Research Data Center to ensure privacy and confidentiality. CCHS participants gave informed consent to Statistics Canada for use of their data in accordance with the Statistics Act and no compensation was provided to the CCHS participants. For these reasons, ethics review was waived by the health research ethics committee of Dalhousie University.

Results

Overview

The 2018 CCHS data were used to develop the prediction models. The demographic and socioeconomic characteristics of the participants in the 2018 CCHS are presented in Table 1. In the 2018 CCHS, 17.82% reported perceived mental health need and 3.81% reported unmet mental health need. In 2019, the proportion of perceived and unmet mental health needs was 18.04% and 3.91%, respectively; in 2020, the proportion of perceived and unmet mental health needs was 18.1% and 3.92%, respectively.

Table . The demographic and socioeconomic characteristics of the participants in the 2018 cross-sectional Canadian Community Health Survey.

Variables	Values
Men, weighted %	49.4
Women, weighted %	50.6
Age (years), mean (SD)	45.7 (0.14)
Married, common law, or partner, weighted %	58.2
Single, weighted %	30.1
Divorced, separated, or widowed, weighted %	11.7
<High school, weighted %	17.4
High school, weighted %	22.6
≥College or university, weighted %	60
Employed, weighted %	67.2
Unemployed, weighted %	22.4
<15 or >75 years old[a], weighted %	10.4
Immigrants, weighted %	27.1
Nonimmigrants, weighted %	72.9
White, weighted %	75.5
Non-White, weighted %	24.5
Rural residence, weighted %	17.2
Urban residence, weighted %	82.8
Have insurance for medication, weighted %	19.7
No, weighted %	80.3

[a] Participants aged <15 or 75+ years were not eligible for the employment question.

Predicting Perceived Mental Health Need

The individual based models with the best performance for predicting perceived mental health need are in supplemental tables (Tables S1 [Atlantic region], S2 [Central region], and S3 [Western region] in Multimedia Appendix 1). Sex, age, self-reported mental health, physician diagnosed mood and anxiety disorders, and self-reported life stress and life satisfaction were the predictors in the 3 regional models. Perceived work stress and household food insecurity were in the models for the Atlantic and Central regions, but not for the Western region. However, immigrants, smoking, or problematic drinking and material deprivation were predictors in the models for the Atlantic and the Western region, not for the Central region. The individual based models had good discriminative power with C statistics over 0.83 (Tables S4-S6 in Multimedia Appendix 1) and good calibration by comparing observed and predicted proportions.

Converting the individual based models into synthetic estimation models, the predicted proportion of perceived mental health need in each province was estimated. The absolute difference between the observed and predicted proportions in 2018 was less than 1% for all provinces (Tables 2-4), except Saskatchewan (3.37%). This is not surprising as the models were developed using the 2018 CCHS data. To validate the models, we applied the synthetic models directly in 2019 and 2020 data. The extent to which the predicted or synthetic estimates agree with observed proportions varied by regions (Tables 2-4). The models had the best performance in Ontario and British Columbia (absolute difference <1% in both 2019 and 2020), followed by Quebec (absolute difference=1.54% in 2019 and 0.54% in 2020). The models for Nova Scotia, New Brunswick, Saskatchewan, and Alberta had 1 absolute difference being less than 1% and the other greater than 2%. The absolute differences between observed and predicted proportions in Newfoundland and Labrador (−4.16% for perceived needs in 2020) and Prince Edward Island (4.58% for perceived needs in 2019) were larger than those in other provinces. When applying the models in the 10 selected health regions, the data showed that the model performed well in health regions in Ontario (first 2 digits with 35) and in Quebec (first 2 digits with 24), supported by the relatively small absolute differences between observed and predicted proportions (Table 5). Health region 1304 had the largest absolute differences in 2019 and 2020.

Table . The observed and predicted proportions of perceived and unmet mental health needs in the Atlantic provinces in the 2018, 2019, and 2020 Canadian Community Health Survey.

		Observed proportion (%)	Predicted proportion (%)	Absolute difference[a] (%)
Newfoundland and Labrador				
Perceived need				
	2018	14.9	14.57	−0.33
	2019	16.24	14.04	−2.2
	2020	19.35	15.19	−4.16
Unmet need				
	2018	2.43	2.07	−0.36
	2019	2.93	1.94	−0.99
	2020	2.43	1.96	−0.47
Prince Edward Island				
Perceived need				
	2018	14.37	14.23	−0.14
	2019	20.92	16.34	−4.58
	2020	16.89	14.93	−1.96
Unmet need				
	2018	3.63	3.96	0.33
	2019	2.74	4.15	1.41
	2020	2.6	3.85	1.25
Nova Scotia				
Perceived need				
	2018	19.81	19.15	−0.66
	2019	23.58	20.44	−3.14
	2020	20.56	21.33	0.77
Unmet need				
	2018	3.43	3.52	0.09
	2019	5.78	3.55	−2.23
	2020	4.23	3.3	−0.93
New Brunswick				
Perceived need				
	2018	17.66	18.44	0.78
	2019	17.22	19.97	2.75
	2020	20	20.83	0.83
Unmet need				
	2018	3.05	2.85	−0.2
	2019	3.85	3.52	−0.33
	2020	3.71	3.42	−0.29

[a] Absolute difference=predicted proportion–observed proportion.

Table . The observed and predicted proportions of perceived and unmet mental health needs in Ontario and Quebec in the 2018, 2019, and 2020 Canadian Community Health Survey.

			Observed proportion (%)	Predicted proportion (%)	Absolute difference[a] (%)
Ontario					
	Perceived need				
		2018	16.93	16.63	−0.30
		2019	17.35	16.94	−0.41
		2020	17.97	17.44	−0.53
	Unmet need				
		2018	3.96	3.98	0.02
		2019	4.08	4.09	0.01
		2020	4	4.29	0.29
Quebec					
	Perceived need				
		2018	17.08	17.65	−0.43
		2019	16.02	17.56	1.54
		2020	16.72	17.26	0.54
	Unmet need				
		2018	3.25	3.26	0.01
		2019	3.01	3.21	0.2
		2020	3.36	3.09	−0.27

[a] Absolute difference=predicted proportion−observed proportion.

Table . The observed and predicted proportions of perceived and unmet mental health needs in the Western provinces in the 2018, 2019, and 2020 Canadian Community Health Survey.

			Observed proportion (%)	Predicted proportion (%)	Absolute difference[a] (%)
Manitoba					
	Perceived need				
		2018	18.8	19.52	0.72
		2019	18.86	19.56	0.7
		2020	16.87	20.13	3.26
	Unmet need				
		2018	3.59	4.18	0.59
		2019	5.11	4.12	−0.99
		2020	4.25	4.43	0.18
Saskatchewan					
	Perceived need				
		2018	16.57	19.94	3.37
		2019	18.61	21.14	2.53
		2020	15.98	21.31	5.33
	Unmet need				
		2018	3.1	4.03	0.93
		2019	2.84	4.23	1.39
		2020	4.07	4.16	0.09
Alberta					
	Perceived need				
		2018	21.73	20.77	−0.96
		2019	20.02	20.91	0.89
		2020	18.92	21.49	2.57
	Unmet need				
		2018	4.52	4.41	−0.11
		2019	3.66	4.43	0.77
		2020	3.72	4.54	0.82
British Columbia					
	Perceived need				
		2018	18.39	18.38	−0.01
		2019	19.14	18.86	−0.28
		2020	20.05	20.01	−0.04
	Unmet need				
		2018	4.28	4.16	−0.12
		2019	4.9	N/A[b]	N/A
		2020	4.87	4.51	−0.36

[a] Absolute difference=predicted proportion–observed proportion.

[b] N/A: household food insecurity data were not collected in British Columbia in 2019.

Table . The observed and predicted proportions of perceived mental health needs by selected health regions in 2018, 2019, and 2020 Canadian Community Health Survey.

Health regions and years		Observed proportion (%)	Predicted proportion (%)	Absolute difference[a] (%)
1301				
	2018	17.42	19.08	1.66
	2019	19.81	22.47	2.66
	2020	22.89	22.48	−0.41
1304				
	2018	11.86	16.12	4.26
	2019	6.59	15.23	8.64
	2020	9.45	19.61	10.16
2404				
	2018	14.22	16.04	1.82
	2019	14.01	15.83	1.82
	2020	10.78	14.72	3.94
3544				
	2018	20.12	19.59	−0.53
	2019	20.4	18.7	−1.7
	2020	20.13	20.36	0.23
3570				
	2018	12.57	14.91	2.34
	2019	16.29	15.36	−0.93
	2020	17.97	15.87	−2.1
3536				
	2018	14.99	15.9	0.91
	2019	17.4	16.16	−1.24
	2020	17.21	16.38	−0.83
5913				
	2018	15.17	18.28	3.11
	2019	N/A[b]	N/A	N/A
	2020	24.29	18.8	−5.49
5921				
	2018	13.75	18.66	4.91
	2019	N/A	N/A	N/A
	2020	21.24	20.54	−0.7
5922				
	2018	15.8	18.44	2.64
	2019	N/A	N/A	N/A
	2020	19.77	19.46	−0.31
5943				
	2018	17.88	17.71	−0.17
	2019	N/A	N/A	N/A
	2020	23.37	17.84	−5.53

[a] Absolute difference=predicted proportion–observed proportion.

[b] N/A: estimates were not available due to lack of data on household food insecurity.

Predicting Unmet Mental Health Need

The individual based models with the best performance for predicting unmet mental health need are in Tables S7 (Atlantic region), S8 (Central region), and S9 (Western region) in Multimedia Appendix 1. As seen from the tables, sex, age, marital status, self-reported mental health, self-reported life stress, and low sense of belonging to the community were the common predictors in the models for the 3 regions. Household income was a predictor only in the model for Central region; professional diagnosed mood disorders, problematic drinking, and smoking and social deprivation were the factors specific for the model for the Atlantic region. The individual based models had good discriminative power with C statistics over 0.77 (Tables S10-S12 in Multimedia Appendix 1) and good calibration by comparing observed and predicted proportions.

The individual based models were converted into synthetic estimation models, and the predicted proportion of perceived mental health need in each province was estimated (Tables 2-4). The data demonstrated that the models for unmet mental health need performed well in most of the provinces with the absolute difference between observed and predicted proportions being less than 1%, except for Prince Edward Island, Nova Scotia, and Saskatchewan. When applying the models in the 10 selected health regions, the data showed that the models calibrated well with data in all selected health regions with absolute difference between observed and predicted proportions being less than 1% (Table S13 in Multimedia Appendix 1).

Discussion

Principal Results

The 3 waves of CCHS showed that the proportion of the Canadian population that perceived the need for mental health care was between 17.82% and 18.1%, and about 4% reported unmet mental health need. This study demonstrated the feasibility of integrating individual and community level data to build informative synthetic models for perceived and unmet mental health care needs at the population level. The models performed well in predicting the outcomes at both the provincial and health regional levels, particularly in populous provinces such as Ontario, Quebec, and British Columbia. The absolute differences between the observed and predicted proportions of perceived and unmet mental health needs in Ontario, Quebec, and British Columbia were less than 1%. For the rest of the provinces, the results showed that the models at provincial and health regional levels predicting unmet mental health needs had better calibration than the models predicting perceived mental health need.

Limitations

This study had several limitations. First, although the predictors in the models were associated with the outcomes, causal inferences cannot be made. The goal of the prediction models is to identify a key set of factors that in combination are best predictive of the outcomes. The models are not to test hypothesis or make inferences about etiology. Second, the relationships between the selected factors and the outcomes are complex. The logistic regression model is a linear function. Although we found no evidence of interactions among the selected predictors, nonlinear relationships between some predictors and the outcomes are still possible. Future studies may test if models using machine learning techniques have better performance. Third, the selection of candidate predictors was limited by what CCHS collected. It is possible that other factors are associated with the outcomes of this study, but were not measured in CCHS, therefore could not be examined in this study. Finally, the models appeared to perform well at regional and provincial levels. Relatively large absolute differences between predicted and observed proportions of perceived and unmet needs were observed in some health regions (eg, health region 1304). Caution should be exercised for using them at the health region level.

Comparison With Prior Work

The CCHS used a nationwide representative sample of the Canadian household population. However, we found that a model developed for the whole population performed well only in the most populous provinces such as Ontario and Quebec, but not in the rest of the provinces. Developing models for specific regions appeared to be a better approach. The regional models have some common predictors, including sex, age, self-reported mental health, physician diagnosed mood and anxiety disorders, self-reported life stress and life satisfaction for perceived mental health need, and sex, age, marital status, self-reported mental health, self-reported life stress, and low sense of belonging to the community for unmet mental health need. These factors have been found to be significantly associated with mental health problems [33] in the literature. On the other hand, the coefficients associated with these factors varied by regional models and these models also contained predictors that are regionally specific. This finding indicates that there are common factors in the mechanisms underlying perceived and unmet mental health needs across geographic regions and populations. The extent to which these factors influence perceived and unmet mental health needs may differ due to the population characteristics of these provinces and regions. The differences in population characteristics and the distribution of the outcomes across the regions may also contribute to the finding that the models contained region specific predictors.

We have not found studies on predicting perceived and unmet mental health needs at the population level. Therefore, a direct comparison with previous studies is not possible. There are several studies predicting county or state level suicide and severe mental illness. For example, Kandula et al [34] modelled county-level suicide risk in the United States using county-level predictors derived from 8 different databases of different sources (government programs, health surveys, or private organizations). For some predictors such as the prevalence of major depressive episode, only state-level estimates were available and these estimates were extrapolated to the counties [34]. Hudson [30] explored the utility of regression synthetic estimation model that incorporated individual data from the National Comorbidity

Survey, census and hospital administrative data to predict state-level prevalence of severe mental illness. The advantages of these population risk prediction models are the use of community level predictors from existing sources or published research and the ability of adapting the models in local context. Notably, our study used the regression synthetic estimation modeling approach. We used CCHS data. The use of a single data source may improve the efficiency of data analysis, data access, and eventually of the decision-making process.

One critical element of building risk prediction models is assessing model performance and model validation. This is to ensure that the developed model is accurate and has good performance in different populations or at different time periods. In this study, we developed the models using the 2018 CCHS data, and validated the models using the data from 2019 and 2020 CCHS. Furthermore, we validated the models in a random sample of health regions. These models were designed to estimate population proportions and to identify regions with high levels of mental health need and unmet mental health need; these models were not to be used by clinicians to identify high-risk individuals. Therefore, the focus of model performance assessment can be different. Kandula et al [34] used symmetric proportional error (observed deaths–predicted deaths)/(observed deaths+predicted deaths) to quantify model calibration. Hudson [30] calculated the absolute difference between the predicted and observed prevalence of severe mental illness. In this study, we calculated the absolute difference between the observed and model predicted proportions of perceived and unmet mental health needs. The absolute difference indicates the extent to which the model prediction deviates from the observed value. However, there is no consensus about what an optimal threshold of absolute difference should be. There may be other indicators that are useful for assessing the performance of population level models. Consultations with policy and decision makers (ie, end users of the models) would be helpful to understand what indicators are informative about model performance, and the level of the model error that is acceptable.

The results of this study are expected to have implications for population mental health planning. Few would deny that resources allocation should be partly driven by needs, and needs assessments typically require the knowledge of potential changes in prevalence estimates and in local population profiles, for example, their demographics, diagnoses, and mental health services use. The prediction models developed by this study will allow decision makers and mental health services planners to forecast the proportions of perceived and unmet mental health needs in the years to come at the provincial (state) and health regional (county) levels based on the potential changes in local population profiles. Such profile changes may be estimated using health administrative data and national population census data. Additionally, region-specific estimates can help categorize health regions—for example, regions with relatively stable mental health needs especially those that remained in the highest or lowest groups, or regions in which the largest year-to-year changes are observed—and hence help identify areas with a greater need of preventive resources, or conversely identify areas where interventions seem to be effective.

Conclusion

In conclusion, accurate forecast of perceived and unmet mental health needs in the population can allow policy and decision makers and mental health services planners to categorize regions or communities that are at high need and to monitor changes so that they may mobilize resources and interventions to the right populations and the right places at the right time. Regularly collected population health data such as those from the CCHS are readily accessible to policy and decision makers and mental health services planners. The models are particularly useful for service planners at the health regional level because population health surveys usually do not contain sufficient numbers of participants at the health regional level. Future studies are needed to identify methods to improve prediction for regions with small numbers of residents.

Acknowledgments

This study was supported by a grant from the Canadian Institutes of Health Research.

Data Availability

The Canadian Community Health Survey (CCHS) data were collected and are maintained by Statistics Canada. The deidentified raw data can only be accessed at a Research Data Center after becoming a “deemed employee” of Statistics Canada. However, Statistics Canada has prepared carefully modified Public Use Microdata Files (PUMFs) of the CCHS which ensure that individuals cannot be identified, and which are in the public domain.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File, 35 KB - publichealth_v11i1e66056_app1.docx]

References

1. Wang J. Perceived barriers to mental health service use among individuals with mental disorders in the Canadian general population. Med Care 2006 Feb;44(2):192-195. [doi: 10.1097/01.mlr.0000196954.67658.95] [Medline: 16434920]
2. Kessler RC, Zhao S, Blazer DG, Swartz M. Prevalence, correlates, and course of minor depression and major depression in the National Comorbidity Survey. J Affect Disord 1997 Aug;45(1-2):19-30. [doi: 10.1016/s0165-0327(97)00056-6] [Medline: 9268772]
3. Bradshaw J. The conceptualization and measurement of need: a social policy perspective. In: Popay J, Williams G, editors. Researching the People's Health: Routledge; 2005:57-70 URL: https://www.taylorfrancis.com/chapters/edit/10.4324/9780203996782-12/conceptualization-measurement-need-social-policy-perspective [accessed 2025-02-05]
4. Wichstrøm L, Belsky J, Jozefiak T, Sourander A, Berg-Nielsen TS. Predicting service use for mental health problems among young children. Pediatrics 2014 Jun;133(6):1054-1060. [doi: 10.1542/peds.2013-3184] [Medline: 24819574]
5. Stevens A, Gabbay J. Needs assessment needs assessment. Health Trends 1991;23(1):20-23. [Medline: 10113881]
6. Roos NP, Fransoo R, Carrière KC, et al. Needs-based planning: the case of Manitoba. CMAJ 1997 Nov 1;157(9):1215-1221. [Medline: 9361640]
7. Vigo D, Jones W, Dove N, et al. Estimating the prevalence of mental and substance use disorders: a systematic approach to triangulating available data to inform health systems planning. Can J Psychiatry 2022 Feb;67(2):107-116. [doi: 10.1177/07067437211006872] [Medline: 33827278]
8. Carr W, Wolfe S. Unmet needs as sociomedical indicators. Int J Health Serv 1976;6(3):417-430. [doi: 10.2190/MCG0-UH8D-0AG8-VFNU] [Medline: 955751]
9. Sunderland A, Findlay LC. Perceived need for mental health care in Canada: results from the 2012 Canadian Community Health Survey-Mental Health. Health Rep 2013 Sep;24(9):3-9. [Medline: 24258361]
10. Urbanoski K, Inglis D, Veldhuizen S. Service use and unmet needs for substance use and mental disorders in Canada. Can J Psychiatry 2017 Aug;62(8):551-559. [doi: 10.1177/0706743717714467] [Medline: 28616935]
11. Wang J. A longitudinal population-based study of treated and untreated major depression. Med Care 2004 Jun;42(6):543-550. [doi: 10.1097/01.mlr.0000128001.73998.5c] [Medline: 15167322]
12. Lasser KE, Himmelstein DU, Woolhandler S. Access to care, health status, and health disparities in the United States and Canada: results of a cross-national population-based survey. Am J Public Health 2006 Jul;96(7):1300-1307. [doi: 10.2105/AJPH.2004.059402] [Medline: 16735628]
13. Chen J, Hou F. Unmet needs for health care. Health Rep 2002;13(2):23-34. [Medline: 12743954]
14. Strengthening the case for investing in Canada's mental health system: economic considerations. Mental Health Commission of Canada. 2017. URL: https://www.mentalhealthcommission.ca/wp-content/uploads/drupal/2017-03/case_for_investment_eng.pdf [accessed 2025-02-07]
15. Canadian community health survey (CCHS) annual component user guide: 2018 and 2017-2018 microdata file. Statistics Canada. 2019. URL: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3226
16. CANUE Data Portal. URL: https://www.canuedata.ca/ [accessed 2025-02-05]
17. Pampalon R, Hamel D, Gamache P, Raymond G. A deprivation index for health planning in Canada. Chronic Dis Can 2009;29(4):178-191. [Medline: 19804682]
18. Matheson FI, Dunn JR, Smith KLW, Moineddin R, Glazier RH. Development of the Canadian Marginalization Index: a new tool for the study of inequality. Can J Public Health 2012 Apr 30;103(8 Suppl 2):S12-S16. [doi: 10.1007/BF03403823] [Medline: 23618065]
19. Meadows G, Harvey C, Fossey E, Burgess P. Assessing perceived need for mental health care in a community survey: development of the Perceived Need for Care Questionnaire (PNCQ). Soc Psychiatry Psychiatr Epidemiol 2000 Sep;35(9):427-435. [doi: 10.1007/s001270050260] [Medline: 11089671]
20. Prins M, Meadows G, Bobevski I, et al. Perceived need for mental health care and barriers to care in the Netherlands and Australia. Soc Psychiatry Psychiatr Epidemiol 2011 Oct;46(10):1033-1044. [doi: 10.1007/s00127-010-0266-3] [Medline: 20686887]
21. Thomas S, Wannell B. Combining cycles of the Canadian Community Health Survey. Health Rep 2009 Mar;20(1):53-58. [Medline: 19388369]
22. Nam B, Wilcox HC, Hilimire M, DeVylder JE. Perceived need for care and mental health service utilization among college students with suicidal ideation. J Am Coll Health 2018;66(8):713-719. [doi: 10.1080/07448481.2018.1434779] [Medline: 29384469]
23. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav 1995 Mar;36(1):1-10. [Medline: 7738325]
24. Fluharty M, Taylor AE, Grabski M, Munafò MR. The association of cigarette smoking with depression and anxiety: a systematic review. Nicotine Tob Res 2017 Jan;19(1):3-13. [doi: 10.1093/ntr/ntw140] [Medline: 27199385]
25. Escelsior A, Belvederi Murri M, Corsini GP, et al. Cannabinoid use and self-injurious behaviours: a systematic review and meta-analysis. J Affect Disord 2021 Jan 1;278:85-98. [doi: 10.1016/j.jad.2020.09.020] [Medline: 32956965]
26. Wang J, Patten SB. Prospective study of frequent heavy alcohol use and the risk of major depression in the Canadian general population. Depress Anxiety 2002;15(1):42-45. [doi: 10.1002/da.1084] [Medline: 11816052]

27. Bailey AP, Hetrick SE, Rosenbaum S, Purcell R, Parker AG. Treating depression with physical activity in adolescents and young adults: a systematic review and meta-analysis of randomised controlled trials. Psychol Med 2018 May;48(7):1068-1083. [doi: 10.1017/S0033291717002653] [Medline: 28994355]
28. Mammen G, Faulkner G. Physical activity and the prevention of depression: a systematic review of prospective studies. Am J Prev Med 2013 Nov;45(5):649-657. [doi: 10.1016/j.amepre.2013.08.001] [Medline: 24139780]
29. Heady P, Clarke P, Brown G, et al. Model-based small area estimation series no. 2. small area estimation project report. : Office for National Statistics; 2003.
30. Hudson CG. Validation of a model for estimating state and local prevalence of serious mental illness. Int J Methods Psychiatr Res 2009 Dec;18(4):251-264. [doi: 10.1002/mpr.294] [Medline: 20014156]
31. Vrieze SI. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychol Methods 2012 Jun;17(2):228-243. [doi: 10.1037/a0027127] [Medline: 22309957]
32. Hilden J, Habbema JD, Bjerregaard B. The measurement of performance in probabilistic diagnosis. II. Trustworthiness of the exact values of the diagnostic probabilities. Methods Inf Med 1978 Oct;17(4):227-237. [Medline: 366335]
33. Arango C, Dragioti E, Solmi M, et al. Risk and protective factors for mental disorders beyond genetics: an evidence-based atlas. World Psychiatry 2021 Oct;20(3):417-436. [doi: 10.1002/wps.20894] [Medline: 34505386]
34. Kandula S, Martinez-Alés G, Rutherford C, et al. County-level estimates of suicide mortality in the USA: a modelling study. Lancet Public Health 2023 Mar;8(3):e184-e193. [doi: 10.1016/S2468-2667(22)00290-0] [Medline: 36702142]

Abbreviations

CANUE: Canadian Urban Environment
CCHS: Canadian Community Health Survey
PNCQ: Perceived Needs for Care Questionnaire

Edited by A Mavragani; submitted 02.09.24; peer-reviewed by LK Chua, W Wang; revised version received 14.01.25; accepted 22.01.25; published 19.02.25.

Please cite as:
Wang J, Orpana H, Carrington A, Kephart G, Vasiliadis HM, Leikin B
Development and Validation of Prediction Models for Perceived and Unmet Mental Health Needs in the Canadian General Population: Model-Based Synthetic Estimation Study
JMIR Public Health Surveill 2025;11:e66056
URL: https://publichealth.jmir.org/2025/1/e66056
doi:10.2196/66056

Identifying Data-Driven Clinical Subgroups for Cervical Cancer Prevention With Machine Learning: Population-Based, External, and Diagnostic Validation Study

Zhen Lu[1], MSc; Binhua Dong[2,3], MSc; Hongning Cai[4], MSc; Tian Tian[5], PhD; Junfeng Wang[6], PhD; Leiwen Fu[7], PhD; Bingyi Wang[8,9], PhD; Weijie Zhang[1], BD; Shaomei Lin[10], BD; Xunyuan Tuo[11], MSc; Juntao Wang[12], MSc; Tianjie Yang[13], MSc; Xinxin Huang[14], MSc; Zheng Zheng[13], MSc; Huifeng Xue[15], MSc; Shuxia Xu[16], BD; Siyang Liu[1], PhD; Pengming Sun[2,3,17], MD, PhD; Huachun Zou[18,19,20], PhD

[1]School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China

[10]Department of Gynecology, Shunde Women's and Children's Hospital of Guangdong Medical University, Foshan, China

[11]Department of Gynecology, Gansu Provincial Maternity and Child-care Hospital, Lanzhou, China

[12]Department of Gynecology, Guiyang Maternal and Child Health Care Hospital, Guiyang, China

[13]Department of Gynecology, Shenzhen Maternity & Child Healthcare Hospital, Shenzhen, China

[14]The Ministry of Health, Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

[15]Center for Cervical Disease Diagnosis and Treatment, Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

[16]Department of Pathology, Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

[17]School of Group Medicine and Public Health, Peking Union Medical College, Beijing, China

[18]School of Public Health, Fudan University, Shanghai, China

[19]Shenzhen Campus, Sun Yat-sen University, Shenzhen, China

[2]Department of Gynecology, Laboratory of Gynecologic Oncology, Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

[20]Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics and Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

[3]Fujian Key Laboratory of Women and Children’s Critical Diseases Research, Fuzhou, China

[4]Department of Gynecology, Maternal and Child Health Hospital of Hubei Province (Women and Children's Hospital of Hubei Province) Wuhan, Wuhan, China

[5]School of Public Health, Xinjiang Medical University, Urumqi, China

[6]Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, Netherlands

[7]Beijing Chest Hospital, Capital Medical University, Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing, China

[8]Institute for HIV/AIDS Control and Prevention, Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China

[9]Department of HIV/AIDS Control and Prevention, Guangdong Provincial Academy of Preventive Medicine, Guangzhou, China

Corresponding Author:
Huachun Zou, PhD
School of Public Health, Fudan University, Shanghai, China

Abstract

Background: Cervical cancer remains a major global health issue. Personalized, data-driven cervical cancer prevention (CCP) strategies tailored to phenotypic profiles may improve prevention and reduce disease burden.

Objective: This study aimed to identify subgroups with differential cervical precancer or cancer risks using machine learning, validate subgroup predictions across datasets, and propose a computational phenomapping strategy to enhance global CCP efforts.

Methods: We explored the data-driven CCP subgroups by applying unsupervised machine learning to a deeply phenotyped, population-based discovery cohort. We extracted CCP-specific risks of cervical intraepithelial neoplasia (CIN) and cervical cancer through weighted logistic regression analyses providing odds ratio (OR) estimates and 95% CIs. We trained a supervised machine learning model and developed pathways to classify individuals before evaluating its diagnostic validity and usability on an external cohort.

Results: This study included 551,934 women (median age, 49 years) in the discovery cohort and 47,130 women (median age, 37 years) in the external cohort. Phenotyping identified 5 CCP subgroups, with CCP4 showing the highest carcinoma prevalence. CCP2–4 had significantly higher risks of CIN2+ (CCP2: OR 2.07 [95% CI: 2.03 - 2.12], CCP3: 3.88 [3.78 - 3.97], and CCP4: 4.47 [4.33 - 4.63]) and CIN3+ (CCP2: 2.10 [2.05 - 2.14], CCP3: 3.92 [3.82 - 4.02], and CCP4: 4.45 [4.31 - 4.61]) compared to CCP1 (P<.001), consistent with the direction of results observed in the external cohort. The proposed triple strategy was validated as clinically relevant, prioritizing high-risk subgroups (CCP3-4) for colposcopies and scaling human papillomavirus screening for CCP1-2.

Conclusions: This study underscores the potential of leveraging machine learning algorithms and large-scale routine electronic health records to enhance CCP strategies. By identifying key determinants of CIN2+/CIN3+ risk and classifying 5 distinct subgroups, our study provides a robust, data-driven foundation for the proposed triple strategy. This approach prioritizes tailored prevention efforts for subgroups with varying risks, offering a novel and scalable tool to complement existing cervical cancer screening guidelines. Future work should focus on independent external and prospective validation to maximize the global impact of this strategy.

(JMIR Public Health Surveill 2025;11:e67840) doi:10.2196/67840

KEYWORDS

cervical cancer; human papillomavirus; screening; machine learning; cervical tumor; cancer; carcinoma; tumor; malignant; ML; phenomapping strategy; logistic regression; regression; population-based; validation study; cancer prevention; validity; usability; algorithm; surveillance; electronic health record; EHR

Introduction

Cervical cancer is the fourth most common cancer among women, with an estimated 660,000 new cases and 350,000 deaths globally in 2022 [1,2]. It is the most common cancer in 25 countries and the leading cause of cancer death in 37 countries. Despite being largely preventable through human papillomavirus (HPV) vaccination [3], the high incidence and persistence of high-risk HPV (hrHPV) remain the primary risk factors for cervical cancer [4] and related diseases [5-7]. In November 2020, the World Health Organization (WHO) launched a global initiative [8] to eliminate cervical cancer as a public health problem, emphasizing a triple intervention strategy: vaccinating at least 90% of girls against HPV by age 15 years, screening 70% of women with a high-performance test by ages 35 and 45 years, and treating at least 90% of detected precancerous lesions and invasive cancers. Yet, globally, an estimated 1.6 billion (67%) of 2.3 billion women aged 20 - 70 years have never been screened for cervical cancer [9], and in China, 5-year screening coverage among women aged 35 - 49 years was only 33%. Establishing high-quality, sustainable, and acceptable cervical cancer prevention (CCP) with broad coverage in resource-limited regions remains a critical challenge [10-12]. While HPV-based screening has demonstrated significant benefits [13-18], leading the WHO to recommend it as the primary method for CCP globally [19], its adoption in low- and middle-income countries (LMICs) is limited due to health inequities, resource constraints, and limited access to affordable, clinically validated HPV tests [9,20]. Addressing these barriers requires personalized CCP strategies.

We hypothesize that data-driven CCP subgroups, identified using machine learning, can provide valuable insights into CCP, enabling personalized strategies to reduce the risk of cervical intraepithelial neoplasia (CIN) and cervical cancer. Personalized CCP strategies depend on the unique phenotypic profile of each individual. The population, with its varying phenotypic diversity, comprises heterogeneous subgroups that reflect multiple underlying behavior patterns and causes of disease. Clinical practice guidelines also recommend that cervical cancer screening should be tailored to an individual's risk profile of HPV infection [21,22]. Therefore, the discovery of CCP subgroups could inform and improve the development of new screening strategies, public health policies, and clinical decision-making, as well as contribute to trial design, despite the complex causal relationships among individual risk factors [4]. To date, no study has identified CCP subgroups. The incomplete understanding of these subgroups across large, diverse populations, coupled with insufficient validation through various methods, has hindered the implementation of the triple intervention strategy. Machine learning-based approaches allow for the phenotyping of entire populations based on individual characteristics, enabling the prediction of phenotypic clusters and disease onset [23,24]. Additionally, by defining a computational phenomap—a mathematical construct of individual phenotypes based on baseline measures [24]—computational methods can assess the heterogeneous risk effects of distinct subpopulations. This approach accounts for the phenotypic diversity within populations and their subsequent histopathological diagnoses.

As such, in this study, we aimed to: (1) test the hypothesis that distinct CCP subgroups exhibit differential risks of cervical precancer or cancer based on their complex phenotypic profiles (development); (2) demonstrate internal validity (within a dataset and across methods), as well as external and diagnostic validity (across datasets; validation); and (3) propose a computational phenomapping strategy with clinical relevance and pathways to improve global access to CCP (impact).

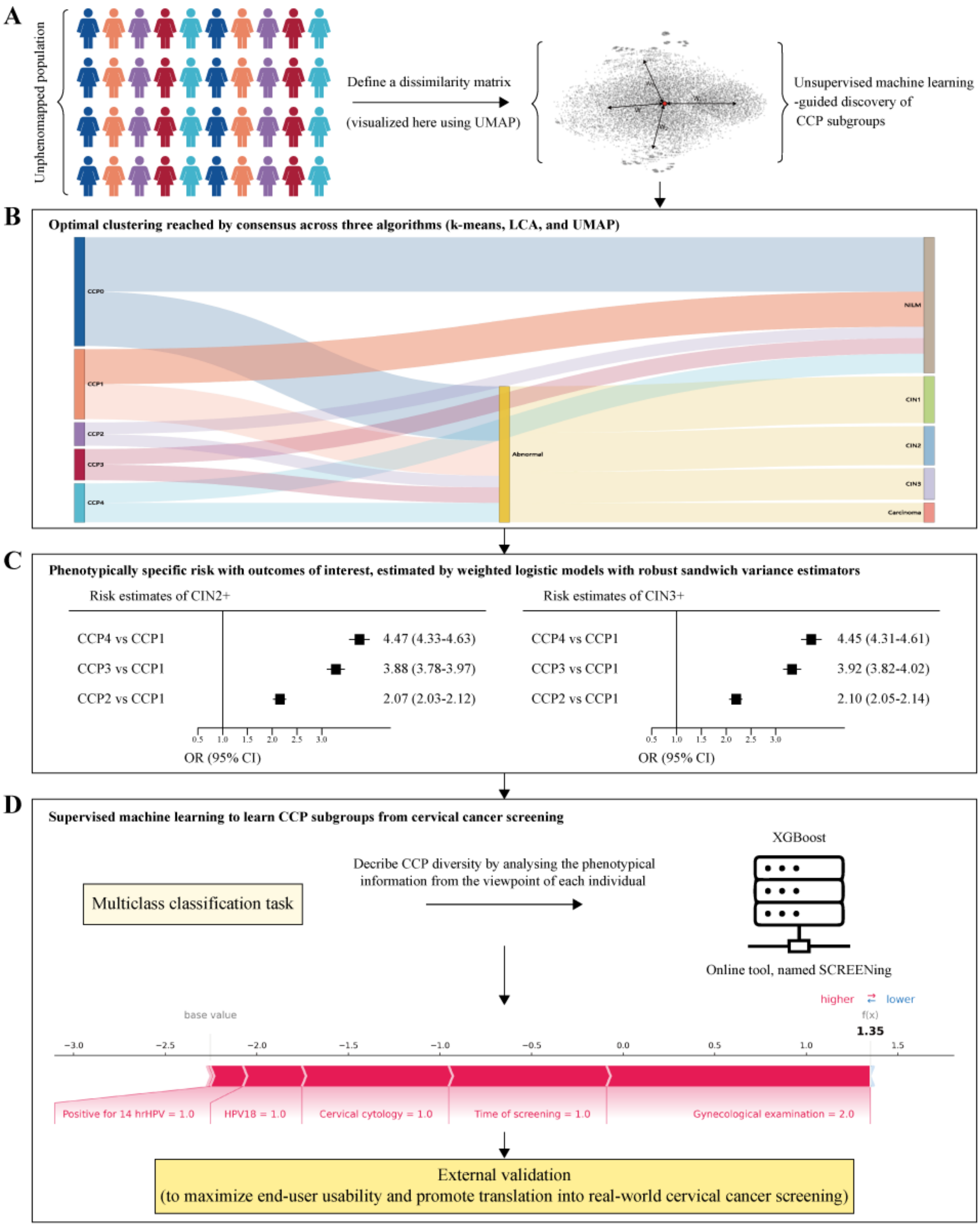
Methods

Data Source

Deidentified data were extracted from electronic health records (EHRs) of the national cervical cancer screening program in China. In summary, our study included eligible women aged

25 - 65 years who participated in the cervical cancer screening. Data from Fujian Province (2014 - 2023) were used to establish a discovery cohort to train the models. Additionally, EHR data from 5 other regions—Shenzhen City, Foshan City, Hubei Province, Gansu Province, and Guizhou Province—were employed as an external cohort to validate the generalizability of the models across diverse populations. Details on the study design, as illustrated in Figure 1, are available in Appendix S1 in Multimedia Appendix 1.

Figure 1. Study design. (**A**) We explored the data-driven cervical cancer prevention (CCP) subgroups by applying unsupervised machine learning to a deeply phenotyped, population-based discovery cohort. In this population-based, external, and diagnostic validation study, we aimed to use different machine learning methods to test our hypothesis (**B**) that individuals within distinct CCP subgroups exhibit differential risks of cervical precancer or cancer based on their complex phenotypic profiles. After identifying CCP subgroups, we extracted CCP-specific risks of outcomes of interest (**C**) conditionally on all predefined and algorithmically selected features through weighted logistic regression analyses providing average odds ratio (OR) estimates. Additionally, we stratified individuals based on their key features to conduct subgroup analyses. Finally, we trained a supervised machine learning model and developed pathways (**D**) to classify individuals using features most consistently linked to CCP subgroups before evaluating its diagnostic validity and usability on an external cohort. Pathologically abnormal diagnoses consisted of CIN1, CIN2, CIN3, and carcinoma. UMAP: uniform manifold approximation and projection; CCP: cervical cancer prevention; LCA: latent class analysis; NILM: negative for intraepithelial lesion or malignancy; CIN1: cervical intraepithelial neoplasia grade 1; CIN2: cervical intraepithelial neoplasia grade 2; CIN3: cervical intraepithelial neoplasia grade 3; CIN2+: cervical intraepithelial neoplasia grade 2 or worse; CIN3+: cervical intraepithelial neoplasia grade 3 or worse; OR: odds ratio; XGBoost: eXtreme gradient boosting; SCREENing: clinical Subgroups for CeRvical cancEr prEvention using computational phenomaps and machine learning; hrHPV: high-risk human papillomavirus; HPV: human papillomavirus.



Ethical Considerations

This study was approved by the Ethics Committee of Fujian Maternal and Child Health Hospital (2023KY141). No additional informed consent was required for this analysis. All participant data were deidentified prior to analysis. No images of identifiable individuals are included in the manuscript or supplementary materials.

Data Preprocessing

In line with the consensus that detecting and treating CIN2 or CIN3 [25,26], key premalignant cervical conditions, can prevent progression to invasive cervical cancer, our primary and secondary outcomes were CIN2+ and CIN3+, respectively. These outcomes were confirmed by histopathology and align with those commonly used in previous studies [26-29]. We selected features based on prior research [22,30-34] and input from clinical, biostatistical, and epidemiological experts. These included demographic characteristics, cervical cancer screening history, HPV infection status, and medical examination results. Details on data preprocessing are provided in Appendix S1 in Multimedia Appendix 1.

Unsupervised Machine Learning

We calculated phenotypic distances between individuals using Gower's distance [35], a dissimilarity metric suitable for mixed continuous and categorical data. To visualize phenotypic variation in the population, we employed uniform manifold approximation and projection (UMAP) [36] to construct a phenomap. This approach enhances interpretability by presenting the distributions of individuals within a multidimensional phenotypic space, capturing the full range of baseline phenotypes. To identify and categorize the underlying CCP subgroups in which individuals exhibit phenotypic similarity, we employed and compared 3 methods to capture and validate CCP diversity: k-means [37], latent class analysis [38], and UMAP [39]. The final number of CCP subgroups was determined by reaching a consensus across all 3 approaches. Further details are provided in Appendix S1 in Multimedia Appendix 1.

Risk Estimates and an Algorithm to Identify Subgroups

To further explore the association between CCP subgroups and outcomes of interest, we calculated CCP-specific risk estimates (odds ratios, ORs) [40] and 95% CIs using inverse probability weights [41]. Additionally, we trained an eXtreme gradient boosting (XGBoost) algorithm [42] to predict 5 CCP subgroups with differential risks of outcomes, as described in our previous work [23]. Briefly, the model was trained and calibrated using an isotonic regression algorithm, and internally validated in the discovery cohort. The SHapley Additive exPlanations (SHAP) method was employed to identify each feature's relative contribution [23,43] and enhance the model's explainability. Model performance was evaluated using OVR AUROC (one-versus-rest area under the receiver operating characteristics curve, extended for multiple classes), Brier score, and calibration curves as primary metrics. In external validation, the model's diagnostic validity was assessed by comparing CCP-specific ORs between datasets to evaluate cross-cluster and cross-dataset risk differences. Further details are provided in Appendix S1 in Multimedia Appendix 1.

Developing Pathways to Improve Impact

We named our computational phenomapping strategy SCREENing (clinical Subgroups for CeRvical cancEr prEvention using computational pheNomaps and machine learnING). To assess real-world usability and effectiveness, we consulted 11 clinical experts and 3 epidemiologists on its clinical relevance, justification, result interpretability, and potential impact on screening strategies and public health policymaking.

Statistical Analysis

Given the importance of menopause in women, we specifically examined CCP-specific risks across age using spline analyses, with interaction tests to assess whether age modified these risks. Subgroup analyses were also performed by stratifying women based on key features. Details are provided in Appendix S1 in Multimedia Appendix 1. We followed the Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research [44] and the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis statement [45]. Data management was performed using the lulab.utils R package [46]. All analyses [47] were conducted using Python (version 3.11.6), SAS Enterprise Guide (version 7.1), and R (version 4.3.2; R Foundation for Statistical Computing).

Results

Characteristics of Cohorts

The study included 551,934 women (median age 49 years [42,48]; 10% infected with hrHPV, 1.2% with HPV-16, and 0.6% with HPV-18) in the discovery cohort (Table 1) and 47,130 women (median age 37 years [32,45]; 16.6% with hrHPV, 3.9% with HPV-16, and 1.4% with HPV-18) in the external cohort (Table S3 in Multimedia Appendix 1). In the discovery cohort, 9932 were pathologically diagnosed with cervical abnormalities, including 533 with carcinoma. In contrast, the external cohort showed a higher prevalence of cervical abnormalities at 11.1% (5251).

Table . Characteristics by cervical cancer prevention (CCP) subgroups in the discovery cohort. Categorical features are summarized as numbers (percentages), and continuous features as median (Q1, Q3), as appropriate.

Characteristics	CCP0[a] (n=542,002)	CCP1 (n=2242)	CCP2 (n=3770)	CCP3 (n=2278)	CCP4 (n=1642)	Total (n=551,934)
Age (years), median (Q1, Q3)	48.00 (42.00, 54.00)	48.00 (42.00, 54.00)	48.00 (43.00, 55.00)	51.00 (44.00, 57.00)	48.00 (43.00, 54.00)	49.00 (42.00, 54.00)
Race/ethnicity, n (%)						
Han	534,023 (98.53)	2214 (98.75)	3707 (98.33)	2236 (98.16)	1607 (97.87)	543,787 (98.52)
Others	4528 (0.84)	15 (0.67)	30 (0.80)	29 (1.27)	20 (1.22)	4622 (0.84)
Missing	3451 (0.64)	13 (0.58)	33 (0.88)	13 (0.57)	15 (0.91)	3525 (0.64)
History of cervical cancer screening, n (%)						
Missing	234 (0.04)	0 (0.00)	2 (0.05)	0 (0.00)	0 (0.00)	236 (0.04)
No	411,029 (75.84)	0 (0.00)	3768 (99.95)	2278 (100.00)	1642 (100.00)	418,717 (75.86)
Yes	130,739 (24.12)	2242 (100.00)	0 (0.00)	0 (0.00)	0 (0.00)	132,981 (24.09)
Time of previous screening, n (%)						
Missing	785 (0.14)	13 (0.58)	2 (0.05)	0 (0.00)	0 (0.00)	800 (0.14)
No previous screening	411,029 (75.84)	0 (0.00)	3768 (99.95)	2278 (100.00)	1642 (100.00)	418,717 (75.86)
Within 3 years from now	65,519 (12.09)	1091 (48.66)	0 (0.00)	0 (0.00)	0 (0.00)	66,610 (12.07)
More than 3 years ago	64,669 (11.93)	1138 (50.76)	0 (0.00)	0 (0.00)	0 (0.00)	65,807 (11.92)
Gynecological examination, n (%)						
Missing	2180 (0.40)	1 (0.04)	1 (0.03)	0 (0.00)	0 (0.00)	2182 (0.40)
Normal	417,921 (77.11)	1508 (67.26)	3763 (99.81)	1568 (68.83)	0 (0.00)	424,760 (76.96)
Abnormal	121,901 (22.49)	733 (32.69)	6 (0.16)	710 (31.17)	1642 (100.00)	124,992 (22.65)
Positive for high-risk HPV[b, c], n (%)						
No	496,234 (91.56)	99 (4.42)	206 (5.46)	27 (1.19)	67 (4.08)	496,633 (89.98)
Yes	45,768 (8.44)	2143 (95.58)	3564 (94.54)	2251 (98.81)	1575 (95.92)	55,301 (10.02)
Positive for low-risk HPV[d], n (%)						
No	534,313 (98.58)	2035 (90.77)	3664 (97.19)	1813 (79.59)	1602 (97.56)	543,427 (98.46)
Yes	7689 (1.42)	207 (9.23)	106 (2.81)	465 (20.41)	40 (2.44)	8507 (1.54)
Positive for possible high-risk HPV[e], n (%)						
No	537,999 (99.26)	2130 (95.00)	3686 (97.77)	1987 (87.23)	1617 (98.48)	547,419 (99.18)
Yes	4003 (0.74)	112 (5.00)	84 (2.23)	291 (12.77)	25 (1.52)	4515 (0.82)
Number of HPV infections	0.00 (0.00, 0.00)	1.00 (1.00, 2.00)	1.00 (1.00, 1.00)	2.00 (2.00, 3.00)	1.00 (1.00, 1.00)	0.00 (0.00, 0.00)
Positive for HPV-16, n (%)						
No	538,375 (99.33)	1684 (75.11)	2745 (72.81)	1494 (65.58)	1199 (73.02)	545,497 (98.83)
Yes	3627 (0.67)	558 (24.89)	1025 (27.19)	784 (34.42)	443 (26.98)	6437 (1.17)
Positive for HPV-18, n (%)						
No	539,716 (99.58)	2002 (89.30)	3442 (91.30)	1887 (82.84)	1508 (91.84)	548,555 (99.39)
Yes	2286 (0.42)	240 (10.70)	328 (8.70)	391 (17.16)	134 (8.16)	3379 (0.61)
Positive for HPV-31, n (%)						
No	540,143 (99.66)	2138 (95.36)	3662 (97.14)	2063 (90.56)	1588 (96.71)	549,594 (99.58)

Characteristics	CCP0[a] (n=542,002)	CCP1 (n=2242)	CCP2 (n=3770)	CCP3 (n=2278)	CCP4 (n=1642)	Total (n=551,934)
Yes	1859 (0.34)	104 (4.64)	108 (2.86)	215 (9.44)	54 (3.29)	2340 (0.42)
Positive for HPV-33, n (%)						
No	539,290 (99.50)	2096 (93.49)	3608 (95.70)	1978 (86.83)	1562 (95.13)	548,534 (99.38)
Yes	2712 (0.50)	146 (6.51)	162 (4.30)	300 (13.17)	80 (4.87)	3400 (0.62)
Positive for HPV-35, n (%)						
No	540,820 (99.78)	2197 (97.99)	3708 (98.36)	2160 (94.82)	1622 (98.78)	550,507 (99.74)
Yes	1182 (0.22)	45 (2.01)	62 (1.64)	118 (5.18)	20 (1.22)	1427 (0.26)
Positive for HPV-39, n (%)						
No	537,675 (99.20)	2092 (93.31)	3638 (96.50)	2021 (88.72)	1600 (97.44)	547,026 (99.11)
Yes	4327 (0.80)	150 (6.69)	132 (3.50)	257 (11.28)	42 (2.56)	4908 (0.89)
Positive for HPV-45, n (%)						
No	540,998 (99.81)	2205 (98.35)	3741 (99.23)	2204 (96.75)	1632 (99.39)	550,780 (99.79)
Yes	1004 (0.19)	37 (1.65)	29 (0.77)	74 (3.25)	10 (0.61)	1154 (0.21)
Positive for HPV-51, n (%)						
No	537,654 (99.20)	2053 (91.57)	3590 (95.23)	1954 (85.78)	1565 (95.31)	546,816 (99.07)
Yes	4348 (0.80)	189 (8.43)	180 (4.77)	324 (14.22)	77 (4.69)	5118 (0.93)
Positive for HPV-52, n (%)						
No	525,904 (97.03)	1622 (72.35)	3019 (80.08)	1396 (61.28)	1276 (77.71)	533,217 (96.61)
Yes	16,098 (2.97)	620 (27.65)	751 (19.92)	882 (38.72)	366 (22.29)	18,717 (3.39)
Positive for HPV-56, n (%)						
No	539,439 (99.53)	2133 (95.14)	3682 (97.67)	2076 (91.13)	1610 (98.05)	548,940 (99.46)
Yes	2563 (0.47)	109 (4.86)	88 (2.33)	202 (8.87)	32 (1.95)	2994 (0.54)
Positive for HPV-58, n (%)						
No	534,764 (98.66)	1877 (83.72)	3297 (87.45)	1691 (74.23)	1413 (86.05)	543,042 (98.39)
Yes	7238 (1.34)	365 (16.28)	473 (12.55)	587 (25.77)	229 (13.95)	8892 (1.61)
Positive for HPV-59, n (%)						
No	540,034 (99.64)	2180 (97.23)	3710 (98.41)	2124 (93.24)	1620 (98.66)	549,668 (99.59)
Yes	1968 (0.36)	62 (2.77)	60 (1.59)	154 (6.76)	22 (1.34)	2266 (0.41)
Positive for HPV-66, n (%)						
No	540,640 (99.75)	2186 (97.50)	3720 (98.67)	2180 (95.70)	1614 (98.29)	550,340 (99.71)
Yes	1362 (0.25)	56 (2.50)	50 (1.33)	98 (4.30)	28 (1.71)	1594 (0.29)
Positive for HPV-68, n (%)						
No	538,250 (99.31)	2124 (94.74)	3654 (96.92)	2018 (88.59)	1604 (97.69)	547,650 (99.22)
Yes	3752 (0.69)	118 (5.26)	116 (3.08)	260 (11.41)	38 (2.31)	4284 (0.78)
Positive for HPV-11, n (%)						
No	541,435 (99.90)	2218 (98.93)	3764 (99.84)	2235 (98.11)	1631 (99.33)	551,283 (99.88)
Yes	567 (0.10)	24 (1.07)	6 (0.16)	43 (1.89)	11 (0.67)	651 (0.12)
Positive for HPV-42, n (%)						
No	540,587 (99.74)	2199 (98.08)	3745 (99.34)	2175 (95.48)	1639 (99.82)	550,345 (99.71)
Yes	1415 (0.26)	43 (1.92)	25 (0.66)	103 (4.52)	3 (0.18)	1589 (0.29)
Positive for HPV-43, n (%)						
No	541,315 (99.87)	2212 (98.66)	3759 (99.71)	2210 (97.01)	1641 (99.94)	551,137 (99.86)

Characteristics	CCP0[a] (n=542,002)	CCP1 (n=2242)	CCP2 (n=3770)	CCP3 (n=2278)	CCP4 (n=1642)	Total (n=551,934)
Yes	687 (0.13)	30 (1.34)	11 (0.29)	68 (2.99)	1 (0.06)	797 (0.14)
Positive for HPV-44, n (%)						
No	540,505 (99.72)	2212 (98.66)	3753 (99.55)	2201 (96.62)	1636 (99.63)	550,307 (99.71)
Yes	1497 (0.28)	30 (1.34)	17 (0.45)	77 (3.38)	6 (0.37)	1627 (0.29)
Positive for HPV-6, n (%)						
No	540,748 (99.77)	2207 (98.44)	3748 (99.42)	2199 (96.53)	1635 (99.57)	550,537 (99.75)
Yes	1254 (0.23)	35 (1.56)	22 (0.58)	79 (3.47)	7 (0.43)	1397 (0.25)
Positive for HPV-81, n (%)						
No	539,354 (99.51)	2178 (97.15)	3747 (99.39)	2136 (93.77)	1632 (99.39)	549,047 (99.48)
Yes	2648 (0.49)	64 (2.85)	23 (0.61)	142 (6.23)	10 (0.61)	2887 (0.52)
Positive for HPV-53, n (%)						
No	538,445 (99.34)	2139 (95.41)	3689 (97.85)	2015 (88.45)	1617 (98.48)	547,905 (99.27)
Yes	3557 (0.66)	103 (4.59)	81 (2.15)	263 (11.55)	25 (1.52)	4029 (0.73)
Cervical cytology examination, n (%)						
NILM[f]	42,746 (7.89)	489 (21.81)	721 (19.12)	438 (19.23)	345 (21.01)	44,739 (8.11)
No examination due to negative for high-risk HPV	487,235 (89.90)	3 (0.13)	7 (0.19)	0 (0.00)	0 (0.00)	487,245 (88.28)
ASC-US[g]	7577 (1.40)	808 (36.04)	1185 (31.43)	721 (31.65)	531 (32.34)	10822 (1.96)
LSIL[h]	2245 (0.41)	488 (21.77)	772 (20.48)	546 (23.97)	286 (17.42)	4337 (0.79)
AGC[i]	99 (0.02)	18 (0.80)	29 (0.77)	13 (0.57)	30 (1.83)	189 (0.03)
Missing but positive for high-risk HPV	779 (0.14)	99 (4.42)	231 (6.13)	159 (6.98)	81 (4.93)	1349 (0.24)
ASC-H[j]	902 (0.17)	147 (6.56)	363 (9.63)	163 (7.16)	145 (8.83)	1720 (0.31)
AGC-FN[k]	7 (0.00)	4 (0.18)	9 (0.24)	3 (0.13)	6 (0.37)	29 (0.01)
HSIL[l]	411 (0.08)	185 (8.25)	443 (11.75)	233 (10.23)	215 (13.09)	1487 (0.27)
Carcinoma	1 (0.00)	1 (0.04)	10 (0.27)	2 (0.09)	3 (0.18)	17 (0.00)
Histopathological diagnosis, n (%)						
NILM	542,002 (100.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	542,002 (98.20)
CIN1[m]	0 (0.00)	1542 (68.78)	2456 (65.15)	1445 (63.43)	948 (57.73)	6391 (1.16)
CIN2[n]	0 (0.00)	22 (0.98)	29 (0.77)	15 (0.66)	16 (0.97)	82 (0.01)
CIN2/3[o]	0 (0.00)	551 (24.58)	1113 (29.52)	716 (31.43)	485 (29.54)	2865 (0.52)
CIN3[p]	0 (0.00)	16 (0.71)	19 (0.50)	13 (0.57)	13 (0.79)	61 (0.01)

Characteristics	CCP0[a] (n=542,002)	CCP1 (n=2242)	CCP2 (n=3770)	CCP3 (n=2278)	CCP4 (n=1642)	Total (n=551,934)
Carcinoma	0 (0.00)	111 (4.95)	153 (4.06)	89 (3.91)	180 (10.96)	533 (0.10)

[a] CCP: cervical cancer prevention.

[b] HPV: human papillomavirus.

[c] high-risk HPV: HPV-16/18/31/33/35/39/45/51/52/56/58/59/66/68.

[d] low-risk HPV: HPV-11/40/42/43/44/6/61/72/81.

[e] possible high-risk HPV: HPV-53/70/73/82/83.

[f] NILM: negative for intraepithelial lesion or malignancy.

[g] ASC-US: atypical squamous cells of undetermined significance.

[h] LSIL: low-grade squamous intraepithelial lesion.

[i] AGC: atypical glandular cells.

[j] ASC-H: atypical squamous cells, cannot exclude high-grade squamous intraepithelial lesion.

[k] AGC-FN: AGC-favor neoplastic.

[l] HSIL: high-grade squamous intraepithelial lesion.

[m] CIN1: cervical intraepithelial neoplasia grade 1.

[n] CIN2: cervical intraepithelial neoplasia grade 2.

[o] CIN2/3 is reflective of CIN2 or CIN3, i.e., HSIL; carcinoma consists of AIS (adenocarcinoma in situ) and cancer.

[p] CIN3: cervical intraepithelial neoplasia grade 3.

Data-Driven Subgroups

The discovery cohort was extensively phenotyped based on 31 features (Figure 2). Visual assessment of the risk phenomaps revealed that nearly all 6 features were heterogeneously distributed in the phenomic space to varying degrees, indicating distinct phenotypic neighborhoods. In the discovery cohort, the CCP subgroups identified across 3 algorithms and all metrics were stable (Figure 3A and Table S4 in Multimedia Appendix 1), with the optimal number of clusters being 4 in the subpopulation of 9932 women with cervical abnormalities. The distribution of cervical abnormalities across subgroups is presented in Figure 3B. Additionally, women with normal screening results were treated as another CCP subgroup (Table 1). Following a detailed analysis of each subgroup's features, we labeled the 5 identified subgroups as follows: (0) healthy, (1) early onset, (2) screening-targeted, (3) late onset, and (4) carcinoma-specific. CCP1 had the highest prevalence of CIN1, while CCP4 had the highest prevalence of carcinoma (Figure 3C). Most features were able to discriminate well between the subgroups. In the external cohort, the subgroups were consistent (Table S3 in Multimedia Appendix 1), though CCP1 was not identified due to missing information on previous screenings. The distribution of subgroups was similar across the discovery and external cohorts, with CCP0 being the most common subgroup and CCP4 having the highest prevalence of carcinoma.

Figure 2. Manifold representations of the phenotypic architecture of 9932 individuals pathologically diagnosed with cervical abnormality from the discovery cohort (N=551,934). A total of 9932 individuals are embedded in the phenotypic space based on dissimilarity metrics (Gower's distance) derived from 31 included phenotypic features; thus, phenotypically similar women tend to be topologically closer. In the subfigures, each dot represents an individual, with coloring based on the value of features. Since the dimensionality reduction is nonlinear, axes have been omitted, and only the comparisons between distances are meaningful. Pathologically abnormal diagnoses consisted of CIN1, CIN2, CIN3, and carcinoma. 14 high-risk HPV: HPV-16/18/31/33/35/39/45/51/52/56/58/59/66/68; low-risk HPV: HPV-11/40/42/43/44/6/61/72/81; possible high-risk HPV: HPV- 53/70/73/82/83.

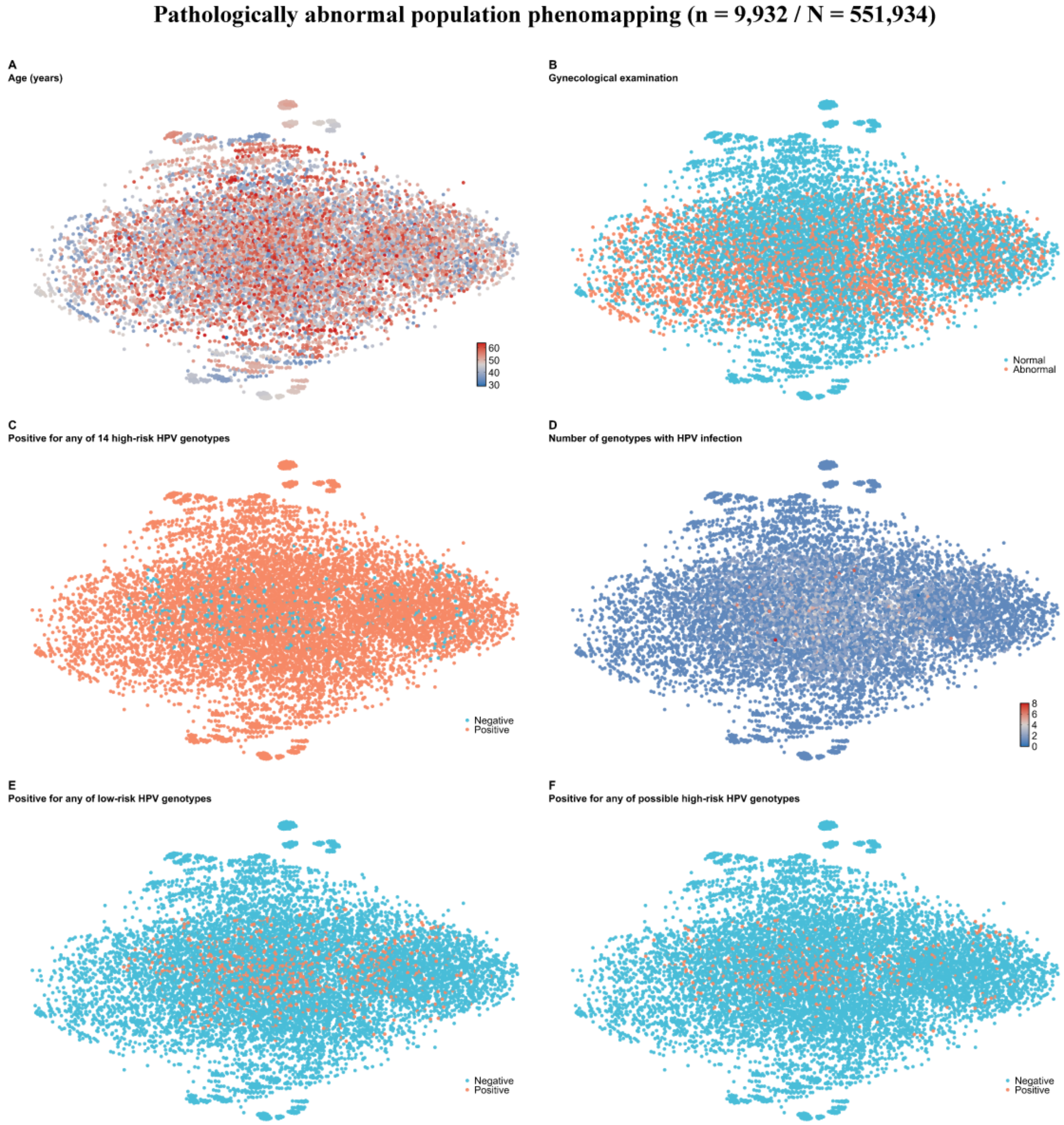
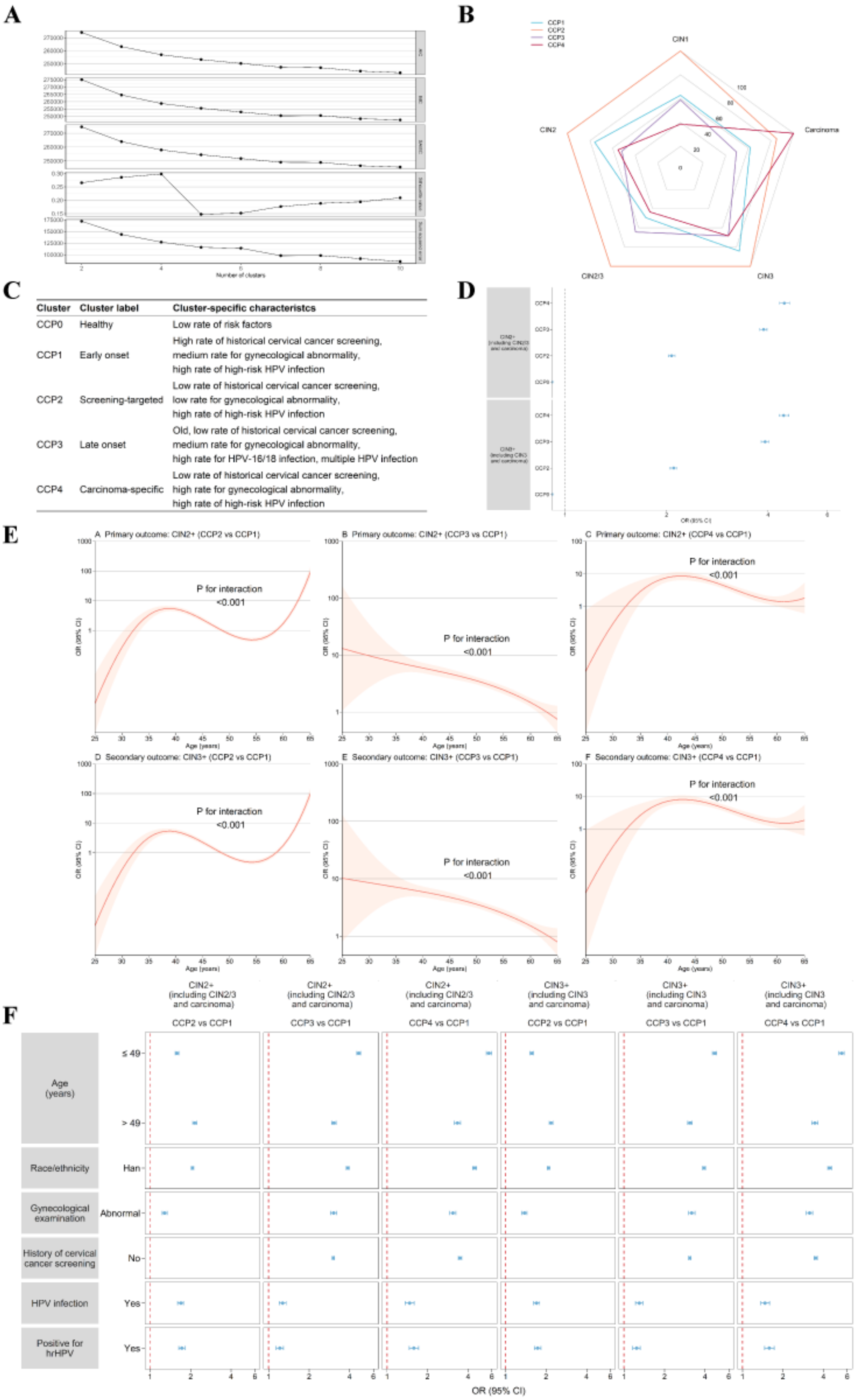


Figure 3. Determination of the 5 CCP subgroups with differential risks of CIN2+/CIN3+. (**A**) Elbow plot to determine the optimal number of clusters. Across 3 algorithms (k-means, LCA, and UMAP), identified CCP subgroups were stable, with the optimal number of clusters being 4 in 9932 women pathologically diagnosed with cervical abnormality. Women with normal screening test findings were considered as one independent CCP subgroup free of cervical cancer (Table 1). (**B**) Relative prevalence of pathologically abnormal diagnoses across CCP subgroups. For each pathologically abnormal diagnosis, the highest prevalence was designated as 100%, and the prevalence in each of the other CCP subgroups was relative to that prevalence (0 ‐ 100). (**C**) Characteristics of CCP subgroups. (**D**) CCP-specific risk estimates of CIN2+/CIN3+ compared with the CCP1 subgroup. Adjusted ORs (dots) and 95% CIs (error bars) are presented. The dashed line marks an OR of 1.00; lower limits of 95% CIs with values greater than 1.00 indicate significantly increased risk. (**E**) The CCP-specific risks of CIN2+/CIN3+ across age compared with the CCP1 subgroup. Age was transformed into a restricted cubic spline function for the analyses. *P* value was based on 2-sided Chi-squared test on the interaction between age and CCP subgroups. A *P* value of <.05 suggests that age modifies the association between CCP subgroups and CIN2+/CIN3+. Light-colored bands represent 95% CIs. (**F**) Subgroup analyses of CCP-specific risk estimates of CIN2+/CIN3+ compared with the CCP1 subgroup. Adjusted ORs (dots) and 95% CIs (error bars) are presented. The dashed line marks an OR of 1.00; lower limits of 95% CIs with values greater than 1.00 indicate significantly increased risk. Empty subfigures suggest insufficient samples for the analysis. AIC: Akaike information criterion; BIC: Bayesian information criterion; SABIC: sample size-adjusted BIC; CCP: cervical cancer prevention; CIN1: cervical intraepithelial neoplasia grade 1; CIN2: cervical intraepithelial neoplasia grade 2; CIN3: cervical intraepithelial neoplasia grade 3; CIN2+: cervical intraepithelial neoplasia grade 2 or worse; CIN3+: cervical intraepithelial neoplasia grade 3 or worse; OR: odds ratio; hrHPV: high-risk human papillomavirus.

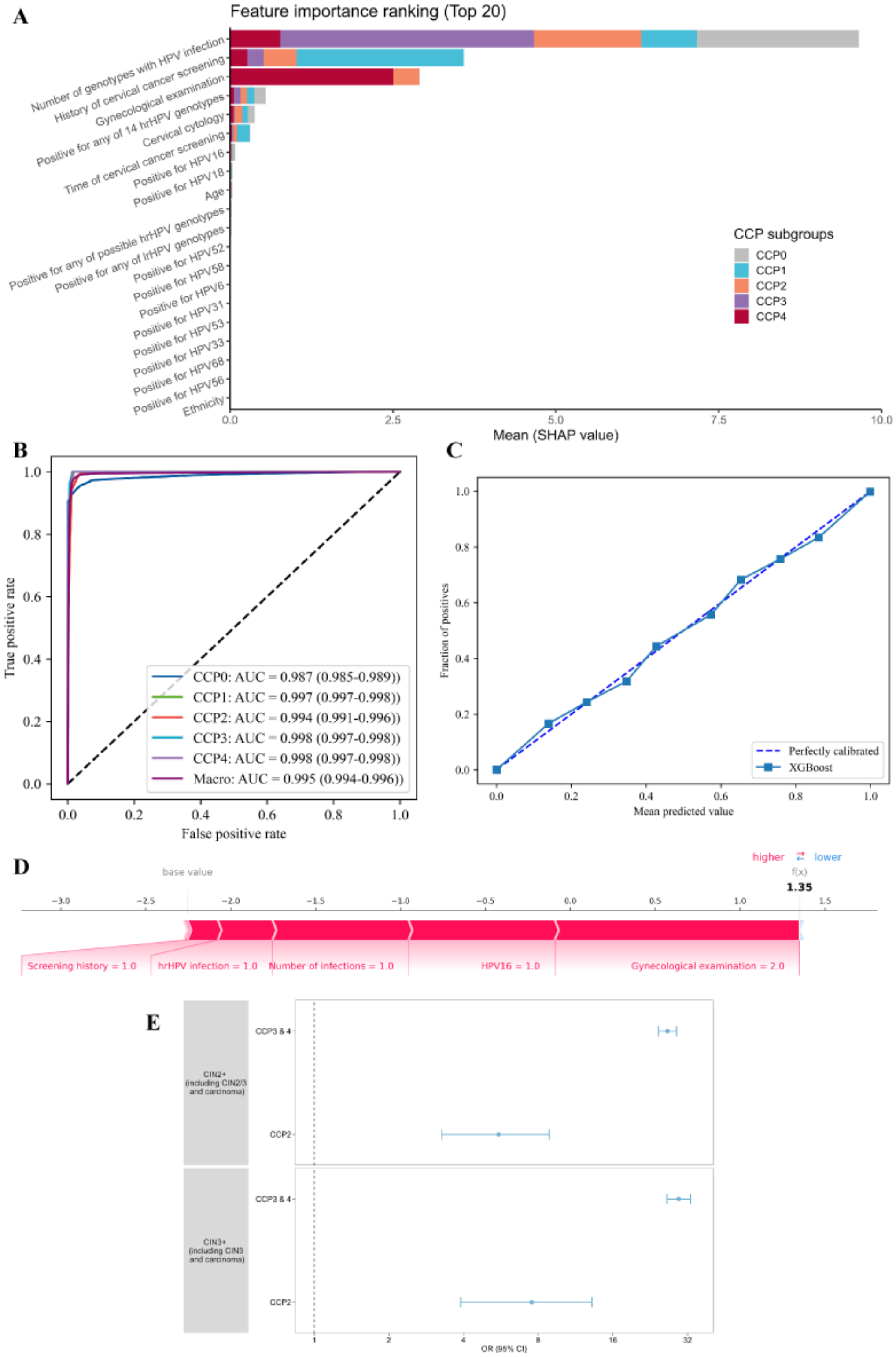


Risk Estimates and Diagnostic Validity

In the diagnostic validity analysis of the discovery cohort (Figure 3D), women in CCP2-4 exhibited a significantly increased risk of both CIN2+ (CCP2: OR 2.07, 95% CI [2.03 ‑ 2.12]; CCP3: 3.88 [3.78 ‑ 3.97]; CCP4: 4.47 [4.33 ‑ 4.63]) and CIN3+ (2.10 [2.05 ‑ 2.14]; 3.92 [3.82 ‑ 4.02]; 4.45 [4.31 ‑ 4.61]) compared to CCP1. Risk analysis across age groups showed that the risks for CIN2+/CIN3+ were evident in the age ranges of 34 ‑ 47 and 60 ‑ 65 in CCP2, 26 ‑ 61 in CCP3, and 35 ‑ 60 in CCP4, respectively (Figure 3E). Interaction analyses between age and subgroups revealed that the risks of CIN2+/CIN3+ increased with age in the 34 ‑ 39 and 60 ‑ 65 age ranges in CCP2, in the 35-43/44 age range in CCP4. Conversely, the risks decreased with age in CCP2 within the 40 ‑ 47 age range and in CCP4 within the 45 ‑ 60 age range (all *P* for interaction <.001). Subgroup analyses indicated that the risks for CIN2+/CIN3+ were present across nearly all subgroups, categorized by age, race, gynecological examination, screening history, hrHPV infection, and number of infections (Figure 3F and Table S5 in Multimedia Appendix 1).

SHAP analysis identified the top 10 key features for prediction, including the number of infections, screening history, gynecological examination, hrHPV infection, cervical cytology, time since previous screening, HPV-16/18 infection, age, and possible hrHPV infection (Figure 4A). The final XGBoost model demonstrated excellent discrimination (OVR AUROC 0.995 [0.994 - 0.996]), with CCP0 showing the lowest AUROC of 0.987 [0.985 - 0.989], and strong calibration (Brier score 0.021 [0.020 - 0.022]) (Figure 4B–C and Table S6 in Multimedia Appendix 1). For the external diagnostic validity of the identified subgroups, women in CCP2-4 showed differential and increased risk for both CIN2+ (CCP2: 5.54 [3.27 - 8.86]; CCP3&4: 26.56 [24.44 - 28.88]) and CIN3+ (7.53 [3.90 - 13.18]; 29.47 [26.46 - 32.86]) compared to CCP0. The cross-cluster risk differences were consistent across the discovery and external cohorts. To illustrate the model's explainability, Figure 4D presents the SHAP plot for a woman in the CCP4 subgroup, who had no previous screening and tested positive for HPV-16 and CIN1 based on cervical cytology. Given this information, our model predicted an elevated risk of CCP4. Additionally, a screenshot of our browser-accessible tool based on the model developed in this study, to clarify how the model can be easily accessed via a browser in clinical settings, is shown in Figure S1 in Multimedia Appendix 1.

Figure 4. Feature importance, discrimination, calibration, explainability, and external diagnostic validity. (**A**) The top 20 features for prediction of CCP subgroups with differential risks of CIN2+/CIN3+ are shown. The y-axis represents the features included in the model development (in descending order of importance), and the x-axis indicates the mean of SHAP values. The 10 most important features were selected to train the final parsimonious model. (**B**) ROC curves in the internal validation from the discovery cohort. The dark purple line represents the macro-average of all 5 CCP subgroups. (**C**) Calibration curve of the alignment between predicted and observed CCP subgroups for the final model. The solid line corresponds to the calibration curve, with the dashed line corresponding to a reference for perfect calibration (ie, perfect alignment between the predicted and observed probabilities). (**D**) Explanation of the SCREENing tool (ie, the inference process of the final model with a woman in the CCP4 subgroup). (**E**) In external validation, to assess cross-cluster and cross-dataset risk differences, the model was measured and compared by diagnostic validity (CCP-specific ORs) between datasets. CCP: cervical cancer prevention; hrHPV: high-risk human papillomavirus; lrHPV: low-risk human papillomavirus; SHAP: SHapley Additive exPlanations; AUC: area under receiver operating characteristics curve; XGBoost: eXtreme gradient boosting.



Developed Pathways

Based on our analyses, we proposed a triple SCREENing strategy to prioritize CCP subgroups with varying CIN2+/CIN3+ risks: (1) top priority for colposcopy referrals to women in CCP3-4 in resource-constrained settings, (2) higher priority given for scaling up organized, population-based HPV screening programs (with adequate follow-up) for women in CCP1-2, and (3) lowest priority for CCP0 women, with large-scale screening limited to resource-rich settings. To assess its impact, we evaluated the real-world usability and effectiveness of SCREENing. Sample clinicians reported that the identified subgroups and included features were clinically relevant for risk-based cervical cancer management, particularly for colposcopy referrals in CCP3-4. They also found the strategy transparent, interpretable, and generalizable across clinical settings, with the browser-accessible model feasible and effective for indicating risk probabilities during consultations. Sample epidemiologists highlighted the strategy's potential to advance screening practices, public health policies, and trial design by enabling phenomapping, estimating subgroup-specific risk profiles, and establishing high-performance, cost-effective CCP for resource-limited regions. They noted its capability for testing effectiveness and cost-effectiveness in well-designed, prospective studies.

Discussion

This study represents one of the largest EHR analyses to date, employing diverse machine learning methods and robust validation approaches to classify subgroups and predict CIN2+/CIN3+ risk. By leveraging comprehensive and interconnected phenotypic features, our study identified 5 CCP subgroups with varying risks. CCP2–4 had significantly higher risks of CIN2+ (CCP2: OR 2.07 [95% CI: 2.03 - 2.12], CCP3: 3.88 [3.78 - 3.97], CCP4: 4.47 [4.33 - 4.63]) and CIN3+ (CCP2: 2.10 [2.05 - 2.14], CCP3: 3.92 [3.82 - 4.02], CCP4: 4.45 [4.31 - 4.61]) compared to CCP1 (*P*<.001), consistent with the direction of results observed in the external cohort. Our findings offer a robust foundation for the proposed triple SCREENing strategy. This tailored approach prioritizes high-risk subgroups, providing actionable insights for cervical cancer prevention, particularly in LMICs.

To this end, various approaches, including single-cell transcriptomic analysis [49], sequence and phylogenetic analysis [50,51], and cluster analysis [48,52-54], have been employed to define genetic and cellular heterogeneity [55] in cervical cancer. These studies suggest that distinct subpopulations exhibit heterogeneous risk effects, linking individual features to varying absolute cervical cancer risk. Our 5 CCP subgroups align with findings from 2 major population-based clustering studies [56,57]. One study used a Poisson regression–based CEM clustering algorithm [56] to identify clusters of Indian states with similar cervical cancer incidence patterns. However, due to its focus on approximating missing data on sexual behavior, HPV prevalence, or cervical cancer incidence [56], the clusters were not ideal for screening purposes. That study did not provide sufficient details to estimate cluster-specific risks, and its reliance on features not typically available in routine screening limited real-world applicability. Additionally, strict model assumptions and the inclusion of relatively few features further constrained its generalizability. Another study performed hierarchical clustering of HPV-related methylation sites to identify subgroups of patients with cervical cancer [57]. While informative for prognosis and clinical management, the 3 clusters were not designed for cervical cancer screening. Instead, they are more suitable for guiding prognosis assessment, refining risk stratification, and optimizing treatment strategies in clinical practice.

In contrast, our study, specifically designed for cervical cancer screening, identified complex interactions among factors such as the number of infections, screening history, gynecological examination, hrHPV and HPV-16/18 infections, cervical cytology, time since previous screening, age, and possible hrHPV infection as key determinants of the 5 CCP subgroups with varying CIN2+/CIN3+ risks. Compared to previous studies [56,57], our design and identified subgroups are more representative of real-world cervical cancer screening. Leveraging the comprehensive and interconnected phenotypic features from EHRs, our algorithm extracted CCP-specific risk estimates, providing a robust foundation for the proposed triple SCREENing strategy. This strategy prioritizes subgroups with differential CIN2+/CIN3+ risks, offering a tailored approach to CCP. We recommend implementing SCREENing as a supplemental tool to existing guidelines [58-60], while accounting for the unique priorities and constraints of LMICs. SCREENing enables clinicians to perform effective, risk-based screening, followed by adequate diagnosis, surveillance, and management, while empowering policy makers to optimize public health policies and resource allocation. This approach has the potential to mitigate resource shortages in LMICs, reduce delays in diagnosis and treatment, and enhance screening efficiency by focusing efforts on high-risk populations, ultimately maximizing population-level benefits. Additionally, our analysis highlights the modifying effect of age on CCP-specific CIN2+/CIN3+ risks, emphasizing the need for greater attention to menopause and age [61] in cervical cancer screening strategies.

Methodologically, our study advances the external validation of machine learning algorithms for identifying subgroups and predicting cervical precancer or cancer risk using large-scale routine EHRs—a rarity in previous studies, which have often been limited to small samples. Our robust, structured framework of internal, external, and diagnostic validation enhances the acceptability and generalizability of unsupervised and supervised machine learning approaches in routine CCP and is adaptable to other diseases. The identified subgroups demonstrated good accuracy and diagnostic validity for CIN2+/CIN3+ both within and across datasets, though performance was lower in scenarios with missing key features. Differences in diagnostic validity across datasets may reflect variations in HPV infection patterns and previous screening histories, which influence the distribution of risk factors. Our findings of 5 subgroups and the proposed strategy are novel, offering a framework for assessing validity in screening, follow-up surveillance, and treatment for cervical cancer. This study signals the potential for more effective and targeted approaches to CCP.

This study represents one of the largest EHR analyses to date, employing multiple machine learning methods, datasets, and validation approaches to classify subgroups and predict CIN2+/CIN3+ risk. However, there are several limitations. First, a major limitation of this study is that our models were not externally validated beyond China. HPV infection patterns and the epidemiology of cervical abnormalities differ across countries and regions, which may impact the generalizability of our models. Given this variability, and although we used data from 5 distinct regions within China, we strongly recommend further validation of our models' performance before applying them in settings not included in this study. We acknowledge that externally validating the model in additional multicenter studies worldwide is crucial for assessing its transferability and applicability across different clinical settings. Additionally, the tool may face challenges in fully adapting to various real-world scenarios when implemented outside the controlled validation environment of this study. To address these concerns, we plan to conduct independent external and prospective validations, as well as pilot implementation across diverse clinical settings in future studies. These efforts will evaluate the model's effectiveness using a wider range of data, particularly in the context of real-world cervical cancer prevention across regions beyond those included in this study. We also encourage independent researchers to validate our model in their own settings, where feasible. While this study marks a significant step in the development of the SCREENing tool, we acknowledge that further research and validation are necessary to establish its effectiveness in real-world applications.

Second, we acknowledge that the use of retrospective EHR data may introduce biases and may not fully capture the real-time challenges of clinical practice. These challenges include issues such as missing data, data errors, and the heterogeneity of EHR systems across different settings. Although we made efforts to control for the quality of the EHR data, it remains inherently difficult to fully address these concerns without significant infrastructure changes, as well as ongoing monitoring and data validation. Therefore, we explicitly caution that the findings of this study should be interpreted with caution. In light of this limitation, future studies should aim to bridge the gap between retrospective analyses and the practical challenges of data collection in clinical settings. Third, while this study focused on routine EHR data, we acknowledge that incorporating multi-omics data—such as genomic, proteomic, and imaging data—could further enhance the model's robustness and performance, providing valuable new insights into cervical abnormalities. We propose this as an area for future research to better understand the progression of cervical cancer and to offer novel perspectives on controlling and eliminating cervical cancer. Fourth, due to the limitations of routinely collected EHR content, we were unable to include several key sexual behavior features, such as the number of sexual partners and oral contraceptive use [62], in our analyses. These factors are known to significantly influence HPV infection and cervical cancer risk. While the model in this study demonstrated high performance, the absence of such data constrained our ability to further analyze the impact of sexual behaviors on findings, such as subgroup characterization, which could provide valuable insights. We recommend that future studies incorporate these data to validate our findings and offer a more comprehensive understanding of this field.

In conclusion, this study underscores the potential of leveraging machine learning algorithms and large-scale routine EHRs to enhance cervical cancer prevention strategies. By identifying key determinants of CIN2+/CIN3+ risk and classifying 5 distinct subgroups, our study provides a robust, data-driven foundation for the proposed triple SCREENing strategy. This approach prioritizes tailored prevention efforts for subgroups with varying risks, offering a novel and scalable tool to complement existing cervical cancer screening guidelines. Future work should focus on independent external and prospective validation to maximize the global impact of this strategy.

Acknowledgments

We thank all the contributors for their efforts in this study. We also thank all funding sources for their funding. This research was funded by Fujian Province's Third Batch of Flexible Introduction of High-Level Medical Talent Teams (TD202307), Merck Investigator Studies Program [100073, 59484], National Natural Science Foundation of China (82271658), Major Scientific Research Program for Young and Middle-aged Health Professionals of Fujian Province, China (2021ZQNZD011), Fujian Province Central Government-Guided Local Science and Technology Development Project (2023L3019), Fujian Provincial Health Technology Project (2024GGA061), and Fujian Provincial Science and Technology Innovation Joint Fund (2021Y9171).

Data Availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Authors' Contributions

ZL conceived and designed the study in consultation with HZ. ZL analyzed the data and interpreted the results. ZL drafted the manuscript with all authors (BD, HC, TT, JW, LF, BW, WZ, TY, XT, JW, SL, XH, ZZ, HX, SX, SL, PS, and HZ) critically reviewing the manuscript. HZ and SL provided administrative and material support, as well as supervision and mentorship.

Each author contributed important intellectual content during drafting or revision of the manuscript and accepts accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved.

All authors had access to all the data in this study and approved the final version of the manuscript. The corresponding authors attests that all the listed authors meet the authorship criteria and that no others meeting the criteria have been omitted. We confirm that no generative AI tools, such as ChatGPT or other language models, were used in the writing or preparation of this manuscript.
S Liu, PS, and HZ are co-corresponding authors.
The complete address for the three co-corresponding authors are as follows:

S Liu: West Building 1 Medical Garden 4th floor, Gongchang Road 66, Sun Yat-sen University (Shenzhen Campus) Shenzhen, Guangming District, 518107, China; liusy99@mail.sysu.edu.cn (SL); +86 135 4329 6396.
PS: 18 Daoshan Road, Fuzhou 350001, Fujian, P.R. China (Phone: +86-591-87558732; Fax: +86-591-87551247; email: fmsun1975@fjmu.edu.cn).
HZ: Room 435, Bld #8, 130 Dongan Road, Xuhui District, Shanghai (200032), China; zouhuachun@fudan.edu.cn (HZ); +86 21 5423 7707.

ZL, BD, HC, TT, and Junfeng Wang are co-first authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data Supplement.
[DOCX File, 380 KB - publichealth_v11i1e67840_app1.docx]

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
2. Chen YC, Chen YY, Su SY, et al. Projected time for the elimination of cervical cancer under various intervention scenarios: age-period-cohort macrosimulation study. JMIR Public Health Surveill 2024 Apr 18;10:e46360. [doi: 10.2196/46360] [Medline: 38635315]
3. Lama Y, Hu D, Jamison A, Quinn SC, Broniatowski DA. Characterizing trends in human papillomavirus vaccine discourse on Reddit (2007-2015): an observational study. JMIR Public Health Surveill 2019 Mar 27;5(1):e12480. [doi: 10.2196/12480] [Medline: 30916662]
4. Walboomers JMM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999 Sep;189(1):12-19. [doi: 10.1002/(SICI)1096-9896(199909)189:1<12::AID-PATH431>3.0.CO;2-F]
5. Lu Z, Zhou X, Fu L, et al. Prognostic prediction models for oropharyngeal squamous cell carcinoma (OPSCC): a protocol for systematic review, critical appraisal and meta-analysis. BMJ Open 2023 Oct 12;13(10):e073375. [doi: 10.1136/bmjopen-2023-073375] [Medline: 37827742]
6. Yang SP, Lin XY, Hu M, Cai CF. The prognostic and predictive effects of human papillomavirus status in hypopharyngeal carcinoma: population-based study. JMIR Public Health Surveill 2022 Dec 16;8(12):e40185. [doi: 10.2196/40185] [Medline: 36525304]
7. Yang L, Lu Z, Bian J, Li F, Zou H. Association between chronic viral infection-related hospitalization and risk of cardiovascular disease: a population-based cohort study. J Med Virol 2024 Jan;96(1):e29350. [doi: 10.1002/jmv.29350] [Medline: 38180233]
8. WHO. Global strategy to accelerate the elimination of cervical cancer as a public health problem. 2020. URL: https://www.who.int/publications/i/item/9789240014107
9. Bruni L, Serrano B, Roura E, et al. Cervical cancer screening programmes and age-specific coverage estimates for 202 countries and territories worldwide: a review and synthetic analysis. Lancet Glob Health 2022 Aug;10(8):e1115-e1127. [doi: 10.1016/S2214-109X(22)00241-8] [Medline: 35839811]
10. Gravitt PE, Silver MI, Hussey HM, et al. Achieving equity in cervical cancer screening in low- and middle-income countries (LMICs): strengthening health systems using a systems thinking approach. Prev Med 2021 Mar;144:106322. [doi: 10.1016/j.ypmed.2020.106322]
11. Arbyn M, Simon M, Peeters E, et al. 2020 list of human papillomavirus assays suitable for primary cervical cancer screening. Clin Microbiol Infect 2021 Aug;27(8):1083-1095. [doi: 10.1016/j.cmi.2021.04.031]
12. Poljak M, Oštrbenk Valenčak A, Gimpelj Domjanič G, Xu L, Arbyn M. Commercially available molecular tests for human papillomaviruses: a global overview. Clin Microbiol Infect 2020 Sep;26(9):1144-1150. [doi: 10.1016/j.cmi.2020.03.033] [Medline: 32247892]
13. Arbyn M, Verdoodt F, Snijders PJF, et al. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. Lancet Oncol 2014 Feb;15(2):172-183. [doi: 10.1016/S1470-2045(13)70570-9] [Medline: 24433684]

14. Mezei AK, Armstrong HL, Pedersen HN, et al. Cost-effectiveness of cervical cancer screening methods in low- and middle-income countries: a systematic review. Int J Cancer 2017 Aug 1;141(3):437-446. [doi: 10.1002/ijc.30695] [Medline: 28297074]
15. Camara H, Nosi S, Munnull G, et al. Women's acceptability of a self-collect HPV same-day screen-and-treat program in a high burden setting in the Pacific. BMC Health Serv Res 2022 Dec 12;22(1):1514. [doi: 10.1186/s12913-022-08842-1] [Medline: 36510192]
16. Ronco G, Dillner J, Elfström KM, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. The Lancet 2014 Feb;383(9916):524-532. [doi: 10.1016/S0140-6736(13)62218-7]
17. Arbyn M, Ronco G, Anttila A, et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. Vaccine (Auckl) 2012 Nov 20;30 Suppl 5:F88-F99. [doi: 10.1016/j.vaccine.2012.06.095] [Medline: 23199969]
18. Bouvard V, Wentzensen N, Mackie A, et al. The IARC perspective on cervical cancer screening. N Engl J Med 2021 Nov 11;385(20):1908-1918. [doi: 10.1056/NEJMsr2030640] [Medline: 34758259]
19. WHO. Guidelines for screening and treatment of precancerous lesions for cervical cancer prevention. 2021. URL: https://www.who.int/publications/i/item/9789240030824
20. Mungo C, Guliam A, Chinula L, et al. Comparison of the ScreenFire and Xpert HPV assays for the detection of human papillomavirus and cervical precancer among women living with HIV in Malawi. medRxiv 2024 Feb 23:2024.02.21.24303142. [doi: 10.1101/2024.02.21.24303142] [Medline: 38617305]
21. WHO. Comprehensive cervical cancer control: a guide to essential practice. 2014.
22. Cheung LC, Egemen D, Chen X, et al. 2019 ASCCP risk-based management consensus guidelines: methods for risk estimation, recommended management, and validation. J Low Genit Tract Dis 2020 Apr;24(2):90-101. [doi: 10.1097/LGT.0000000000000528] [Medline: 32243306]
23. Zou H, Lu Z, Weng W, et al. Diagnosis of neurosyphilis in HIV-negative patients with syphilis: development, validation, and clinical utility of a suite of machine learning models. EClinicalMedicine 2023 Aug;62:102080. [doi: 10.1016/j.eclinm.2023.102080] [Medline: 37533423]
24. Oikonomou EK, Spatz ES, Suchard MA, Khera R. Individualising intensive systolic blood pressure reduction in hypertension using computational trial phenomaps and machine learning: a post-hoc analysis of randomised clinical trials. Lancet Digit Health 2022 Nov;4(11):e796-e805. [doi: 10.1016/S2589-7500(22)00170-4] [Medline: 36307193]
25. Melnikow J, Henderson JT, Burda BU, Senger CA, Durbin S, Weyrich MS. Screening for cervical cancer with high-risk human papillomavirus testing: updated evidence report and systematic review for the US Preventive Services Task Force. J Am Med Assoc 2018 Aug 21;320(7):687-705. [doi: 10.1001/jama.2018.10400] [Medline: 30140883]
26. Xiao T, Wang C, Yang M, et al. Use of virus genotypes in machine learning diagnostic prediction models for cervical cancer in women with high-risk human papillomavirus infection. JAMA Netw Open 2023 Aug 1;6(8):e2326890. [doi: 10.1001/jamanetworkopen.2023.26890] [Medline: 37531108]
27. van der Waal D, Bekkers RLM, Dick S, et al. Risk prediction of cervical abnormalities: the value of sociodemographic and lifestyle factors in addition to HPV status. Prev Med 2020 Jan;130:105927. [doi: 10.1016/j.ypmed.2019.105927]
28. Rothberg MB, Hu B, Lipold L, et al. A risk prediction model to allow personalized screening for cervical cancer. Cancer Causes Control 2018 Mar;29(3):297-304. [doi: 10.1007/s10552-018-1013-4]
29. Lee CH, Peng CY, Li RN, et al. Risk evaluation for the development of cervical intraepithelial neoplasia: development and validation of risk-scoring schemes. Int J Cancer 2015 Jan 15;136(2):340-349. [doi: 10.1002/ijc.28982] [Medline: 24841989]
30. Bouvard V, Baan R, Straif K, et al. A review of human carcinogens--Part B: biological agents. Lancet Oncol 2009 Apr;10(4):321-322. [doi: 10.1016/s1470-2045(09)70096-8] [Medline: 19350698]
31. Schreiberhuber L, Barrett JE, Wang J, et al. Cervical cancer screening using DNA methylation triage in a real-world population. Nat Med 2024 Aug;30(8):2251-2257. [doi: 10.1038/s41591-024-03014-6] [Medline: 38834848]
32. Marcus JZ, Cason P, Downs LS, Einstein MH, Flowers L. The ASCCP Cervical Cancer Screening Task Force Endorsement and Opinion on the American Cancer Society updated cervical cancer screening guidelines. J Low Genit Tract Dis 2021 Jul 1;25(3):187-191. [doi: 10.1097/LGT.0000000000000614] [Medline: 34138787]
33. Moy LM, Zhao FH, Li LY, et al. Human papillomavirus testing and cervical cytology in primary screening for cervical cancer among women in rural China: comparison of sensitivity, specificity, and frequency of referral. Int J Cancer 2010 Aug 1;127(3):646-656. [doi: 10.1002/ijc.25071] [Medline: 19960441]
34. Saslow D, Solomon D, Lawson HW, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. Am J Clin Pathol 2012 Apr;137(4):516-542. [doi: 10.1309/AJCPTGD94EVRSJCG] [Medline: 22431528]
35. Gower JC. A general coefficient of similarity and some of its properties. Biometrics 1971 Dec;27(4):857. [doi: 10.2307/2528823]
36. McInnes L, Healy J, Saul N, Großberger L. UMAP: uniform manifold approximation and projection. arXiv. Preprint posted online on 2018arXiv:180203426.
37. Reese JT, Blau H, Casiraghi E, et al. Generalisable long COVID subtypes: findings from the NIH N3C and RECOVER programmes. eBioMedicine 2023 Jan;87:104413. [doi: 10.1016/j.ebiom.2022.104413] [Medline: 36563487]

38. Zhou W, He MM, Wang F, Xu RH, Wang F, Zhao Q. Latent class analysis-derived classification improves the cancer-specific death stratification of molecular subtyping in colorectal cancer. NPJ Precis Oncol 2023 Jun 23;7(1):60. [doi: 10.1038/s41698-023-00412-w] [Medline: 37353681]
39. Faghri F, Brunn F, Dadu A, et al. Identifying and predicting amyotrophic lateral sclerosis clinical subgroups: a population-based machine-learning study. Lancet Digit Health 2022 May;4(5):e359-e369. [doi: 10.1016/S2589-7500(21)00274-0] [Medline: 35341712]
40. Lu Z, Fu L, Yang L, et al. Hesitancy to undergo SARS-CoV-2 rapid antigen testing in China: nationwide cross-sectional study. JMIR Public Health Surveill 2023 May 1;9:e43555. [doi: 10.2196/43555] [Medline: 36888911]
41. Xu E, Xie Y, Al-Aly Z. Long-term neurologic outcomes of COVID-19. Nat Med 2022 Nov;28(11):2406-2415. [doi: 10.1038/s41591-022-02001-z]
42. Chen T, Guestrin C, editors. XGBoost: a scalable tree boosting system. 2016 Presented at: KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.
43. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2020 Jan;2(1):56-67. [doi: 10.1038/s42256-019-0138-9] [Medline: 32607472]
44. Luo W, Phung D, Tran T, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. J Med Internet Res 2016 Dec 16;18(12):e323. [doi: 10.2196/jmir.5870] [Medline: 27986644]
45. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD statement. Ann Intern Med 2015 Jan 6;162(1):55-63. [doi: 10.7326/M14-0697]
46. Lu Z. lulab.utils: Supporting Functions Maintained by Zhen Lu. 2024. URL: https://CRAN.R-project.org/package=lulab.utils
47. Lu Z. SCREENing. URL: https://github.com/Leslie-Lu/SCREENing
48. Rosty C, Sheffer M, Tsafrir D, et al. Identification of a proliferation gene cluster associated with HPV E6/E7 expression level and viral DNA load in invasive cervical carcinoma. Oncogene 2005 Oct 27;24(47):7094-7104. [doi: 10.1038/sj.onc.1208854] [Medline: 16007141]
49. Li Y, Wang C, Ma A, et al. Identification of HPV oncogene and host cell differentiation associated cellular heterogeneity in cervical cancer via single ‐ cell transcriptomic analysis. J Med Virol 2023 Aug;95(8). [doi: 10.1002/jmv.29060]
50. Sait K, Turki R, Abuzenadah AM, Jiffiri OH, Bohmaidah A, Sohrab SS. Genetic diversity and phylogenetic analysis of HPV 16 & 18 variants isolated from cervical specimens of women in Saudi Arabia. Saudi J Biol Sci 2019 Feb;26(2):317-324. [doi: 10.1016/j.sjbs.2018.05.005] [Medline: 31485171]
51. Balanda M, Fernández J, Vergara N, et al. Genetic variability of human papillomavirus type 66 L1 gene among women presenting for cervical cancer screening in Chile. Med Microbiol Immunol 2019 Dec;208(6):757-771. [doi: 10.1007/s00430-019-00621-w] [Medline: 31165237]
52. Wan F, Miao X, Quraishi I, Kennedy V, Creek KE, Pirisi L. Gene expression changes during HPV-mediated carcinogenesis: a comparison between an in vitro cell model and cervical cancer. Int J Cancer 2008 Jul 1;123(1):32-40. [doi: 10.1002/ijc.23463] [Medline: 18398830]
53. Vázquez-Ortíz G, Ciudad CJ, Piña P, et al. Gene identification by cDNA arrays in HPV-positive cervical cancer. Arch Med Res 2005;36(5):448-458. [doi: 10.1016/j.arcmed.2005.04.016] [Medline: 16099320]
54. Widschwendter A, Gattringer C, Ivarsson L, et al. Analysis of aberrant DNA methylation and human papillomavirus DNA in cervicovaginal specimens to detect invasive cervical cancer and its precursors. Clin Cancer Res 2004 May 15;10(10):3396-3400. [doi: 10.1158/1078-0432.CCR-03-0143] [Medline: 15161694]
55. Shiri Aghbash P, Shirvaliloo M, Khalo Abass Kasho A, Alinezhad F, Nauwynck H, Bannazadeh Baghi H. Cluster of differentiation frequency on antigen presenting-cells: the next step to cervical cancer prognosis? Int Immunopharmacol 2022 Jul;108:108896. [doi: 10.1016/j.intimp.2022.108896] [Medline: 35640377]
56. Man I, Georges D, Bonjour M, Baussano I. Approximating missing epidemiological data for cervical cancer through footprinting: a case study in India. Elife 2023 May 25;12:e81752. [doi: 10.7554/eLife.81752] [Medline: 37227260]
57. Yang S, Wu Y, Wang S, et al. HPV-related methylation-based reclassification and risk stratification of cervical cancer. Mol Oncol 2020 Sep;14(9):2124-2141. [doi: 10.1002/1878-0261.12709] [Medline: 32408396]
58. Li M, Wei L, Sui L, et al. Guidelines for cervical cancer screening in China. Gynecol Obstet Clin Med 2023 Dec;3(4):189-194. [doi: 10.1016/j.gocm.2023.10.005]
59. Sawaya GF, Smith-McCune K, Kuppermann M. Cervical cancer screening: more choices in 2019. J Am Med Assoc 2019 May 28;321(20):2018-2019. [doi: 10.1001/jama.2019.4595] [Medline: 31135834]
60. Wang W, Arcà E, Sinha A, Hartl K, Houwing N, Kothari S. Cervical cancer screening guidelines and screening practices in 11 countries: a systematic literature review. Prev Med Rep 2022 Aug;28:101813. [doi: 10.1016/j.pmedr.2022.101813] [Medline: 35637896]
61. Hickey M, Basu P, Sassarini J, et al. Managing menopause after cancer. The Lancet 2024 Mar;403(10430):984-996. [doi: 10.1016/S0140-6736(23)02802-7]

62. Gierisch JM, Coeytaux RR, Urrutia RP, et al. Oral contraceptive use and risk of breast, cervical, colorectal, and endometrial cancers: a systematic review. Cancer Epidemiol Biomarkers Prev 2013 Nov 1;22(11):1931-1943. [doi: 10.1158/1055-9965.EPI-13-0298]

Abbreviations

CCP: cervical cancer prevention
CIN: cervical intraepithelial neoplasia
EHR: electronic health record
HPV: human papillomavirus
hrHPV: high-risk HPV
LMIC: low- and middle-income country
OR: odds ratio
OVR AUROC: one-versus-rest area under the receiver operating characteristics curve
SCREENing: clinical Subgroups for CeRvical cancEr prEvention using computational pheNomaps and machine learnING
SHAP: SHapley Additive exPlanations
UMAP: uniform manifold approximation and projection
WHO: World Health Organization
XGBoost: eXtreme gradient boosting

Edited by A Mavragani; submitted 22.10.24; peer-reviewed by G Guerrero-Contreras, ZP Liu; revised version received 29.01.25; accepted 29.01.25; published 19.03.25.

Please cite as:
Lu Z, Dong B, Cai H, Tian T, Wang J, Fu L, Wang B, Zhang W, Lin S, Tuo X, Wang J, Yang T, Huang X, Zheng Z, Xue H, Xu S, Liu S, Sun P, Zou H
Identifying Data-Driven Clinical Subgroups for Cervical Cancer Prevention With Machine Learning: Population-Based, External, and Diagnostic Validation Study
JMIR Public Health Surveill 2025;11:e67840
URL: https://publichealth.jmir.org/2025/1/e67840
doi:10.2196/67840

A Call for Action: Lessons Learned From a Pilot to Share a Complex, Linked COVID-19 Cohort Dataset for Open Science

Clara Amid[1*], PhD; Martine Y van Roode[1*], MSc; Gabriele Rinck[2], PhD; Janko van Beek[1], PhD; Rory D de Vries[1], PhD; Gijsbert P van Nierop[1], PhD; Eric C M van Gorp[1], Prof Dr; Frank Tobian[3], MSc; Bas B Oude Munnink[1], PhD; Reina S Sikkema[1], PhD; Thomas Jaenisch[3,4], MD, PhD; Guy Cochrane[2], PhD; Marion P G Koopmans[1,5], DVM, Prof Dr

[1]Department of Viroscience, Erasmus Medical Center (Erasmus MC), Dr Molewaterplein 40, Rotterdam, Netherlands

[2]European Molecular Biology Laboratory (EMBL), European Bioinformatics Institute (EBI), Wellcome Genome Campus, Hinxton, Cambridgeshire, United Kingdom

[3]Heidelberg Institute of Global Health (HIGH), Heidelberg University Hospital (UKHD), Heidelberg, Germany

[4]Center for Global Health, Colorado School of Public Health, Aurora, CO, United States

[5]Pandemic and Disaster Preparedness Center (PDPC), Rotterdam, Netherlands

[*]these authors contributed equally

Corresponding Author:
Martine Y van Roode, MSc
Department of Viroscience, Erasmus Medical Center (Erasmus MC), Dr Molewaterplein 40, Rotterdam, Netherlands

Abstract

The COVID-19 pandemic proved how sharing of genomic sequences in a timely manner, as well as early detection and surveillance of variants and characterization of their clinical impacts, helped to inform public health responses. However, the area of (re)emerging infectious diseases and our global connectivity require interdisciplinary collaborations to happen at local, national and international levels and connecting data to understand the linkages between all factors involved. Here, we describe experiences and lessons learned from a COVID-19 pilot study aimed at developing a model for storage and sharing linked laboratory data and clinical-epidemiological data using European open science infrastructure. We provide insights into the barriers and complexities of internationally sharing linked, complex cohort datasets from opportunistic studies for connected data analyses. An analytical timeline of events, describing key actions and delays in the execution of the pilot, and a critical path, defining steps in the process of internationally sharing a linked cohort dataset are included. The pilot showed how building on existing infrastructure that had previously been developed within the European Nucleotide Archive at the European Molecular Biology Laboratory-European Bioinformatics Institute for pathogen genomics data sharing, allowed the rapid development of connected "data hubs." These data hubs were required to link human clinical-epidemiological data under controlled access with open high dimensional laboratory data, under FAIR (Findable, Accessible, Interoperable, Reusable) principles. Based on our own experiences, we call for action and make recommendations to support and to improve data sharing for outbreak preparedness and response.

(JMIR Public Health Surveill 2025;11:e63996) doi:10.2196/63996

KEYWORDS

data sharing; data management; open science; COVID-19; emerging infectious disease; global health

Introduction

The COVID-19 pandemic proved in a real-case scenario how rapid FAIR (Findable, Accessible, Interoperable, Reusable) [1] sequence data sharing was used to inform public health response policies [2]. An international collaboration, the Reconciliation of Cohort data in Infectious Diseases (ReCoDID) [3], a 4-year project that began in January 2019, was funded by the European Commission's Horizon 2020 programme with a focus on the global response to emerging pathogens. Building on existing infrastructures and partnerships, the consortium was funded to develop a sustainable model for the storage, curation, and analyses of complex datasets collected from infectious disease related cohorts to facilitate and speed up outbreak research. In 2020, a pilot study was developed within this project using COVID-19 data from Erasmus Medical Center (Erasmus MC) as a use-case. The incentive was that in an ongoing global COVID-19 pandemic, linking data for connected data analysis with the potential to better understand the disease would be of public interest and of public health importance. The aim of the pilot was to test the potential for sharing linked laboratory data and clinical-epidemiological (CE) data in the evolving European infrastructure for open science in a centralized model. The centralized model, as opposed to a decentralized or federated data sharing model, was chosen for this pilot because we wanted to build on existing infrastructure [4,5]. The model has the potential of hosting and curation of large datasets, the benefit of standardized data input and output leading to interoperability, easing re-analysis by others necessary for improved prediction and preparedness on global, EU and national level, as was needed for a newly emerged outbreak. A decentralized or federated data sharing model has the benefit of not needing to move the data, but the consequences often manifest themselves in no sharing or limited sharing and linking of datasets, difficulty to apply federated data analyses across a wide range of data formats and sources, with predictions and preparedness happening more at local or regional level. Given the diversity of datasets and the complexity of sharing data from clinical research and in line with the rapidly evolving situation during the COVID-19 pandemic, we wanted to learn what it would take to release linked data from opportunistic studies, rather than pre-approved cohort data.

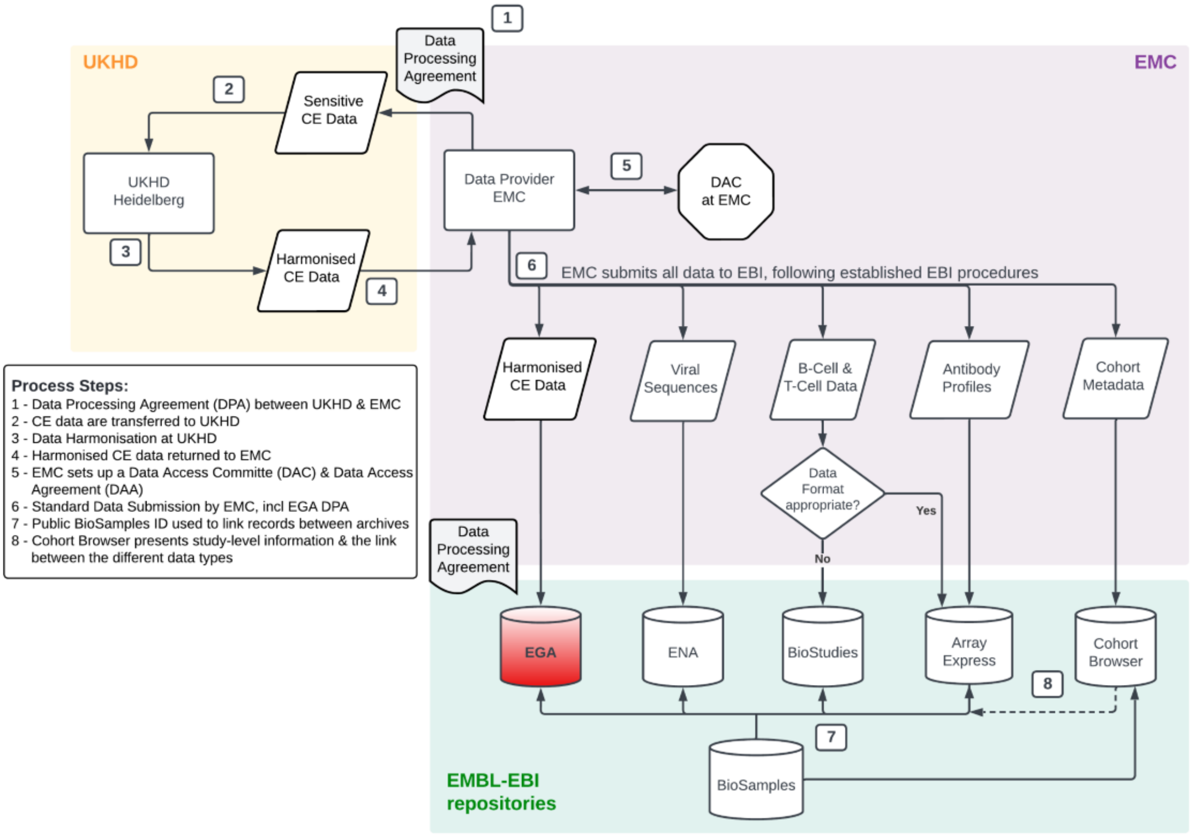
Here, we describe experiences and lessons learned from this pilot, providing insights into the complexities of internationally sharing linked, cohort datasets for connected data analyses.

Pilot Set Up

The following sections have references to the timeline (Multimedia Appendix 1) in which the steps taken to set up and complete a pilot with a small but complex, linked dataset from opportunistic COVID-19 clinical and laboratory research studies as done in (academic) hospital settings are shown. We divided these steps into the following categories: data identification and acquisition, ethics and legal requirements (data protection), data harmonization, and data storage and delivery. Starting in November 2020, meetings took place with basic researchers and clinical scientists who had generated COVID-19 data as part of pandemic related research [6,7]. The drafted data map illustrated how complex the seemingly “simple” datasets (clinical, genomic, pathogen genomic, and high-dimensional laboratory) were. Different data sharing rules and governance for different parts of the datasets existed, and different identifiers and platforms for data capture and storage had been used. It became apparent that these procedures were not unique to COVID-19 pandemic and were common practice at the hospital. Finally, after a further meeting with the involved scientists to discuss datasets that could be linked in a pilot, a small cohort of 151 patients (intensive care unit and non–intensive care unit) who had tested SARS-CoV-2 positive (reverse transcription polymerase chain reaction) between May 2020 and May 2021 was selected for inclusion in the pilot, for whom several data types were partially available, including SARS-CoV-2 viral whole genomes, protein-microarray serological readouts, T- and B-cell data as well as CE data (data identification and acquisition).

Due to the extensive legal discussions around data sharing (ethics and legal requirements), the implementation of the long-term model of the ReCoDID data workflow was brought to a standstill (see “Enablers and Barriers Affecting the Execution of the Pilot”). Therefore, a slightly deviated dataflow model for the pilot was agreed on. The deviation was solely focused on CE data and meant that instead of Heidelberg University Hospital (UKHD) transferring harmonized CE data directly to the European Genome-Phenome Archive (European Molecular Biology Laboratory-European Bioinformatics Institute [EMBL-EBI]), the harmonized CE data were transferred back from UKHD to Erasmus MC, and Erasmus MC carried out the CE data submission to the European Genome-Phenome Archive, including a further data processing agreement (Figure 1). This change was necessary to circumvent one of the barriers (”Adherence to GDPR, and differences in its interpretation” in Table 1) and to move forward with the pilot and to facilitate the development of the connected datasets platform. Furthermore, the Erasmus MC’s legal and privacy offices and the institute’s COVID-19 clinical database team (established in June 2020) were contacted to discuss the aim of the pilot, the access procedures to the CE data, as well as all other steps and legal requirements for sharing the CE data internationally. For this purpose, a Data Protection Impact Assessment was carried out. A data access request for CE data though the Erasmus MC’s COVID-19 clinical database was also required (ethics and legal requirements).

Figure 1. Overview of Erasmus Medical Center (EMC)'s pilot dataflow, linkage and presentation for EMC's connected dataset pilot. The roles of the participating institutes within this pilot were as follows. EMC: mobilization and consolidation of the datasets (data identification and acquisition) by the EMC team at the source within EMC, participant-level and derived BioSample record registration to include metadata about biological samples, linking and submission of all data to repositories within the European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) infrastructure including deposition of clinical-epidemiological (CE) data under controlled access at European Genome-Phenome (EGA), depositions of viral sequences at the European Nucleotide Archive (ENA), B-cell and T-cell data, as well as antibody profiles at BioStudies and ArrayExpress repositories, respectively, and setting up a data access committee (DAC) and data access agreement (DAA) (data storage and delivery); University Hospital Heidelberg (UKHD): harmonization of the CE data and data dictionary at UKHD according to international standards (data harmonization); EMBL-EBI: provided the process for data linking on a participant level, including linking of the CE data to omics data types, using the hierarchical BioSamples database model and presentation of EMC's cohort study in the Reconciliation of Cohort data in Infectious Diseases Cohort Browser (data storage and delivery). All steps taken in this process (numbered arrows) are also numbered in the key entitled "Process Steps."



Process Steps:
1 - Data Processing Agreement (DPA) between UKHD & EMC
2 - CE data are transferred to UKHD
3 - Data Harmonisation at UKHD
4 - Harmonised CE data returned to EMC
5 - EMC sets up a Data Access Committe (DAC) & Data Access Agreement (DAA)
6 - Standard Data Submission by EMC, incl EGA DPA
7 - Public BioSamples ID used to link records between archives
8 - Cohort Browser presents study-level information & the link between the different data types

Table . Enablers and barriers that affected the Erasmus Medical Center (Erasmus MC) pilot's execution.

Description of enablers and barriers	Stakeholders	Category of actions
Enabler Description		
Active support from legal advisors or GDPR[a] specialists. Navigating the complexity of regulations governing data sharing, such as the GDPR, and to apply these accordingly, requires expert privacy and legal knowledge. The active support from legal or GDPR specialists was therefore pivotal in defining the process of dataflow within ReCoDID[b].	Academic hospitals or institutes	Ethics and legal requirements
Building forward on existing infrastructure for sharing (biomedical) data, allowed the rapid development of the connected data hubs to include controlled access sensitive data for/in this pilot.	Researchers; European Union and national funding agencies	Data storage and delivery
Building forward on existing collaborations, as developed in ReCoDID and COMPARE (Collaborative Management Platform for Detection and Analyses of (Re)emerging and Food-borne Outbreaks in Europe) [8], facilitated the steps in the pilot and helped to further build trust necessary for data sharing between participants. The collaboration between the data provider or controller (Erasmus MC) and the respective EMBL-EBI[c] repositories, as well as between the EMBL-EBI repositories, was needed in this pilot to drive further developments and to extend the existing infrastructures to host and present connected datasets.	Researchers; European Union and national funding agencies	All
Dedicated data scientist championing the pilot. Acquiring data from different sources and systems and collected by various specialist teams, linking such data, and ensuring the (legal) requirements are fulfilled before sharing such data, is cumbersome and time-consuming. A dedicated data scientist, supported by encouragement and advocacy of the principal investigator for data sharing amongst researchers, ensured the pilot kept moving ahead.	Academic hospitals or institutes	All
Barrier Description		
Adhering to GDPR, and differences in its interpretation. The data workflow gave rise to legal issues in the context of the GDPR, with regards to GDPR roles (eg, controller vs processor), potential liabilities, different interpretations of Article 49 (d) (Derogations for specific situations, important reasons of public interest), also in regard to the transfer of data to international organizations (here: EMBL-EBI), and resulted in long-lasting legal discussions to define the process of dataflow within ReCoDID.	European Union and national governments; supervisory authorities[d]	Ethics and legal requirements

Description of enablers and barriers	Stakeholders	Category of actions
Insufficient FAIRification[e] and sharing of initial data. Lack of both mandatory FAIRification and sharing rules and specialized staff in inter-pandemic times led to a situation during the pandemic, when time and resources were naturally scarce, often differently prioritized and competing priorities existed, where FAIRification and data sharing was not the main priority. This also led to a situation where harmonization of CE[f] data according to international standards (WHO-ISARIC[g] format) [9] was delayed or not completed.	Academic hospitals or institutes; researchers; European Union and national funding agencies	Data identification and acquisition; data storage and delivery
Insufficient interoperability of the data. Establishing linkages between CE data and corresponding high-dimensional lab data was challenging, and the availability of overlapping data types was limited. Various identifiers were used for CE data and derived high-dimensional laboratory data, and metadata was not fully harmonized, making it difficult to establish linkages. Furthermore, as high-dimensional data types were generated in separately funded projects, a limited amount of overlap in high-dimensional data types existed for the patients.	Academic hospitals or institutes; researchers; European Union and national funding agencies	Data identification and acquisition
Issues contributing to difficulty in CE data acquisition		

Description of enablers and barriers	Stakeholders	Category of actions
• Competing priorities: there was a plethora of clinical studies to be entered into the institute's COVID-19 clinical database. Consequently, there were difficulties in getting approval for prioritization of data entry and data access for this project. Finally, over time there was a reduced interest with decreasing burden of COVID-19 in completing the data entry of CE data. • Time consuming and resource intensive handling and storage of CE data: CE data was stored at different specialized hospital health record systems, that allowed only partial automation of data entry into the institute's COVID-19 clinical database, and records needed to be administered individually. During a time when clinical services were overburdened with COVID-19 cases this resulted in a backlog of CE data that needed to be entered into the institute's database. • Discontinued institutional funding and support for its COVID-19 clinical database: after one year (in 2021), the database was discontinued and subsequently placed under different management. This delay also meant that additional requirements, eg, related to data protection and ethics, had to be met before CE data could be accessed and shared.	Academic hospitals or institutes	Data identification and acquisition

[a] GDPR: General Data Protection Regulation.

[b] ReCoDID: Reconciliation of Cohort data in Infectious Diseases.

[c] EMBL-EBI: European Molecular Biology Laboratory-European Bioinformatics Institute.

[d] The supervisory authorities (barrier “Adhering to GDPR, and difference in its interpretation”) refer to the independent public authorities in each Member State responsible for monitoring the application of the GDPR.

[e] FAIR: Findable, Accessible, Interoperable, Reusable.

[f] CE: clinical-epidemiological.

[g] WHO-ISARIC: World Health Organization-International Severe Acute Respiratory and emerging Infection Consortium.

Linking of Different Data Types

At the time of conducting this pilot, submission of SARS-CoV-2 viral genomes to repositories such as EMBL-EBI's European Nucleotide Archive (for raw sequences) and GISAID [10] (for consensus sequences) was done routinely, other laboratory datasets (protein-microarrays, T- and B-cell data) were not shared. To link SARS-CoV-2 genomes to the other datasets a check was done to verify whether viral genomes had been submitted from the 151 individuals. This exercise required iterative rounds of mapping 4 different identifiers. The internal mapping exercise further revealed the partiality of all other datasets for the individuals included in this cohort. Serum antibody profiles generated using protein-microarrays were available for 40 patients, T-cell phenotype and reactivity data for 28 patients and clonal antibody cross-reactivity data using B-cell profiling for 17 patients out of 151 selected patients in the cohort. For 7 patients out of 151 all data types were available.

Configuration of Data Infrastructure and Data Delivery

With the exception of the viral genomes that were routinely submitted to known repositories (refer to “Linking of different data types” section), for storage and linkage of the dataset it was needed to identify suitable respositories for all other laboratory datasets. This meant establishing where new pilot data types could be hosted, and how the EMBL-EBI system could be leveraged to link all pilot datasets within the EMBL-EBI system, so that the datasets could be eventually also “presented” through a Cohort browser [11]. Although the EMBL-EBI repositories were not fully designed to host one of our data types (T-cell) it was possible to submit the outputs with some adjustments to the format as proof of concept. This leaves room for discussion between the respective researchers and the EMBL-EBI to allow for a more routine application.

The upload of the last Erasmus MC pilot datasets was completed in November 2022. To link all datasets EMBL-EBI's BioSample database [12] was used with a hierarchical model allowing linkage between individuals and their derived samples. With the completion of linking all data types on participant level and presentation of all datasets through the Cohort browser, the pilot was considered as completed (data storage and delivery) in December 2022 [13]. Study level metadata and links to the datasets in the corresponding repositories can be accessed through the Cohort browser and the BioSamples browser at EMBL-EBI [14].

Enablers and Barriers Affecting the Execution of the Pilot

Several enablers and barriers were identified that either facilitated or hampered the execution of this pilot, respectively (Table 1). The main category of actions affected by these barriers and enablers was also identified, as well as the main stakeholder(s) to influence the enablers and barriers.

Several barriers significantly delayed the sharing of a linked dataset in this pilot. The barriers associated with complying with ethics and legal requirements, and with data identification and acquisition, caused the most evidential delays. Some of these barriers tended to reinforce each other, particularly due to shared underlying issues, such as the siloed nature of clinical and complex laboratory data collection and storage. At the start of the pilot, CE data was in different hospital systems, and not yet integrated into the (newly established) institute's database for COVID-19 patient related CE data (”Issues contributing to difficulty in CE data acquisition” in Table 1). Derived high dimensional laboratory data was stored in different systems, separate from patient and diagnostic databases (“Insufficient FAIRification and sharing of initial data”). This was a result of data being gathered by multiple different specialist teams, often for separately funded projects, as well as the lack of linked data infrastructures capturing both patient and (raw) laboratory data, which is typical for hospital systems. Furthermore, the COVID-19 pandemic was characterized by a fragmented funding landscape. Many projects were funded, often focusing on a specific research question or field, thereby reinforcing the siloed nature of specialist teams' work. In addition, for each of these projects, patient sample materials were needed (eg, swabs, blood, and serum). Consequently, high-dimensional lab data were generated for only a subset of patients, based on availability of sample material and requirements of the specific research topic (”Insufficient interoperability of the data”).

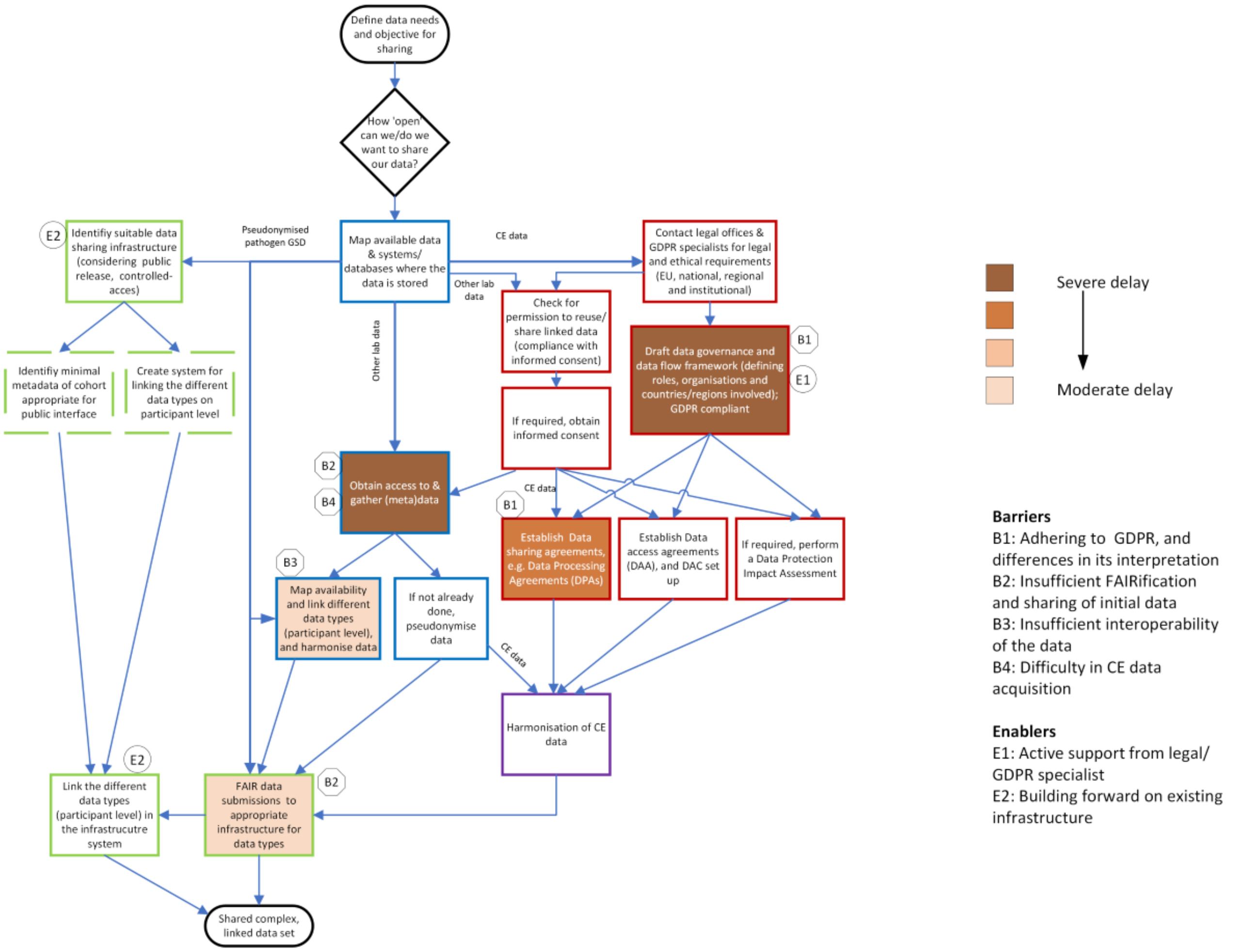
Interestingly, the newly established institute's COVID-19 clinical database could have in principle worked as an enabler for data sharing. It was designed with the aim for CE data to be collected, stored and retrieved from one central database, and potentially address the issue of CE data being siloed in different existing hospital IT systems and databases and improve the FAIR aspects of such data. Yet, a variety of issues, described in Table 1, caused significant delays in CE data acquisition and sharing.

Several factors contributed to or facilitated the sharing of a linked dataset in this pilot, which showed the willingness of the involved teams to overcome the barriers in the sharing of a complex, linked dataset.

Critical Path for Sharing a Complex, Linked Dataset

The critical path analysis (Figure 2) defined the steps in the process of sharing a complex, linked cohort dataset, showing the necessary order of steps or tasks to complete the submission and sharing of such datasets through an appropriate infrastructure, based on our pilot experiences. These steps were classified into the same 4 categories of actions introduced above, data identification and acquisition (4 steps), ethics and legal requirements (7 steps), external data harmonization (1 step), and data storage and delivery (5 steps). Most of the steps in this process could be worked on concurrently, depending on the availability of resources (eg, time of legal staff; availability of appropriate, and standardized templates). As shown in Figure 2, pathogen genetic sequence data (GSD) required a minimum number of steps (1 step), followed by other laboratory data (6 steps), whereas CE data required at least 10 steps. This also corresponds to the time for completion of all steps for each data type in our pilot, with the least amount of time for pathogen GSD. Figure 2 also shows how the identified barriers and enablers affected the process for this pilot.

Figure 2. Critical path for sharing a complex, linked dataset. Two rounded boxes mark the start and endpoint; the diamond box represents a decision. Steps were classified into 4 categories of actions (color coding corresponds to timeline): data identification and acquisition (4 steps, shown in text boxes with blue margins), ethics and legal requirements (7 steps, shown in red), external data harmonization (1 step, shown in purple), and data storage and delivery (5 steps, shown in green). Dashed green boxes represent steps that were needed for this first pilot to build the system (European Molecular Biology Laboratory-European Bioinformatics Institute). Different types of data required different steps in this process, as shown for pseudonymized pathogen genetic sequence data (GSD), other laboratory datasets (other lab data), and clinical-epidemiological (CE) data. The identified barriers and enablers are indicated at the specific steps in the process that were mainly affected for our pilot; 2 enablers are not shown as these did not affect specific steps but acted as general facilitators of this process for our pilot. DAC: data access committee; EU: European Union; FAIR: Findable, Accessible, Interoperable, Reusable; GDPR: General Data Protection Regulation;



Mission (Im)possible: A Call for Action

The pilot demonstrated that sharing a linked cohort dataset using a centralized approach can succeed, but several crucial barriers (as described in Table 1) exist that cause long delays. Although these barriers were not unique to COVID-19 pandemic and have been extensively described previously [15-17], they persist today, as currently the burden to overcome or circumvent such barriers is placed on the individual researchers, projects, or institutes. However, the root causes of these barriers are systemic in nature. To alleviate these barriers, stakeholders at institutional, national and international or EU level (eg, governments, health care, research, journals, funders, and legal) need to work jointly toward building a model that not only supports but motivates data exchange by simplifying steps in the process of data sharing, taking privacy and legal aspects into account, and to reward researchers in new ways [18]. Below we describe some immediate actions to achieve this based on our experiences.

Support With Privacy Questions and General Data Protection Regulation

We recommend creating a roadmap amongst EU countries for how to deal with common issues, to avoid (recurrent) lengthy legal discussions. Harmonization of regional and national implementation of GDPR, and alleviation of the disproportionate burden in legal compliance activities for data sharing with international organizations, needs to be achieved [19]. Finally, additional support should be provided by the European Union to their funded projects, for example, by providing legal guidance or a top up for projects to hire legal experts.

Build and maintain FAIR Data Sharing During “Interpandemics”

Mandatory FAIRification of data in funded projects with funding for personnel to build and maintain FAIR data collections and FAIR sharing, and monitoring of and consequences for noncompliance. At the same time, more data repositories submitting to evaluation-based initiatives and accreditation processes [20-22], could help with a cultural change toward more FAIR data sharing. Furthermore, FAIRification and sharing of data should become standard practice during the education of bachelor, master, and PhD students.

Decrease Fragmented Funding Landscape

Consideration for funding interdisciplinary projects to enable combined data sharing and, continued funding support for existing data infrastructures is key to sustaining the efforts of data FAIRification and sharing, both nationally and internationally. One clear example is the COVID-19 open data portal [23], an initiative launched at the explicit wish from the European Commission. It provides access to open data from a range of areas of expertise for use across sciences (biomedical and social sciences). A second example is the newly launched World Health Organization (WHO) pandemic hub, to develop “collaborative surveillance,” a new area that seeks to bring together data collected by different areas of expertise. This includes the classical epidemiological data that are the core of public health surveillance, but also data on climate parameters, and data on factors that influence disease dynamics.

Willingness and Commitment Building (Culture Change)

We call for establishment of a common policy about data sharing allowing culture change at the EU or international level and for provision of resources to do so. One way of supporting culture change can be achieved by reflecting how researchers are recognized, rewarded, or assessed beyond publishing results in a journal. On short term, acknowledging data providers when reusing their datasets, and the use of citable data Digital Object Identifiers (DOI) pointing to their data could help (eg, as used for one of the datasets used in this pilot [24]). At institutional level, common policies about data sharing should also be established, and resources provided to implement these, to improve existing sharing norms and cultures.

Standardization and Interoperability

Institutes need to prioritize standardization or reconciliation of existing databases and systems to tackle data siloes. Overarching infrastructures at University Hospitals that would allow and enhance standardization, harmonization, and linking of different data types locally and nationally could also increase interoperability at EU and international level [25].

Discussion

Without action, barriers will persist, especially for sharing CE data, which hamper FAIR data sharing for infectious disease preparedness and response. The actions described here also contribute toward the commitment from closed to open research information, as put forward in the recent Barcelona Declaration [26]. The first and uppermost question needs to be answered at institutional and national level: Do people want to share data and what are they prepared to invest in terms of a collaborative effort to do so, if there was funding, GDPR and other support, for data sharing to move away from a felt “burden” to being a seen “benefit.” On the path toward more open data sharing researchers’ and other stakeholders’ concerns about sharing “their” data need to be addressed [15,27,28].

Meanwhile, decentralized or federated approaches are often preferred, as these circumvent barriers [16,29,30] and are less resource- and time intensive for (clinical) researchers and institutes. Federated networks, where decentralized but interconnected nodes allow data to be queried or analyzed by other nodes in the network without the data leaving its location, have been proposed as a solution and are being worked on to address the siloing of health data and current barriers to data sharing [31]. However, common obstacles to broader uptake of federated networks include the absence of data standards or limited adherence to existing standards, the complexity of designing, implementing and deploying federated solutions that preserve privacy [32]. The challenges must be addressed before federated networks can be implemented more widely, especially across national borders [31], but the question remains if full federation will ever be possible. Other concerns with federated approaches, such as the forthcoming European Health Data Space [33], are about significant delays in the authorization by national health data access bodies for reuse or secondary use of health data [34], and introduction of paywalls [35], potentially leading to inequity.

There is a need to streamline the timeliness of steps for sharing CE and other data, and to ensure no additional barriers are created in sharing a linked cohort dataset. Ongoing discussions to include digital sequence information, including pathogen GSD, under the Nagoya protocol [36], or the WHO CA+ [37], could potentially result in new barriers due to additional steps to complete legal procedures before pathogen GSD could be shared. This may considerably delay the prompt sharing and access to pathogen GSD, which has proven crucial in early detection, diagnostics, and identifying variants during the COVID-19 pandemic [2], while its impact on improving access and benefit sharing mechanisms remains uncertain [38].

The multidisciplinary nature of infectious disease research and collaborations reflects the complexity of research and public health questions which need to be understood and answered around (re-) emerging diseases. While certain answers can be found in isolation within a specific area of research and small numbers of patients, larger complex questions require a joint outlook and a combination of data sources to be examined together [39].

Conclusions

Here we have shown the complexities in sharing a small but complex, linked dataset from opportunistic COVID-19 clinical and laboratory research studies as done in (academic) hospital settings. We also provided actionable elements to shorten the timeline to go through this process. Actions that are taken now to improve data sharing for outbreak preparedness and response will improve our ability to detect and respond to emerging threats in the future. We call upon governments, funders, global and regional organizations, scientists and their institutes, journals and industries to tackle known barriers hampering data sharing for infectious disease preparedness and response.

Acknowledgments

The work described here would not have been possible without many other people within the Erasmus Medical Center (Erasmus MC) and the Reconciliation of Cohort data in Infectious Diseases (ReCoDID) consortium; we are grateful for your dedication, effort, and for supporting the performance of this study. Specifically, we would like to thank the following people. At Erasmus MC: Pieter Fraaij and Jeroen van Kampen for support with the data landscape; Divyae Prasad and Marie-Catherine Bouquieaux for SARS-CoV-2 data submissions; Mickey Haasnoot and Laurens Groenendijk for support with the legal requirements; Tamara Van Huuksloot, Viola Woeckel, Anne-Lotte Masson and Kenneth Davids for support with the institutional database. At European Molecular Biology Laboratory-European Bioinformatics Institute: Nadim Rahman and Zahra Waheed for support with the European Nucleotide Archive; Dipayan Gupta for support with BioSample; and the ArrayExpress/BioStudies and European Genome-Phenome Archive teams. At Heidelberg University Hospital (UKHD; Germany): Fruzsina Molnar-Gabor for legal support concerning the General Data Protection Regulation. At McMaster University (Canada): Priscilla Cesar and Claudia Emerson for the overall RECoDID Data Governance framework. Financial support for this work was provided by the European Union’s Horizon 2020 research and innovation program (grant agreements 825746 [ReCoDID] and 874735 [VEO]). The funders had no role in any aspect pertinent to the study, nor in the decision to submit the manuscript. A preprint of this manuscript has been published by Zenodo.

Authors' Contributions

Although the datasets themselves and their connected data analysis are not the focus of this manuscript, the pilot study, to share a complex linked COVID-19 cohort dataset, would not have been possible without the respective data providers. Data providers were BBOM, RDdV, GPvN, RSS, JvB, and ECMvG. FT carried out clinical-epidemiological data harmonization. Funding acquisition was done by TJ, MPGK, MYvR, and GC. CA and GR carried out the project administration (management and coordination responsibility for the pilot planning and execution at Erasmus Medical Center and European Molecular Biology Laboratory-European Bioinformatics Institute respectively). The project was conceptualized and the manuscript conceived by MYvR, CA, and MPGK. Writing-original draft was carried out by CA and MYvR, and Writing-review and editing by all authors. Figures were designed by MYvR and GR. Visualization (preparation, creation and presentation of work at consortium meetings) was done by CA and MYvR. CA and MYvR contributed equally to this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary material - timeline.
[DOCX File, 456 KB - publichealth_v11i1e63996_app1.docx]

References

1. Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 2016 Mar 15;3:160018. [doi: 10.1038/sdata.2016.18] [Medline: 26978244]
2. Oude Munnink BB, Nieuwenhuijse DF, Stein M, et al. Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. N Med 2020 Sep;26(9):1405-1410. [doi: 10.1038/s41591-020-0997-y] [Medline: 32678356]
3. Reconciliation of Cohort data in infectious diseases. ReCoDID. URL: https://recodid.eu [accessed 2024-04-29]
4. Amid C, Pakseresht N, Silvester N, et al. The COMPARE data hubs. Database (Oxford) 2019 Jan 1;2019:baz136. [doi: 10.1093/database/baz136] [Medline: 31868882]
5. Rahman N, O’Cathail C, Zyoud A, Sokolov A, Oude Munnink B, Gruning B, et al. Mobilisation and analyses of publicly available SARS-CoV-2 data for pandemic responses. Microb Genom 2024;10(2). [doi: 10.1101/2023.04.19.537514]
6. Aguilar-Bretones M, Westerhuis BM, Raadsen MP, et al. Seasonal coronavirus-specific B cells with limited SARS-CoV-2 cross-reactivity dominate the IgG response in severe COVID-19. J Clin Invest 2021 Nov 1;131(21):e150613. [doi: 10.1172/JCI150613] [Medline: 34499051]
7. Lu L, Sikkema RS, Velkers FC, et al. Adaptation, spread and transmission of SARS-CoV-2 in farmed minks and associated humans in the Netherlands. Nat Commun 2021 Nov 23;12(1):6802. [doi: 10.1038/s41467-021-27096-9] [Medline: 34815406]
8. COMPARE. Collaborative management platform for detection and analyses for (re-)emerging and foodborne outbreaks in Europe. URL: https://www.compare-europe.eu/ [accessed 2024-01-17]
9. International severe acute respiratory and emerging infection consortium. ISARIC. URL: https://isaric.org/research-2/covid-19-clinical-research-resources/covid-19-crf/ [accessed 2024-01-09]
10. GISAID. URL: https://gisaid.org [accessed 2024-01-16]

11. Cohort Browser Infectious disease cohort studies shared at EMBL-EBI. URL: https://www.pathogensportal.org/cohorts?activeTab=Dashboard&acc=BSC0000005 [accessed 2024-04-29]
12. Courtot M, Gupta D, Liyanage I, Xu F, Burdett T. BioSamples database: FAIRer samples metadata to accelerate research data management. Nucleic Acids Res 2022 Jan 7;50(D1):D1500-D1507. [doi: 10.1093/nar/gkab1046] [Medline: 34747489]
13. Stroe O. First linked dataset on Pathogens Portal: EMBL-EBI. 2022. URL: https://www.ebi.ac.uk/about/news/updates-from-data-resources/pathogens-portal-linked-dataset/ [accessed 2022-12-20]
14. BioSamples, EMBL-EBI. Search results for ReCoDID COVID-19 pilot study. URL: https://www.ebi.ac.uk/biosamples/samples?text=ReCoDID%20COVID-19%20pilot%20study [accessed 2023-12-20]
15. Ribeiro CDS, Roode MY, Haringhuizen GB, Koopmans MP, Claassen E, Burgwal LHM. How ownership rights over microorganisms affect infectious disease control and innovation: A root-cause analysis of barriers to data sharing as experienced by key stakeholders. PLoS ONE 2018;13(5):e0195885. [doi: org/10.1371/journal.pone.0195885]
16. Bernier A, Molnár-Gábor F, Knoppers BM, et al. Reconciling the biomedical data commons and the GDPR: three lessons from the EUCAN ELSI collaboratory. Eur J Hum Genet 2024 Jan;32(1):69-76. [doi: 10.1038/s41431-023-01403-y] [Medline: 37322132]
17. Vlahou A, Hallinan D, Apweiler R, et al. Data sharing under the general data protection regulation: time to harmonize law and research ethics? Hypertension 2021 Apr;77(4):1029-1035. [doi: 10.1161/HYPERTENSIONAHA.120.16340] [Medline: 33583200]
18. European Commission, Directorate-General for Research and Innovation, Maxwell L. Maximising investments in health research – FAIR data for a coordinated COVID-19 response– Workshop report: Publications Office of the European Union. 2022 URL: https://data.europa.eu/doi/10.2777/726950
19. Science europe feedback on the implementation of the GDPR. Science Europe. 2024 Feb 16. URL: https://scienceeurope.org/media/5blavcwx/se_response_ec_consultation_gdpr_implementation_final.pdf
20. Lin D, Crabtree J, Dillo I, et al. The TRUST Principles for digital repositories. Sci Data 2020 May 14;7(1):144. [doi: 10.1038/s41597-020-0486-7] [Medline: 32409645]
21. ELIXIR core data resources: ELIXIR Europe. ELIXIR. URL: https://elixir-europe.org/platforms/data/core-data-resources/
22. Resources, GCB. Global Core Biodata Resources: Global Core Biodata Resources. URL: https://globalbiodata.org/what-we-do/global-core-biodata-resources
23. COVID-19 Data Portal - accelerating scientific research through data. URL: https://www.covid19dataportal.org/ [accessed 2024-11-25]
24. Oude Munnink B. ND, Sikkema R, Schapendonk C, et al. On behalf of the Dutch national COVID-19 response team SARS-CoV-2 whole-genome sequencing for informed public health decision making and outbreak tracking and tracing in the Netherlands. URL: https://www.ebi.ac.uk/biostudies/studies/S-BSST6622021. https://doi.org/10.6019/S-BSST662
25. Rinaldi E, Stellmach C, Rajkumar NMR, et al. Harmonization and standardization of data for a pan-European cohort on SARS- CoV-2 pandemic. NPJ Digit Med 2022 Jun 14;5(1):75. [doi: 10.1038/s41746-022-00620-x] [Medline: 35701537]
26. Barcelona declaration on open research information. 2024. URL: https://barcelona-declaration.org/ [accessed 2024-06-01]
27. Federer LM, Lu YL, Joubert DJ, Welsh J, Brandys B. Biomedical data sharing and reuse: attitudes and practices of clinical and scientific research staff. PLoS ONE 2015;10(6):e0129506. [doi: 10.1371/journal.pone.0129506] [Medline: 26107811]
28. van Roode MY, Dos S Ribeiro C, Farag E, et al. Six dilemmas for stakeholders inherently affecting data sharing during a zoonotic (re-)emerging infectious disease outbreak response. BMC Infect Dis 2024 Feb 12;24(1):185. [doi: 10.1186/s12879-024-09054-0] [Medline: 38347527]
29. Peñalvo JL, Mertens E, Ademović E, et al. Unravelling data for rapid evidence-based response to COVID-19: a summary of the unCoVer protocol. BMJ Open 2021 Nov 18;11(11):e055630. [doi: 10.1136/bmjopen-2021-055630] [Medline: 34794999]
30. Tacconelli E, Gorska A, Carrara E, et al. Challenges of data sharing in European Covid-19 projects: a learning opportunity for advancing pandemic preparedness and response. Lancet Reg Health Eur 2022 Oct;21:100467. [doi: 10.1016/j.lanepe.2022.100467] [Medline: 35942201]
31. Hallock H, Marshall SE, ’t Hoen PAC, et al. Federated networks for distributed analysis of health data. Front Public Health 2021;9:712569. [doi: 10.3389/fpubh.2021.712569] [Medline: 34660512]
32. Casaletto J, Bernier A, McDougall R, Cline MS. Federated analysis for privacy-preserving data sharing: a technical and legal primer. Annu Rev Genomics Hum Genet 2023 Aug 25;24:347-368. [doi: 10.1146/annurev-genom-110122-084756] [Medline: 37253596]
33. European Parliamentary Research Service (EPRS). European health data space: European parliament. 2023 URL: https://www.europarl.europa.eu/RegData/etudes/ATAG/2023/754642/EPRS_ATA(2023)754642_EN.pdf
34. Lähteenmäki J, Vuorinen AL, Pajula J, et al. Integrating data from multiple Finnish biobanks and national health-care registers for retrospective studies: Practical experiences. Scand J Public Health 2022 Jun;50(4):482-489. [doi: 10.1177/14034948211004421] [Medline: 33845693]
35. FINDATA social and health data permit authority pricing: FINDATA social and health data permit authority. URL: https://findata.fi/en/pricing [accessed 2024-05-24]

36. Convention on Biological Diversity (United Nations) Nagoya protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilization to the convention on biological diversity. URL: https://www.cbd.int/abs/doc/protocol/nagoya-protocol-en.pdf [accessed 2023-12-22]
37. World Health Organization (WHO). Bureau's text of the WHO convention, agreement, or other international instrument on pandemic prevention, preparedness and response (WHO CA+) for the consideration of the Intergovernmental Negotiating Body at its fifth meeting: World Health Organization (WHO). URL: https://apps.who.int/gb/inb/pdf_files/inb5/A_INB5_6-en.pdf [accessed 2023-06-02]
38. DSI Scientific Network Using digital sequence information (DSI) to design an mRNA vaccine: COVID-19 case study: DSI Scientific Network. 2023. URL: https://www.dsiscientificnetwork.org/wp-content/uploads/2023/04/DSI-covid-vaccines-factsheet.pdf
39. Eckhardt M, Hultquist JF, Kaake RM, Hüttenhain R, Krogan NJ. A systems approach to infectious disease. Nat Rev Genet 2020 Jun;21(6):339-354. [doi: 10.1038/s41576-020-0212-5] [Medline: 32060427]

Abbreviations

CE: clinical-epidemiological
EMBL-EBI: European Molecular Biology Laboratory-European Bioinformatics Institute
Erasmus MC: Erasmus Medical Center
FAIR: Findable, Accessible, Interoperable, Reusable
GDPR: General Data Protection Regulation
GSD: genetic sequence data
ReCoDID: Reconciliation of Cohort data in Infectious Diseases
UKHD: Universitätsklinikum Heidelberg (University Hospital Heidelberg)
WHO: World Health Organization

Edited by A Mavragani; submitted 05.07.24; peer-reviewed by LN Mayigane, S Pesälä; revised version received 28.11.24; accepted 01.12.24; published 11.02.25.

Please cite as:
Amid C, van Roode MY, Rinck G, van Beek J, de Vries RD, van Nierop GP, van Gorp ECM, Tobian F, Oude Munnink BB, Sikkema RS, Jaenisch T, Cochrane G, Koopmans MPG
A Call for Action: Lessons Learned From a Pilot to Share a Complex, Linked COVID-19 Cohort Dataset for Open Science
JMIR Public Health Surveill 2025;11:e63996
URL: https://publichealth.jmir.org/2025/1/e63996
doi:10.2196/63996

Ethical Considerations for Wastewater Surveillance Conducted by the US Department of Defense

Hunter Jackson Smith[1,2], MD, MPH, MBE; Richard T Agans[3], PhD; William J Kowallis[4,5], PhD

[1]Global Emerging Infections Surveillance Branch, Armed Forces Health Surveillance Division, 11800 Tech Rd Suite #200, Silver Spring, MD, United States
[2]Johns Hopkins Berman Institute of Bioethics, Baltimore, MD, United States
[3]Applied Technology and Genomics Division, Defense Centers for Public Health, Dayton, OH, United States
[4]Department of Emerging Biological Threats, Defense Centers for Public Health, Aberdeen, MD, United States
[5]Department of Computational and Chemical Sciences, Carlow University, Pittsburgh, PA, United States

Corresponding Author:
Hunter Jackson Smith, MD, MPH, MBE
Global Emerging Infections Surveillance Branch, Armed Forces Health Surveillance Division, 11800 Tech Rd Suite #200, Silver Spring, MD, United States

Abstract

The US Department of Defense (DoD) is establishing its wastewater surveillance capacities to support national security objectives and promote the public health and medical readiness of US service members. Wastewater surveillance is an emerging technology that has traditionally been leveraged for detecting infectious diseases. However, its potential future applications could yield a vast and unpredictable amount of information that could be used for a wide variety of both health- and nonhealth-related purposes. The US military also serves an inimitable role for the country and its citizens, and exercises significant levels of control over its service members compared to civilian organizations. Further, its present and potential wastewater surveillance activities may reach far beyond just military installations. These factors raise unique ethical considerations that must be accounted for by leaders and policymakers to ensure the DoD implements a wastewater surveillance network in a manner that is both impactful in supporting public health and appropriate to the scope and population under surveillance. This paper explores important ethical features in conducting wastewater surveillance that are both specific to the DoD experience and applicable to wider public health initiatives.

(JMIR Public Health Surveill 2025;11:e67145) doi:10.2196/67145

KEYWORDS

wastewater; surveillance; ethics; military; Department of Defense

Introduction

Wastewater represents a promising source for obtaining valuable public health data to inform decision-making and response efforts [1]. Wastewater testing includes the tools and technologies used for various activities (eg, research, law enforcement, public health), while wastewater surveillance (WWS) is the activity of assessing biological, chemical, and analyte signatures for determining population-level trends to inform public health decision-making [2-5]. Traditionally, WWS has been leveraged to monitor for select infectious disease pathogens, such as poliovirus and more recently, SARS-CoV-2 during the COVID-19 pandemic [2-4,6]. However, wastewater testing has been used in recent years to detect numerous other targets, ranging from dietary patterns to the genetic makeup of communities [2,3,6-10]. Wastewater surveillance is an appealing public health instrument due to its scalability, cost-effectiveness compared to clinical monitoring, and the ability to detect surges in infectious disease cases, days before they become clinically relevant [11,12]. Recognizing the potential utility and value of WWS, the Centers for Disease Control and Prevention led the development of the National Wastewater Surveillance System. This system allows environmental scientists and public health departments across the country to track specific pathogens, including SARS-CoV-2, influenza A, respiratory syncytial virus, and mpox and report data into the National Wastewater Surveillance System to inform public health decision-making [13]. This system currently collects data from over 1000 sites across the United States, representing 35% of the US population, and it continues to add collaborating sites and other pathogens of public health concern to its surveillance list [13]. There are also plans by researchers and private corporations to further expand the detection capabilities of wastewater testing technologies to include illicit drug use, noncommunicable diseases, genetic information, and other targets. There are meaningful differences in the ethical considerations between wastewater testing and WWS, partly because wastewater testing has broader applications, such as research, law enforcement, and espionage [14,15]. However, this paper focuses on WWS, addressing wastewater testing only as it pertains to WWS as a public health activity.

Recently, the Department of Defense (DoD) has started prioritizing wastewater testing and surveillance, as highlighted in the emerging capabilities and technologies section of the 2023 Biodefense Posture Review [16]. It currently oversees multiple individual WWS projects within the United States and overseas, partners with other government agencies and private industries to explore and expand its wastewater testing technologies, and is building the capacity to develop and manage a DoD WWS network. It has already invested tens of millions of dollars toward advancing its WWS efforts. The DoD plays a distinctive role in ensuring social security and maintains a unique relationship with its ervice members (SM). A DoD WWS network has a potentially vast global scope, covering diverse civilian and military populations [17]. Given (1) the DoD's unique population and security considerations, (2) the rapid evolution and promising applications of wastewater testing technologies, and (3) the significant unknowns regarding the potential future uses of these technologies, there is an urgent need to address the ethical considerations inherent to implementing a WWS network for the DoD. These efforts are imperative to ensure it is executed appropriately and has suitable policies in place for management.

It is critical that environmental scientists, policymakers, and those engaged in WWS must be aware of the current and potential ethical issues related to wastewater testing and surveillance. Policy development, particularly for emerging technologies, must be focused on ethical implications and should aim to provide appropriate guidance and boundaries. This paper explores the ethical features of WWS and offers recommendations to address these concerns— that are important for any group planning to implement a WWS system. While many ethical considerations discussed in this paper are universal to any WWS system, their application within the DoD (whether in identity, scope, or importance) compared to other contexts is unique and deserves particular attention.

Targets of Wastewater Surveillance

The scientific evidence and methodology for wastewater sampling and testing must be established for each pathogen or target under consideration for surveillance. Pursuing large-scale sampling without a sufficient understanding of the target would be both ineffective and an unethical use of resources. For infectious diseases, this might involve pathogen-specific knowledge of (1) sampling and detection methods; (2) laboratory testing, genomic sequencing, and analysis; (3) relationship between detection levels in wastewater and clinical detection in humans; and (4) methods to mitigate the spread of the disease. A key component of conducting ethical public health activities includes ensuring the surveillance and responses to the surveillance findings are done in effective, scientifically sound, and evidence-based ways. The negative opportunity costs of inefficiently utilizing public health funds could harm other health areas where funds could have otherwise been allocated. Addressing this concern should include pilot testing of new pathogen targets or wastewater testing technologies as well as cost-effectiveness analyses prior to full integration into a WWS system [18]. For example, if a pilot project discovered that a new potential feature of the WWS system only marginally reduced infectious disease transmission at a significant cost, then those funds might be better directed to more effective public health activities.

Kass [19] describes an ethical framework for the assessment of public health programs and emphasizes that one of the key aspects of an ethically conducted program includes its effectiveness in achieving the stated goals. Relatedly, monitoring and program evaluation should be incorporated into the workflow of a public health program to ensure alignment with its goals [20]. For WWS, this could involve monitoring to determine whether WWS programs are effective at limiting infectious disease transmission among communities where they are being implemented, compared to those without such programs. Additionally, as WWS technologies and methodologies are developed and refined, efforts should be made to assess for concordance between WWS findings and clinical or on-ground surveillance data. Thus, an ethical WWS program must employ scientifically sound sampling and analysis methods and demonstrate effectiveness in achieving its goals, supported by clinical and epidemiologic evidence.

Currently, most WWS programs are focused on infectious diseases. However, there are other potential targets that have been explored beyond infectious pathogens (eg, noncommunicable diseases such as cardiovascular disease and cancer). While it is difficult to predict what technologies will become available in the future to assess WWS [21,22], many of the core ethical issues at present will remain relevant regardless of the future technological capability. The DoD leadership should be cognizant of preventing excessive mission creep (ie, broadening of the original objective) [23,24]. Using WWS for purposes beyond promoting and preserving health can be dangerous and cause ethically fraught scenarios. For example, while opioids and other substances of abuse (eg, methamphetamines, cocaine, nicotine) could be monitored via WWS, it raises equity concerns and risks of overpolicing [25]. Although the DoD already implements randomized drug screening for its SMs, WWS could introduce additional equity concerns. For instance, many higher-ranking personnel and officers typically live off-base, but the DoD would more accessibly conduct WWS in on-base areas where lower-ranking personnel primarily reside. This disparity might result in military leadership focusing punitive actions on lower-ranking enlisted personnel in an inequitable, unjustified manner, despite the possibility that higher-ranking individuals living off-base might use illicit substances in equal measures but remain undetected. This inequity would be particularly problematic considering that WWS captures all individuals using the sanitation system, including SMs, families, civilian employees, contractors, and visitors. In the future, WWS could also be used to detect medication-related metabolites (eg, reproductive health medications), human genetic content (eg, racial or ethnic composition of communities), and other potentially sensitive or politically contentious targets. It is important to thoughtfully consider what may become possible for WWS capabilities and to ensure that DoD leaders and policymakers remain vigilant in keeping its focus on appropriate health targets and public health–promoting outcomes.

Populations and Areas of Wastewater Surveillance

Where should the DoD focus its wastewater surveillance efforts? This question is practically, legally, and ethically important because there are few areas where WWS could be conducted with only DoD SMs present. Understanding the size and vulnerability of the populations targeted by WWS can influence the interpretation of and responses to the data produced by it [26]. Thus, exploring where the DoD should conduct WWS naturally leads to the question: on whom should the DoD focus its WWS efforts?

On-base DoD locations and DoD transport vessels are obvious areas to focus WWS efforts. Four primary location types can be considered for on-base WWS: (1) facilities, (2) clinics and hospitals, (3) recruit and trainee sites, and (4) housing. On-base DoD locations refer to areas both within the United States and abroad, including permanent international installations and temporary bases established during conflicts. On-base facilities (both medical and nonmedical) and recruit and trainee sites are the least ethically problematic locations, as these areas typically constitute force health protection–relevant activities and have the highest concentration of SMs. While DoD civilian and contractor personnel also work in these areas, conducting WWS at these locations is fully within the DoD's responsibility and legal jurisdiction to support and promote the health of DoD SMs as well as civilian and contractor personnel. Military transport vessels (eg, large aircraft and naval ships) may prove to be particularly important catchment points for sentinel surveillance, especially for tracking the transmission of infectious diseases [27-29].

On-base housing is another force health protection–relevant area, as these are locations where SMs and those closest to them spend most of their off-duty time. An SM's family often lives on-base with them, and the family's health can directly impact the health of SMs and their units. Further, infectious diseases present among SM families are also present among SMs; therefore, conducting surveillance of beneficiaries can better inform the epidemiologic picture of infectious disease prevalence and transmission within military units. Thus, there is a high probability of detecting pathogens of DoD significance in these areas. Moreover, the DoD has a precedent of assuming some fiduciary responsibility for promoting the health and well-being of SM families, through a variety of clinical and social support systems (eg, Tricare health care coverage, paternity/maternity leave policy, education options, free counseling services, housing support, wellness centers). It is also impossible to extricate on-base housing WWS sampling of SMs from family members who live with them. Therefore, it is within the DoD's purview to promote SM families' health, and if the WWS is conducted in a minimally invasive, deidentified, and ethically sound manner, WWS could be a valuable way to further support these health objectives.

The DoD also has a vested interest in WWS abroad; these international locations can be grouped into three categories: (1) joint military exercises, (2) ally military installations, and (3) foreign areas of mutual interest with the host nation. There should be two key underlying requirements for all international locations (beyond DoD installations located abroad): (1) the DoD must obtain the permission of the host nation in which they are guests to conduct WWS and (2) the WWS should provide some benefit to that host nation. Ensuring these criteria are fulfilled will promote an equitable relationship between nations and help avoid exploitative or colonialist interpretations of mutually beneficial public health activities. Another critical consideration for the DoD in conducting international WWS is that the host nation should be empowered to exercise as much sovereign control over the public health activities occurring in its own territory as it sees fit. The DoD and the host nation may both have interests in conducting WWS in certain locations in that country; however, the host nation should decide if and how WWS should proceed. The DoD's role should be supportive, which can manifest in a variety of ways depending on the context and needs of the host nation, including leading the WWS project, funding host nation researchers, providing methodologic input and subject matter expertise, or developing technical and analytical capabilities. Challenges may arise if an ally nation requests WWS support that the DoD is not positioned to support, either on principle or due to resource constraints. International scientific collaborations require agreement from all involved parties. If the DoD does not want to engage, then they are not required to do so, particularly if the DoD has ethical reservations regarding an expansion of the ally nation's surveillance activities to include controversial targets or populations. If the DoD is supportive of the request but is resource constrained, then it can offer support in other ways (eg, technical consultation, references to other potential collaborators). If the foreign nation's request involves surveillance targets that are of importance to them but not of interest to the DoD, then the DoD can consider offering low-cost support (eg, information, consultation) in the spirit of preserving and strengthening the relationship. Generally, approaching host nations from the perspective of support, providing the host nation opportunities to lead decision-making, and emphasizing the mutually beneficial aspects of the collaboration will support an equitable partnership.

Joint military exercises serve as collaborative training opportunities for two or more militaries to build cohesiveness, promote mission readiness, and grow camaraderie among allies. In such scenarios, it is acceptable for either the DoD or other militaries to lead WWS efforts, provided all the participating militaries whose wastewater is sampled endorse their permission for WWS to be conducted during the exercise, and for the data to be shared with the sampled parties that participated, if requested. This practice supports the participating military's autonomy by providing an opportunity for group informed consent, whereby a nation's military can provide permission or denial for participation of its SMs. For WWS at ally military health care facilities and installations, it is appropriate for the ally nation's military to lead, with the DoD providing support these efforts as defined by what is agreed upon between the two militaries. The flow of data should be bidirectional and beneficial for both parties for the endeavor to be equitable. For international civilian settings, such as health care facilities, the ally nation's relevant health organizations (eg, ministries of health, health research institutes) should serve as key partners, and the DoD should support and coordinate with them in the conduct of the surveillance and sharing of data. The host nation's organization should make efforts to ensure awareness of WWS efforts within the communities involved, especially in specific locations such as health care facilities, and that the facility's leadership is aware, and approves of the WWS.

Management of Wastewater Surveillance Data

Once the WWS has been conducted, how should data be managed? The DoD faces important ethical considerations regarding the type of data to be collected and its management. A balance must be struck between allowing beneficial data access and ensuring data protection. WWS can generate a variety of data, including molecular and biological signatures, health and disease markers, biochemical indicators of illicit activities, and associated identifiable demographic metadata [17]. While possessing large amounts of data is typically beneficial for determining appropriate responses, course of actions, and preventive planning, protecting these data is crucial, especially as the DoD arguably experiences more frequent and powerful attacks on its information security infrastructure compared to other organizations. Therefore, managing the focus on which data to protect (eg, personally identifiable information vs aggregate data) and how to protect it (eg, cloud or servers vs hardline firewalls) will be crucial for ensuring minimized risk of data weaponization or leakage. Wastewater surveillance can produce highly granular data that can be used by public health researchers and officials for improving the quality of the communities they serve; however, this level of detail is not always essential when communicating with end users. Establishment of well-defined data-sharing agreements which incorporate ethical guidelines and treat WWS data with the same scrutiny as Controlled Unclassified Information or Classified determinations, wherein raw information sharing is conducted only as needed, will aid in reducing improper use of WWS information. Data security can be handled in a similar manner by escalating access and storage restrictions corresponding to data aggregation and specificity. For instance, aggregated, community-level detection data of public health targets could be stored on shared drives with minimal restrictions; however, individual molecular and genetic data could be stored behind hardline firewalls on dedicated IT systems. Anonymization and the use of beacons are additional potential methods that could help in ensuring individual human subjects cannot be identified by WWS; however, both methods have their inherent limitations [30]. Nonetheless, it is likely that data security approaches will not lend to one size fits all solutions and will be dictated by the surveillance effort, parties involved, and required outcomes. The DoD must take strong measures to ensure only authorized personnel access appropriate levels of WWS data.

As WWS data becomes more granular and specific, the ethical responsibility to ensure data security and appropriate use and access increases proportionally. This escalation of responsibility in proportion to the level of granularity of the data is derived from the increasing probability of it being personalized and potentially damaging to individuals or small groups. For example, depending on where the wastewater is sampled, it could be possible to pinpoint illicit drug use or the location of an individual to a specific household. In fact, analyzing local sewage for genetic markers was one of the methods potentially used by the United States to confirm Osama bin Laden's location prior to him being killed in 2011 [15]. It may also become possible in the future to obtain personal genetic information through WWS and to associate it with personally identifiable information; the DoD must be prepared to either dispose or strongly protect this data. There are significant risks if foreign adversaries gain access to this level of data, as well as real concerns regarding how this information could be misused domestically in politically motivated ways. One way to combat these concerns is to establish clear sampling strategies with requirements to destroy unused sample portions once the appropriate predetermined targets have been tested or detected.

There are also important considerations regarding the sharing of WWS data when it has broader implications beyond the DoD, and is relevant to other public health entities, such as US government groups (eg, local health departments, the Centers for Disease Control and Prevention), ally nations, or nongovernmental organizations (eg, the Bill and Melinda Gates Foundation). There are certain ethical responsibilities in the sharing of appropriate levels of data when it can also be of benefit to others, and the DoD should take steps to collaborate with these groups to promote health security. Thus, taking all these considerations into account, one overarching goal for a DoD WWS network is to ensure the data is maintained at an aggregate level as much as possible for monitoring of community and population-level trends rather than specific information on individuals or small groups.

Utilization and Response to Wastewater Surveillance Data

Public health surveillance data should not be collected unless there is a potential for a meaningful and actionable response. If surveillance findings are not translated to public health action or policy changes, the benefits accrued from the surveillance efforts are minimal [31]. Collecting data only for the sake of collection is an inappropriate use of resources; therefore, there should be focused and organized plans in place for using the WWS systems and information derived from them prior to their implementation. Additionally, there should be rigorous discussions about how military leaders and public health officials respond to such information as it becomes available to them. There are innumerable ways in which the leadership could respond to WWS data when it is received in near real-time. While this information offers an opportunity to ethically use that data, there is a risk of inappropriate response to it [32]. Notably, military commanders make the final decisions regarding public health actions or interventions, particularly those of consequence. It is likely that a medical or public health professional will not be making the decision, and the military commander may have no background in science, medicine, or public health. While making the intervention decisions in response to WWS data, these commanders and their leadership teams should lean heavily upon the expert advice of medical and public health professionals to inform their decisions to help ensure they take the most appropriate actions. The key consideration here is that the responses—whether determined by military command or leadership from preventive medicine, public health, or clinical groups—should be appropriate, proportional, and context- and community-specific based on the information received through the WWS system.

General Ethical Considerations

Finally, there are unique ethical considerations for the DoD that civilian organizations do not necessarily face to the same extent. The military has substantial authority, control, and access over the lives of its SMs. It can dictate aspects ranging from what to wear, where to live, how to behave, and nearly every aspect of SMs' lives. Given the extent of this control and its durability and depth, the DoD has an elevated ethical responsibility toward the health and well-being of its SMs [33]. Further, there is a public health ethics principle, which argues that a public health program should use the least invasive or minimally burdensome (to the population being targeted by the program) methods to accomplish its goals [16,34]. Service members already forgo significant autonomy afforded to civilians; therefore, any public health program that might additionally strip away their autonomy must be seriously deliberated before implementation. Although WWS may not directly interfere with the lives of SMs while the data is collected, the actions related to WWS (eg, how leadership responds to the information, what data is collected, and how the data is managed) could infringe upon their privacy and autonomy. However, if conducted appropriately, the potential health benefits for the SMs would likely outweigh minor privacy or autonomy concerns. Additionally, the implications extend beyond military populations when considering the health of SMs. The mission of the DoD and SMs is to protect US citizens, interests, and national security. If SMs are unwell, experience disease outbreaks, or are targeted using bioweapons, it places the country at risk, as they will be less able to support their security objectives by failing to achieve optimal operational capabilities and mission readiness. Furthermore, SMs and military installations do not exist in isolation. If an infectious disease emerges among a military population in a certain area, it is likely there is a similar problem in the surrounding civilian community, or at least presents a risk to them. Therefore, these considerations add to the elevated responsibility of the DoD to support and promote the health of its SMs.

There are also several broad ethical considerations for how the DoD conducts WWS. There is a tension underlying many public health actions, between preserving personal freedom and autonomy of individuals and maximizing the public health benefits for the group. This tension is manifested in several ways, including the amount and types of data to be collected, proportionality and appropriateness of response to the data, level of invasiveness of the surveillance, populations, and areas selected for surveillance [35,36]. DoD leadership should balance protecting the freedom and autonomy of its SMs while also prioritizing the health of the force from a population perspective [33]. Additional resources providing ethical guidance for the operation of public health surveillance systems that are applicable for WWS are available from the American Public Health Association's Public Health Code of Ethics and the World Health Organization's Guidelines on Ethical Issues in Public Health Surveillance [34,37].

Justice is another key ethical concept, and when appraising a DoD WWS program, two relevant perspectives must be considered: (1) distributive justice and (2) procedural justice. Distributive justice is an ethical principle concerned with the fair distribution of benefits and burdens. In the DoD WWS context, it involves ensuring that the distribution of resources, outcomes, and interventions based on the surveillance findings is done in an equitable way (eg, distributing vaccines based on WWS findings of acute elevations in vaccine-preventable diseases). It also includes the fair distribution of resources for conducting WWS throughout the United States and international installations. Procedural justice is an ethical principle concerned with fair processes and methods. It includes features such as inclusion, transparency, neutrality, respect, and trust [38]. For DoD WWS, it might involve transparent communications and decision-making, community considerations, stakeholder participation, access to relevant data for the surveilled communities, among others.

There are a variety of other general ethical considerations that are applicable to WWS, which are briefly mentioned here and should be explored further. First, there are differences between conducting public health surveillance versus research and the varying regulatory, oversight, and ethical requirements. Second, there may be differences in what actions are permissible when considering baseline surveillance versus those needed during public health emergencies. Finally, there are inevitable future possibilities beyond our current capabilities or knowledge regarding what might become possible through conducting WWS. What might WWS be capable of in the future, and how do we anticipate and plan for that? Although there are myriad ethical considerations for the appropriate conduct of WWS by the DoD both currently and in the future, it is undoubtedly a burgeoning area of public health that the DoD should pursue and be engaged in a substantial, effective, and meaningful manner, as there are innumerable potential benefits to a DoD WWS network.

Conclusions

Wastewater surveillance offers tremendous promise and opportunities to advance public health by enhancing disease surveillance, informing public health and clinical decision-making, and improving public health response times to mitigate negative health impacts to populations [39,40]. The DoD should adopt WWS and position itself to take advantage of the current and potential future benefits. It should also recognize the practical and ethical differences between using wastewater testing for public health surveillance, research, or other potential military purposes. With over 1.3 million active-duty SMs, 1.1 million National Guard and Reserve members, and 700,000 civilian personnel, the DoD is the largest organization in the United States, and has responsibilities to their health and readiness [41]. Further, the Military Health

System serves a total of approximately 9.5 million beneficiaries, which includes both SMs and their families [42]. WWS has the potential to benefit not only these populations, but others in the United States and overseas as well. However, these opportunities also present a potential for misuse, which could be detrimental to those being surveilled, raising important ethical questions about how and why WWS should be conducted. The DoD also has elevated ethical obligations to the well-being of its SMs due to the breadth and depth of control it exercises over them. Although ethical issues may arise related to WWS, these issues can and should be addressed through proper deliberation and policy implementation, as the potential public health benefits could be substantial. The DoD has responsibilities to protecting the national security of its citizens broadly, which implies a crucial role in infectious disease surveillance for outbreaks as well as bioengineered and emerging pathogen threats. Therefore, the DoD should pursue wastewater surveillance in an ethically appropriate, scientifically robust, and thoughtfully planned and executed manner to guide public health decision-making, protect the well-being of its SMs, and support US government health security objectives.

Acknowledgments

The authors would like to express their gratitude to Col Brett Nishikawa, CDR Matthew Kasper, Lt. Col Mark Lehman, Dr. Lindsay Morton, Dr. Seth Faith, Dr. Anthony Fries, Dr. Blake Stamps, Ms. Genny Cook, Dr. Liza Dawson, MAJ Jennifer Jackiw, MAJ Matthew Angel, Ms. Heidi Taylor, Dr. Subrahmanyam Yerramilli, Dr. Georgia Millward, and 1LT Daniel Evans for their thoughtful reviews and feedback on the paper.

Disclaimer

The Defense Health Agency has reviewed this material, and there is no objection to its publication. The contents, views or opinions expressed in this publication are those of the authors and do not necessarily reflect official policy or position of the Armed Forces Health Surveillance Division, the Defense Centers for Public Health-Aberdeen, the Defense Centers for Public Health-Dayton, the Defense Health Agency, the Department of Defense, or Departments of the Army, Navy, or Air Force. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

Conflicts of Interest

None declared.

References

1. Diamond MB, Keshaviah A, Bento AI, et al. Wastewater surveillance of pathogens can inform public health responses. Nat Med 2022 Oct;28(10):1992-1995. [doi: 10.1038/s41591-022-01940-x]
2. Huizer M, Ter Laak TL, de Voogt P, van Wezel AP. Wastewater-based epidemiology for illicit drugs: a critical review on global data. Water Res 2021 Dec 1;207:117789. [doi: 10.1016/j.watres.2021.117789] [Medline: 34731667]
3. Kilaru P, Hill D, Anderson K, et al. Wastewater surveillance for infectious disease: a systematic review. Am J Epidemiol 2023 Feb 1;192(2):305-322. [doi: 10.1093/aje/kwac175] [Medline: 36227259]
4. Luo J, Bello D, Pagsuyoin S. Long-term wastewater-based surveillance and impacts of the COVID-19 pandemic on drug use trends in a U.S. northeast rural town. Science of The Total Environment 2023 Jun;877:162806. [doi: 10.1016/j.scitotenv.2023.162806]
5. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division; Division on Earth and Life Studies, Board on Population Health and Public Health Practice, Water Science and Technology Board, Committee on Community Wastewater-based Infectious Disease Surveillance. Wastewater-Based Disease Surveillance for Public Health Action: Washington (DC): National Academies Press (US); 2023. [doi: 10.17226/26767] [Medline: 37184191]
6. Bowes DA, Driver EM, Savic S, et al. Integrated multiomic wastewater-based epidemiology can elucidate population-level dietary behaviour and inform public health nutrition assessments. Nat Food 2023 Mar;4(3):257-266. [doi: 10.1038/s43016-023-00717-w] [Medline: 37118274]
7. Clark M, Severn M. Wastewater surveillance for communicable diseases: emerging health technologies. In: Canadian Journal of Health Technologies 2, 2023, Vol. 3. [doi: 10.51731/cjht.2023.575] [Medline: 38502748]
8. Choi PM, Tscharke B, Samanipour S, et al. Social, demographic, and economic correlates of food and chemical consumption measured by wastewater-based epidemiology. Proc Natl Acad Sci U S A 2019 Oct 22;116(43):21864-21873. [doi: 10.1073/pnas.1910242116] [Medline: 31591193]
9. Evans ED, Dai C, Isazadeh S, Park S, Ratti C, Alm EJ. Longitudinal wastewater sampling in buildings reveals temporal dynamics of metabolites. PLoS Comput Biol 2020 Jun;16(6):e1008001. [doi: 10.1371/journal.pcbi.1008001] [Medline: 32598361]
10. Yang Z, Xu G, Reboud J, Kasprzyk-Hordern B, Cooper JM. Monitoring genetic population biomarkers for wastewater-based epidemiology. Anal Chem 2017 Sep 19;89(18):9941-9945. [doi: 10.1021/acs.analchem.7b02257] [Medline: 28814081]

11. Farkas K, Hillary LS, Malham SK, McDonald JE, Jones DL. Wastewater and public health: the potential of wastewater surveillance for monitoring COVID-19. Curr Opin Environ Sci Health 2020 Oct;17:14-20. [doi: 10.1016/j.coesh.2020.06.001] [Medline: 32835157]
12. Hart OE, Halden RU. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: feasibility, economy, opportunities and challenges. Sci Total Environ 2020 Aug 15;730:138875. [doi: 10.1016/j.scitotenv.2020.138875] [Medline: 32371231]
13. National wastewater surveillance system (NWSS). Centers for Disease Control and Prevention. 2023. URL: https://www.cdc.gov/nwss/wastewater-surveillance.html [accessed 2024-01-17]
14. National wastewater drug monitoring program reports. Australian Criminal Intelligence Commission. 2024. URL: https://www.acic.gov.au/publications/national-wastewater-drug-monitoring-program-reports [accessed 2024-10-15]
15. Allison G. How it went down. TIME. 2012. URL: https://time.com/archive/6641671/how-it-went-down [accessed 2024-10-15]
16. 2023 Biodefense Posture Review. US Department of Defense. 2023. URL: https://media.defense.gov/2023/Aug/17/2003282337/-1/1/1/2023_BIODEFENSE_POSTURE_REVIEW.PDF [accessed 2024-05-20]
17. Goldberg Z, Linder AG, Miller LN, Sorrell EM. Wastewater collection and sequencing as a proactive approach to utilizing threat agnostic biological defense. Health Secur 2024;22(1):11-15. [doi: 10.1089/hs.2023.0075] [Medline: 37856169]
18. Sanjak JS, McAuley EM, Raybern J, et al. Wastewater surveillance pilot at US military installations: cost model analysis. JMIR Public Health Surveill 2024 Sep 6;10:e54750. [doi: 10.2196/54750] [Medline: 39240545]
19. Kass NE. An ethics framework for public health. Am J Public Health 2001 Nov;91(11):1776-1782. [doi: 10.2105/ajph.91.11.1776] [Medline: 11684600]
20. German RR, Lee LM, Horan JM, et al. Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep 2001 Jul 27;50(RR-13):1-35. [Medline: 18634202]
21. Amin V, Bowes DA, Halden RU. Systematic scoping review evaluating the potential of wastewater-based epidemiology for monitoring cardiovascular disease and cancer. Sci Total Environ 2023 Feb 1;858(Pt 3):160103. [doi: 10.1016/j.scitotenv.2022.160103] [Medline: 36370774]
22. Ladyzhets B. What toilets can reveal about COVID, cancer and other health threats. Nature New Biol 2024 Apr;628(8008):492-494. [doi: 10.1038/d41586-024-01092-7] [Medline: 38632479]
23. Doorn N. Wastewater research and surveillance: an ethical exploration. Environ Sci (Camb) 2022 Oct 27;8(11):2431-2438. [doi: 10.1039/d2ew00127f] [Medline: 36353217]
24. Rinde M, Science History Institute. The murky ethics of wastewater surveillance. Distillations Magazine. 2023 Feb. URL: https://sciencehistory.org/stories/magazine/the-murky-ethics-of-wastewater-surveillance [accessed 2024-04-23]
25. Hrudey SE, Silva DS, Shelley J, et al. Ethics guidance for environmental scientists engaged in surveillance of wastewater for SARS-CoV-2. Environ Sci Technol 2021 Jul 6;55(13):8484-8491. [doi: 10.1021/acs.est.1c00308] [Medline: 34101444]
26. Coffman MM, Guest JS, Wolfe MK, et al. Preventing scientific and ethical misuse of wastewater surveillance data. Environ Sci Technol 2021 Sep 7;55(17):11473-11475. [doi: 10.1021/acs.est.1c04325] [Medline: 34431294]
27. Ahmed W, Bertsch PM, Angel N, et al. Detection of SARS-CoV-2 RNA in commercial passenger aircraft and cruise ship wastewater: a surveillance tool for assessing the presence of COVID-19 infected travellers. J Travel Med 2020 Aug 20;27(5):116. [doi: 10.1093/jtm/taaa116] [Medline: 32662867]
28. Traveler-based genomic surveillance for early detection of new SARS-cov-2 variants. Centers for Disease Control and Prevention. 2023. URL: https://wwwnc.cdc.gov/travel/page/travel-genomic-surveillance [accessed 2024-03-05]
29. Li J, Hosegood I, Powell D, et al. A global aircraft-based wastewater genomic surveillance network for early warning of future pandemics. Lancet Glob Health 2023 May;11(5):e791-e795. [doi: 10.1016/S2214-109X(23)00129-8] [Medline: 37061316]
30. Jacobs D, McDaniel T, Varsani A, Halden RU, Forrest S, Lee H. Wastewater monitoring raises privacy and ethical considerations. IEEE Trans Technol Soc 2021;2(3):116-121. [doi: 10.1109/TTS.2021.3073886]
31. Lee LM, Thacker SB. Public health surveillance and knowing about health in the context of growing sources of health data. Am J Prev Med 2011 Dec;41(6):636-640. [doi: 10.1016/j.amepre.2011.08.015] [Medline: 22099242]
32. Gable L, Ram N, Ram JL. Legal and ethical implications of wastewater monitoring of SARS-CoV-2 for COVID-19 surveillance. J Law Biosci 2020;7(1):lsaa039. [doi: 10.1093/jlb/lsaa039] [Medline: 32793373]
33. Smith HJ. Ethical responsibilities of a military to the social determinants of health of its service members. Mil Med 2022 Aug 25;187(9-10):252-256. [doi: 10.1093/milmed/usab554] [Medline: 35389484]
34. Public health code of ethics. American Public Health Association. 2019. URL: https://www.apha.org/-/media/files/pdf/membergroups/ethics/code_of_ethics.ashx [accessed 2024-05-30]
35. Marckmann G, Schmidt H, Sofaer N, Strech D. Putting public health ethics into practice: a systematic framework. Front Public Health 2015;3:23. [doi: 10.3389/fpubh.2015.00023] [Medline: 25705615]
36. Schröder-Bäck P, Duncan P, Sherlaw W, Brall C, Czabanowska K. Teaching seven principles for public health ethics: towards a curriculum for a short course on ethics in public health programmes. BMC Med Ethics 2014 Oct 7;15:73. [doi: 10.1186/1472-6939-15-73] [Medline: 25288039]
37. WHO guidelines on ethical issues in public health surveillance. World Health Organization. 2017. URL: https://www.who.int/ethics/publications/public-health-surveillance/en [accessed 2024-05-30]

38. Bowes DA, Darling A, Driver EM, et al. Structured ethical review for wastewater-based testing in support of public health. Environ Sci Technol 2023 Sep 5;57(35):12969-12980. [doi: 10.1021/acs.est.3c04529] [Medline: 37611169]
39. Keshaviah A, Diamond MB, Wade MJ, Scarpino SV, Global Wastewater Action Group. Wastewater monitoring can anchor global disease surveillance systems. Lancet Glob Health 2023 Jun;11(6):e976-e981. [doi: 10.1016/S2214-109X(23)00170-5] [Medline: 37202030]
40. Levy JI, Andersen KG, Knight R, Karthikeyan S. Wastewater surveillance for public health. Science 2023 Jan 6;379(6627):26-27. [doi: 10.1126/science.ade2503] [Medline: 36603089]
41. The executive branch. White House Archives. URL: https://obamawhitehouse.archives.gov/1600/executive-branch [accessed 2024-02-15]
42. Patients by beneficiary category. Health.mil: The official website of the Military Health System. 2023. URL: https://www.health.mil/Military-Health-Topics/MHS-Toolkits/Media-Resources/Media-Center/Patient-Population-Statistics/Patients-by-Beneficiary-Category [accessed 2024-05-30]

Abbreviations

DoD: Department of Defense
SM: Service member
WWS: wastewater surveillance

Edited by A Mavragani; submitted 13.11.24; peer-reviewed by C Mansfeldt, M Gutierrez, W Hall; revised version received 13.12.24; accepted 15.12.24; published 06.02.25.

Please cite as:
Smith HJ, Agans RT, Kowallis WJ
Ethical Considerations for Wastewater Surveillance Conducted by the US Department of Defense
JMIR Public Health Surveill 2025;11:e67145
URL: https://publichealth.jmir.org/2025/1/e67145
doi:10.2196/67145

The Promise and Perils of Artificial Intelligence in Advancing Participatory Science and Health Equity in Public Health

Abby C King[1], PhD; Zakaria N Doueiri[2], BS; Ankita Kaulberg[2], BS, MBA; Lisa Goldman Rosas[3], MPH, PhD

[1]Departments of Epidemiology & Population Health and of Medicine (Stanford Prevention Research Center), Stanford University School of Medicine, 1701 Page Mill Road, Palo Alto, CA, United States

[2]Department of Epidemiology & Population Health, Stanford University School of Medicine, Palo Alto, CA, United States

[3]Departments of Epidemiology & Population Health and of Medicine (Division of Primary Care and Population Health), Stanford University School of Medicine, Stanford University, Palo Alto, CA, United States

Corresponding Author:
Abby C King, PhD
Departments of Epidemiology & Population Health and of Medicine (Stanford Prevention Research Center), Stanford University School of Medicine, 1701 Page Mill Road, Palo Alto, CA, United States

Abstract

Current societal trends reflect an increased mistrust in science and a lowered civic engagement that threaten to impair research that is foundational for ensuring public health and advancing health equity. One effective countermeasure to these trends lies in community-facing citizen science applications to increase public participation in scientific research, making this field an important target for artificial intelligence (AI) exploration. We highlight potentially promising citizen science AI applications that extend beyond individual use to the community level, including conversational large language models, text-to-image generative AI tools, descriptive analytics for analyzing integrated macro- and micro-level data, and predictive analytics. The novel adaptations of AI technologies for community-engaged participatory research also bring an array of potential risks. We highlight possible negative externalities and mitigations for some of the potential ethical and societal challenges in this field.

(JMIR Public Health Surveill 2025;11:e65699) doi:10.2196/65699

KEYWORDS

digital health; artificial intelligence; community-based participatory research; citizen science; health equity; societal trends; public health; viewpoint; policy makers; public participation; information technology; micro-level data; macro-level data; LLM; natural language processing; machine learning; language model; Our Voice

Introduction

> *"While the future might indeed be bright for AI, it wouldn't be so by accident. We will have to earn it, together."* [Dr. Fei-Fei Li] [Founding Director, Stanford Institute for Human-Centered AI] [From The Worlds I See: Curiosity, Exploration, and Discovery at the Dawn of AI 1]

The last several decades have witnessed a growing mistrust in science among both policy makers and the public at large on an unprecedented scale [2]. The reasons for this trend appear to be many and varied, including a lack of understanding about how science actually operates, partly at least due to an absence of participatory educational opportunities and direct engagement in meaningful science activities [3]. This, in turn, has been linked with poor self-ratings of scientific literacy among school-aged adolescents and adults alike [3]. A general confusion about what the public should reasonably expect from research in terms of scientific veracity or "truth" has been an additional contributor to the public's skepticism about science. These problems have been amplified to a significant extent by social media, which has often worked, either intentionally or unintentionally, to confuse or obfuscate the truth [4]. Unfortunately, such efforts arguably have been abetted by scientific institutions themselves as well as scientific journal outlets which, in their haste to grab the day's headlines, have at times stripped scientific results of their complexity, nuance, and context; or, shockingly, filled untold numbers of journals with "fake science" [5]. Well-documented historical injustices also have contributed to a lack of trust in academic and scientific institutions, particularly among communities of color. For example, the US Public Health Service Untreated Syphilis Study at Tuskegee has generated mistrust in health care institutions and research for generations [6]. More recently, the University of Arizona was found guilty of using DNA samples from the Havasupai Tribe for research without their consent [7]. Recent controversies and public health crises (eg, the COVID-19 pandemic) also have given rise to increased questioning of public health authorities and have sown doubts regarding scientific validity and reliability, which have exacerbated feelings of mistrust.

This alarming trend of mistrust in science threatens public health efforts in multiple ways, which include the following recent examples: (1) misinformation concerning vaccines, including outright dismissal of the extensive multigenerational evidence base supporting the development and use of vaccines to prevent a myriad of dangerous and at times fatal diseases that have threatened the public's health; (2) skepticism surrounding the wide-ranging, multidisciplinary evidence base establishing the considerable effects of climate change that threaten the health of the planet and all of its living inhabitants; and (3) the growing distrust of health care systems as well as pharmaceutical companies (ie, "big Pharma"), which has become especially acute among marginalized groups, including minoritized and under-resourced communities [8].

The mistrust of science coincides with other pernicious societal trends reflecting disconnection and skepticism in other facets of daily life. These include growing concerns in the health and mental health fields around the current levels of civic disconnectedness and community disengagement among significant sections of the population. Recent data indicate, for example, that civic participation, social engagement, and community connection among segments of today's youth and young adults are declining, with this trend having been worsened by the COVID-19 pandemic [9].

The confluence of these trends with the emergence of an often-bewildering pace of advances in the worldwide IT sector has served to exacerbate feelings of confusion and mistrust among broad segments of the population. This has included, in particular, the unexpectedly disruptive force of the current artificial intelligence (AI) revolution across all sectors, including health care, education, and research. Threats regarding the spread of misinformation, increasing systematic biases against already marginalized populations, data privacy issues, and potential job displacement are among the leading concerns of the public, in addition to other impacts that the widespread acceptance and use of AI might have on various segments of society, from individuals through governments and policy makers [10].

The above issues have led to the following questions: (1) how can science in general be made more accessible to the public in ways that can reduce confusion and mistrust and increase the meaningfulness of scientific inquiry to directly benefit individuals and their communities, and (2) how can the potential of AI be harnessed to bolster the public's engagement in public health science while mitigating the distrust and confusion about both research and AI itself?

Increasing Public Participation in Scientific Research

Public participation in science research (PPSR) is one way researchers have been addressing the increasing mistrust among decision makers and the public toward science. A growing number of studies have included PPSR methods, such as citizen science, that have utilized educational outreach efforts to increase public knowledge and awareness of scientific principles and best practices [11]. Important, systematic efforts have been employed to actively involve a greater number and variety of community members in the scientific process [12]. Such participatory research paradigms can have substantial positive impacts for increasing transparency, knowledge, and trust in scientific inquiry, particularly when they involve generating actionable data of direct relevance to those involved [13].

Within the broader umbrella of PPSR frameworks, citizen science—generally defined as involving lay persons in the research process to advance science—is an increasingly popular approach to engage community members in research activities [14]. A strength behind citizen science is its underlying principle that scientific inquiry can often benefit from including diverse members of the community from varying educational, occupational, social, and cultural backgrounds. Various forms of citizen science have occurred over several centuries and have involved diverse fields, including the life, environmental, health, social, and behavioral sciences [15]. Citizen science can directly benefit research by providing extra pairs of "helping hands" in the data collection and problem-solving aspects of science [14]. Beyond democratizing data acquisition, citizen scientists also can contribute to the fuller research process, including problem definition, data interpretation and prioritization, data-driven solution generation, and results dissemination [16-18]. This more fully participatory "by the people" form of citizen science can advance scientific inquiry as well as promote meaningful knowledge and solutions for the local communities that are involved [13,16,17].

An example of this type of citizen science in the public health arena is the technology-enabled "Our Voice" Global Citizen Science Research Initiative [19], where residents from diverse and often under-resourced communities learn how to partner with researchers and community organizations in collecting and interpreting aggregated and anonymized contextual data (eg, photos, text, and audio). These data identify features of residents' local environments that help or hinder their health [16,17]. They then learn how to effectively communicate their data to decision makers and work with them to activate health-benefiting changes in their local environments [16,17]. Over nearly a decade, the multi-generational "Our Voice" Initiative has produced demonstrable community-driven physical and social environmental impacts across 100+ projects involving over 25 countries spanning 6 continents [16,17].

Leveraging the Potential of AI to Promote More Powerful Citizen Science for Advancing Public Health

While the current focus of health-oriented AI applications has been primarily at the individual patient level, the thoughtful and strategic application of AI at the community level offers potentially powerful tools to augment participatory science activities across the fuller research process. At the same time, it is important that both researchers and participants recognize the potential risks of AI and are committed to implementing multilayered mitigation tactics aimed at addressing them (see Table 1, Figure 1, and the subsequent section focused on risk mitigation).

Table . Examples of potential artificial intelligence (AI) uses within community-engaged public health citizen science research, along with potential risks and risk mitigation strategies.

Potential use of AI within community-engaged research	Potential risks	Risk mitigation strategies
Conversational LLMs[a] (ChatGPT or similar) for onboarding participants, offering personalized guidance on engaging with technology, and asking thought-provoking questions	• Biased or culturally insensitive conversations emerge • AI "hallucinations" where factually inaccurate commentary is made	• Dedicated time spent in prompt engineering to determine language best used to minimize these biases in this context, and then shared as best practices to the public • Public education on AI, the training data that were used, and discussion of shortcomings
Text-to-image Generative AI tools (DALL-E, Midjourney, or similar) for helping the community visualize their data (eg, turning text descriptions into relevant images)	• Biased or culturally insensitive pictures generated and shared • Impossible or impractical to execute ideas presented	• Human (expert) review and gatekeeping before images shared broadly to the public • Prompt engineering and similar efforts aimed at community data collection are shared with participants in advance to enhance the potential relevance and feasibility of solutions
Data Mining (Descriptive Analytics) for analyzing macro- and microlevel data, generating integrated data reports, visualizations, etc.	• Inaccuracies or AI hallucinations injected into the dataset • Loss of control, research rigor, and cognitive competence by scientists who defer too much to AI	• Intentional efforts made to feed AI inclusive and community representative data • Recursive AI use: feed output back to AI to check for inaccuracies • Require scientists to review all data generated for accuracy and meet field standards for scientific literacy or competence
Data Mining (Predictive Analytics) for reviewing large datasets, providing data aggregation reports, and generating visual aids to help the community interpret data	• Data presented may result in wrong conclusions, resulting in further inequities • Loss of control, research rigor, and cognitive competence by scientists who defer too much to AI	• Human (expert) review and governance of all generated findings • Recursive AI use: feed output back to AI to check for inaccuracies • Require scientists to review all data and methods for accuracy and meet field standards for scientific literacy or competence • Empirical testing of AI to determine when it helps or hurts scientific skills and competencies

[a] LLM: large language model

Figure 1. Artificial intelligence (AI) integration and risk mitigation strategies for participatory action citizen science: example of “Our Voice” Method and workflow. LLM: large language model; OV: Our Voice.

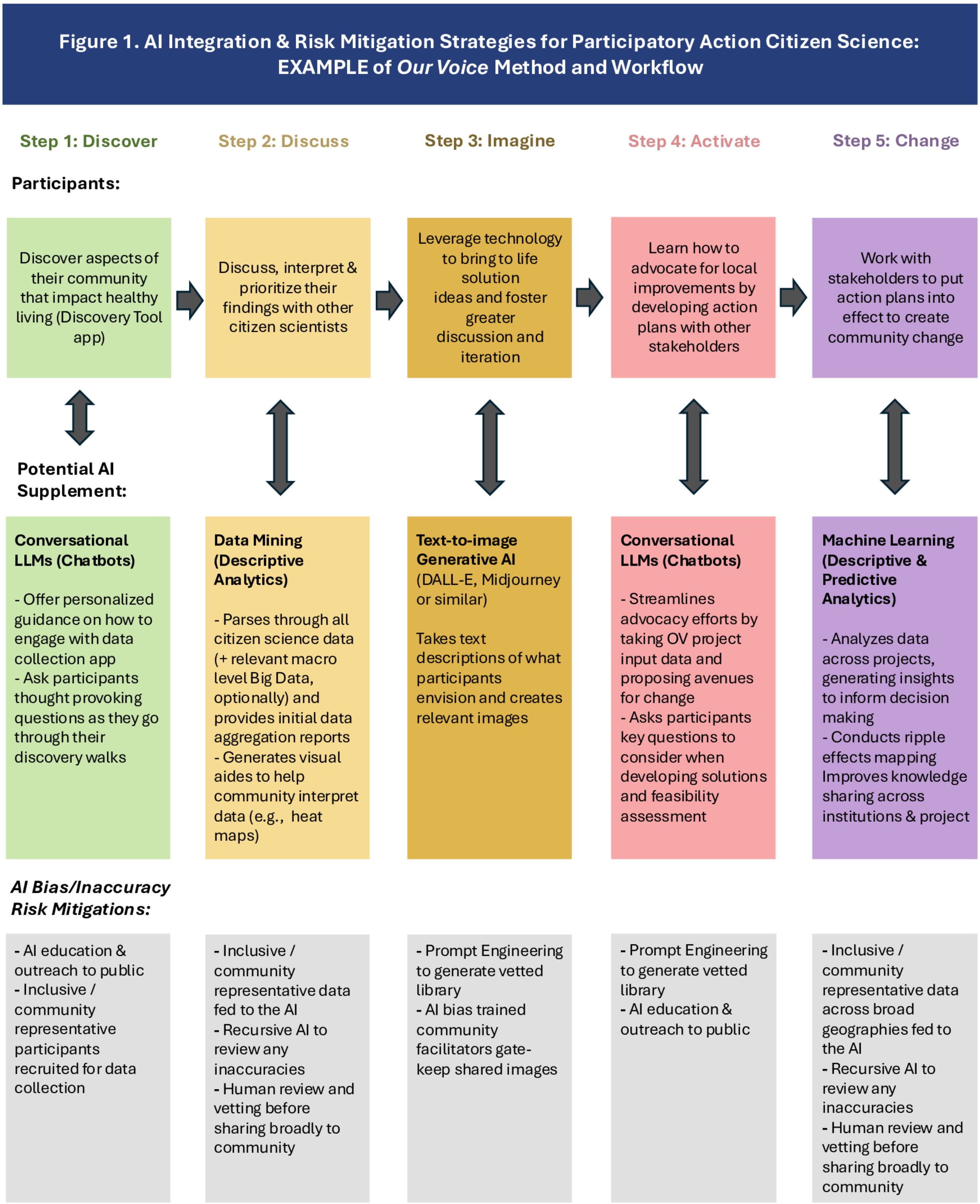


Figure 1. AI Integration & Risk Mitigation Strategies for Participatory Action Citizen Science: EXAMPLE of *Our Voice* Method and Workflow

	Step 1: Discover	Step 2: Discuss	Step 3: Imagine	Step 4: Activate	Step 5: Change
Participants:	Discover aspects of their community that impact healthy living (Discovery Tool app)	Discuss, interpret & prioritize their findings with other citizen scientists	Leverage technology to bring to life solution ideas and foster greater discussion and iteration	Learn how to advocate for local improvements by developing action plans with other stakeholders	Work with stakeholders to put action plans into effect to create community change
Potential AI Supplement:	**Conversational LLMs (Chatbots)** - Offer personalized guidance on how to engage with data collection app - Ask participants thought provoking questions as they go through their discovery walks	**Data Mining (Descriptive Analytics)** - Parses through all citizen science data (+ relevant macro level Big Data, optionally) and provides initial data aggregation reports - Generates visual aides to help community interpret data (e.g., heat maps)	**Text-to-image Generative AI** (DALL-E, Midjourney or similar) Takes text descriptions of what participants envision and creates relevant images	**Conversational LLMs (Chatbots)** - Streamlines advocacy efforts by taking OV project input data and proposing avenues for change - Asks participants key questions to consider when developing solutions and feasibility assessment	**Machine Learning (Descriptive & Predictive Analytics)** - Analyzes data across projects, generating insights to inform decision making - Conducts ripple effects mapping Improves knowledge sharing across institutions & project
AI Bias/Inaccuracy Risk Mitigations:	- AI education & outreach to public - Inclusive / community representative participants recruited for data collection	- Inclusive / community representative data fed to the AI - Recursive AI to review any inaccuracies - Human review and vetting before sharing broadly to community	- Prompt Engineering to generate vetted library - AI bias trained community facilitators gate-keep shared images	- Prompt Engineering to generate vetted library - AI education & outreach to public	- Inclusive / community representative data across broad geographies fed to the AI - Recursive AI to review any inaccuracies - Human review and vetting before sharing broadly to community

Using the “Our Voice” multistep citizen science method as an example, potentially valuable AI applications can occur at each step (see Figure 1). The steps include resident-engaged data collection using a multilingual mobile app (“Discover” step); facilitated analysis, discussion, and prioritization of relevant local issues (“Discussion” step); leveraging of visualization technologies to bring potential solutions into clearer focus (“Imagine” step); communicating relevant community issues and brainstorming feasible solutions with local decision makers (“Activate” step); and continued partnering with stakeholders in bringing action plans to fruition (“Change” step). In addition to these integral steps, the above “Our Voice” method has been supplemented at times with an increasingly proactive and thoughtful approach to engage the community prior to embarking on citizen science activities, as well as the thoughtful dissemination of findings and insights to researchers, decision makers, and the public (a "Share" step). The careful integration of relevant AI applications into these different citizen science processes may assist in supplementing and enriching the different step outputs, as well as helping to address at least some of the challenges with public trust, given the active role that community members play in this type of participatory science.

Some examples of potentially relevant AI tools that can be explored are given below for each step.

Discover (Data Collection) Step

A challenge to address in the discover (data collection) step is that the inclusion of AI in this step should not supersede the real, on-the-ground data collected by local citizen scientists. Any additional data presented by the AI at this step is intended to be an augmentation that goes through human review before it is included in project data capture.

Conversational large language models (LLMs) can be employed in a strategic and transparent way to personally guide the use of citizen science data collection apps such as the "Our Voice Discovery Tool" app, along with "real-time" dynamic prompts to more fully engage residents as they go through their data collection walks. In addition, utilizing computer vision, hearing, and other "real-world" perceptual sensing tools (eg, smell) may help to produce more robust and comprehensive data capture and classifications to augment the contextually focused citizen science data [20]. The goal of using such sensing tools along with LLMs would be to enrich citizen scientists' data collection activities to produce more nuanced and detailed data to analyze and combine with other relevant data sources (eg, geographic information systems and social network models).

Discuss Step

A challenge to address in the discuss step is that AI should not absolve researchers and participants from the task of ensuring analytical accuracy and rigor, nor replace efforts in interpreting the data and gleaning insights. As AI is known at times to experience "hallucinations" and commit errors when it comes to data analysis and interpretation, whatever is generated by it should be seen as a first-pass draft of something that helps researchers and community members enrich their own interpretations and iterations.

ChatGPT and similar conversational LLMs can be engaged during community meetings to support residents' group discussions during data interpretation and solution-building activities. For example, LLMs may be able to support thoughtful community-level discussion and advocacy, not as an omniscient "black box" but as an attentive community-level support system and Socratic thinking aide. Rather than attempting to "replace" community members' perspectives and lived experiences, LLMs may be employed to play the role of "devil's advocate" in a more neutral way, to help residents better understand the concerns of marginalized groups as well as policy makers. LLMs also may allow participants to better assess the feasibility of their data-driven ideas for positive change by supplementing knowledge or experience gaps that the average citizen scientist may have (eg, "Considering the financial constraints of X and timeframe to solve this issue by Y, which of these 5 solutions are the most feasible to address …..?"). In presenting participants with more detailed background information regarding the implementation of different data-driven solutions or strategies, residents can be further informed about the potential complexities that come with improving their communities. AI also could be leveraged to generate and describe multilevel insights by combining diverse sources of data, from citizen scientist–generated microscale data that capture local contexts and individual-level biosensing outputs (eg, heart rate and physiological stress responses) through macroscale population-based social determinants and health data. This, in turn, could result in the development of interactive maps from such mixed-methods data sources that arguably could foster faster insights and public health responses from policy makers. For example, the initial efforts to visually combine citizen scientist–generated microscale data describing residents' neighborhood-lived experiences with macroscale epidemiological data about their larger environmental contexts have yielded richer results than would have been obtained from either data source alone [21]. Leveraging AI in this manner to assist with combining and visualizing different forms of data could actually free up researchers and citizen scientists to engage in higher-order thinking and analysis that is required for creatively addressing some of society's most pressing issues. For instance, in California, low participation in community-supported public health initiatives that provide opportunities for affordable healthy food access is often driven by diverse factors that can vary by locale. The multifactorial issues driving low program uptake often are challenging for researchers, governmental agencies, and community-based organizations to understand sufficiently to be able to intervene effectively. It is possible that AI tools could help by aiding the visualization of how microlevel "lived experience" data vary with respect to macrolevel contextual data. Such multilevel visualizations would allow researchers and community partners to spend more of their time understanding and solving the problem, as opposed to gathering and synthesizing relevant data to describe the problem.

Imagine Step

A challenge to address in the imagine step is that AI should not have free reign in determining what the best solutions for the community are. Using culturally tailored inputs, AI's creativity can be used as a supplemental aide to the collective community's brainstorming while not superseding it.

Several AI tools may be useful in helping residents, decision makers, and researchers better imagine what scientifically generated health-enhancing changes can actually look like in their own communities and environments (eg, through generative AI text-to-image tools such as DALL-E and Midjourney). For example, Stanford's Center on Human-Centered Artificial Intelligence awarded a seed grant to "Our Voice" that will allow citizen scientists to take images collected during local "discovery walks" and transform them into reimagined futures through such generative AI tools (eg, an image of an unsafe intersection to get to school could be "morphed" into images of different solutions such as a crosswalk, pedestrian bridge, etc). The early testing of this form of AI with school children living in the El Pozón neighborhood of Cartagena, Colombia—where the majority of inhabitants live below the poverty line and lack access to essential services—is showing how the thoughtful, culturally tailored use of text-to-image AI tools can generate meaningful and actionable solutions to local environmental problems (eg, stagnant water and uncollected garbage). These rapid visual renderings could in turn lead to enhanced and richer discussions of trade-offs and foster new ideas among community members and researchers.

Activate and Change Steps

A challenge to address in the activate and change steps is that, as noted above, an optimal goal of AI use is to supplement and enrich such conversations, as opposed to driving them. Sharing the above types of visualization tools and information with local decision makers may help to jump-start relevant solution-building efforts. Generative AI could be used to aid in the creation of presentation materials, help community members role-play conversations with local decision makers in preparation for scheduled meetings with them, and identify issues and suggest modifications to the action plan created during the community meetings.

Share Step

A challenge to address in the share step is that it is recommended that researchers avoid risking diminished scientific rigor and accuracy through outsourcing major data interpretation to AI. As noted earlier, people are in the best position to fully understand and evaluate what the data mean, including, in this case, the citizen scientists who have collected those data.

Machine learning tools also can be used to aid researchers in the aggregation of data across projects and expedite their ability to glean useful insights across different citizen science projects and research groups.

In addition to the above types of citizen science-oriented strategies, through utilizing AI-generated virtual reality simulations, citizen scientists along with other community members can experience physical, cultural, and social situations that they might not otherwise encounter in their daily lives [22]. Such AI-generated virtual reality simulations also can provide community members and decision makers with hands-on opportunities to interact with virtual characters representing differing cultural backgrounds that may lead to increased cultural sensitivity [22]. These AI-driven virtual or simulated environments can additionally allow community members to try out citizen science data gathering and similar tools before real-world use [20].

Additional Recommendations on Mitigating Potential Risks of AI

Amidst the potentially promising uses of AI for the public health and citizen science fields noted above, including its potential to expand the breadth and utility of the data being collected and interpreted, a number of ethical challenges and concerns have been raised, some of which have been touched on earlier and in other articles [23]. These include the different types of risks, as described in this paragraph. AI could diminish, distort, or replace data being collected in the field with output that is less accurate (ie, AI "hallucinations") and less representative of community member experiences and responses. AI output may be based on information and sources that are not culturally relevant or appropriate for a given community or context. Potential challenges emerge when AI is presented with contrasting ideas from citizen scientists, researchers, and other stakeholders and autonomously makes prioritization choices among them. The introduction of AI tools into the public health research process could actually serve to increase the "depersonalization" of the research process, leading to a greater rather than a lower distrust of science. Finally, major concerns have been raised that through potentially oversimplifying and directing aspects of the scientific process, LLMs and similar AI tools could significantly diminish scientists' data processing skills and cognitive competencies.

The publicly available use of ChatGPT and similar generative AI tools as "short cuts" for increasing time efficiencies at the potential expense of thought and deliberation has served to heighten many of these concerns.

In response to such threats, some potential solutions have begun to emerge (see Table 1), as described below.

Thoughtfulness and Transparency

A potential solution is being thoughtful and transparent in determining when IA (information augmentation to support human intelligence) versus AI (artificial intelligence, which is often used as a replacement for some human activities) is best suited. For example, in places where the radius for impact is large and missteps would be difficult to rectify or could cause grave harm, AI, which to date generally does not have community participation or checks for bias mitigation, ought to be used with great caution. In such instances, using technology instead to aggregate and synthesize data findings, and leaving the interpretation of those findings to people, could be wiser.

A mitigation example could be the data collection of citizen scientists' lived experiences in the context of community determinants of health that could be synthesized to share summary statistics, generate heat maps, or word clouds; however, any interpretation of their data, along with solution-building and implementation, would not be outsourced to AI without human supervision and community-level interpretations and buy-in.

Employing Explainable AI Concepts and Strategies

Another potential solution is the use of explainable AI (XAI) that involves processes and methods that allow human users to understand and better trust the results and output created by machine learning and similar algorithms. It can be used to describe an AI model, its expected uses and impacts, and potential biases (eg, based on sex or race).

A mitigation example is the use of a set of vetted chat prompts and outlined best practices that could be provided to users ("prompt engineering"), instead of assuming that the results of ChatGPT are correct and accurate. These best practices could include examples of wording to enter into ChatGPT to improve results and ideally catch AI “hallucinations.” Here is one such example: “For any and all responses, end each response with a bulleted list that includes (1) any assumptions you are making, (2) citations of any sources you have used to determine the response to my prompt, and (3) any biases or concerns you want to review if you had access to unlimited data or personnel.” In addition, the responses generated by the AI system can be fed back to itself with the prompt “Check this response for any factual inaccuracies or state anything you cannot back up with reputable sources.”

Launching Active Educational and Outreach Initiatives

There are growing concerns related to the potential of AI, including generative AI, to promote biased results and intensify health inequities through the lack of diverse data that are representative of different societal groups. Therefore, it is important to launch active educational and outreach initiatives to better inform the public across all walks of life about AI, its uses, ethical implications, and the precautions about which people should be aware. Educational efforts aimed specifically at diverse communities and populations may help to increase transparency concerning both the promise and limitations of this emerging field. It also may help ensure that key community norms and values are taken into account [24]. Such outreach-based multidisciplinary and community-driven educational efforts tailored to the needs of different populations may additionally help to empower currently under-represented groups to actively participate in data collection and information sharing to help mitigate such biases [24]. Involving community members in the development of AI methodologies from the beginning is important to ensure transparency and reduce mistrust among participants [24]. While such proactive educational efforts have been focused primarily on health care, they could benefit public health research and activities as well.

A mitigation example is described here. Zainab Garba-Sani and colleagues have developed the A.C.C.E.S.S. AI model, which lays a framework for involving communities in the AI-rollout process as follows: A = “Affirm your aims,” C = “Consider your communities,” C = “Cultivate your conversations,” E = “Embrace your essentials,” S = “Specify your scope,” and S = “Scrutinize your spaces” [24]. Through training sessions and workshops with diverse groups, A.C.C.E.S.S. serves the dual purpose of not only educating people about AI and its potential applications but also gathering their input on how AI should be thoughtfully designed for their unique contexts. Another example of such an initiative is the nonprofit “AI4ALL” movement that aims to advance an increasingly human-centered, inclusive AI discipline [25]. Through proactive participatory approaches such as these, we can begin to demystify AI and ensure that diverse communities are brought to the table when implementing AI tools.

Increasing Efforts to Improve Data Upon Which AI is Modeled and Trained

As AI is merely holding up a mirror to the society or data upon which it is trained, the biases inherent in AI responses are likely due to the missing perspectives and incomplete data upon which it was built. Leveraging collaborations across institutions, social or civic interest groups, and diverse communities to improve the training data should help AI become more aware of differing perspectives and hopefully reduce if not completely minimize the bias. Outreach-based multidisciplinary and community-driven educational efforts as described below that are tailored to the needs of different populations may additionally help to empower currently under-represented groups to actively participate in data collection and information sharing to help mitigate such biases [24].

A mitigation example is the need for clear guidelines and incentives from scientific organizations, universities, journals, and granting organizations (eg, specific grant announcements from the NIH in this area) concerning expanding the population representativeness of the datasets upon which AI is based; this can send a clear message regarding the importance of this issue to the field as well as society at large.

Setting AI-Specific Ethical Standards and Having Expected AI-Influenced Outputs Verified by Experts

It is important to set AI-specific ethical standards across universities, other research organizations, scientific outlets, and funding sources that include the ethics review of study methods as well as have the expected AI-influenced outputs verified by experts trained in the relevant ethics fields. Publishers of scientific journals, books, and other scientific communication channels should be encouraged to provide specific guidelines related to the publication of research investigations that include AI. Demanding transparency of AI use in scientific research more generally is recommended, as a growing collection of policy makers, technology developers, and scientific organizations are currently doing. Setting up appropriate “guardrails” prior to AI use is similarly important to help ensure that flawed, offensive, or inaccurate information does not unexpectedly find its way into AI-driven conversations or other outputs.

A mitigation example is described here. As part of institutional review board policies and activities, some grant-making organizations have included, in addition to the review of a grant proposal’s scientific merit, a separate ethics review involving experts in ethics and similar fields to identify areas of potential ethical concern and methods for the mitigation of such concerns. An example is the Stanford Human-Centered Artificial Intelligence Center’s process of including a separate ethics review of AI grant proposals to help ensure that appropriate AI-relevant guardrails are in place prior to funding.

Regulating Transparent Data Ownership and Use

It is important to regulate transparent data ownership and use to ensure that access to citizen science data is not limited by commercial interests [20]; this helps to ensure that AI tools and resources are available to everyone, which can democratize its use and diminish public suspicion and skepticism. For example, many users of ChatGPT and similar platforms may not understand that their data currently may be sent to the parent company OpenAI to subsequently power future AI learning, and then be ultimately used by unnamed sources for a variety of undisclosed purposes.

A mitigation example is that, instead of relying on public GPTs whose data ownership and use are ambiguous, the Stanford

School of Medicine and other academic institutions have developed their own custom GPT that ensure data security, HIPAA compliance, and similar outcomes. Additional AI resources that are being developed at Stanford include a user-friendly “AI Playground” platform, built on open-source technologies, that allows the Stanford community to try out various AI models and tools as well as compare the utility of different tools in a head-to-head fashion.

Human Governance and Oversight Over AI Output

It is imperative when undertaking community-engaged forms of research that scientists work to ensure that AI-generated contributions to the research endeavor do not supersede or otherwise hamper the credibility of the contributions from community members and their representatives. While AI tools are often used to obtain meaning from large datasets, in the participatory public health research field, it is typically the residents themselves who can best provide meaning and context to the data being collected and interpreted.

A mitigation example is that the addition of community advisory experts, panels, and similar oversight bodies can help ensure that the use of AI enriches but does not overshadow either the scientific process itself or scientists and the public’s participation in that process. It will be particularly valuable to include the expertise of ethicists familiar with AI who can identify potential areas of concern and help build appropriate safeguards.

Ensuring That Scientific Competencies are Enhanced Rather Than Diminished as a Consequence of the AI “Revolution”

With the rapid expansion and ease of use of ChatGPT and other LLM tools across a range of scientific and educational fields, there have been growing concerns related to the detrimental effects that the overuse of such tools can have on current and future generations of scientists and the public. These concerns include the real danger that regular substitution of LLM-generated written and oral communication in science, as well as across other facets of life, could significantly diminish cognitive competencies and the creative and innovative thinking that is the foundation of the scientific process itself. It is clear, at this early stage of generative AI use in participatory science as well as in other scientific areas, that carefully created guidelines for AI usage in all aspects of scientific education and training should be developed and put into place to ensure LLMs and other AI tools support, rather than replace, the facilitation and participation of community-engaged research endeavors.

A mitigation example is described here. A year-long deliberative process by Stanford University’s Department of Epidemiology & Population Health has resulted in a clear set of early principles and guidelines for the use of generative AI tools in one of its major education programs. The guidelines have helped to diminish ambiguity among students and faculty while reinforcing the importance of human-centered educational activities in growing the cognitive competencies of vital importance to the advancement of science. In addition to clear guidelines for AI use and requirements that scientists at all levels of training and experience review their data and methods for accuracy, other suggestions include the further development of field-specific scientific standards for scientific literacy and competence independent of AI usage.

It is also important to ensure that scientific data collection and analysis activities themselves are not in some way diminished or curtailed through overdependence on AI tools. As noted earlier, the judicious use of certain AI tools in the participatory science area may help to enrich and expand the insights and impacts gained through participation by residents, decision makers, and scientists working together. However, it is critical that such AI tools undergo thorough empirical testing to determine when, how, and with whom they can enhance rather than impede scientific knowledge and insights. As one example, the “Our Voice” research team is conducting initial experiments to better determine how the use of “text-to-image” generative AI may affect the ability of citizen scientists to better visualize and understand the proposed health-promoting environmental changes drawn from their data. Among the populations for which such tools are being systematically tested are older US women and school children living in the El Pozón neighborhood of Cartagena, Colombia, which was highlighted earlier. In a second example, a recent “Our Voice” citizen science project conducted in Bogotá, Colombia has illustrated how resident-collected data that were augmented via virtual reality tools can ignite compelling discussions and solution building among participants, scientists, and local stakeholders [22].

The multidimensional mixed-methods data collection methods, sources, and analyses required by the types of participatory scientific methods described in this Viewpoint also may serve as a means for preventing the diminution of cognitive competencies. For example, we have found that the increased complexity of the multilevel, mixed-methods data acquired when microscale (eg, citizen scientists’ “lived experiences”) and macroscale data sources are integrated requires an increased interdisciplinary understanding of both these different forms of data and more intricate and multifaceted data analysis methods, which can serve to enhance cognitive competencies [21].

Discussion and Conclusions

Citizen science has been shown to be effective in positively engaging diverse groups of community members in meaningful, action-oriented research [17]. Such participatory methods can help to lessen both mistrust in science and civic disengagement, which are trends that threaten to impair advances in public health. Can the judicious applications of various AI tools serve to amplify the potential of participatory science for reducing these trends while broadening meaningful scientific advances in public health? Given that it is clear that the AI “genie” is out of the bottle and will likely have substantial impacts across virtually all aspects of science, it is critical that we meet the AI challenge head on. The inevitable integration of AI into participatory research methods must be thoughtfully managed in order to amplify, and not replace, the collective documentation of human lived experiences, and to ensure that the applications of such data to affect change truly reflect the priorities of local communities. We believe that the time is ripe to systematically evaluate the potential of AI beyond current individual-level applications by exploring its utility as a community aid for advancing public health. At the same time, it is imperative that we continually keep in mind the potential risks accompanying its use, especially among underserved, under-resourced communities. We call on the research community to consider such challenges proactively and empirically over the coming years, finding innovative ways to mitigate the risks while testing the impacts of AI on trust in science, civic engagement, knowledge advancement, and community-driven change.

Acknowledgments

The authors thank Ann Banchoff, MSW, MPH and Steven Lovett, PhD for their valuable input related to this manuscript. We also thank Eduardo De La Vega Sr, PhD for his innovative early community-engaged research activities in Colombia using generative large language models, and Arjan Walia, MS for his leadership in the initial systematic evaluation of generative artificial intelligence models in community-engaged research with older US women. Dr. King received partial support from US Public Health Services grants 5R01AG07149002 (King) and 1R44AG071211 (M Fauci) from the National Institute on Aging. King, Doueiri, and Kaulberg received partial support from a seed grant provided by the Stanford Institute for Human-Centered Artificial Intelligence at Stanford University. The authors declare no potential conflicts of interest with respect to the research, authorship, or publication of this article. No copyrighted material, surveys, instruments, or tools are included in this manuscript.

Conflicts of Interest

None declared.

References

1. The Worlds I See: Curiosity, Exploration, and Discovery at the Dawn of AI: Macmillan Publishers; 2024:4.
2. Kennedy B, Tyson A. Americans' Trust in Scientists, Positive Views of Science Continue to Decline.: Pew Research Center; 2023. URL: https://www.pewresearch.org/science/2022/02/15/americans-trust-in-scientists-other-groups-declines [accessed 2023-11-14]
3. Achterberg P, de Koster W, van der Waal J. A science confidence gap: Education, trust in scientific methods, and trust in scientific institutions in the United States, 2014. Public Underst Sci 2017 Aug;26(6):704-720. [doi: 10.1177/0963662515617367] [Medline: 26644011]
4. Epstein Z, Sirlin N, Arechar A, Pennycook G, Rand D. The social media context interferes with truth discernment. Sci Adv 2023 Mar 3;9(9):eabo6169. [doi: 10.1126/sciadv.abo6169] [Medline: 36867704]
5. Nasr N. Overcoming the discourse of science mistrust: how science education can be used to develop competent consumers and communicators of science information. Cult Stud Sci Educ 2021;16(2):345-356. [doi: 10.1007/s11422-021-10064-6] [Medline: 34178180]
6. Alsan M, Wanamaker M. Tuskegee and the health of black men. Q J Econ 2018 Feb;133(1):407-455. [doi: 10.1093/qje/qjx029] [Medline: 30505005]
7. Garrison NA. Genomic justice for native Americans: impact of the Havasupai case on genetic research. Sci Technol Human Values 2013;38(2):201-223. [doi: 10.1177/0162243912470009] [Medline: 28216801]
8. Bazargan M, Cobb S, Assari S. Discrimination and medical mistrust in a racially and ethnically diverse sample of California adults. Ann Fam Med 2021;19(1):4-15. [doi: 10.1370/afm.2632] [Medline: 33431385]
9. Kannan VD, Veazie PJ. US trends in social isolation, social engagement, and companionship nationally and by age, sex, race/ethnicity, family income, and work hours, 2003-2020. SSM Popul Health 2023 Mar;21:101331. [doi: 10.1016/j.ssmph.2022.101331] [Medline: 36618547]
10. Byrne M. The disruptive impacts of next generation generative artificial intelligence. Comput Inform Nurs 2023 Jul 1;41(7):479-481. [doi: 10.1097/CIN.0000000000001044] [Medline: 37417716]
11. Haywood BK, Besley JC. Education, outreach, and inclusive engagement: Towards integrated indicators of successful program outcomes in participatory science. Public Underst Sci 2014 Jan;23(1):92-106. [doi: 10.1177/0963662513494560] [Medline: 23887249]
12. Mintz E, Couch J. Biomedical citizen science at the National Institutes of Health. Citiz Sci 2022 Dec 15;7(1):16651505. [doi: 10.5334/cstp.543]
13. Skarlatidou A, Haklay M, Hoyte S, van Oudheusden M, Bishop IJ. How can bottom-up citizen science restore public trust in environmental governance and sciences? Recommendations from three case studies. Environ Sci Policy 2024 Oct;160:103854. [doi: 10.1016/j.envsci.2024.103854]
14. King AC, Winter SJ, Chrisinger BW, Hua J, Banchoff AW. Maximizing the promise of citizen science to advance health and prevent disease. Prev Med 2019 Feb;119:44-47. [doi: 10.1016/j.ypmed.2018.12.016] [Medline: 30593793]
15. Allf BC, Cooper CB, Larson LR, et al. Citizen science as an ecosystem of engagement: implications for learning and broadening participation. Bioscience 2022 Jul;72(7):651-663. [doi: 10.1093/biosci/biac035] [Medline: 35769502]
16. King AC, King DK, Banchoff A, et al. Employing participatory citizen science methods to promote age-friendly environments worldwide. Int J Environ Res Public Health 2020 Feb 27;17(5):1541. [doi: 10.3390/ijerph17051541] [Medline: 32121001]

17. King AC, Odunitan-Wayas FA, Chaudhury M, et al. Community-based approaches to reducing health inequities and fostering environmental justice through global youth-engaged citizen science. Int J Environ Res Public Health 2021 Jan 21;18(3):892. [doi: 10.3390/ijerph18030892] [Medline: 33494135]
18. English PB, Richardson MJ, Garzón-Galvis C. From crowdsourcing to extreme citizen science: participatory research for environmental health. Annu Rev Public Health 2018 Apr 1;39:335-350. [doi: 10.1146/annurev-publhealth-040617-013702] [Medline: 29608871]
19. King AC. Our Voice Global Citizen Science Research Initiative. 2024. URL: https://ourvoice.stanford.edu [accessed 2025-01-27]
20. Ceccaroni L, Bibby J, Roger E, et al. Opportunities and risks for citizen science in the age of artificial intelligence. Citiz Sci 2019 Nov 28;4(1):1-14. [doi: 10.5334/cstp.241]
21. Soltani S, Hinman JA, Blanco-Velazquez I, et al. Bringing micro to the macro: how citizen science data enrich geospatial visualizations to advance health equity. J Maps 2023;19(1):2216217. [doi: 10.1080/17445647.2023.2216217] [Medline: 37448978]
22. Guevara-Aladino P, Sarmiento OL, Rubio MA, et al. Urban care for unpaid caregivers: community voices in the care block program, in Bogotá, Colombia. J Urban Health 2024 Dec;101(6):1113-1127. [doi: 10.1007/s11524-024-00899-z] [Medline: 39316308]
23. Hsu YC, Verma H, Mauri A, Nourbakhsh I, Bozzon A. Empowering local communities using artificial intelligence. Patterns (N Y) 2022 Mar 11;3(3):100449. [doi: 10.1016/j.patter.2022.100449] [Medline: 35510187]
24. Garba-Sani Z, Farincacci-Roberts C, Essien A. AI: A new framework for advancing health equity in health care AI. Health Aff 2024. [doi: 10.1377/forefront.20240424.369302]
25. AI4ALL webpage. : Stanford University URL: https://ai-4-all.org [accessed 2024-11-11]

Abbreviations

AI: artificial intelligence
LLM: large language model
PPSR: public participation in science research

Edited by A Mavragani; submitted 22.08.24; peer-reviewed by FA Causio, R Feltrero; revised version received 21.11.24; accepted 28.11.24; published 14.02.25.

Please cite as:
King AC, Doueiri ZN, Kaulberg A, Goldman Rosas L
The Promise and Perils of Artificial Intelligence in Advancing Participatory Science and Health Equity in Public Health
JMIR Public Health Surveill 2025;11:e65699
URL: https://publichealth.jmir.org/2025/1/e65699
doi:10.2196/65699

Recent Use of Novel Data Streams During Foodborne Illness Cluster Investigations by the United States Food and Drug Administration: Qualitative Review

Michael C Bazaco[1], PhD; Christina K Carstens[1], PhD; Tiffany Greenlee[1], MPH; Tyann Blessington[1], PhD; Evelyn Pereira[1], JD; Sharon Seelman[1], MS; Stranjae Ivory[1], MPH; Temesgen Jemaneh[1], PhD; Margaret Kirchner[1], PhD; Alvin Crosby[1], MBA; Stelios Viazis[1], PhD; Sheila van Twuyver[2], BS; Michael Gwathmey[3], MS; Tanya Malais[3], BS; Oliver Ou[1], PhD; Stephanie Kenez[1], MPH; Nichole Nolan[1], MPH; Andrew Karasick[1], MD; Cecile Punzalan[1], MD; Colin Schwensohn[4], MPH; Laura Gieraltowski[4], PhD; Cary Chen Parker[1], MPH; Erin Jenkins[1], MPH; Stic Harris[1], DVM

[1]US Food and Drug Administration, College Park, MD, United States
[2]US Food and Drug Administration, Minneapolis, MN, United States
[3]US Food and Drug Administration, Silver Spring, MD, United States
[4]US Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:
Michael C Bazaco, PhD
US Food and Drug Administration, College Park, MD, United States

Abstract

Foodborne illness is a continuous public health risk. The recognition of signals indicating a cluster of foodborne illness is key to the detection, mitigation, and prevention of foodborne adverse event incidents and outbreaks. With increased internet availability and access, novel data streams (NDSs) for foodborne illness reports initiated by users outside of the traditional public health framework have emerged. These include, but are not limited to, social media websites, web-based product reviews posted to retailer websites, and private companies that host public-generated notices of foodborne illnesses. Information gathered by these platforms can help identify early signals of foodborne illness clusters or help inform ongoing public health investigations. Here we present an overview of NDSs and 3 investigations of foodborne illness incidents by the US Food and Drug Administration that included the use of NDSs at various stages. Each example demonstrates how these data were collected, integrated into traditional data sources, and used to inform the investigation. NDSs present a unique opportunity for public health agencies to identify clusters that may not have been identified otherwise, due to new or unique etiologies, as shown in the 3 examples. Clusters may also be identified earlier than they would have been through traditional sources. NDSs can further provide investigators supplemental information that may help confirm or rule out a source of illness. However, data collected from NDSs are often incomplete and lack critical details for investigators, such as product information (eg, lot numbers), clinical or medical details (eg, laboratory results of affected individuals), and contact information for report follow-up. In the future, public health agencies may wish to standardize an approach to maximize the potential of NDSs to catalyze and supplement adverse event investigations. Additionally, the collection of essential data elements by NDS platforms and data-sharing processes with public health agencies may aid in the investigation of foodborne illness clusters and inform subsequent public health and regulatory actions.

(JMIR Public Health Surveill 2025;11:e58797) doi:10.2196/58797

KEYWORDS

foodborne illness surveillance; novel data streams; outbreak investigations; novel data; foodborne illness; foodborne; illness; United States; public health; prevention; outbreaks; social media; product review; cluster; product information; surveillance; epidemiology

Introduction

An estimated 1 in 6 people experiences an episode of foodborne illness in the United States annually, placing an estimated US $77.7 billion burden on the US economy [1,2]. Most of these illnesses are considered sporadic; however, limitations in foodborne outbreak surveillance may hinder the detection of outbreak-associated cases. The identification and investigation of foodborne illness clusters requires a high degree of coordination across state and local jurisdictions and, in some cases, federal agencies [3].

At the federal level, outbreaks are investigated when they involve cases in multiple states, implicate imported products, or when federal support is requested. Most of these outbreaks that involve known pathogens are traditionally initiated through laboratory-based surveillance systems with the identification of illness clusters by molecular subtyping of clinical isolates or case reports from ill individuals with a common exposure and illness who seek medical care. These clusters can then be linked to a common vehicle based on supporting epidemiologic information, traceback findings, and laboratory testing results from food and environmental samples. This information is usually first received by state or local health departments (LHDs) before it is passed on to federal partners, when deemed necessary. Outbreaks that involve known pathogens may also be identified through matching the genetic similarity of historic and contemporary isolates from product or environmental samples to clinical isolates via molecular subtyping techniques, such as whole genome sequencing [4].

In situations where an etiology is unknown or has not yet been identified, consumer complaint systems provide a mechanism for consumers to self-report adverse health events, including foodborne illness. This allows public health entities to conduct disease surveillance to help mitigate this primary limitation of laboratory-based surveillance. Consumer complaint–based systems allow for the collection and storage of consumer reports that can be submitted in a range of formats, including in-person, phone, e-mail, and web form [5]. A survey of LHDs in the United States revealed that a majority (81%) use consumer complaint–based systems and attributed the detection of 69% of foodborne outbreaks in the past year to consumer complaints [6]. State health departments that have conducted self-assessments of their consumer-complaint systems, such as Minnesota and Rhode Island, have credited the detection of 79% and 80% of foodborne outbreaks to consumer complaints, respectively [7,8]. Of the surveyed LHDs in the United States, 75% reported the ability to receive consumer complaints via email, and 40% reported the availability of a web-based reporting form [6]. The Integrated Food Safety Centers for Excellence help states to continually improve these complaint-based surveillance systems as well [9]. In the United States, the various consumer complaint systems create a patchwork that helps to capture information on foodborne illnesses that may not be reported to medical professionals. However, holistically, this patchwork can pose challenges as data from the various consumer complaint systems are stored separately. Recognition of signals depends on a sufficient number of complaints to come to the same complaint system.

Complaint-based surveillance is also used at the federal level to identify potential clusters of foodborne illness. The US Food and Drug Administration (FDA) collects complaints or reports submitted from consumers, health care providers, industry members, public health officials, or other submitters via 2 web-based mechanisms using the MedWatch or Safety Reporting Portal as well as by phone, email, fax, and physical mail [10-13]. A complainant may submit similar reports via multiple mechanisms, and it may not be possible to distinguish if they are duplicate reports or are separate cases. A key element of identifying illness clusters via this approach is the availability of complete and reliable epidemiologic data, such as illness onset date and contact information, to link complainant reports. This information is also needed to narrow down potential etiology and food vehicle [5].

Disease surveillance and outbreak information received by the FDA from various internal and external sources, including federal, state, and local health partners, is evaluated by the FDA's Coordinated Outbreak Response & Evaluation (CORE) Network, which was established in 2011 and charged with coordinating efforts to find, stop, and prevent illnesses linked to FDA Center for Food Safety and Applied Nutrition (CFSAN) regulated products [14]. These products include all human foods not falling under the jurisdiction of the US Department of Agriculture Food Safety Inspection Service and account for 75% of the food supply in the United States [15]. Once an actionable signal has been identified, further investigational activities and public health actions (eg, product traceback, recalls, and communication) are coordinated across local, state, and federal partners, a process that has been described elsewhere [3].

Traditional investigative tools, such as laboratory-based surveillance, do face challenges including but not limited to underdiagnosis and underreporting of illnesses, varying illness reporting structures, the increased use of culture-independent diagnostic testing, and resource availability [16,17]. While some of these challenges are mitigated by public health consumer complaint systems described earlier, the increased accessibility and use of novel data streams (NDSs) provide new information that may further aid in the detection and investigation of foodborne illness clusters. NDSs include nontraditional sources of information such as social media, web-based customer reviews, multimedia news (comment sections from news stories about clusters of illnesses), and private participatory reporting platforms (PPRPs) that actively collect data on food poisoning.

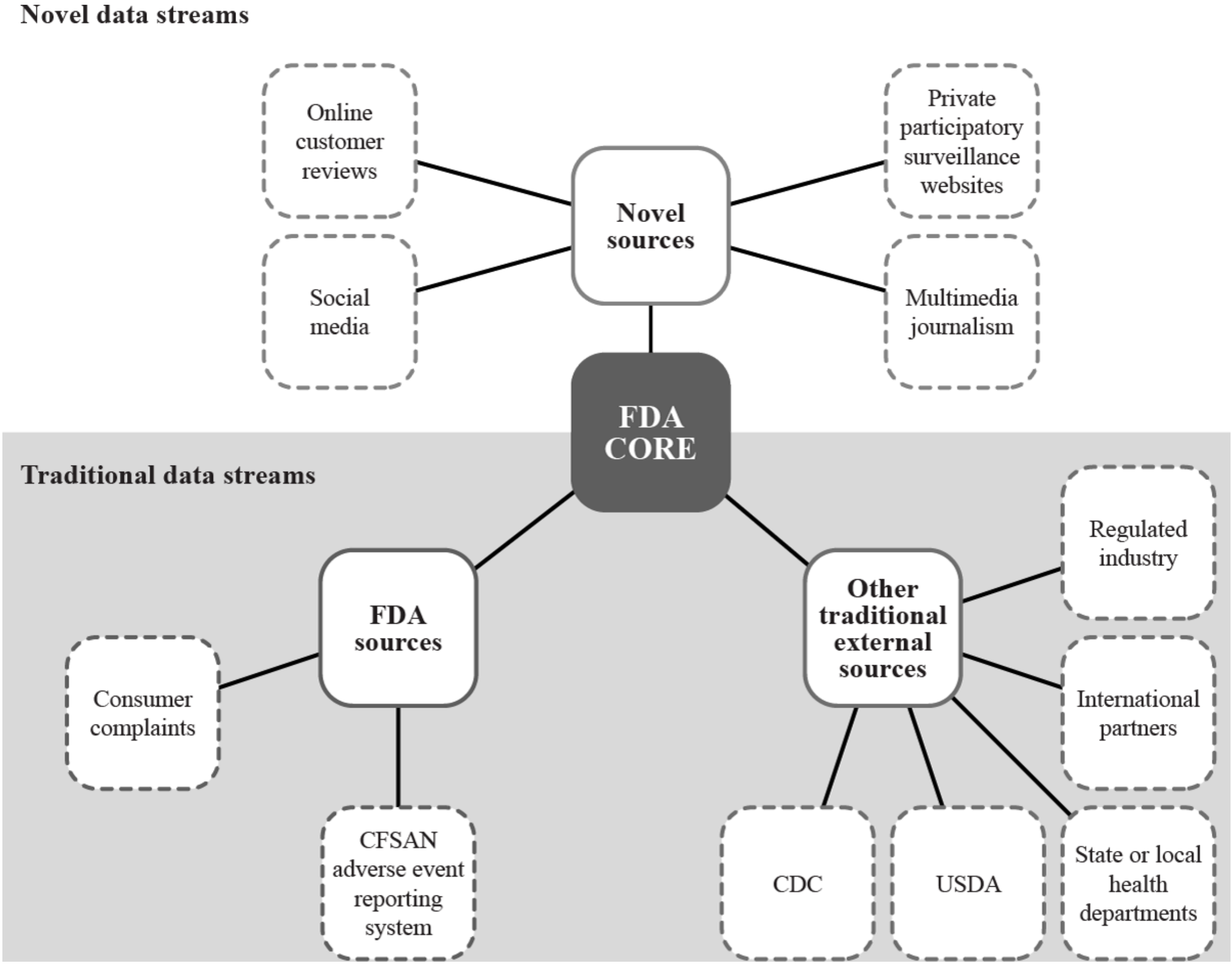
The availability of crowdsourced web-based information provides an opportunity for public health entities to potentially augment existing traditional foodborne illness surveillance systems or identify clusters not detected by traditional tools. The evolution of cell phones to more readily allow users to access the internet has made these NDSs even more accessible as a tool for the submission of consumer complaints. These NDSs, in the context of foodborne illness, consist of reports of consumer information related to food products that are relayed digitally through social media, search histories, customer reviews (ie, product and restaurant reviews), and PPRPs [18]. Some PPRP sites have the express goal of collecting data on people who believe that they have gotten sick from food. Information from these sources may be shared directly with public health agencies or may be mined by agencies in response to a signal.

NDSs can be integrated with traditional public health surveillance programs to catalyze or supplement investigation, and several examples of this exist at the local and state health department level. For instance, the Chicago Department of Public Health has used social media posts generated within jurisdictional limits to address consumer complaints of food poisoning and to target health inspections [19]. The New York

City Department of Health and Mental Hygiene has similarly evaluated the use of customer-submitted restaurant reviews to detect outbreaks and target health inspections [20]. These are just a few documented examples of the use of NDSs by state and LHDs. No systematic use of these data sources has been documented at the federal level at this time.

Reports of foodborne illness clusters can come from a variety of traditional and novel data sources (Figure 1). In this review, we (1) highlight key details of 3 recent adverse event clusters that were investigated by the FDA and were catalyzed or supplemented by NDSs, (2) discuss challenges and limitations associated with the current use of NDSs by the FDA, and (3) propose opportunities for improvement and the future use of NDS applications during foodborne illness surveillance and investigational activities conducted by the FDA.

Figure 1. Examples of adverse event and illness cluster data sources evaluated by the US Food and Drug Administration Coordinated Outbreak Response & Evaluation (FDA CORE) Network. Multimedia journalism may also be considered a traditional data source in some cases. CFSAN: Center for Food Safety and Applied Nutrition; CDC: Centers for Disease Control and Prevention; USDA: US Department of Agriculture.



Three Case Studies

Overview

To illustrate the use of NDSs during FDA investigations of foodborne illness clusters, the following 3 case studies describe recent incidents linked to dry breakfast cereal, ketogenic shakes, and plant-based crumbles, respectively. Each of these incidents provides a unique look at how these data streams were used by the FDA to identify or supplement investigations. There may be duplicate reports included in the analyses of these case studies, as reports may have been submitted by the same complaint through multiple points in FDA and to a given PPRP or other form of NDS. The following case studies are meant to highlight examples of this use and are not meant to be comprehensive reviews of the entire incidents.

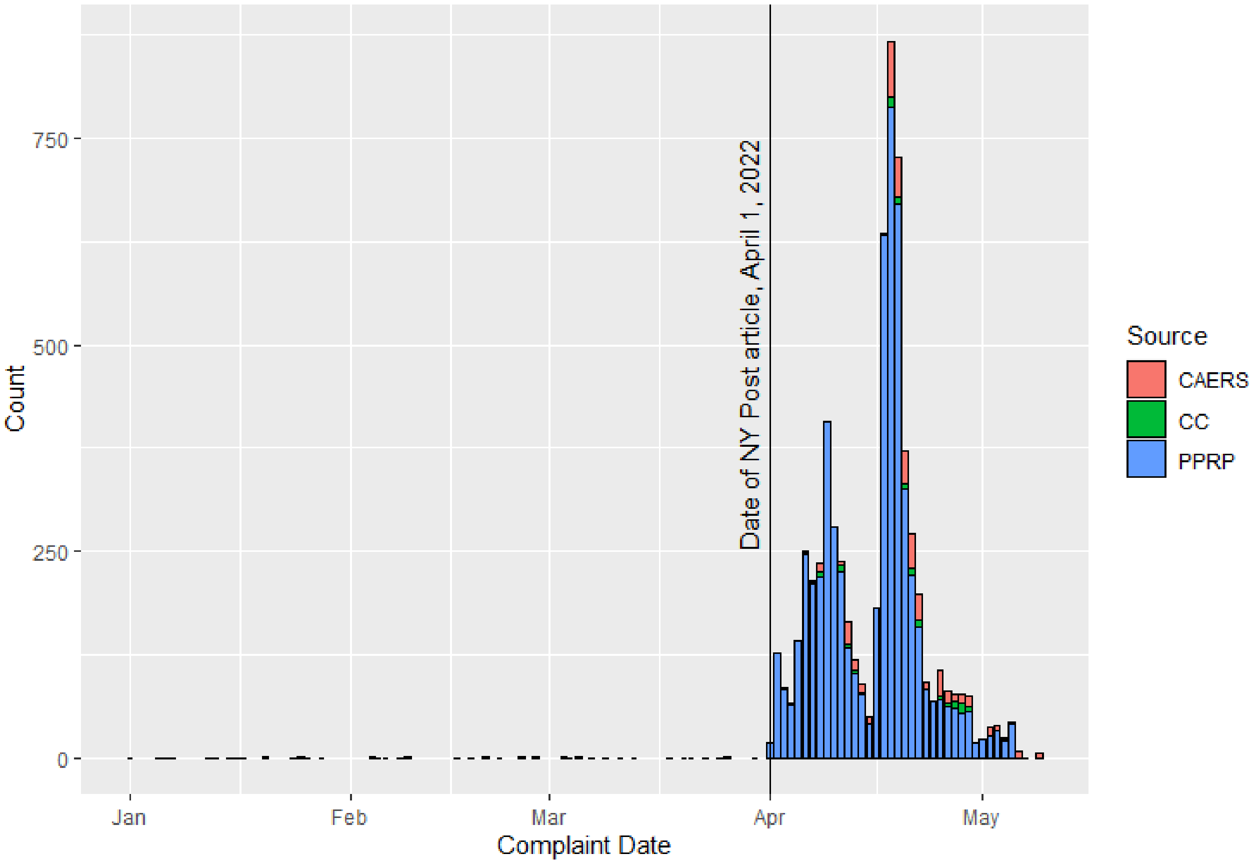
Dry Breakfast Cereal

On April 1, 2022, the New York Post published an article that highlighted an increase in illness reports mentioning a dry breakfast cereal product on a PPRP [21]. The article stated that at least 139 consumers reported vomiting and diarrhea after consuming the dry breakfast cereal of interest, which prompted an FDA investigation. After FDA CORE initiated an investigation, the PPRP shared complaint data with the FDA, which was reviewed by FDA CORE and used for an initial assessment of the incident. FDA's medical evaluation of 470 PPRP reports received from January 1, 2022, to April 5, 2022, identified symptom onset data for 349 (74%) reports, most of which (96%) suggested nonrapid symptom onset (exceeding 30 min since consumption). In some cases, due to the nature of the reporting tool, these data sometimes omitted necessary epidemiologic information such as the date the person first reported feeling ill, food exposure date, and illness duration.

Thus, the investigation instead focused on reports directly submitted to the FDA. FDA consumer complaint reports collected between June 1, 2021, and May 10, 2022, that mentioned exposure to the dry breakfast cereal product of interest and explicit gastrointestinal (GI) symptoms were reviewed. Only one of the 120 consumer complaint reports submitted between June 1, 2021, and May 10, 2022, was received in 2021. The remaining 119 were submitted after April 1, 2022, the day the NY Post article was published. Of these 120 consumer complaints, many reported GI symptoms.

Between June 1, 2021, and May 10, 2022, there were a total of 438 Center for Food Safety and Applied Nutrition Adverse Event Reporting System (CAERS) reports associated with the dry breakfast cereal product, 437 of which were submitted after April 1, 2022 (Figure 2). The FDA conducted an analysis of 406 CAERS reports that contained suspected cases with exposure to the breakfast cereal, explicit GI symptoms, and were received prior to April 30, 2022. Of 406 CAERS reports, only 160 had information on illness onset, most of which suggested symptom onset occurred 1‑24 hours after consumption. Overall, 558 FDA adverse event reports (120 consumer complaint and 438 CAERS reports) and 6272 PPRP reports of GI illness were associated with this incident.

Figure 2. Distribution of adverse event reports received by various reporting platforms concerning dry breakfast cereal (2022). CAERS: Center for Food Safety and Applied Nutrition Adverse Event Reporting System; CC: Food and Drug Administration consumer complaints; NY: New York; PPRP: private participatory reporting platform.



In response to illness reports related to the dry breakfast cereal product, the FDA conducted preventive control inspections at 2 manufacturing facilities that produced the dry breakfast cereal of interest and at 1 product supplier in order to document manufacturing operations and identify potential deficiencies. An FDA investigation was also conducted at the corporate offices of the dry breakfast cereal brand in order to collect records pertaining to manufacturing processes, facilities, products, suppliers, quality control practices, and consumer complaint data. Concurrently, the Michigan Department of Agriculture and Rural Development initiated a preventive control inspection at another supplier to the brand. No significant observations were found during these inspections. FDA chemical and microbial analyses were completed on product and environmental samples. FDA and state partners also analyzed clinical and consumer product samples. However, no causative agent for this incident was identified. Evidence from this investigation indicated that this was, in fact, not a microbial outbreak related to dry cereal.

Ketogenic Shakes

In March 2022, a PPRP shared with the FDA an analysis of complaints they received about Brand A protein or meal replacement shakes, a generic brand produced for Retailer A. The analysis demonstrated that there was an increase in complaints associated with Brand A when compared to other protein or meal replacement shake brands starting in January 2021. The PPRP provided all complaints to the FDA for Brand A protein or meal replacement shakes submitted between

January 2021 and March 2022. A review of this information found that 28 of 32 reports specifically identified the ketogenic shake from Brand A, and the FDA initiated an investigation into this product. During the investigation, 6 additional complaints were received, culminating in a total of 34 complaints specifically identifying Brand A ketogenic shakes reported to the PPRP between January 2, 2021, and March 30, 2022. These PPRP reports were challenging to use as they lacked complete epidemiologic information, and specific product details, such as lot numbers, were not consistently available. Further, some of the data that was provided was unstructured and required manual review and extraction.

Additionally, reviews for the ketogenic shake product were extracted on March 24, 2022, from Retailer A's website. This data also required manual review and data extraction. Between January 9, 2021, and February 2, 2022, 54 reviews contained text with complaints indicating an adverse event. Finally, FDA systems were queried for complaints related to the Brand A ketogenic shake over the previous 2 years. Between June 6, 2021, and April 8, 2022, 2 consumer complaints and 4 CAERS reports were received by the FDA. Complaints from all 3 data sources (FDA, consumer reviews, and PPRP complaints) were compiled, and a total of 93 were reviewed. Symptoms were varied but included mostly GI distress, and many reported a rapid onset of symptoms (<30 min). The number of reports received for Brand A was compared to the number for Brand B ketogenic shake to ascertain whether the amount of complaints was unusual. The Brand B ketogenic shake had nearly equivalent product labeling to the Brand A ketogenic shake, but the 2 brands did not share a manufacturer. Both brands were nationally distributed in the United States. The FDA CAERS database and the PPRP were also queried for complaints that specified Brand B during the same time frame. These Brand B queries returned no complaints from the PPRP and 1 complaint from FDA CAERS.

Closed retail samples of 2 flavors of Brand A ketogenic shakes and open and closed samples from a complainant's home were collected. Chemical and sensory analyses were completed on all product samples collected by the FDA, and no unusual findings were identified. No public communications or regulatory actions resulted from this investigation. However, details of the investigation were shared with Retailer A, who had already discontinued distribution of the product based on poor web-based reviews and who further, in response to FDA notification, made the remaining product in their stores unavailable for purchase. Ultimately, laboratory testing did not detect any specific undeclared ingredients or toxins as a causative agent for these adverse events, but the investigation did highlight potential consumer intolerances to some declared ingredients used in ketogenic food products.

Plant-Based Crumble Product

In June 2022, the FDA received consumer complaints and adverse event reports referencing GI illness suggesting abnormal liver function (eg, jaundice and dark urine) that may have been attributable to the consumption of a plant-based crumble product produced by Brand C. As of June 23, 2022, 20 CAERS reports and 13 consumer complaints had been evaluated. Although some complainants reported several Brand C products, all complaints specified the plant-based crumble product. Most reported GI and liver-associated symptoms. This incident became heavily publicized as reports pertaining to the adverse events were reported digitally. Additionally, consumers reported adverse health outcomes to various other sources, such as state reporting mechanisms, a PPRP, and directly to Brand C. As of October 18, 2022, this incident included 393 (177 consumer complaints and 216 CAERS reports) adverse event reports in 39 states with 133 hospitalizations and zero deaths. Estimated illness onset dates ranged from April 19, 2022, to September 4, 2022.

Closed and open samples of Brand C plant-based crumbles product were sampled from complainants' homes. Closed and undistributed product samples were also obtained from Brand C's warehouses. Both chemical and microbiological analyses were performed, but there were no significant findings. However, in 2024, the FDA made a public posting noting that tara flour, an ingredient used in this product, did not meet the Generally Recognized as Safe standard, so it was an unapproved food additive [22]. The FDA issued a web post on the investigation and posted the company recall on the FDA recalls page [23,24]. The FDA further performed several domestic inspections related to the incident and 1 foreign inspection, though none of these actions provided evidence of contamination.

During this investigation, the FDA also reached out to the PPRP to examine any reports they had received regarding this product. The website had received 24 complaints associated with Brand C products from April 4, 2022, to October 18, 2022, 7 of which mentioned the product of interest (plant-based crumbles). Of these 7 complaints, 4 reported GI symptoms and 3 reported various liver-associated ailments. Records from the PPRP were obtained to rule out other vehicles of interest, such as other Brand C products and products from a separate company, Brand D. Although 8 other products were mentioned, none were mentioned more than twice. Further, the PPRP had only received a single complaint for Brand D, which described nausea and vomiting. Based on these reports and the data from FDA systems, other Brand C and Brand D products were excluded as not likely vehicles of illness for this incident.

Discussion

Overview

The use of NDSs to identify adverse event signals or supplement FDA investigations of foodborne illness clusters may offer opportunities such as earlier detection of clusters and vehicle confirmation as seen in the case studies described earlier. However, data collection from NDSs also presents challenges that may reduce their utility during ongoing incident investigations. The advantages and challenges associated with the use of NDS data are discussed in the following section, along with a summary of future steps to better harmonize public health efforts.

Advantages

Traditional, laboratory-based methods of foodborne disease surveillance are largely reliant on the detection of foodborne illnesses via information collected at the clinical level, such as laboratory tests and health care provider reports [25]. These signals are dependent on ill persons seeking medical care, yet some research indicates that fewer than 15% of individuals with self-suspected foodborne illness seek medical attention or report to an LHD [25]. The most recent Foodborne Disease Active Surveillance Network (FoodNet) population survey supports this, noting that only 16.9% of respondents with an acute GI illness sought medical care [26]. Further, most episodes of foodborne illness are self-limiting, and only a small percentage of illnesses in the United States are estimated to result in hospitalization [2]. Once a case of foodborne illness is detected, associated epidemiologic information is gathered by trained epidemiologists, which ensures collected information is detailed and accurate. However, national mandatory reporting of foodborne illnesses is limited to a select set of diseases [27]. While this list includes most well-established foodborne pathogens, it is not as useful for illnesses without a known etiology. Notification of the public health system to new illnesses is also slow, as lag times often range from 2 to 4 weeks or more, depending on the etiology, time of year, and laboratory testing capability. As discussed, some of these limitations are mitigated by various consumer complaint-based systems that are not limited by specific etiology and do not require a person to seek medical care. This is particularly important for localized outbreak detection.

NDSs have the potential to address gaps that exist in traditional foodborne illness surveillance methods and empower consumers. For example, data from these systems can alert public health agencies to foodborne illnesses in situations where medical care is not obtained (eg, due to financial reasons, lack of perceived necessity). NDSs may also provide early signals for foodborne illness clusters that would otherwise be slowed due to expected lag times in case reporting or unforeseen delays. Common barriers to traditional reporting, such as cost, time, access, or awareness, may also be avoided by consumers who use NDSs. In the United States, mechanisms to report an adverse event directly to a public health entity vary across the state and federal levels, while NDS companies have varying user-friendly formats for accepting consumer complaints. Therefore, the generation of an adverse event report on a PPRP or a review on a product webpage may be more intuitive, simpler, and less time-consuming for the public in some situations.

The broad reach of NDSs strengthens their ability to catalyze public health investigations. The larger number of associated complaints can also contextualize reported symptoms, which is useful in situations where the pathogen or hazard is emerging and case symptomology is unknown. One example of this is the ketogenic shake investigation, in which complainants described a rapid onset of short-lived symptoms after consuming the product, making it more unlikely for complainants to seek care under these circumstances. Despite the low volume of related complaints reported to the FDA, which typically would not rise to the level of an investigation, an FDA investigation was triggered due to the number of complaints reporting a similar product or symptomology within product reviews and PPRP reports.

A signal received via NDS of a suspected illness related to an FDA-regulated product may allow the FDA to evaluate consumer risk using investigative tools such as traceback if the incoming data includes necessary details about the product. Traceback helps ascertain if illnesses are related to a common source, which leads to further investigation and, potentially, product actions. Details within NDS reports may help to target firm inspections, product sample collection, and product testing, all of which can inform compliance actions. Retailers may also use information from NDSs as an early alert to a poor-quality product or one that may be causing adverse reactions among consumers, such as in the example provided with ketogenic shakes. This unsolicited information allows retailers to internally initiate actions to protect consumers prior to a manufacturer-initiated product recall or the attention of regulatory and public health agencies. Traceback information obtained from NDSs may also help to rule out products not linked to adverse events. Although NDS data can inform industry and regulatory investigations, the source for NDS reports may be anonymous; therefore, the veracity of the data should be considered during decision-making.

Challenges

An overarching limitation of NDSs stems from the quality of available data. Consumer complaints made to the FDA use forms that incorporate standardized epidemiologic data elements, including those identified by the Council to Improve Foodborne Outbreak Response [27]. These forms also collect details used to assess complaint trends and detect potential food safety issues. NDSs do not reliably capture key data elements essential for public health investigations such as those outlined in Table 1, which includes an expanded list of data elements in addition to those put forth by the Council to Improve Foodborne Outbreak Response. As NDSs often obtain information for other purposes, data gaps are expected and can hinder public health investigations. For example, epidemiologic information is used by investigators to identify a common etiologic agent, which is critical in the absence of clinical validation. Without epidemiologic information, the cause of a food safety issue is more challenging to identify, and the scope of product testing is significantly more resource intensive. Product and exposure information are also necessary to identify a common source of illness and inform traceback investigations. A lack of actionable data inhibits public health response, which was illustrated during the 3 investigations described here.

Another limitation associated with NDS systems is that the data are often unstructured. Information from complainants is typically in the form of free-text entry (eg, social media posts, product reviews, web-based comments), which currently require manual review to extract key data elements. This process is resource-intensive, subject to human error, and has limited feasibility for large numbers of posts or comments. Automated data review may reduce the burden of manual review in the future, but specialized tools are expensive investments for ad hoc analyses, and varying data sources may necessitate different technical approaches. Similar to traditional outbreak data sources, NDS reports may also be duplicated across reporting platforms, as individuals may submit complaints on multiple NDS platforms, multiple complaints on one platform, and directly to the FDA (Figure 1). Some NDS platforms do minimize report duplication with requirements for users to sign into an account, to have verified purchases to leave a review, or by tracking the IP address of the user. However, evaluation of complaint uniqueness for enumeration across platforms may require manual deduplication or simply not be possible.

Table . List of key data elements for actionable public health response for collection by novel data systems, expanded and modified from the Council to Improve Foodborne Outbreak Response [27].

Data type	Data element
Epidemiologic	• Illness onset date and time • Symptoms • Duration of symptoms • Care-seeking behavior • Suspected exposure of interest • Exposure date • Other recent exposures (past 7 days)
Product	• Photo of product packaging • Brand • Best buy date or best if used by date • Lot code • Product description on label • Purchase location and purchase date • Manufacturer, distributor information
Complainant’s contact information	• Name • Phone number • Email address
Affected individual’s contact information (if not the same as the complainant’s)	• Name • Phone number • Email address
Affected individual’s demographic information	• Age • Gender • State of affected individual’s residence
Method of report	• Date of complaint • Type of platform • Method of complaint

Ultimately the structure, breadth, and utility of information available in an NDS are dependent on the purpose of the platform. Certain NDSs, such as web-based reviews and social media posts, share information publicly, though not explicitly for public health goals. Due to this, food safety complaints made on these platforms usually lack critical data elements but have the potential to provide contextual information that can catalyze or supplement an investigation. PPRPs do have a public health goal, and compared to official government complaint systems, they may be easier to find and use, which makes them attractive options to complainants. However, although these sites may be easier to use, the data captured may be incomplete or of lower quality than data collected from traditional complaint systems. Public health and regulatory officials must keep these limitations in mind when considering NDSs in an investigational context.

When a signal for an adverse event is received, public health investigators review the reported symptoms, in conjunction with laboratory diagnoses, to determine an etiology and define cluster criteria (ie, develop a case definition) [27]. With NDSs, limited epidemiologic information may result in nonspecific symptomology. Alternatively, investigators may rely on medical records and laboratory documentation to accumulate evidence in search of a possible etiology, but complainant contact information is not often included in NDS reports, which prevents investigators from obtaining follow-up information. Thus, reports from NDSs must be grouped based on available epidemiologic features, which results in a broad case definition, extensive record requests, and resource-intensive review by investigators. Therefore, biological plausibility and medical review should be considered when refining a case definition and evaluating likely exposures when using NDSs.

Evidence has shown that media attention, particularly TV news coverage, can lead to an increase in adverse event reporting during a health scare [28,29]. This would also affect reporting to NDSs, which could see greater impacts given their ease of reporting, particularly if an NDS uses tools such as social media to share information on the incident. This effect was illustrated during the dry cereal incident, as complaints to both NDSs and FDA increased dramatically after newspaper reporting (Figure 2). One possible further complication of media attention is the amplification of unreliable and spurious signals that leads to stress on public health response entities, who already have limited resources.

Many actors are involved in multistate investigations, each with unique roles and challenges. One of these challenges is identifying proper medical and epidemiologic experts to evaluate data collected from PPRP or other NDS sources. CDC, for example, has investigational programs aligned to specific pathogens, so it can be difficult to identify which group would provide expertise during an investigation involving an unknown etiology.

As NDS platforms are not managed by established public health systems, information sharing with entities that have authority to take public health action is limited. Although some public health agencies may partner with third-party NDS platforms, subscribe to their services, or survey their data, NDS data may not be automatically escalated to a public health authority, which may not always be clear to the public when submitting their report. Also, some NDS companies may require public health agencies to pay to access complaint data. In addition, NDS platforms vary in the user information they collect, which can limit the ability to conduct follow-up with ill persons. It is also important to note that these NDS systems may divert ill people away from submitting reports to established public health reporting and surveillance systems. People who report their illnesses on these NDS platforms may assume that their reports will automatically be shared with authorities, which is not the case. In addition, submitting complaints to NDS systems may give people the false expectation that the information they provide will be actionable by public health agencies.

One reason that laboratory results can be confirmed relatively quickly in traditional outbreaks is that the etiological agent is known, and multiple laboratories are often able to perform and analyze testing results. Conversely, NDS-detected outbreaks often have an unknown etiological agent, which requires broad, instead of targeted, testing and often takes longer to return results.

Similar to limitations discussed regarding epidemiologic data, information collected from consumers through NDSs may lack critical elements needed for product traceback, such as product and purchase information. Although photos can be shared of a product, photos and product descriptions obtained from consumers do not consistently provide information on lot numbers, purchase dates, or purchase locations. Using a phone to submit complaints, while convenient, may limit the amount of information that people are able to submit. This information can aid regulatory authorities to conduct traceback, as some products may have the same name or label but have been produced or grown at different locations. Purchase information may also inform traceback investigations through identification of the manufacturer; however, these details are not consistently captured from NDSs or from traditional outbreak data sources. Ultimately, the lack of consistent product and purchase information may hinder the identification of a specific implicated product for further industry or regulatory action.

Opportunities

While NDSs have been helpful at the local level in identifying establishment-based illness clusters or targeting limited resources for inspections, foodborne investigations at the federal level focus on multistate outbreaks or vehicles that enter interstate commerce [7,20,28]. In these situations, etiologic agents are generally not related to ill workers or conditions at a retail establishment, but rather issues further up the supply chain. Similar to traditional outbreak data sources, NDSs reports may be of assistance during FDA investigations to identify signals related to branded or packaged products that are easily identified by the consumer, as opposed to loose items, such as produce, with no clear or easily identifiable brand or manufacturer. NDSs are also most likely to identify incidents that involve a rapid onset of symptoms, as with the ketogenic shakes investigation, or unusual symptoms, as in the plant-based crumble product investigation.

Signal detection and data collection through NDSs will likely become more prevalent. Therefore, it is incumbent upon public health agencies to determine the best ways to use this data, integrate it into established signal detection systems, and potentially collaborate with platform hosts. There is also a need to establish a process, once a signal is detected via NDS, where ill individuals can be interviewed with a standardized questionnaire that can be used to further evaluate potential exposures and etiologies. NDS providers should consider adding the noted critical data elements (Table 1) to existing data captured by their sites or provide public health officials with contact information to follow up with complainants. CDC and state health departments or LHDs can play an important role in coordinating patient interviews and epidemiological data collection, including medical records. In the future, the further development of tools for natural language processing, website scraping, and optical character recognition technology (software to extract text from an image) may also be useful when applied to NDSs for signal detection or to supplement ongoing outbreak investigations.

Conclusions

NDSs may enable quicker signal detection in certain situations compared to traditional surveillance methods for foodborne illness clusters attributable to unknown agents. Recently, NDSs have provided supplementary data during traditional investigations and sparked investigations into previously undetected signals, which may or may not have provided enough information for an actionable public health response. While PPRPs and product reviews are the focus in this paper, future signals may come from other NDSs, such as web-based social media platforms. To effectively integrate these data into public health surveillance and response, working protocols could be developed to help agencies better collect, surveil, and assess this data. NDSs with a focus on public health should also do their best to collect critical epidemiologic, product, and clinical data, enhancing their utility as a resource for public health agencies.

Acknowledgments

The response efforts discussed in this manuscript included numerous public health officials at local and state health and regulatory departments and public health laboratories in the United States, who serve as the backbone of any multi-state foodborne illness outbreak investigation. We also appreciate the assistance provided by US Food and Drug Administration (FDA) Emergency Response Coordinators who coordinated the response between FDA and state partners. The findings and conclusions of this report are those of the authors and do not necessarily represent the official position of the FDA and the Centers for Disease Control and Prevention (CDC).

Conflicts of Interest

None declared.

References

1. Scharff RL. Economic burden from health losses due to foodborne illness in the United States. J Food Prot 2012 Jan;75(1):123-131. [doi: 10.4315/0362-028X.JFP-11-058] [Medline: 22221364]
2. Scallan E, Griffin PM, Angulo FJ, Tauxe RV, Hoekstra RM. Foodborne illness acquired in the United States--unspecified agents. Emerg Infect Dis 2011 Jan;17(1):16-22. [doi: 10.3201/eid1701.091101p2] [Medline: 21192849]
3. Seelman S, Viazis S, Merriweather SP, Cloyd TC, Aldridge M, Irvin K. Integrating the Food and Drug Administration Office of the Coordinated Outbreak Response and Evaluation Network’s foodborne illness outbreak surveillance and response activities with principles of the National Incident Management System. J Emerg Manag 2021;19(2):131-141. [doi: 10.5055/jem.0567] [Medline: 33954963]
4. Wellman A, Bazaco MC, Blessington T, et al. An overview of foodborne sample-initiated retrospective outbreak investigations and interagency collaboration in the United States. J Food Prot 2023 Jun;86(6):100089. [doi: 10.1016/j.jfp.2023.100089] [Medline: 37024093]
5. Kim TN, Edmundson AR, Hedberg CW. Foodborne Illness Complaint Systems Detect, and Restaurant Inspection Programs prevent restaurant-associated foodborne illness outbreaks. Foodborne Pathog Dis 2024 Feb;21(2):92-98. [doi: 10.1089/fpd.2023.0086] [Medline: 38010953]
6. Li J, Shah GH, Hedberg C. Complaint-based surveillance for foodborne illness in the United States: a survey of local health departments. J Food Prot 2011 Mar;74(3):432-437. [doi: 10.4315/0362-028X.JFP-10-353] [Medline: 21375880]
7. Li J, Smith K, Kaehler D, Everstine K, Rounds J, Hedberg C. Evaluation of a statewide foodborne illness complaint surveillance system in Minnesota, 2000 through 2006. J Food Prot 2010 Nov;73(11):2059-2064. [doi: 10.4315/0362-028x-73.11.2059] [Medline: 21219718]
8. Yousaf A, Viveiros B, Caron G. Rhode Island Department of Health Foodborne Illness Complaint System: a descriptive and performance analysis. J Food Prot 2019 Sep;82(9):1568-1574. [doi: 10.4315/0362-028X.JFP-19-135] [Medline: 31429595]
9. Integrated Food Safety Centers of Excellence. URL: https://foodsafetycoe.org [accessed 2025-02-25]
10. Punzalan C, Budnitz DS, Chirtel SJ, et al. Swallowing problems and dietary supplements: data from U.S. Food and Drug Administration Adverse Event Reports, 2006-2015. Ann Intern Med 2019 Nov 19;171(10):771-773. [doi: 10.7326/M19-0947] [Medline: 31426091]
11. MedWatch online voluntary reporting form. US Food and Drug Administration. 2024. URL: https://www.accessdata.fda.gov/scripts/medwatch/index.cfm [accessed 2025-02-25]
12. Safety Reporting Portal. 2024. URL: https://www.safetyreporting.hhs.gov/SRP2/en/Home.aspx?sid=09abe76b-57ab-49a2-9ff9-22cd52eb8e32 [accessed 2025-02-25]
13. Consumer complaint coordinators. US Food and Drug Administration. 2024. URL: https://www.fda.gov/safety/report-problem-fda/consumer-complaint-coordinators [accessed 2025-02-25]
14. About the CORE network. US Food and Drug Administration. 2024. URL: https://www.fda.gov/food/outbreaks-foodborne-illness/about-core-network [accessed 2025-02-25]
15. FDA at a glance: regulated products and facilities. US Food and Drug Administration. 2024. URL: https://www.fda.gov/media/115824/download#:~:text=FDA%20regulates%20about%2075%20percent%20of%20the%20U.S.%20food%20supply [accessed 2025-02-25]
16. Scallan E, Hoekstra RM, Angulo FJ, et al. Foodborne illness acquired in the United States--major pathogens. Emerg Infect Dis 2011 Jan;17(1):7-15. [doi: 10.3201/eid1701.p11101] [Medline: 21192848]
17. White AE, Tillman AR, Hedberg C, et al. Foodborne illness outbreaks reported to National Surveillance, United States, 2009-2018. Emerg Infect Dis 2022 Jun;28(6):1117-1127. [doi: 10.3201/eid2806.211555] [Medline: 35608555]
18. Deng X, Cao S, Horn AL. Emerging applications of machine learning in food safety. Annu Rev Food Sci Technol 2021 Mar 25;12:513-538. [doi: 10.1146/annurev-food-071720-024112] [Medline: 33472015]
19. Harris JK, Mansour R, Choucair B, et al. Health department use of social media to identify foodborne illness - Chicago, Illinois, 2013-2014. MMWR Morb Mortal Wkly Rep 2014 Aug 15;63(32):681-685. [Medline: 25121710]

20. Harrison C, Jorder M, Stern H, et al. Using online reviews by restaurant patrons to identify unreported cases of foodborne illness - New York City, 2012-2013. MMWR Morb Mortal Wkly Rep 2014 May 23;63(20):441-445. [Medline: 24848215]
21. Fickenscher L. Lucky Charms cereal causing vomiting, diarrhea: report. New York Post. URL: https://nypost.com/2022/04/01/lucky-charms-cereal-causing-vomiting-diarrhea-report/ [accessed 2025-02-25]
22. FDA update on the post-market assessment of tara flour. US Food and Drug Administration. 2024. URL: https://www.fda.gov/food/hfp-constituent-updates/fda-update-post-market-assessment-tara-flour [accessed 2025-02-25]
23. Daily harvest issues voluntary recall of french lentil + leek crumbles due to potential health risk. US Food and Drug Administration. 2022. URL: https://www.fda.gov/safety/recalls-market-withdrawals-safety-alerts/daily-harvest-issues-voluntary-recall-french-lentil-leek-crumbles-due-potential-health-risk [accessed 2025-02-25]
24. Investigation of adverse event reports: French lentil & leek crumbles (June 2022). US Food and Drug Administration. 2022. URL: https://www.fda.gov/food/outbreaks-foodborne-illness/investigation-adverse-event-reports-french-lentil-leek-crumbles-june-2022 [accessed 2025-02-25]
25. Fein SB, Lin CTJ, Levy AS. Foodborne illness: perceptions, experience, and preventive behaviors in the United States. J Food Prot 1995 Dec;58(12):1405-1411. [doi: 10.4315/0362-028X-58.12.1405] [Medline: 31159043]
26. Foodborne diseases active surveillance network - population survey. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/foodnet/surveys/population.html2024 [accessed 2025-02-25]
27. CIFOR Guidelines for Foodborne Disease Outbreak Response, 3rd edition: The Council to Improve Foodborne Outbreak Response (CIFOR); 2022. URL: https://cifor.us/downloads/clearinghouse/CIFOR-Guidelines-Complete-third-Ed.-FINAL.pdf [accessed 2025-02-25]
28. Faasse K, Gamble G, Cundy T, Petrie KJ. Impact of television coverage on the number and type of symptoms reported during a health scare: a retrospective pre-post observational study. BMJ Open 2012;2(4):e001607. [doi: 10.1136/bmjopen-2012-001607] [Medline: 22904334]
29. Philen RM, Kilbourne EM, McKinley TW, Parrish RG. Mass sociogenic illness by proxy: parentally reported epidemic in an elementary school. Lancet 1989 Dec 9;2(8676):1372-1376. [doi: 10.1016/s0140-6736(89)91976-4] [Medline: 2574312]

Abbreviations

CAERS: Center for Food Safety and Applied Nutrition Adverse Event Reporting System
CFSAN: Center for Food Safety and Applied Nutrition
CORE: Coordinated Outbreak Response & Evaluation
FDA: Food and Drug Administration
GI: gastrointestinal
LHD: local health department
NDS: novel data stream
PPRP: private participatory reporting platform

Edited by A Mavragani; submitted 26.03.24; peer-reviewed by AE White, C Hedberg, ES Walter, K Pogreba-Brown; revised version received 12.11.24; accepted 13.11.24; published 28.02.25.

Please cite as:
Bazaco MC, Carstens CK, Greenlee T, Blessington T, Pereira E, Seelman S, Ivory S, Jemaneh T, Kirchner M, Crosby A, Viazis S, van Twuyver S, Gwathmey M, Malais T, Ou O, Kenez S, Nolan N, Karasick A, Punzalan C, Schwensohn C, Gieraltowski L, Chen Parker C, Jenkins E, Harris S
Recent Use of Novel Data Streams During Foodborne Illness Cluster Investigations by the United States Food and Drug Administration: Qualitative Review
JMIR Public Health Surveill 2025;11:e58797
URL: https://publichealth.jmir.org/2025/1/e58797
doi:10.2196/58797

Original Paper

Leveraging Administrative Health Databases to Address Health Challenges in Farming Populations: Scoping Review and Bibliometric Analysis (1975-2024)

Pascal Petit[1], PhD; Nicolas Vuillerme[1,2], PhD, Prof Dr

[1]Laboratoire AGEIS, Université Grenoble Alpes, La Tronche Cedex, France

[2]Institut Universitaire de France, Paris, France

Corresponding Author:
Pascal Petit, PhD
Laboratoire AGEIS
Université Grenoble Alpes
Bâtiment Jean Roget, UFR de Médecine
Domaine de La Merci
La Tronche Cedex, 38706
France
Phone: 33 4 76 63 71 04
Email: pascal.petit@univ-grenoble-alpes.fr

Abstract

Background: Although agricultural health has gained importance, to date, much of the existing research relies on traditional epidemiological approaches that often face limitations related to sample size, geographic scope, temporal coverage, and the range of health events examined. To address these challenges, a complementary approach involves leveraging and reusing data beyond its original purpose. Administrative health databases (AHDs) are increasingly reused in population-based research and digital public health, especially for populations such as farmers, who face distinct environmental risks.

Objective: We aimed to explore the reuse of AHDs in addressing health issues within farming populations by summarizing the current landscape of AHD-based research and identifying key areas of interest, research gaps, and unmet needs.

Methods: We conducted a scoping review and bibliometric analysis using PubMed and Web of Science. Building upon previous reviews of AHD-based public health research, we conducted a comprehensive literature search using 72 terms related to the farming population and AHDs. To identify research hot spots, directions, and gaps, we used keyword frequency, co-occurrence, and thematic mapping. We also explored the bibliometric profile of the farming exposome by mapping keyword co-occurrences between environmental factors and health outcomes.

Results: Between 1975 and April 2024, 296 publications across 118 journals, predominantly from high-income countries, were identified. Nearly one-third of these publications were associated with well-established cohorts, such as Agriculture and Cancer and Agricultural Health Study. The most frequently used AHDs included disease registers (158/296, 53.4%), electronic health records (124/296, 41.9%), insurance claims (106/296, 35.8%), population registers (95/296, 32.1%), and hospital discharge databases (41/296, 13.9%). Fifty (16.9%) of 296 studies involved >1 million participants. Although a broad range of exposure proxies were used, most studies (254/296, 85.8%) relied on broad proxies, which failed to capture the specifics of farming tasks. Research on the farming exposome remains underexplored, with a predominant focus on the specific external exposome, particularly pesticide exposure. A limited range of health events have been examined, primarily cancer, mortality, and injuries.

Conclusions: The increasing use of AHDs holds major potential to advance public health research within farming populations. However, substantial research gaps persist, particularly in low-income regions and among underrepresented farming subgroups, such as women, children, and contingent workers. Emerging issues, including exposure to per- and polyfluoroalkyl substances, biological agents, microbiome, microplastics, and climate change, warrant further research. Major gaps also persist in understanding various health conditions, including cardiovascular, reproductive, ocular, sleep-related, age-related, and autoimmune diseases. Addressing these overlooked areas is essential for comprehending the health risks faced by farming communities and guiding public health policies. Within this context, promoting AHD-based research, in conjunction with other digital data sources (eg, mobile health, social health data, and wearables) and artificial intelligence approaches, represents a promising avenue for future exploration.

(JMIR Public Health Surveill 2025;11:e62939) doi:10.2196/62939

KEYWORDS

farming population; digital public health; digital epidemiology; administrative health database; farming exposome; review; bibliometric analysis; data reuse

Introduction

Background

Public health research seeks to identify and understand the factors that influence population health to effectively prevent diseases and promote health and well-being for all [1,2]. A broad range of environmental determinants can impact health across the life span. One of the core areas of public health research, known as the exposome, investigates how cumulative environmental influences contribute to disease etiology and pathogenesis [3-18]. The exposome, which complements genomic research, refers to the comprehensive examination of all environmental exposures experienced throughout an individual’s lifetime, including physical, chemical, biological, psychosocial, and behavioral factors, from conception to death [3-18]. The exposome classically includes 3 overlapping domains: the general external exposome (eg, climate and built environment); the specific external exposome (eg, chemical exposure, lifestyle, and occupations); and the internal exposome (eg, aging, oxidative stress, metabolism, and gut microbiome) [8,14,16,17,19,20]. Understanding the exposome is crucial for enabling both population-wide and precision prevention [3,21-23]. However, fully describing the exposome is challenging due to the vast diversity and the temporal and spatial variability of environmental factors [3]. Public health research in this area requires data on both risk factors and adverse health outcomes to progress effectively [3,14,24,25].

The volume of data collected has grown exponentially as the world becomes increasingly reliant on technology and digitization [26,27]. Data are omnipresent in our everyday lives, leading science toward data-driven research [27,28], in particular in the health field. The digital transformation in health care has enabled unprecedented data availability, collection, storage, and analysis capabilities, leading to a paradigm shift in health care systems, with entire care pathways becoming digitized [29,30]. Health-related data now represent approximately 6% of all digital data globally, a figure that continues to rise [31]. This explosion of data has transformed research, providing new opportunities, especially in public health, to enhance disease understanding and evaluate intervention effectiveness [27,28,32-35]. The integration of digital technologies and digital data in public health has led to the emergence of “digital public health,” an evolving field focused on using digital data to achieve public health goals [33,36-41]. Public health research is moving from isolated data systems to more integrated, accessible, and reusable data resources [42]. Reusing data allows researchers to explore various health determinants, including environmental, occupational, behavioral, and organizational factors, fostering a holistic approach to disease prevention and health promotion strategies [14].

Within the digital public health framework, 2 main types of data are being used, namely primary and secondary data. Primary data are tailor-made, designed for a specific purpose, and often used once or repeatedly for the same goal [43-45]. Primary data are the cornerstone of traditional public health policy and decision-making. These data are derived from several types of studies [46-50], in particular observational cohorts (eg, the Framingham Cohort study [51,52]) [46-50,53,54], case-control studies [46-50], cross-sectional surveys (eg, the China Health and Retirement Longitudinal Study [55,56]) [46-50,53], and experimental studies [46-50]. Primary data have many advantages [46-50]. They are rich, of high quality, and are designed to answer specific research questions for public health and epidemiological purposes. Primary data are usually available at the individual level and are derived from studies that control for certain biases. By contrast, they are cumbersome, time-consuming, and costly to set up and maintain [53,54,57]. The representativeness of primary data is also limited in size, geographic scope, and temporal coverage and can erode with time [46-50,53]. Primary data are not free from bias, such as selection, healthy worker, recall, or prevarication biases [53,58].

Unlike traditional public health, digital public health does not rely solely on primary data but takes advantage of the myriad of existing digital data that have not been generated originally for research purposes (ie, secondary data) to overcome some limitations intrinsic to primary data and complement them [28,43,44,53,59-68]. Indeed, some data can have an additional impact when used beyond the context for which they were originally created [68,69]. Secondary data are collected for purposes other than public health or epidemiology and include contextual data (eg, air quality and climate data) [14,24,26,29,70-74], person-generated data (eg, social media, crowdsourcing, and mobile health) [2,24,26,31,43,61,62,73,75-86], synthetic data (eg, digital twin) [87-91], and administrative health databases (AHDs) [26,64,68,81,92-102].

AHD is a broad term encompassing a wide range of routinely collected data on individuals’ health and sociodemographic information collected for registration, billing, record keeping, and other administrative purposes [26,64,81,93,95,98,100-102]. For this review, based on previous works [93,95,96,103-105], AHDs included population registers, claims databases, disease registers, electronic health or medical records, and hospital discharge databases that were collected at a local, regional, national, or international level (Table 1) [26,61,62,93,95,96,100,103-110].

Table 1. Definition and characteristics of administrative health databases included in this review.

	Population register	Claims database	Disease register	Electronic health or medical record	Hospital discharge database
Definition	Digital sociodemographic information on the residents of a country	Routinely collected digital information on individual data regarding reimbursement, records of health services, medical procedures, prescriptions, and medical diagnoses	A continuous and exhaustive digital collection of individual data regarding 1 or more health events in a geographically defined population	Systematized digital record of a patient's medical information collected in real time	Digital records of service use with information about patients, their care, and their stay in the hospital
Source	Local or national authorities	Insurance programs or schemes and health care providers	Health care institutions (eg, hospitals)	Hospitals, physicians, health care centers, and institutions	Hospitals
Population	All individuals residing in a country	All individuals covered by an insurance program or scheme	All individuals diagnosed with a specific health event in a population on a geographically defined scale	All patients using the health care system	All patients from a hospital
Purpose or finality	For the administrative purposes of government: to provide reliable information	To store financial and administrative information for medical insurers' and providers' use	For clinical and research purposes: to collect information about people diagnosed with a specific health event	For clinical and billing purposes: to document patients' clinical condition	For billing or accounting purposes
Health event	None	Health events covered by insurance or a health care provider	Specific health events (eg, cancer)	Health events requiring care that are reported in medical records	Health events from hospital admission

AHDs offer many advantages for research. Such data are collected as part of routine administrative processes, reducing additional costs for researchers. Therefore, AHDs offer relatively inexpensive access to a large number of individuals who can be tracked with time for several years, guaranteeing the representativeness of the populations studied [26,54,78,93,95,104,105,111-114]. Data recorded within AHDs are structured, coded in a standardized way, and less affected by participation and recall biases [54,58,95,106,113,115]. AHDs enable the study of rare events and populations underrepresented in studies using only primary data [95,111-113]. AHDs have limitations inherent to their nature, such as the absence of some confounding factors, the limited granularity of certain information, the data complexity, and confidentiality issues [73,78,93,95,115-125].

Rationale

AHDs are increasingly used in population-based health research due to their complementarity with traditional sources of public health and epidemiological data (ie, primary data) [42,59,64,87,93,95,96,126-128]. The reuse of AHDs, referring to their application beyond their original or intended purpose, holds major potential to advance public health and epidemiological research, offering insights that can guide public health decision-making [42,59,64,87,93,95,96,105,126-131]. Although several reviews have previously explored the general use of AHDs in research [42,95,107,129,132-135], others have focused on their application within specific countries [96,136], examined individual AHDs [108,137], or investigated their role in studying specific diseases and adverse health outcomes [44,93,104,138-142]. However, to the best of our knowledge, no study has synthesized how AHDs are reused for epidemiological and public health research within a specific population group.

To address this gap, we conducted a comprehensive scoping review and bibliometric analysis aimed at identifying how AHDs are used to address health issues in a specific population. We selected farming populations as an illustrative example because they present unique health and disease patterns [143-147]. Globally, approximately 27% of the workforce is engaged in occupational farming, and this group is exposed to numerous risk factors (ie, exposomes), including pesticides, biological agents, and limited access to health care [148]. These exposures put them at heightened risk for a wide range of adverse health outcomes [143,145,147,149]. Although agricultural safety and health have become a major public health issue in recent decades, most research on the health of farming populations has relied on traditional epidemiological and community-based studies, which often face limitations in terms of sample size, geographic scope, temporal coverage, and the range of health events examined [145,150,151].

In this context, AHDs offer valuable opportunities to enhance public health and epidemiological research in farming populations by providing broader insights, identifying at-risk subgroups, and informing health services and policy development [152]. The primary objectives of this scoping review were two-fold: (1) to summarize the current state of AHD-based research in farming populations by examining which types of AHDs are used and why, whether AHDs are integrated with other data sources, which farming populations have been studied, and what exposures and health outcomes have been explored and (2) to identify key areas of interest and potential research gaps and unmet needs in this field.

Methods

Overview

This scoping review was conducted and reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) and evidence maps guidelines (Table S1 in Multimedia Appendix 1) [153] following a single screening approach. The protocol of this study was not registered. A 7-step procedure was used: research question formulation, identifying relevant publications, title review, abstract review, full-text review, data extraction, and data analysis.

To formulate our research question, we followed the Joanna Briggs Institute guidelines, using the population, concept, and context criteria framework [154]. Our population included all individuals engaged in farming and all individuals exposed to farming-related exposures. The concepts included all possible public health and epidemiological research works that involved the study of a health outcome of interest. The context was the use, in any setting, of at least one of the AHDs, as defined in Table 1.

Search Strategy and Selection Criteria

To develop and validate the search strategy, previous reviews that examined the reuse of AHDs for population-based research were identified and refined [93,96,103,104]. Our initial search revealed that electronic health records (EHRs) are often interchangeably referred to as electronic medical records (EMRs). A distinction between EHR and EMR is sometimes made, with EMR describing patients' care from only 1 practice (eg, specific encounters in hospitals), which is contrary to EHR [105]. In that case, EMR serves as a data source for EHR. This distinction was not considered in this paper. In addition, to ensure comprehensiveness, the search terms were broadened by searching for their synonyms. For example, search terms such as "electronic health record," "digital health record," "electronic medical record," "EHR," or "EMR" were used as synonyms for electronic health or medical records. A total of 72 terms pertaining to 2 categories (farming and AHDs) were used (Textbox 1). The search terms were reflective of our research topic and question.

Textbox 1. Search terms.

Farming

- husbandry* OR agriculture* OR farming OR farm* OR agricultural* OR farmworker*

Administrative health databases (combined using AND)

- "health record" OR "health records" OR "digital record" OR "digital records" OR "health administrative register" OR "health administrative registry" OR "health register" OR "health registry" OR "medical register" OR "medical registry" OR "electronic health record" OR "electronic health records" OR "EHR" OR "EMR" OR "electronic medical record" OR "electronic medical records" OR "digital medical record" OR "digital medical records" OR "digital health record" OR "digital health records" OR "health administrative data" OR "health administrative database" OR " health administrative dataset" OR " health administrative datasets" OR "health administrative databases" OR "administrative health data" OR "administrative health database" OR "administrative health dataset" OR "administrative health datasets" OR "administrative health databases" OR "insurance data" OR "insurance database" OR "insurance databases" OR "insurance dataset" OR "insurance claim" OR "insurance claims" OR "cancer registry" OR "cancer register" OR "health insurance" OR "health surveillance program" OR "health surveillance programs" OR "Mutualite Sociale Agricole" OR "MSA" OR "health insurance system" OR "record-linkage" OR "population register" OR "population registry" OR "insurance scheme" OR "social security scheme" OR "hospital discharge" OR "administrative claim" OR "administrative claims" OR "medical claims" OR "medical claim" OR "electronic claim" OR "electronic claims" OR "mortality register" OR "mortality registry" OR "mortality record" OR "mortality records" OR "disease register" OR "disease registry" OR "illness register" OR "illness registry" OR "disorder register" OR "disorder registry"

To develop the eligibility criteria, an initial search of the literature was conducted on PubMed, with a review of the first 100 articles that used AHDs for public health and epidemiological research. In our pilot run, disease and morbidity registers were initially not considered as AHD because they were created for clinical and research purposes [47-50,53,76,155,156]. However, because disease registers contain some information derived from medical records, we decided to consider them as AHD for this review. The eligibility criteria are presented in Textbox 2. The search was restricted to original peer-reviewed records (all types were included) written in English or French but not constrained by the year of publication [93,106,157]. Publications that examined partly farming populations, with, for instance, studies reporting health risks for various sectors of activity, were included.

Textbox 2. Eligibility criteria for selection of publications.

Inclusion criteria for articles

- Data had to originate at least partly from the administrative health database (AHD)
- The study had to pertain at least partly to the farming population
- The study had to relate to public health or epidemiological research
- Original peer-reviewed publications
- Publications in English or French

Exclusion criteria for articles

- Publications not describing the use of an AHD
- Animal or in vitro studies
- Publications not in English or French

The final literature search was conducted on both PubMed and Web of Science Core Collection databases. Regarding the Web of Science Core Collection, a topic search was performed. To reduce the bias induced by daily database changes, all data collection (literature retrieval and data download) was conducted and completed on the same day, that is, April 15, 2024. Titles, abstracts, and full-text publications were screened based on pre-established inclusion and exclusion criteria. The inclusion criteria for each phase of the literature search are provided in Table S1 in Multimedia Appendix 1. When abstracts did not contain enough information about correspondence to inclusion or exclusion criteria, the article was considered for full-text review. Reference lists of included publications were not searched, although they might have also yielded new relevant studies.

Data Collection and Processing

A total of 29 metadata were extracted from each publication included in the scoping review (Table 2).

The data underwent rigorous manual validation, cleaning, and harmonization following a structured 5-step process. First, duplicate items (eg, keywords and institutions) were removed. Second, leading and trailing white spaces were eliminated. Third, items were standardized by converting text to lower case, with only the first letter capitalized. In the fourth step, items were harmonized to either singular or plural forms consistently. Finally, synonyms or terms with similar meanings (eg, “illness” and “disease”) were unified under a single term. For instance, “Pesticide,” “Pesticide exposure,” and “Pesticide use” were standardized to “Pesticide,” while “Pulmonary disease copd,” “Copd,” and “Chronic obstructive pulmonary disease” were unified as “COPD.” For cancer-related keywords, the International Classification of Diseases, eleventh revision, was used to consolidate varied terms (eg, “lung cancer,” “lung cancer risk,” “lung and bronchus cancer,” “lung tumor,” “lung tumour,” “lung neoplasm,” and “basal cell carcinoma of the lung”) into standard categories (eg, lung cancer). Quality appraisals were not performed because they were beyond the aim of this review [106,157].

Table 2. List of metadata of interest to collect from the literature search.

Metadata	Fictional example
Publication year	2024
Publication type	Article
Study name	Project X
Goal of the study	To study the association between farming and health outcome
Study type	Ecological study
Is the study nationwide?	Yes
Digital data used	Insurance claims
Goal of the digital data used	To identify farmers
Is active data used?	Yes
Active data used	Clinical examination
Farming exposure considered	Farming activity and pesticide compounds
Farming activities studied, n	10
Pesticide compounds studied, n	29
Population	Adults
Sex	Female
Participants included, n	100 to 1000
Country	France
Oldest data used (year)	1991
Most recent data used (year)	2020
Data follow-up period (years)	4
Years between the most recent data used and publication year, n	7
Disease or health events	Parkinson disease
Authors' names	Gauthier J
Authors' keywords	Pesticide
Authors' country	France
Authors' institution	Université Grenoble Alpes
Journal	Environmental Health Perspectives
Funding body	MIAI@Grenoble Alpes[a]
Citations, n	14

[a] MIAI@Grenoble Alpes: Multidisciplinary Institute in Artificial Intelligence at the Université Grenoble Alpes.

Data Analysis

To analyze the research directions (ie, hot spots and gaps) on the use of AHDs for public health and epidemiological research in farming populations, a bibliometric approach was conducted [158-160]. This analysis examined the number of publications, countries of publications, most active journals, institutions, authors, funding bodies, subject areas, citations of publications, and keywords of publications. Seven bibliometric metrics were computed, including the h-, g-, m-, and Y-indices; dominance factor; annual growth rate (AGR); and fractionalized frequency (Table S3 in Multimedia Appendix 1). The h-index attempts to measure both the productivity and citation impact of the published body of work of an entity (eg, author, institution, and journal) [161,162]. It refers to the total number of publications by a particular entity with at least the same number of citations. The m-index is calculated by dividing the h-index by the number of years of an entity's productive life (eg, researcher) [161]. The g-index of an entity corresponds to the largest number g such that the top g publications have at least $\geq g^2$ citations together [162]. The Y-index refers to the sum of both the total number of first-authored publications and the total number of corresponding author publications [163]. The dominance factor refers, for a particular researcher, to the proportion of multiauthored publications as a specific author's rank to the total number of multiauthored publications [164]. The fractionalized frequency intends to reflect an author's contribution. The AGR refers to the variable's change in percentage as a year-over-year statistic [165]. The most up-to-date journals' impact factors and ranks were retrieved manually using the Journal Citation Report in April 2024.

Spearman correlations were calculated to examine the association between the number of publications and gross domestic product (GDP); population size [166]; and the total labor force, the number of researchers in research and development (per million people), fertilizer consumption (in both % of fertilizer production and kilograms per hectare of arable land), agricultural land (km^2), agricultural land (% of land area), land under cereal production (hectares), permanent cropland (% of land area), cereal production (metric tons), crop production index, food production index, livestock production index, cereal yield (kilogram per hectare), female individual employment in agriculture (% of female employment), male individual employment in agriculture (% of male employment), employment in agriculture (% of total employment); and agriculture, forestry, and fishing, value added (% of GDP). These country characteristics were obtained from the World Bank. The most recent country characteristic (eg, GDP) was used when available.

Research directions, including hot spots and gaps, were investigated with keyword frequency, co-occurrence (counting of paired keywords), and thematic mapping analyses. Thematic mapping and keyword co-occurrence network are 2 complementary but distinct approaches that serve different purposes and offer different insights. In summary, thematic mapping focuses more on the strategic positioning of research themes within a field, while keyword co-occurrence networks emphasize the relationships and connections between specific keywords in the literature [158,167]. Both methods complement each other and are usually used to provide a more comprehensive understanding of research landscapes. The co-occurrence of 2 keywords was defined by the frequency with which they appear together in publications and was quantified using association strength (AS) or equivalent index, calculated as , where c_{ij} is the number of publications in which keywords i and j co-occur, while c_i and c_j are the number of publications in which each keyword appears, respectively [158,167]. AS measures how close 2 keywords are to each other. An AS value of 1 indicates keywords always appear together, while 0 indicates they never co-occur. These keyword co-occurrences can be visualized using a co-occurrence network graph, where a vertex or node represents a keyword, the size of the node represents the keyword frequency, and the edge represents the association between 2 keywords [158,167]. On the basis of the keyword co-occurrence network graph, a community detection procedure can be used to identify groups of words highly associated with each other [158,167]. In other words, equivalent keywords based on AS can be grouped together to identify research themes [158,167]. A strategic diagram or thematic map is based on Callon centrality (x-axis) and Callon density (y-axis) [158,167]. Callon centrality measures the degree of interaction of a theme with other themes. It is defined as , where k is a keyword belonging to a theme and h is a keyword belonging to another theme [158,167]. Callon centrality can be interpreted as an indicator of the importance of a particular topic within the broader research landscape. Callon density measures the internal strength of a theme. It is defined as , where i and j are keywords belonging to the same theme and w is the total number of keywords in a theme [158,167]. Callon density serves as a metric for assessing the progression and maturation of that topic [158,167]. A strategic diagram is divided into 4 quadrants according to Callon centrality and density values, which correspond to 4 types of topics. Hot spots or hot topics are defined by both high density and high centrality values (upper-right quadrant), while basic topics are defined by high centrality but low density values (lower-right quadrant). Peripheral topics are defined by both low centrality and low density values (lower-left quadrant), while niche topics are defined by low centrality and high density values (upper-left quadrant) [158,167].

To focus on agricultural or farming exposome research, a bibliometric profile of the "farming exposome" was constructed, which restricts the exposome concept to environmental exposures specific to farming populations [152,168]. This bibliometric farming exposome picture examined co-occurrences between keywords related to potential risk factors and specific health events (eg, cancers and reproductive disorders).

The bibliometric analysis was conducted and reported according to the preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO; Table S4 in Multimedia Appendix 1) [169]. All analyses were performed using R software (version 4.3.2; R Foundation for Statistical Computing) for Windows 10 (Microsoft Corporation). The bibliometric analysis was performed using the *bibliometrix* R package (version 4.1.4) [170].

Results

Overview

After excluding 4485 irrelevant records, 296 publications were analyzed (Figure 1). The majority were articles (293/296, 98.9%), with a small number of reviews (2/296, 0.7%) and editorial materials (1/296, 0.3%; Table 3). Only one-third of the publications (107/296, 36.1%) were open access (Table 3 and Figure S1 in Multimedia Appendix 1).

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flowchart depicting the literature search and the evaluation process for finding relevant records. The search, conducted on April 15, 2024, in PubMed and Web of Science, had no date restrictions. AHD: administrative health database.

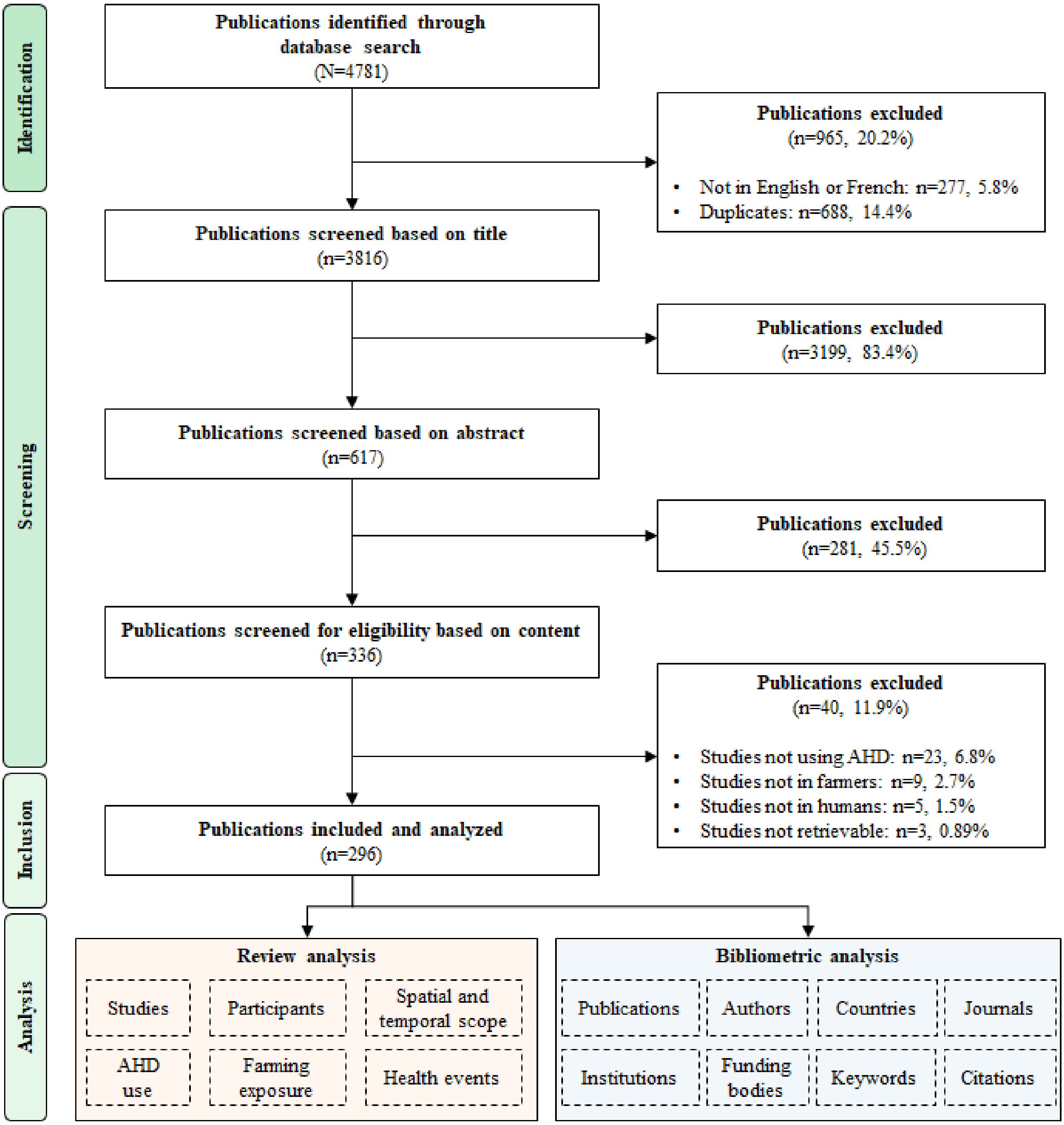


Table 3. Main characteristics of the included publications (N=296).

Description	Results
Timespan	1975 to 2024
Publication type, n (%)	
Article	293 (99)
Review	2 (0.7)
Editorial material	1 (0.3)
Open-access publications, n (%)	107 (36.1)
Document age (y), mean (SD)	14.2 (11.8)
Annual growth rate (%)	5.2
Publication citations	
Total, n[a]	9379
Average citations per publication	31.7
Average citations per year per publications	2.02
References, n	8814
Journals	
Total, n	118
Average number of publications per journal	1.86
Average number of citations per journal	79.5
Authors	
Total, n	1225
Single-author publications, n	4
Author appearances, n	1882
Average number of coauthors per publication	6.36
Average number of publications per author	0.24
International coauthorships (%)	24.3
Author's keywords, n	576
Author's country	
Total, n	34
Average number of publications per country	2.86
Average number of citations per country	436.0
Author's institution	
Total, n	338
Average number of publications per institution	3.11
Average number of citations per institution	101.3
Author's funding body	
Total, n	181
Average number of publications per funding body	2.48
Average number of citations per funding body	77.7

[a] Total, n indicates that the respective parameter has been cited n number of times, as in 296 publications have been cited 9379 times.

The average publication age was 14.2 (SD 11.8) years, ranging from the oldest in 1975 [171] to the most recent in April 2024 [152]. From 1975 onward, there has been a steady increase in publications using AHDs to address health issues in farming populations, with an AGR of 5.2%. Notably, almost one-third of these articles (91/296, 30.7%) were published in the last 5 years, highlighting the rising interest in AHD-based public health research in this population (Figure S2 in Multimedia

Appendix 1). Collectively, the publications received 9379 citations, averaging 31.7 citations per publication. Figure S3 in Multimedia Appendix 1 presents the historical direct citation network. The body of work involved 1225 authors from 338 institutions, with 1882 author appearances and an average of 6 authors per paper (Table 3). Four (1.4%) out of the 296 publications were single-author publications. On average, each paper cited 30 references.

Studies were led by authors from 34 countries, predominantly high-income nations, with 24.3% (72/296) of studies involving multicountry collaborations (Figure S4 in Multimedia Appendix 1). US-based authors contributed the most publications (91/296, 30.7%), followed by authors based in France (71/296, 24%) and Finland (35/296, 11.8%). US authors also had the most citations (3495/9379, 37.2%), with France and Finland ranking second and third, respectively.

Of 296 publications, the 25 (8.4%) most cited ones, appearing in 17 different journals, received between 83 (83/9379, 0.9%) and 485 (485/9379, 5.2%) citations (Table S5 in Multimedia Appendix 1) [150,172-196]. Of these 25 publications, 10 (40%) were published before 2000, another 10 (40%) between 2000 and 2010, and 5 (20%) after 2010. Most of these studies focused on cancer risk (16/25, 64%), while others investigated neurodegenerative disorders (5/25, 20%); respiratory conditions (2/25, 8%); and multiple health outcomes, such as sleep disorders, mental health disorders, and musculoskeletal disorders (2/25, 8%).

Table S6 in Multimedia Appendix 1 provides details on the most productive countries, prolific authors, active journals, institutions, and funding bodies.

Study Characteristics

Table 4 provides an overview of the included publications. Longitudinal study designs were the most common, including retrospective cohorts (129/296, 43.6%) and prospective cohorts (56/296, 18.9%). Case-control studies (62/296, 20.9%), cross-sectional studies (39/296, 13.2%), and ecological studies (17/296, 5.7%) were less common (Multimedia Appendix 2). A few studies (10/296, 3.4%) used multiple study designs [188,194,197-204].

The median follow-up period was 9.5 (IQR 5-17) years. On average, there was a 7-year gap (90% CI 3-14) between the most recent data used and the year of publication, with considerable variation depending on the publication year (Figure 2). The oldest data were from 1801 [205], and the most recent data were from 2022 [206]. Notably, one-third of the data used (98/296, 33.1%) were from before 2000, while nearly three-quarters (214/296, 72.3%) were from before 2015 (Figure S2 in Multimedia Appendix 1). Of 296 studies, only 10 (3.4%) used data from the last 5 years (from 2020), while 80 (27%) used data from the last 10 years (from 2015).

Studies were conducted in all continents, but most participants were from Europe (249/296, 84.1%), followed by North America (85/296, 28.7%), Asia (24/296, 8.1%), Oceania (17/296, 5.7%), Africa (4/296, 1.4%), and Central and South America (4/296, 1.4%). France (70/296, 23.6%) and the United States (67/296, 22.6%) were the most represented countries, followed by Finland (36/296, 12.2%), Sweden (32/296, 10.8%), Denmark (28/296, 9.5%), and Norway (26/296, 8.8%; Figure 3 and Figure S5 in Multimedia Appendix 1). Most studies had a regional or local scope (177/296, 59.8%), in particular, traditional epidemiological studies, such as Agriculture and Cancer (AGRICAN) [207] and Agricultural Health Study (AHS) [172], which used AHDs to either identify potential individuals for inclusion or enrich their cohorts.

Table 4. Characteristics of the included studies (1975 to 2024; N=296).

Characteristic	Values
Research goal, n (%)	
Study the association between farming and a health event	156 (52.7)
Study the association between individual characteristics and a health event	131 (44.3)
Other research goals	9 (3)
Study design, n (%)	
Retrospective cohort	129 (43.6)
Case-control study	62 (20.9)
Prospective cohort	56 (18.9)
Cross-sectional study	39 (13.2)
Ecological study	17 (5.7)
Multiple designs	10 (3.4)
Review	2 (0.7)
Perspective	1 (0.3)
Geographic scope, n (%)	
Nationwide	117 (39.5)
Regional or local	176 (59.5)
Temporal scope (y)	
Follow-up period, median (IQR)	9.50 (5-17)
Follow-up period, mean (SD)	12.8 (14.0)
Gap between the latest data used and publication year, median (IQR)	7.21 (5-9)
Gap between the latest data used and publication year, mean (SD)	7.21 (4.67)
Population, n (%)	
Adult	265 (89.5)
Adult and child	19 (6.4)
Child	8 (2.7)
Not reported	1 (0.3)
Sex, n (%)	
Female	130 (43.9)
Male	169 (57.1)
Female and male	188 (63.5)
Not specified	108 (36.5)
Participants, n (%)	
>1,000,000	50 (16.9)
100,001 to 1,000,000	53 (17.9)
10,001 to 100,000	65 (22)
1001 to 10,000	67 (22.6)
101 to 1000	47 (15.9)
10 to 100	8 (2.7)
Not reported	3 (1)
AHD[a] type, n (%)	
Disease register	158 (53.4)
Electronic health or medical record	124 (41.9)

Characteristic	Values
Insurance claim	106 (35.8)
Population register	95 (32.1)
Hospital discharge databases	41 (13.9)
AHD use, n (%)	
Obtain information on sociodemographics	272 (91.9)
Obtain information on a health event	269 (90.9)
Identify a farmer	147 (49.7)
Identify an individual	140 (47.3)
Obtain information on occupations	117 (39.5)
Exposure assessment	57 (19.3)
Obtain information on a farming activity	43 (14.5)
Other uses	14 (4.7)

[a]AHD: administrative health database.

Figure 2. Number of years between the most recent data used and publication for all included articles (1975-2024). Points refer to the average number of years or gap between the most recent data used and publication (x-axis) for each publication year (y-axis). Error bars refer to the 90% CI of the number of years between the most recent data used and publication.

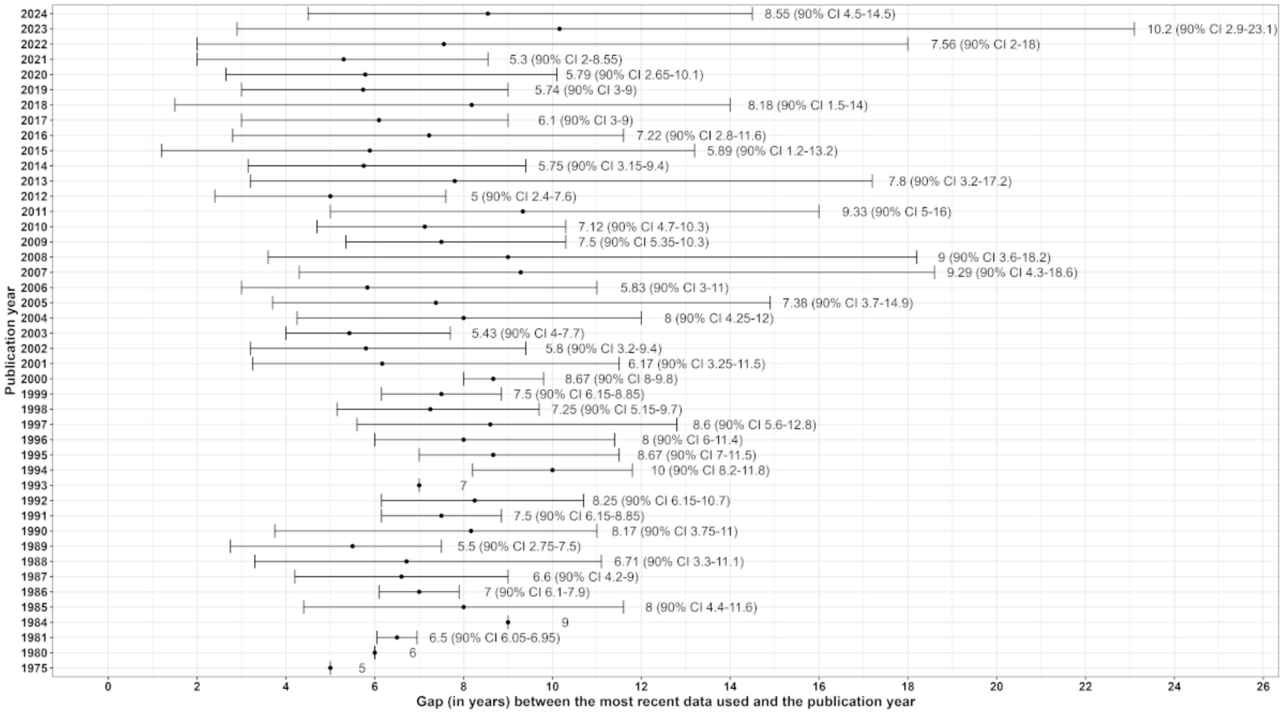
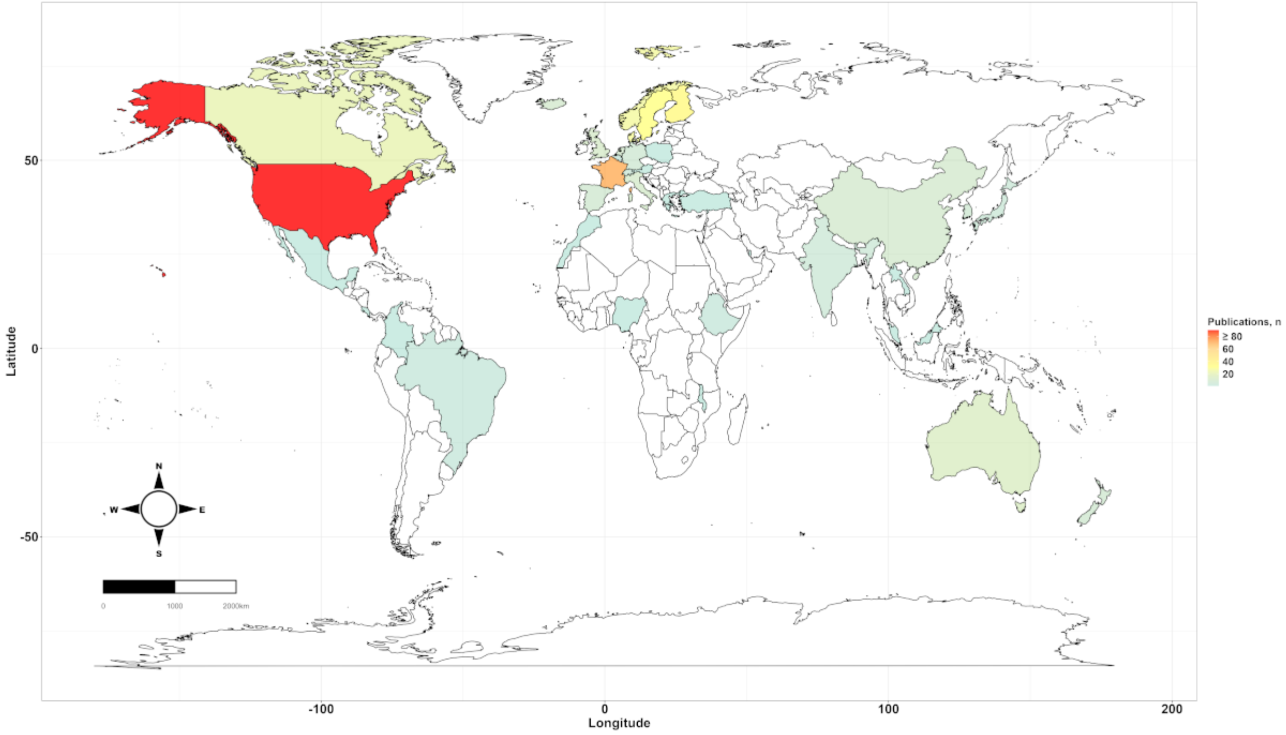


Figure 3. World map of the number of publications per country of the farming population studied between 1975 and 2024.



Most studies included 1001 to 10,000 participants (67/296, 22.6%), followed by studies with 10,001 to 100,000 participants (65/296, 22%) and 100,001 to 1 million participants (53/296, 17.9%; Table 4). Larger studies (>1 million participants) accounted for 16.9% (50/296) of the included publications. Smaller studies, with 100 to 1000 participants, were less common (47/296, 15.9%), and very few (8/296, 2.7%) had <100 participants. Most studies included adult participants (284/296, 95.9%). Of 296 studies, 169 (57.1%) examined male individuals, 130 (43.9%) examined female individuals, and 188 (63.5%) examined both sexes, but 108 (36.5%) did not specify the participants' sex.

More than half of the studies (156/296, 52.7%) aimed to explore the relationship between farming activities (eg, dairy farming) and health events, while 131 studies (44.3%) focused on individual characteristics, such as occupation, age, sex, and socioeconomic status (Figure S6 in Multimedia Appendix 1). Among those studies examining individual characteristics, farming was often considered broadly and compared to other occupations (95/131, 72.5%). Conversely, in studies investigating health outcomes specifically related to farming activities, agriculture was treated as a broad category in only 27.6% (43/156) of the cases. Most studies (277/296, 93.6%) used the general population or other nonfarming groups as the reference category without differentiating farmers by job role (eg, farm managers vs farm workers). Descriptive statistics and multivariable regression were the most commonly used methods. Notably, only 2 studies (2/296, 0.7%) incorporated artificial intelligence (AI) in their analysis [208,209].

Few studies investigated health outcomes in farmers' family members or nonfarmers exposed to farming. Of 296 studies, only 3 (1%) focused on health events in farmers' partners [177,210,211], 5 (1.7%) on farmers' children [179,212-215], and 6 (2%) on nonfarmers exposed to farming-related risks [209,216-221]. There were 11 (3.7%) studies that explored health risks in migrant workers.

Some publications reported findings from the same cohorts (Figure S7 in Multimedia Appendix 1). The 10 most prolific cohorts included France-based AGRICAN (18/296, 6.1%) [207], the US-based AHS (17/296, 5.7%) [172], Nordic Occupational Cancer Study (NOCCA; 12/296, 4.1%) from Nordic countries (Finland, Denmark, Norway, Sweden, and Iceland) [189], France-based Tracking and Monitoring Occupational Risks in Agriculture (TRACTOR; 7/296, 2.4%) [222], and Cancer in the Norwegian Agricultural Population (7/296, 2.4%) [182] cohorts. Other notable cohorts included the US-based United Farm Workers of America (6/296, 2.0%) [223], France-based BALISTIC (5/296, 1.7%) [224], the international (29 countries) consortium agricultural cohort (AGRICOH; 4/296, 1.4%) [150,225], AIRBAg (4/296, 1.4%) from France [226], and the US-based National Agricultural Workers Survey (3/296, 1%) [227]. Among these top 10 cohorts, only NOCCA, United Farm Workers of America, and TRACTOR exclusively used AHDs.

AHD Use

There was high heterogeneity in the coding systems used and the granularity of the information available regarding health events (outcomes), population, and exposure determinants, depending on the AHD and study considered. Regardless of the publication reviewed, AHDs and other datasets were never reported as adhering to the findable, accessible, interoperable, and reusable (FAIR) data principles [228-230]. In addition, none of them could be considered as FAIR data because, with a few exceptions [222], most AHDs were not precisely described, and data availability statements were rare. Furthermore, mainly due to privacy concerns, AHDs were not available for open and free access.

The most commonly used AHDs were disease registers, used in more than half of the studies (158/296, 53.4%), followed by electronic health or medical records (124/296, 41.9%), insurance claims (106/296, 35.8%), population registers (95/296, 32.1%), and hospital discharge databases (41/296, 13.9%; Table 4). Among disease registers, cancer (120/158, 75.9) and mortality registers (75/158, 47.5%) were the most frequently used (Figure S8 in Multimedia Appendix 1 and Multimedia Appendix 2). Nearly one-third of the studies (91/296, 30.7%) relied on a single AHD, with disease registers being the most common (38/91, 42%), followed by insurance claims (29/91, 32%), electronic health or medical records (18/91, 20%), population registers (5/91, 5%), and hospital discharge databases (1/91, 1%). Other types of digital data were used less frequently, including pesticide registration records (13/296, 4.4%), job-exposure matrices (JEMs; 12/296, 4.1%), crop-exposure matrices (11/296, 3.7%), pesticide use records (8/296, 2.7%), climate data (7/296, 2.4%), and air quality data (2/296, 0.7%). While contextual data were sometimes used (9/296, 3.0%), person-generated data, smart agriculture data, and omics were never used.

The AHDs and other digital data were primarily used to obtain sociodemographic information (272/296, 91.9%) and health event data (269/296, 90.9%). They were also used to identify farmers (147/296, 49.7%) or individuals (140/296, 47.3%), gather occupational information (117/296, 39.5%), assess exposure (57/296, 19.3%), obtain data on farming activities (43/296, 14.5%), or track climate conditions (7/296, 2.4%).

Nearly two-thirds of the studies (181/296, 61.1%) relied exclusively on digital data (AHDs or other), while more than one-third (112/296, 37.8%) incorporated self-reported information/active data (requiring active participant involvement) as part of epidemiological cohorts. A total of 111 (37.5%) out of 296 studies used participant-completed questionnaires (paper or electronic) to gather sociodemographic data and confounding factors (98/296, 33.1%), assess exposure (96/296, 32.4%), or collect health information (83/296, 28%). Some information was obtained through interviews (44/296, 14.9%) or clinical examinations (32/296, 10.8%). Biological monitoring (24/296, 8.1%) and airborne monitoring (2/296, 0.7%) were sometimes used, whereas no study reported dermal monitoring (Figure S9 in Multimedia Appendix 1).

Among all the AHDs used, the Mutualité Sociale Agricole (MSA) is a singularity. To the best of our knowledge, it is the only AHD specifically dedicated to the entire farming population of a country. Indeed, MSA is the French national insurance scheme that covers the entire farming workforce (5% of the overall French population) [115,128]. MSA was used in 60 studies (60/296, 20.3%). These studies were often part of cohorts with multiple publications, such as AGRICAN (18/60, 30%), TRACTOR (7/60, 12%), BALISTIC (5/60, 8%), AIRBAg (4/60, 7%), Aging Multidisciplinary Investigation (2/60, 3%) [151], BM3R (2/60, 3%) [231], FERMA (risk factors of the rural environment and allergic and respiratory disease; 1/60, 2%) [232], and Phytoner (1/60, 2%) [233]. Of these, TRACTOR was the only cohort using exclusively MSA data [222].

Farming Exposure

A variety of exposure proxies were used to assess farming-related exposure. The most common proxy was a job title, which generally referred to whether the individual was a farmer (184/296, 62.2%). Other proxies included specific farming activities, such as dairy or crop farming (111/296, 37.5%), general pesticide exposure (yes or no; 62/296, 20.9%), and exposure to specific pesticide compounds (eg, glyphosate or paraquat) or pesticide classes (eg, insecticides; 51/296, 17.2%; Figure S10 in Multimedia Appendix 1 and Multimedia Appendix 2). The number of farming activities studied ranged from just 1 [226] to 78 [222], with an average of 8 farming activities per study. Similarly, the number of pesticide compounds assessed ranged from 1 [234] to 943 [235], with an average of 42 pesticides per study. Only 1 study investigated the mixture effect of exposure to multiple pesticide combinations on human health [236]. Investigations into other chemical exposures were also limited, with only 2 papers each addressing silica exposure [237,238] and air pollution [194,217] (2/296, 0.7%). Notably, no studies examined exposure to per- and polyfluoroalkyl substances or micro- and nanoplastics. Research on the broader farming exposome was rare (5/296, 1.7%) and typically used farming activities as proxies [152].

Of 296 studies, few explored exposure to physical agents, with 5 studies (1.7%) focusing on radiation [187,218,239-241]. No studies investigated the effects of climate change on farmers' health. Exposure to biological agents was rarely studied as well, with just 3 (1%) out of 296 papers addressing mycotoxins [241-243]. Finally, only 3 studies (1%) examined psychological factors related to farming exposure [244-246].

Health Events

The most frequently studied health events were cancer (142/296, 48%), followed by mortality (44/296, 14.9%), injuries (38/296, 12.8%), workplace accidents (32/296, 10.8%), respiratory disorders (30/296, 10.1%), neurodegenerative diseases (28/296, 9.5%), and mental health issues (26/296, 8.8%; Figure S11 in Multimedia Appendix 1 and Multimedia Appendix 2). Less frequently studied conditions included cardiovascular diseases (16/296, 5.4%), autoimmune disorders (11/296, 3.7%), musculoskeletal disorders (11/296, 3.7%), reproductive disorders (3/296, 1.0%), sleep disorders (1/296, 0.3%), and frailty (1/296, 0.3%). Notably, no studies explored the farming microbiome.

Among cancers, lung cancer was the most commonly investigated cancer (43/142, 30.3%), followed by prostate cancer (38/142, 26.8%), leukemia (37/142, 26.1%), colorectal cancer (35/142, 34.6%), multiple myeloma (35/142, 34.6%), non-Hodgkin lymphoma (35/142, 34.6%), bladder cancer (31/142, 21.8%), and brain cancer (31/142, 21.8%; Figure S12 in Multimedia Appendix 1). Respiratory disorders were primarily focused on asthma (15/30, 50%) and COPD (chronic obstructive pulmonary disease; 14/30, 47%). Parkinson disease was the most studied neurodegenerative condition (16/28, 57%), followed by multiple sclerosis (6/28, 21%). Fewer publications examined Alzheimer disease (2/28, 7%) and amyotrophic lateral sclerosis (2/28, 7%; Figure S13 in Multimedia Appendix 1). In the mental health field, suicide (12/26, 46%) and depression

(8/26, 31%) were the most investigated issues (Figure S14 in Multimedia Appendix 1).

Keyword Analysis

Overview

Following an initial extraction of 1259 authors' keywords, manual harmonization was performed. Duplicate keywords were removed through singular or plural standardization (130/1259, 10.3%) and synonym unification and grouping of cancer-related terms (553/1259, 43.9%), yielding a final set of 576 (45.8%) harmonized keywords, which were all used in subsequent analyses to prevent selection bias.

On average, each publication included 8.90 keywords (90% CI 0-17), although 35 (11.8%) out of 296 publications lacked any keywords, in line with the journal guidelines. Keyword analysis confirmed prior findings regarding farming exposure and health outcomes. It also provided deeper insights into emerging research hot spots, directions, and gaps.

Of the total 576 keywords, 301 (52.3%) appeared only once, while 68 (11.8%) were mentioned at least 10 times. More frequently used keywords included 39 that appeared at least 20 times (39/576, 6.8%) and 11 that featured in at least 50 publications (11/576, 1.9%). The 50 most frequently used keywords were mentioned in at least 17 (5.7%) out of 296 publications, while the top 10 appeared in at least 51 publications (17.2%; Table S7 in Multimedia Appendix 1). The most frequently cited keyword was "cancer" (150/296, 50.7%), followed by "mortality" (96/296, 32.4%), "pesticide" (88/296, 29.7%), "occupation" (82/296, 27.7%), "farmer" (77/296, 26.0%), "agriculture" (74/296, 25%), "exposure" (57/296, 19.3%), and "epidemiology" (57/296, 19.3%).

In terms of overall citations, "cancer" (5766/9379, 61.5%), "pesticide" (3569/9379, 38.1%), and "mortality" (3097/9379, 33%) were the most cited keywords. During the past decade, the frequency of the top 5 keywords has drastically increased (Figure S15 in Multimedia Appendix 1). Notably, keywords such as "cancer," "mortality," "occupation," "pesticide," "agriculture," and "farmer" have been consistently present in publications spanning at least 30 years (not necessarily consecutively; Figure S16 in Multimedia Appendix 1). In the last decade, emerging keywords, such as "big data," "administrative health database," "dust," and "BMI," have gained prominence (Figure S17 in Multimedia Appendix 1).

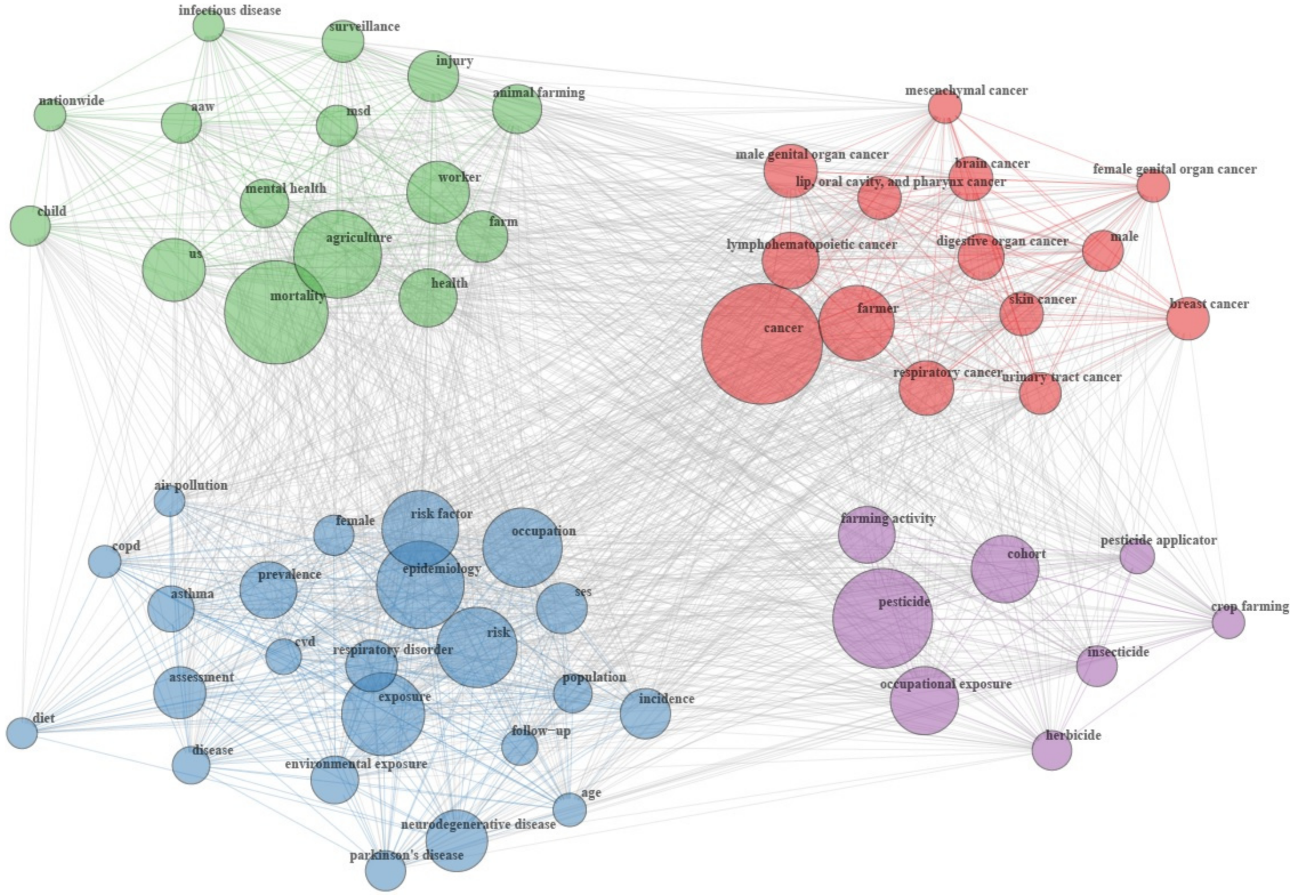
Keyword Co-Occurrence

A keyword co-occurrence network illustrating the frequency of keyword co-mentions in publications was constructed, thereby revealing relationships and conceptual connections (Figure 4). In this network, nodes or vertices represent keywords, with their sizes indicating frequency, while edges denote co-occurrences. The network's density and arrangement reveal topic interconnectivity, with larger vertices representing more frequently mentioned keywords. The network visualization helps identify clusters of related topics and highlights core research areas.

Using a community detection algorithm (spin-glass model with simulated annealing), 4 distinct clusters or communities of keywords were identified. Each cluster groups keywords that are often mentioned together, with stronger internal associations and weaker connections to keywords in other clusters.

The most frequently used keywords for each cluster were "cancer" (red cluster), "pesticide" (purple cluster), "mortality" (green cluster), and "exposure" (blue cluster). The red cluster highlights associations between various types of cancer, reflecting the fact that studies investigating cancer risks often examine multiple types of cancer. The green cluster links "mortality" with terms such as "mental health," "injury," and "animal farming," explained by the association between workplace accidents, mental health issues (eg, suicide), animal farming, and mortality. In the purple cluster, "pesticide" connects with "occupational exposure" and "farming activity," emphasizing that pesticide exposure is primarily studied in occupational settings across different types of farming. The blue cluster connects "exposure" to terms such as "neurodegenerative disease," "respiratory disorder," "cardiovascular disorder," "risk factor," "air pollution," "age," and "diet," indicating the study of various risk factors in relation to several health events. These clusters highlight current research hot spots that focus on 4 main interconnected themes: the associations between risk factors, pesticide exposure, farming activities, and a range of diseases.

Figure 4. Keyword co-occurrence network of the 296 articles published between 1975 and 2024. Each vertex or node represents a keyword, while edges represent the co-occurrence between keywords. Two keywords are connected when they co-occur in the same publication, and the size of each vertex indicates the frequency of a keyword: larger vertices represent more frequently mentioned keywords. Keywords with the same color (cluster) represent a research area. AAW: workplace accident; COPD: chronic obstructive pulmonary disease; CVD: cardiovascular disorder; MSD: musculoskeletal disorder; SES: socioeconomic status.



Thematic Mapping: Research Hot Spots

Figure 5 presents a thematic map that illustrates current research directions. Thematic mapping visualizes the relationship between research themes or topics, enabling the identification of directions, emerging areas, and gaps in the literature. The result is a strategic diagram that shows how themes relate to each other and their relevance within a specific field. The graph is divided into 4 quadrants, categorizing topics based on their relevance (x-axis, Callon centrality) and maturity (y-axis, Callon density) within the broader research landscape. Each circle represents a theme or topic (ie, a cluster of equivalent keywords), with the circle size corresponding to the frequency of the keywords associated with that theme.

The upper-right quadrant represents “hot topics,” which are both highly relevant and mature in the research landscape. Four key hot topics drive AHD-based public health research in farming populations. These include 1 topic focused on cancer research; another on respiratory disorders; and a third encompassing neurodegenerative diseases, workplace accidents, injuries, and mental health issues. The final hot topic involves large-scale studies in France and Europe using big data and insurance claims.

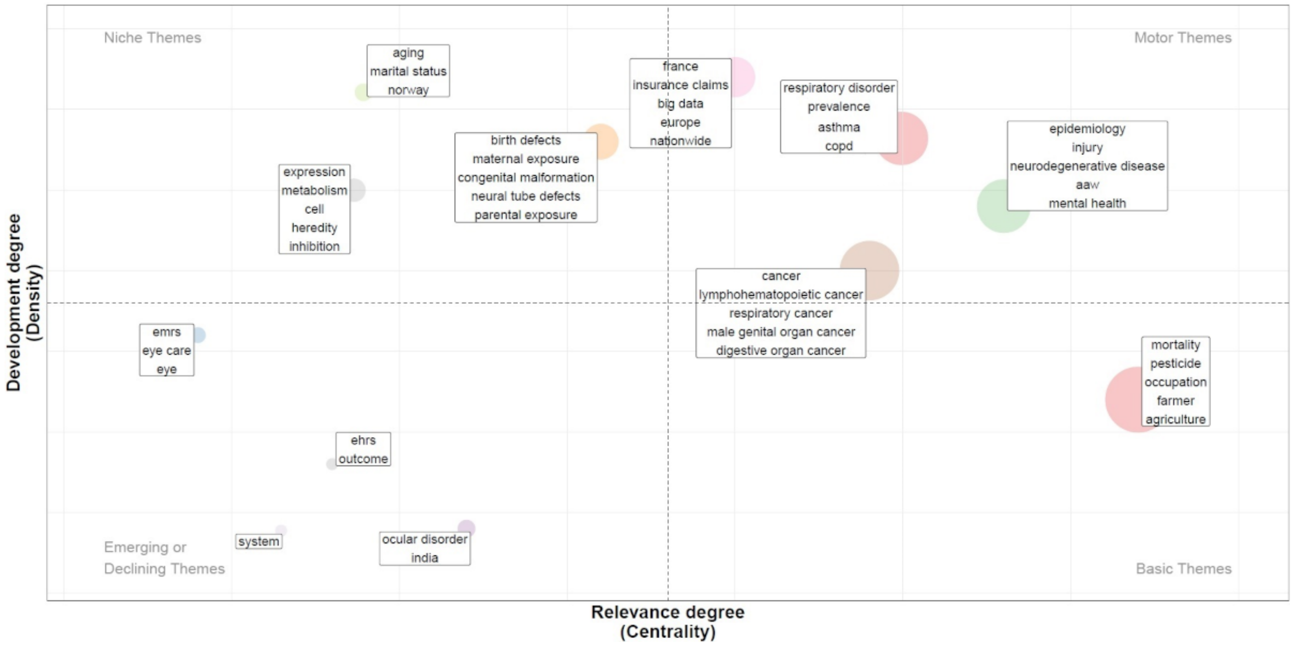
The lower-right quadrant contains “basic topics,” which are relevant but not yet mature in the research landscape. Only 1 such theme emerged: research related to pesticide exposure, mortality, and farming.

In the upper-left quadrant, “niche themes” refer to mature research topics that have not yet achieved full relevance. Three niche themes were identified: the first involves studies examining aging and research conducted in Norway; the second focuses on reproductive disorders and parental exposure, a theme poised to potentially evolve into a hot topic; and the final niche theme covers genetics and metabolism.

Finally, the lower-left quadrant contains “peripheral topics,” which represent either emerging or declining themes with low relevance and maturity. Four peripheral topics were observed, of which 2 (50%) were primarily centered on research on ocular disorders, 1 (25%) on the use of electronic health or medical records, and 1 (25%) on studies conducted in India.

This thematic map helps highlight both well-established and emerging areas of research, as well as gaps that may be ripe for future investigation.

Figure 5. Thematic mapping: research hot spots based on keywords from the 296 articles published between 1975 and 2024. The graph shows how themes relate to each other and their relevance within a specific field. This graph is divided into 4 quadrants, categorizing topics based on their relevance (x-axis and Callon centrality) and maturity (y-axis and Callon density) within the broader research landscape. Each circle represents a theme or topic (ie, a cluster of equivalent keywords), with the circle size corresponding to the frequency of the keywords associated with that theme. AAW: accident at work; COPD: chronic obstructive pulmonary disease; EHR: electronic health record; EMR: electronic medical record.



Bibliometric Farming Exposome

To identify research directions and gaps in the farming exposome literature, a bibliometric keyword co-occurrence analysis was conducted to explore the farming exposome by examining the co-occurrence between keywords associated with potential risk factors and specific health outcomes. This analysis was restricted to exposome-related and health event–related keywords. Of 576 keywords, 130 (22.6%) were related to the exposome, among which 93 (16.1%) were related to the specific external exposome (eg, pesticide), 19 (3.3%) to the general external exposome (eg, climate), and 18 (3.1%) to the internal exposome (eg, oxidative stress). Furthermore, there were 70 (12.2%) health event–related keywords (eg, brain cancer).

The results of this analysis are provided in Tables 5 and 6 and Multimedia Appendix 3, with each cell representing the percentage of occurrences of an exposome-related keyword (eg, air pollution) in all publications mentioning a specific health event–related keyword (eg, Alzheimer disease). For example, a value of 33.3 indicates that an exposome-related keyword appeared in 33.3% of all publications mentioning a specified health event–related keyword. To facilitate interpretation and ease the reading of Tables 5 and 6, exposome-related keywords were categorized into 19 groups (eg, chemical agent) and health event–related keywords into 20 groups.

Table 5. Co-occurrence between keywords related to internal exposomes and health event categories among the articles published between 1975 and 2024. Each cell refers to the number of times (%) a keyword related to an exposome category (eg, chemical agent) was mentioned among all publications in which a keyword related to a health event category (eg, cancer) appeared (N=296). Please note that the absolute value for each row is provided in parentheses with the row header and remains the same for all the parameters in that row.

Health event, n	Internal exposome (%)										
	Age	Sex	BMI	BP[a]	Heredity	Ethnicity	Hormone	Menopause	Metabolism	OS[b]	Inflammation
Cardiovascular disease (n=17)	5.88	11.8	5.88	5.88	0	5.88	0	0	0	0	0
Work-related disease (n=5)	0	20	0	0	0	0	0	0	0	0	0
Autoimmune disease (IBD[c], RA[d], vasculitis, and NR[e]; n=9)	0	0	11.1	0	0	0	0	0	0	0	0
Cancer (n=150)	3.62	27.5	2.17	0	0.73	1.45	3.62	0.73	2.9	0.73	0
Dental health (n=2)	0	0	0	0	0	0	0	0	0	0	0
Ocular disorder (n=6)	16.7	0	0	0	0	0	0	0	0	0	0
Frailty (n=2)	50	0	50	0	0	0	0	0	0	0	0
Anemia (n=1)	100	100	0	0	0	0	0	0	100	0	0
Infectious disease (malaria, Lyme disease, tuberculosis, toxoplasmosis, and NR; n=14)	0	7.14	0	0	0	7.14	0	0	0	0	0
Injury (including workplace accident and disability; n=40)	4.26	8.51	2.13	0	0	2.13	0	2.13	0	0	0
Chronic kidney disease (n=3)	0	0	0	0	0	0	0	0	0	0	0
Mental health disorder (depression, suicide, and NR; n=25)	9.52	4.76	4.76	0	0	0	0	0	0	0	0
Metabolic disorder (diabetes, dysthyroidism, and NR; n=9)	11.1	0	11.1	11.1	0	11.1	0	0	0	0	0
Mortality (n=75)	4.17	17.7	0	0	1.04	0	1.04	2.08	0	0	0
Musculoskeletal disorder (arthritis, low-back pain, and NR; n=14)	7.14	14.3	7.14	0	0	0	0	0	0	0	0
Neurodegenerative disease (AD[f], ALS[g], MND[h], MS[i], PD[j], and NR; n=33)	6.06	0	0	0	3.03	6.06	0	0	12.1	6.06	0
Sensory impairment (n=1)	0	0	0	0	0	0	0	0	0	0	0
Reproductive disorder (birth defects, infertility, spontaneous abortion, and NR; n=24)	0	28.6	14.3	0	0	7.14	7.14	0	7.14	0	0
Respiratory disorder (allergy, asthma, COPD[k], pneumonia, sarcoidosis, and NR; n=39)	5.88	11.8	8.82	2.94	0	5.88	0	0	5.88	0	2.94
Skin disorder (dermatitis and NR; n=2)	0	50	0	0	0	0	0	0	0	0	0

[a] BP: blood pressure.

[b] OS: oxidative stress.

[c] IBD: inflammatory bowel disease.

[d] RA: rheumatoid arthritis.

[e] NR: not reported.

[f] AD: Alzheimer disease.

[g] ALS: amyotrophic lateral sclerosis.

[h] MND: motor neuron disease

[i] MS: multiple sclerosis.

[j] PD: Parkinson disease.

[k] COPD: chronic obstructive pulmonary disease.

Distinct keyword exposome profiles were developed for each health event–related keyword (Figures S18-S43 in Multimedia Appendix 1), as illustrated in Figure 6 for mental health disorders. Most exposome-related keywords associated with keywords related to mental health disorders pertained to the type of occupations as well as chemical, lifestyle, socioeconomic, and psychological factors. Cancer-related keywords were associated mostly with keywords related to the internal (sex) and specific external exposome (chemical agents, lifestyle, and type of occupations). Autoimmune disease–related keywords co-occurred mostly with external exposome–related keywords (chemical agents, lifestyle, type of occupations, and socioeconomic factors). Neurodegenerative disease–related keywords were associated mostly with keywords related to the specific external exposome (lifestyle, chemical agents, and type of occupations). Reproductive disorders co-occurred mostly with internal (sex and BMI) and specific external exposome–related keywords (chemical agents and type of occupations). Keywords related to both musculoskeletal disorder and injury were associated with keywords from all 3 exposome components, in particular sex, type of occupations, lifestyle, biomechanical factors, chemical agents, and psychological factors. Infectious disease–related keywords co-occurred with specific external exposome–related keywords (biological agents and type of occupations), while respiratory disorder–related keywords were associated mostly with internal (sex) and specific external exposome–related keywords (lifestyle, chemical and biological agents, and type of occupations). Cardiovascular disorder–related keywords were associated with keywords from all 3 exposome components, in particular the sex, type of occupations, lifestyle, chemical agents, and socioeconomic and psychological factors.

Table 6. Co-occurrence between keywords related to specific external and general external exposomes and health event categories among the articles published between 1975 and 2024. Each cell refers to the number of times (%) a keyword related to an exposome category (eg, chemical agent) was mentioned among all publications in which a keyword related to a health event category (eg, cancer) appeared (N=296). Please note that the absolute value for each row is provided in parentheses with the row header and remains the same for all the parameters in that row.

Health event, n	Specific external exposome (%)					General external exposome (%)		
	Lifestyle	CA[a]	BA[b]	BF[c]	Occupation	PA[d]	SF[e]	PF[f]
Cardiovascular disease (n=17)	11.8	29.4	0	11.8	23.5	0	29.4	17.6
Work-related disease (n=5)	0	20	20	20	40	0	20	0
Autoimmune disease (IBD[g], RA[h], vasculitis, and NR[i]; n=9)	22.2	44.4	0	0	55.6	0	22.2	0
Cancer (n=150)	17.4	51.4	7.25	1.45	58.7	5.8	8.7	0
Dental health (n=2)	50	0	0	0	0	0	50	0
Ocular disorder (n=6)	16.7	16.7	0	0	0	0	0	16.7
Frailty (n=2)	100	50	0	0	0	0	0	50
Anemia (n=1)	0	0	0	0	0	0	0	0
Infectious disease (malaria, Lyme disease, tuberculosis, toxoplasmosis, and NR; n=14)	21.4	21.4	50	7.14	28.6	0	14.3	7.14
Injury (including workplace accident and disability; n=40)	12.8	14.9	2.13	14.9	23.4	2.13	4.26	6.38
Chronic kidney disease (n=3)	0	100	66.7	0	0	33.3	0	0
Mental health disorder (depression, suicide, and NR; n=25)	28.6	23.8	0	4.76	33.3	4.76	23.8	28.6
Metabolic disorder (diabetes, dysthyroidism, and NR; n=9)	33.3	44.4	0	11.1	44.4	11.1	0	22.2
Mortality (n=75)	17.7	55.2	8.33	4.17	54.2	5.21	17.7	2.08
Musculoskeletal disorder (arthritis, low-back pain, and NR; n=14)	14.3	21.4	7.14	14.3	42.9	7.14	0	21.4
Neurodegenerative disease (AD[j], ALS[k], MND[l], MS[m], PD[n], and NR; n=33)	21.2	48.5	6.06	0	36.4	0	9.09	3.03
Sensory impairment (n=1)	0	0	0	0	100	0	0	0
Reproductive disorder (birth defects, infertility, spontaneous abortion, and NR; n=24)	7.14	35.7	7.14	0	64.3	0	7.14	0
Respiratory disorder (allergy, asthma, COPD[o], pneumonia, sarcoidosis, and NR; n=39)	47.1	35.3	11.8	2.94	35.3	0	8.82	5.88
Skin disorder (dermatitis and NR; n=2)	50	50	0	0	50	0	0	0

[a] CA: chemical agent.

[b] BA: biological agent.

[c] BF: biomechanical factor.

[d] PA: physical agent.

[e] SF: socioeconomic factor.

[f] PF: psychological factor.

[g] IBD: inflammatory bowel disease.

[h] RA: rheumatoid arthritis.

[i] NR: not reported.

[j] AD: Alzheimer disease.

[k] ALS: amyotrophic lateral sclerosis.

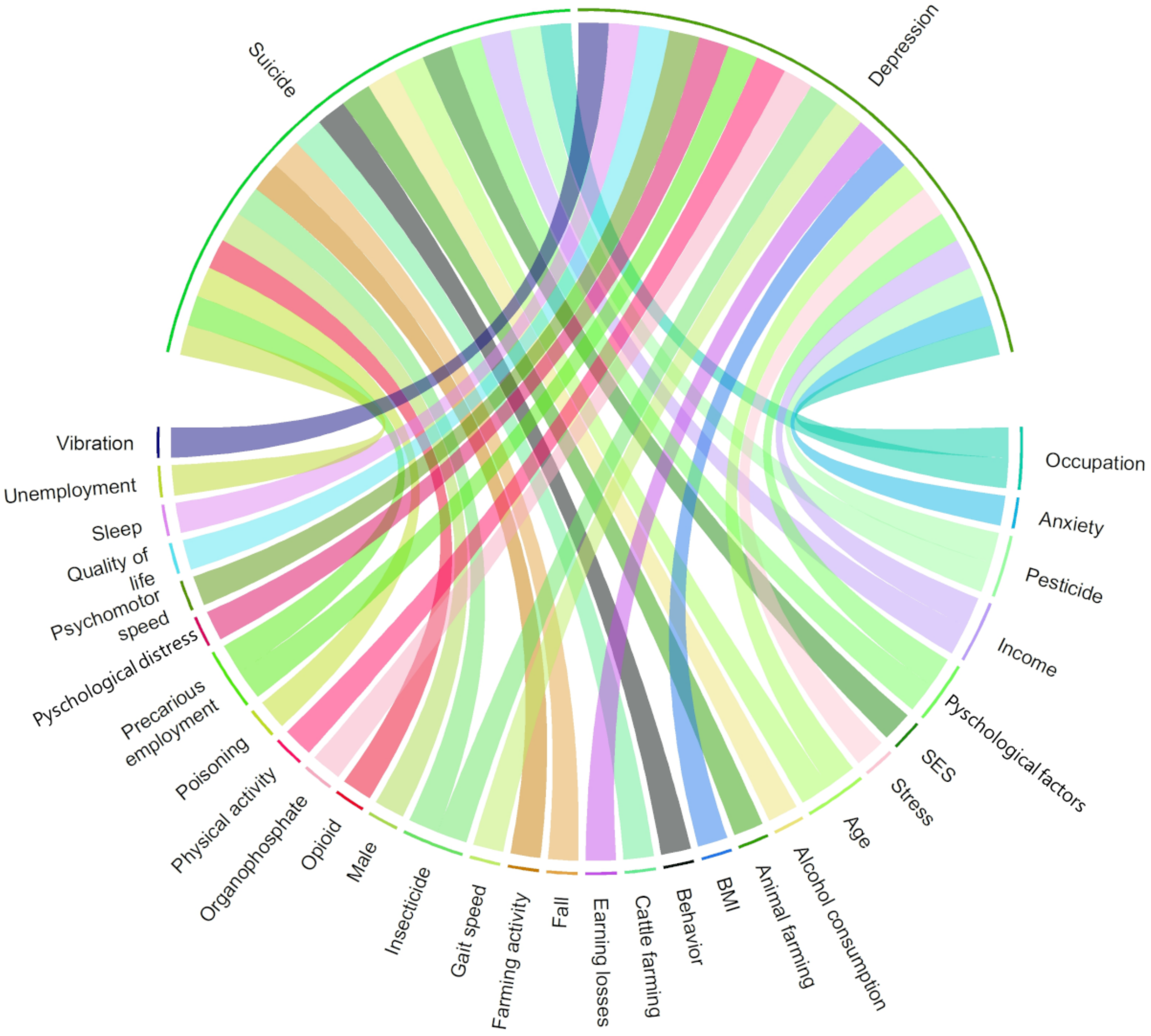
[l] MND: motor neuron disease

[m] MS: multiple sclerosis.

[n] PD: Parkinson disease.

[o] COPD: chronic obstructive pulmonary disease.

Figure 6. Chord diagram of keyword co-occurrence between potential risk factors and mental health disorder keywords from the 296 articles published between 1975 and 2024. Disease-related keywords are displayed on the top half of the chord diagram, while exposome-related keywords are displayed on the bottom half. Each chord or link indicates that an exposome-related keyword was mentioned with a disease-related keyword (co-occurrence) at least once in the same publication. The chord color differs from one exposome-related keyword to another. SES: socioeconomic status.



Discussion

Principal Findings

This review provides the first comprehensive and objective synthesis of research on the use of AHDs to address health issues in farming populations. It identifies major contributors, key publications, and existing research gaps while also suggesting future directions for leveraging AHDs to study health issues in farming populations. Overall, findings indicate that only a small part of the exposome and a limited range of health events have been examined within farming populations through the reuse of AHDs.

Current Directions in AHD Use for Public Health Research in Farming Populations

Research using AHDs in farming populations has been predominantly conducted in developed countries [150,225], with the United States [172,223,227]; France [207,222,224]; Canada [145,186,195]; and Scandinavian nations [182,189], including Denmark, Finland, Norway, and Sweden, leading the field. This dominance is linked to considerable funding from these regions and international collaborations. Scandinavian countries are particularly advanced in AHD use, offering databases that are highly complete, accessible, and well-integrated into public health research. AHDs from Denmark, Sweden, Canada, and France also provide comprehensive data on a patient's digital trajectory within their respective health systems [93,98,104,108,113,115,128,247]. France stands out further, with an AHD dedicated specifically to the entire farming population (MSA). This may explain the frequency of large-scale and long-term studies from these countries, some of which included >100,000 participants. However, many studies still had a regional focus, partly due to the use of AHDs by traditional epidemiological studies such as AGRICAN [207] and AHS [172], which rely on limited resources [47-50,53,54,57]. These studies often used AHDs to identify farming populations for inclusion or to supplement cohort data. The international AGRICOH consortium, initiated by the US National Cancer Institute and the International Agency for Research on Cancer, includes 11 (38%) of the 29 cohorts identified in this review [150,225]. However, several cohorts in AGRICOH were not identified, potentially due to lack of publications, language barriers, or limited use of AHDs. There were many publications associated with these well-established cohorts, for which many of the most prolific authors were working [172,182,189,207,222-227].

The most frequently used AHDs in farming population health research were disease registers, followed by electronic health or medical records and insurance claims. More than two-thirds of the studies used disease registers, in particular, cancer and mortality registers. This is not surprising because disease registers are created for clinical and research purposes with a continuous, exhaustive, and optimal digital collection of individual data regarding ≥1 health event in a geographically defined population [53,105]. The coding systems and the granularity of information related to health outcomes, populations, and exposure determinants varied widely across studies. Most studies (291/296, 98.3%) used AHDs to collect sociodemographic and health event information [222].

There was no consensus on the best methods or proxies to assess farming exposure. A variety of exposure proxies and determinants were used across studies, with indirect methods being the most common. Many studies (237/296, 80.1%) dichotomized proxy, for example, classifying individuals as farmers or nonfarmers or as pesticide-exposed versus nonexposed. In nearly two-thirds of the included studies, job title (ie, "being a farmer") served as the primary exposure proxy. About one-third of studies took into account specific farming activities (eg, dairy farming and crop farming) to reflect the diversity of farming practices. This approach is a valuable proxy for agricultural exposure, offering a broader representation of the farming exposome, which involves multiple stressors beyond just pesticides [147,152,188,201,248,249]. Farming activity information was often derived from digitalized data, such as agricultural censuses or self-reported data from mandatory insurance enrollments [152]. Many studies (111/296, 37.5%) combined AHDs with self-reported data (eg, questionnaires) [172,207], which allowed for more comprehensive data collection but tended to restrict the scope to regional studies due to resource constraints. These studies typically yielded high-quality data, with more potential confounders considered compared to studies relying solely on AHDs. Most studies (68/111, 61.3%) using self-reported data focused on single exposures, mainly pesticides, with only 1 study addressing multiexposure to pesticides [236]. Biological monitoring and airborne chemical sampling were rarely conducted, likely due to practical and financial constraints and the short half-lives of most pesticides [250]. Dermal chemical monitoring has not been reported, even though it is the main exposure route for pesticides [251]. The high number of studies investigating exposure to pesticides may be explained by the fact that AHS focuses on pesticide applicators and their spouses [172] and because many pesticides have adverse health effects on humans, such as neurotoxicity or endocrine disruption [252-256]. Beyond pesticides, farmers face exposure to other chemicals [257], such as air pollution; micro- and nanoplastics [258-263]; biological agents (eg, endotoxins and zoonoses) [264-266]; physical agents (eg, UV radiation, noise, and vibrations) [187,267]; biomechanical factors (eg, repetitive movements, heavy load, and working posture) [198,268,269]; and psychosocial stressors [270-274], which have been less studied than pesticides. Despite these multiple exposures, the broader farming exposome remains understudied.

In addition to AHDs, some studies integrated other secondary data, such as climate data [187,275], air quality data [194,217], JEMs [212,276,277], or crop-exposure matrices [278]. JEMs provide exposure level estimates for various chemicals and stressors based on job categories [250,279]. Although JEMs can provide valuable exposure data, they often lack the specificity of individual-level data, making it difficult to account for task-specific risks, temporal variations, and the inclusion of specific worker subgroups such as female individuals [250,279-282]. The lack of a universal standard for JEMs further complicates their application, which may explain why many studies still rely on self-reported data for more accurate exposure assessment despite the risk of recall bias [250,279-282].

The health outcomes studied were predominantly cancer [145,150,151,171-175,177-179,182,183,185,189,190,196,207,210-212,223], mortality [173,186,194,195,202,205], workplace injuries [198,208], respiratory disorders (eg, asthma and COPD) [151,180,181,213,224,226], neurodegenerative diseases (eg, Parkinson disease) [111,151,176,184,187,188,193,197,201,248,249,283], and mental health issues [151,244,273], which represent focal points within the research field. This is not surprising because many well-established cohorts centered on cancer research, in particular AGRICAN [207], Cancer in the Norwegian Agricultural Population [182], and NOCCA [189]. In addition, arsenic and inorganic arsenic compounds are classified as carcinogenic to humans by the International Agency for Research on Cancer [252,253,284], while malathion, glyphosate, diazinon, dichlorodiphenyltrichloroethane, and occupational exposures in spraying and application of nonarsenical insecticides are classified as probably carcinogenic to humans (group 2A), and several other pesticides are ranked as possibly carcinogenic to humans (group 2B), such as tetrachlorvinphos and parathion. Regarding mortality, it is often cancer and suicide mortalities that are investigated [186,195,202,244]. Furthermore, several pesticides are neurotoxic [252,255], but existing studies focused mainly on Parkinson disease and multiple sclerosis, with a paucity of data on Alzheimer disease and other neurodegenerative diseases [147,283]. In contrast, certain areas, such as cardiovascular diseases [151,194,203,227,285,286], autoimmune disorders [168,219,237], musculoskeletal disorders [192,204,245,287,288], reproductive disorders [242,289,290], sleep disorders [191], aging-related conditions [151,291], hearing impairment [267,292], and the microbiome [8,293-298], remain underexplored, despite their potential relevance to farming populations.

Challenges of Reusing AHD for Public Health Research in Farming Populations

Each AHD presents unique advantages and limitations. For example, large sample sizes and a large number of available health events are strengths, while generalizability and the absence of key confounders are challenges [64,93,95,105,115]. Access to AHDs is frequently restricted by a variety of factors, including governance and technical barriers, such as language, data structure, interoperability, and coding systems. Additional challenges stem from the type of AHD (eg, insurance claims or cancer registers), inadequate documentation (eg, absence of a data dictionary), limited accessibility due to costs or conditions, and jurisdictional and legal constraints [30,62,64,81,113,115]. Identifying the optimal AHD for a given research question is also complex, especially when considering the heterogeneity in coding systems and country-specific data structures. In countries such as Scandinavia, Canada, and France, individual identifiers facilitate data linkage across multiple AHDs, enhancing research opportunities [93,104,113,115,128,247]. However, many AHDs are not ready for research and require significant processing, cleaning, and understanding before they can be analyzed [93,105,113,222,299]. Another major challenge is the long lag between data access, analysis, and research publication. On average, studies used data that were 7 years old at the time of publication, largely due to delays in data access, administrative approvals, and the need to prepare complex datasets for analysis [223,300]. For instance, the TRACTOR project took 2 years to clean and prepare its dataset for research use [222]. These delays are compounded by the time required to conduct statistical analyses and prepare manuscripts for publication, as well as review and publication times (delay from submission to acceptance and from acceptance to publication). Another limitation of AHDs is the lack of detailed exposure data. AHDs rarely include exposure information due to their administrative focus, requiring researchers to supplement with additional data sources, such as JEMs or self-reported data. When exposure information is recorded in AHDs, it is often too generalized, typically only reflecting broad job classifications, such as farming, without specifying detailed activities or stressors. There are some exceptions, such as MSA data, which capture a wider range of specific farming activities (eg, dairy farming) [222]. However, exposure to specific stressors (eg, chemical compounds) is largely absent from AHDs.

The reference populations used in farming studies vary, which precludes direct comparisons and limits the generalizability of the findings. For example, AGRICAN used the general population as a reference [207], while TRACTOR used a farming population [152,168]. Furthermore, studies differ in their focus on specific farming populations, such as the entire agricultural workforce [207], farm managers [152], or pesticide applicators [172], which may lead to distinct exposure profiles that influence health outcomes because these farming populations have different socioeconomic status, experiences, and behaviors. Hence, to avoid or lessen bias, some studies focused on specific farming populations [152,172]. Moreover, the scope of farming populations included in studies is often limited, omitting subgroups, such as farm families, nearby residents, or consumers exposed to agricultural products, which limits the broader application of the findings. In addition, farming practices can vary significantly between countries and studies, and there is no international standardized classification for farming activities. In some cases, farming categories are derived from legal or administrative sources, as seen in the MSA data [152,168,283]. This lack of standardization limits the comparability and generalizability of findings across studies. In addition, the generalizability of the findings to other countries when using AHDs may be limited because of the differences in health care systems [93].

There are several well-known limitations of AHDs that complicate the investigation of health outcomes [301]. Health events captured in AHDs are typically limited to those requiring medical attention, which may not reflect the true incidence of diseases. In addition, the level of detail varies across diseases, even within the same AHD [92]. Although diagnostic accuracy is generally higher in disease registers, these are often geographically limited and cover only a subset of health conditions. For example, in France, cancer registers only cover 23 (24%) out of 96 administrative regions [155]. In addition, certain conditions, such as depression, are not covered by any registers. Identifying health outcomes in AHDs often requires complex algorithms that combine data from multiple sources, such as drug reimbursements, disease declarations, or medical procedures [104,105,128,152,302-304]. In addition, inconsistencies in case definitions and algorithms across studies and countries hinder the ability to compare and pool risk estimates [104,302,304]. Some AHDs also lack critical clinical information, such as laboratory results and genetic data [115,128,305], and the recorded diagnosis or treatment date may not correspond to the actual onset of the disease. Furthermore, diagnosis codes are not always indicative of a confirmed diagnosis.

Emerging Opportunities and Research Needs

While AHDs are well-used in certain countries, there are underexplored opportunities in regions such as the United Kingdom, where AHDs exist but are underused for research [113]. For low- and middle-income countries, the development and access to AHDs remain limited, and international support is needed to expand this research infrastructure. As already reported by Habib et al [306], there is a notable lack of sex-specific data, even though occupational exposures and health outcomes can vary significantly between sexes due to genetic, physiological, psychological, and behavioral factors [307-311]. Future research should address these disparities to provide a more comprehensive understanding of health risks in farming communities. Although there are inherent delays in using AHDs due to the time required for data generation, consolidation, and access, we advocate for the continued publication of studies, even those using older data. Historical data remain vital for better understanding long-term health patterns, particularly for diseases such as cancer, where tumor initiation can span decades [312,313]. Editors should encourage the publication of studies using older datasets, especially when addressing long-term health outcomes (eg, cancer and neurodegenerative diseases) or when recent data are not available [312].

None of the analyzed AHDs fully adhere to the FAIR principles, possibly because most were developed before the establishment of these principles in 2016 [228-230]. Moreover, the assessment of FAIR compliance of AHDs relied solely on information presented in publications, which may not provide a comprehensive evaluation. Nevertheless, there is a critical need to advocate for the integration of FAIR principles within AHDs to enhance public health research [228-230]. Currently, the landscape is favorable for data reuse, particularly with initiatives such as the forthcoming European Health Data Space [314-316]. Data reuse extends beyond mere access; it encompasses data discovery, a fundamental aspect of FAIR principles that involves recognizing the existence of databases [228-230]. To facilitate this, the creation of data catalogs is essential [228]. Numerous data repositories, such as Re3Data [317], Zenodo [318], CANUE [319], Figshare [320], “Epidémiologie – France” [321], data.gouv [322], Dataverse [320], or Data Europa [320], already exist. In addition, specialized multidisciplinary open-access and peer-reviewed journals such as *Scientific Data* and *Data in Brief* publish datasets [318]. A dataset search can also be conducted using the Google (Google LLC) platform [323]. However, the documentation and access conditions for datasets can highly vary across inventories, complicating the selection process for researchers. The absence of indicators or scores for data reusability further hampers efforts to identify the most suitable datasets for specific research questions [45,69]. To our knowledge, no comprehensive catalog of AHDs currently exists to date. A web-based inventory of AHDs, modeled after existing resources, such as OccupationalCohorts.net [324], OccupationalExposureTools.net [325], and Toxicological and Exposure Database Inventory [10], could greatly enhance research endeavors. The motivation for analyzing AHDs often stems from the data they contain. Consequently, as data availability increases, researchers will be better positioned to formulate research questions and engage in a parallel process of “datagraphy” or “datagraphic search” akin to traditional bibliographic research. The objective of datagraphy would be to determine which datasets are best suited for addressing specific research questions, highlighting the need for accessible catalogs to support this goal.

There is also an opportunity to integrate other secondary data, such as person-generated data (eg, mobile health, social media, digital footprints, and wearable sensors); contextual data (eg, climate and air quality); and smart agriculture data [2,83,84,326-328]. These datasets, largely untapped in farming population research, could provide new insights into health outcomes and environmental exposures [101].

Nationwide studies using big data were a hot spot. AI, such as machine learning, is particularly useful for analyzing big data and holds substantial promise for future research [329-331], particularly for identifying predictors of health outcomes in farming populations [332]. To date, AI has been underused, with only 2 studies using it, 1 identifying occupational injuries in agriculture [208] and 1 reviewing the development of chronic kidney disease risk prediction models [209]. Incorporating AI, along with cohort enrichment and interdisciplinary expert interpretation, could open new avenues for research.

Many studies continue to examine agriculture as a broad category, highlighting the need for more detailed investigations into specific farming activities and tasks [147,152,188,201,248,249]. Our analysis reveals major research gaps in understanding environmental and occupational exposures among farming populations, particularly with regard to emerging concerns such as per- and polyfluoroalkyl substances, biological agents, micro- and nanoplastics, and the impact of climate change. Climate change is a critical issue for agriculture, as it may drive shifts in pests, diseases, and farming practices [274,333-339]. Parental exposure appears to be a theme that will soon become a hot topic. Furthermore, research is needed to explore the farming exposome, particularly focusing on the “mixture effects” of multiple simultaneous exposures [340,341]. Omics data, which have not been used in farming population studies to date, represent a promising avenue for future research because genetics and metabolism were found to be a niche theme. Omics data refer to the large-scale datasets generated from various omics technologies that analyze biological molecules (eg, genomics, transcriptomics, proteomics, and metabolomics), which provide comprehensive insights into different biological layers and processes [11,342,343].

To enhance the characterization of farming exposome research using keyword analysis, there is a pressing need for standardized keyword reporting. We advocate for the development of a standardized approach to reporting keywords in scientific journals, including defining a minimum set of information (eg, study type, health outcome, population studied, data sources, and positive, negative, or null associations) and adopting a list of standardized terms. Although challenging, this approach would improve literature searches, make data more comparable and FAIR [228-230], and lead to more efficient, frugal (less time and energy spent to identify relevant information), and accurate synthesis of the scientific literature, such as in reviews and bibliometric analyses.

The prominence of topics such as cancer, neurodegenerative diseases, mortality, injuries, and mental health issues underscores the need for targeted prevention strategies. The thematic map analysis indicates that reproductive disorders (eg, birth defects, endometriosis, and infertility) are on the verge of becoming a central research focus. Emerging and understudied health conditions, including ocular disorders, autoimmune diseases (eg, inflammatory bowel disease and rheumatoid arthritis), sleep disorders (eg, sleep apnea), cardiovascular diseases, and musculoskeletal conditions (eg, low-back pain), warrant increased attention and further research. Aging-related health issues, such as frailty, also represent promising avenues for future research, particularly given the growing aging population and associated health care challenges [24].

Limitations

The findings of this review should be considered in view of their limitations. Because of time and resource constraints, a single screening approach was used, with only 1 author (PP) conducting the review and bibliometric analysis. While single screening is an efficient use of time and resources, there is a higher risk of missing relevant studies than when using dual screening [344,345]. However, when completed by an experienced reviewer familiar with the research topic, the proportion of missed studies is limited and estimated to be around 3% [344]. Therefore, we cannot exclude the possibility that some studies may have been missed. Nevertheless, we are confident that none of these methodological limitations would change the overall conclusions of this work. Our restriction on articles published in English and French may have inadvertently excluded potentially relevant publications. We cannot exclude the possibility that publications using AHDs for addressing health issues in farming populations may have been missed if there was no mention of AHD in the publications’ titles and abstracts. However, it is important to mention that our search strategy was similar to recent reviews that specifically examined the use of AHDs for population-based research [93,96,103,104]. We further broadened our search by including synonyms to improve the comprehensiveness of our literature search. Some details and specificities on the AHDs and other digital data used may be limited because only information reported in each study was used. Shortcomings inherent to bibliometric analysis cannot be excluded. Some authors may have duplicate names, and namesakes could exist. This limitation could not be prevented as a unique author identifier (eg, Open Researcher and Contributor ID number) was not available. Self-citation could not be identified.

While our keyword analysis helped map the farming exposome in AHD-based public health research, this profile is incomplete and potentially biased. Because our review focused on AHD-based studies, we likely missed relevant epidemiological studies not using AHDs, leaving gaps in our understanding of the complete farming exposome across public health. In addition, the variability in keyword reporting practices across journals introduced bias into our keyword analysis. Some journals limit the number of keywords, and the lack of standardized keyword ontologies adds further variability. To mitigate this bias, we manually harmonized the keywords (eg, the use of 1 unique term for a given entity). While this approach is time-consuming, it allows for a more accurate analysis. For instance, if this approach was not performed, the same entity could be designated by various terms that would have been considered separate entities or terms, potentially resulting in underestimates (eg, in the number of publications). Despite these challenges, the findings from our scoping review were consistent with the keyword analysis.

Notwithstanding the aforementioned limitations, most of which are inherent to all scoping reviews and bibliometric analyses [93,96,103,104,158-160], we are confident that our findings can provide a comprehensive picture of what has been published until now (the current state of research and general directions) regarding the use of AHDs for addressing health issues in farming populations. This study may lay the groundwork for researchers to quickly identify research priorities and emerging research directions investigating health issues in farming populations using AHDs.

Conclusions

Technological advancements have greatly increased the volume of research data available, positioning AHDs as critical resources for population-based public health studies [41]. Our review underscores the broad public health implications of AHDs, providing actionable insights for researchers, physicians, and policy makers (Textbox 3). Addressing the identified research gaps is crucial to comprehensively understanding health risks in farming populations.

The insights derived from AHDs can inform meaningful recommendations for policy makers and guide future research directions, ultimately aiding health services and health policy development. Our findings underscore the necessity of comprehensive, interdisciplinary approaches to better understand and mitigate the health risks encountered by farming populations. Such efforts will improve data comparability and research quality while also supporting the formulation of targeted prevention strategies. This, in turn, will enhance health outcomes for farming populations and promote the sustainability of agriculture in an increasingly dynamic environment. The findings from this review offer insights that are not only relevant to farming populations but also potentially generalizable to other populations.

Textbox 3. Take-home messages.

Farming population

Research focusing on low- and middle-income countries, as well as on underrepresented subgroups within farming communities (eg, women, children, and contingent workers), remains insufficiently developed. These areas warrant further investigation to ensure more comprehensive insights.

Administrative health database (AHD) use

The use of AHDs in public health research among farming populations is expanding, offering major potential to enhance epidemiological studies and inform public health decisions. Promoting AHD-based research alongside the integration of other secondary data and artificial intelligence approaches represents a promising direction for future exploration. There is also a need to promote findable, accessible, interoperable, and reusable principles. Creating an AHD catalog or inventory could be a solution that would allow researchers to conduct a “datagraphic search” akin to traditional bibliographic research.

Farming exposure

Published studies on farming-related exposures often rely on broad proxies, such as job titles, neglecting to capture the nuances of specific agricultural tasks. While pesticide exposure remains a predominant research focus, emerging concerns, such as per- and polyfluoroalkyl substances, biological agents, micro or nanoplastics, and the effects of climate change, require urgent attention. The farming exposome remains underexplored despite its potential to uncover important associations between risk factors and a diverse range of health outcomes.

Health outcomes

Cancer, respiratory diseases, neurodegenerative disorders, and mental health issues are among the most frequently studied health outcomes in farming populations. However, significant gaps exist in understanding other critical conditions, such as cardiovascular diseases, reproductive disorders, ocular conditions, autoimmune diseases, musculoskeletal disorders, age-related health issues, and microbiome impacts. Addressing these overlooked areas is essential for a more complete understanding of the health risks faced by farming communities.

Acknowledgments

This work was partly funded by MIAI@Grenoble Alpes (Multidisciplinary Institute in Artificial Intelligence at Grenoble Alpes; grant ANR-19-P3IA-0003, 2019) and the French National Research Agency within the framework of the Investissements d’avenir program (grants ANR‐10‐AIRT‐05, 2012; ANR‐15‐IDEX‐02, 2016). The funding sources had no role in the study design and conduct of the study; collection, management, analysis, and interpretation of data; writing of the report; preparation, review, or approval of the paper; or decision to submit the paper for publication. The authors were not precluded from accessing data in the study, and all authors took final responsibility for the decision to submit for publication.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request and also in the multimedia appendices.

Authors' Contributions

PP was involved in the conceptualization, methodology, software selection, validation, formal analysis, investigation, data curation, writing the original draft and reviewing and editing, visualization, project administration, and funding acquisition of the study. NV was involved in the conceptualization, reviewing and editing, resource collection, and funding acquisition of the study.
All authors meet all 4 criteria for authorship in the International Committee of Medical Journal Editors recommendations. All authors affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any potential discrepancies from the study as originally planned (and, if relevant, registered) have been explained.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 8778 KB - publichealth_v11i1e62939_app1.docx]

Multimedia Appendix 2
Publications included and analyzed.
[XLSX File (Microsoft Excel File), 80 KB - publichealth_v11i1e62939_app2.xlsx]

Multimedia Appendix 3
Co-occurrence between exposome-related keywords and health event-related keywords.
[XLSX File (Microsoft Excel File), 39 KB - publichealth_v11i1e62939_app3.xlsx]

Multimedia Appendix 4
PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 186 KB - publichealth_v11i1e62939_app4.pdf]

Multimedia Appendix 5
BIBLIO checklist for reporting the bibliometric reviews of the biomedical literature.
[PDF File (Adobe PDF File), 270 KB - publichealth_v11i1e62939_app5.pdf]

References

1. Azari R, Borisch B. What is public health? A scoping review. Arch Public Health 2023 May 10;81(1):86 [FREE Full text] [doi: 10.1186/s13690-023-01091-6] [Medline: 37165370]
2. Velmovitsky PE, Kirolos M, Alencar P, Leatherdale S, Cowan D, Morita PP. Leveraging mHealth technologies for public health. JMIR Public Health Surveill 2024 Sep 12;10:e49719 [FREE Full text] [doi: 10.2196/49719] [Medline: 39265164]
3. Barouki R, Lévi Y. Rapport 24-02. L'exposome, une contribution majeure pour la prévention. Bull Acad Natl Med 2024 May;208(5):523-535. [doi: 10.1016/j.banm.2024.03.004]
4. Barouki R, Audouze K, Becker C, Blaha L, Coumoul X, Karakitsios S, et al. The exposome and toxicology: a win-win collaboration. Toxicol Sci 2022 Feb 28;186(1):1-11 [FREE Full text] [doi: 10.1093/toxsci/kfab149] [Medline: 34878125]
5. Bennett B, Workman T, Smith MN, Griffith WC, Thompson B, Faustman EM. Characterizing the neurodevelopmental pesticide exposome in a children's agricultural cohort. Int J Environ Res Public Health 2020 Feb 25;17(5):1479 [FREE Full text] [doi: 10.3390/ijerph17051479] [Medline: 32106530]
6. Goodrich JA, Wang H, Jia Q, Stratakis N, Zhao Y, Maitre L, et al. Integrating Multi-Omics with environmental data for precision health: a novel analytic framework and case study on prenatal mercury induced childhood fatty liver disease. Environ Int 2024 Aug;190:108930 [FREE Full text] [doi: 10.1016/j.envint.2024.108930] [Medline: 39128376]
7. Haddad N, Andrianou XD, Makris KC. A scoping review on the characteristics of human exposome studies. Curr Pollution Rep 2019 Nov 28;5(4):378-393. [doi: 10.1007/S40726-019-00130-7]
8. Merra G, Gualtieri P, La Placa G, Frank G, Della Morte D, De Lorenzo A, et al. The relationship between exposome and microbiome. Microorganisms 2024 Jul 08;12(7):1386 [FREE Full text] [doi: 10.3390/microorganisms12071386] [Medline: 39065154]
9. Pandics T, Major D, Fazekas-Pongor V, Szarvas Z, Peterfi A, Mukli P, et al. Exposome and unhealthy aging: environmental drivers from air pollution to occupational exposures. Geroscience 2023 Dec 09;45(6):3381-3408 [FREE Full text] [doi: 10.1007/s11357-023-00913-3] [Medline: 37688657]
10. Petit P. Toxicological and exposure database inventory: a review. Int J Hyg Environ Health 2022 Sep;246:114055. [doi: 10.1016/j.ijheh.2022.114055] [Medline: 36279787]
11. Price EJ, Vitale CM, Miller GW, David A, Barouki R, Audouze K, et al. Merging the exposome into an integrated framework for "omics" sciences. iScience 2022 Mar 18;25(3):103976 [FREE Full text] [doi: 10.1016/j.isci.2022.103976] [Medline: 35310334]
12. Santos S, Maitre L, Warembourg C, Agier L, Richiardi L, Basagaña X, et al. Applying the exposome concept in birth cohort research: a review of statistical approaches. Eur J Epidemiol 2020 Mar 27;35(3):193-204 [FREE Full text] [doi: 10.1007/s10654-020-00625-4] [Medline: 32221742]
13. Siroux V, Agier L, Slama R. The exposome concept: a challenge and a potential driver for environmental health research. Eur Respir Rev 2016 Jun 31;25(140):124-129 [FREE Full text] [doi: 10.1183/16000617.0034-2016] [Medline: 27246588]
14. Tagliaferro S, Maio S, Pirona F, Stanisci I, Sarno G, Silvi P, EarlyFOOD, HEALS EXHES, et al. Assessing external exposome by implementing an environmental data management system using open data. Sci Rep 2024 Jul 26;14(1):17142 [FREE Full text] [doi: 10.1038/s41598-024-62924-0] [Medline: 39060268]
15. Vineis P, Barouki R. The exposome as the science of social-to-biological transitions. Environ Int 2022 Jul;165:107312 [FREE Full text] [doi: 10.1016/j.envint.2022.107312] [Medline: 35635963]
16. Wei X, Huang Z, Jiang L, Li Y, Zhang X, Leng Y, et al. Charting the landscape of the environmental exposome. Imeta 2022 Dec 02;1(4):e50. [doi: 10.1002/imt2.50] [Medline: 38867899]
17. Wild CP. The exposome: from concept to utility. Int J Epidemiol 2012 Feb;41(1):24-32. [doi: 10.1093/ije/dyr236] [Medline: 22296988]
18. Zhang P, Carlsten C, Chaleckis R, Hanhineva K, Huang M, Isobe T, et al. Defining the scope of exposome studies and research needs from a multidisciplinary perspective. Environ Sci Technol Lett 2021 Oct 12;8(10):839-852 [FREE Full text] [doi: 10.1021/acs.estlett.1c00648] [Medline: 34660833]

19. Ibanez A, Melloni L, Świeboda P, Hynes W, Ikiz B, Ayadi R, et al. Neuroecological links of the exposome and one health. Neuron 2024 Jun 19;112(12):1905-1910 [FREE Full text] [doi: 10.1016/j.neuron.2024.04.016] [Medline: 38723637]
20. Liang KH, Colombijn JM, Verhaar MC, Ghannoum M, Timmermans EJ, Vernooij RW. The general external exposome and the development or progression of chronic kidney disease: a systematic review and meta-analyses. Environ Pollut 2024 Oct 01;358:124509 [FREE Full text] [doi: 10.1016/j.envpol.2024.124509] [Medline: 38968981]
21. Canfell OJ, Littlewood R, Burton-Jones A, Sullivan C. Digital health and precision prevention: shifting from disease-centred care to consumer-centred health. Aust Health Rev 2022 Jun 10;46(3):279-283. [doi: 10.1071/AH21063] [Medline: 34882538]
22. Mess F, Blaschke S, Schick TS, Friedrich J. Precision prevention in worksite health-a scoping review on research trends and gaps. PLoS One 2024 Jun 10;19(6):e0304951 [FREE Full text] [doi: 10.1371/journal.pone.0304951] [Medline: 38857277]
23. Wang Y, Zhu M, Ma H, Shen H. Polygenic risk scores: the future of cancer risk prediction, screening, and precision prevention. Med Rev (2021) 2021 Dec;1(2):129-149 [FREE Full text] [doi: 10.1515/mr-2021-0025] [Medline: 37724297]
24. Kalia V, Belsky DW, Baccarelli AA, Miller GW. An exposomic framework to uncover environmental drivers of aging. Exposome 2022;2(1):osac002 [FREE Full text] [doi: 10.1093/exposome/osac002] [Medline: 35295547]
25. Chung MK, House JS, Akhtari FS, Makris KC, Langston MA, Islam KT, Members of the Exposomics Consortium. Decoding the exposome: data science methodologies and implications in exposome-wide association studies (ExWASs). Exposome 2024;4(1):osae001 [FREE Full text] [doi: 10.1093/exposome/osae001] [Medline: 38344436]
26. Andreu-Perez J, Poon CC, Merrifield RD, Wong ST, Yang GZ. Big data for health. IEEE J Biomed Health Inform 2015 Jul;19(4):1193-1208 [FREE Full text] [doi: 10.1109/JBHI.2015.2450362] [Medline: 26173222]
27. Li C, Huang M. Environmental sustainability in the age of big data: opportunities and challenges for business and industry. Environ Sci Pollut Res Int 2023 Dec 03;30(56):119001-119015. [doi: 10.1007/s11356-023-30301-5] [Medline: 37919492]
28. Pastorino R, De Vito C, Migliara G, Glocker K, Binenbaum I, Ricciardi W, et al. Benefits and challenges of Big Data in healthcare: an overview of the European initiatives. Eur J Public Health 2019 Oct 01;29(Supplement_3):23-27 [FREE Full text] [doi: 10.1093/eurpub/ckz168] [Medline: 31738444]
29. Flahault A, Geissbuhler A, Guessous I, Guérin P, Bolon I, Salathé M, et al. Precision global health in the digital age. Swiss Med Wkly 2017 Apr 07;147(1314):w14423 [FREE Full text] [doi: 10.4414/smw.2017.14423] [Medline: 28421566]
30. Manaouil C, Chamot S, Petit P. Le médecin confronté à l'IA (Intelligence artificielle): Éthique et responsabilité. Méd Droit 2024 Jun;2024(186):50-66. [doi: 10.1016/j.meddro.2024.02.001]
31. Samuel G, Lucassen AM. The environmental sustainability of data-driven health research: a scoping review. Digit Health 2022 Jul 05;8:20552076221111297 [FREE Full text] [doi: 10.1177/20552076221111297] [Medline: 35847526]
32. Adam NR, Wieder R, Ghosh D. Data science, learning, and applications to biomedical and health sciences. Ann N Y Acad Sci 2017 Jan 25;1387(1):5-11. [doi: 10.1111/nyas.13309] [Medline: 28122121]
33. Iyamu I, Xu AX, Gómez-Ramírez O, Ablona A, Chang H, Mckee G, et al. Defining digital public health and the role of digitization, digitalization, and digital transformation: scoping review. JMIR Public Health Surveill 2021 Nov 26;7(11):e30399 [FREE Full text] [doi: 10.2196/30399] [Medline: 34842555]
34. Mooney SJ, Pejaver V. Big data in public health: terminology, machine learning, and privacy. Annu Rev Public Health 2018 Apr 01;39:95-112 [FREE Full text] [doi: 10.1146/annurev-publhealth-040617-014208] [Medline: 29261408]
35. Wesson P, Hswen Y, Valdes G, Stojanovski K, Handley MA. Risks and opportunities to ensure equity in the application of big data research in public health. Annu Rev Public Health 2022 Apr 05;43(1):59-78 [FREE Full text] [doi: 10.1146/annurev-publhealth-051920-110928] [Medline: 34871504]
36. Iyamu I, Gómez-Ramírez O, Xu AX, Chang H, Watt S, Mckee G, et al. Challenges in the development of digital public health interventions and mapped solutions: findings from a scoping review. Digit Health 2022;8:20552076221102255. [doi: 10.1177/20552076221102255] [Medline: 35656283]
37. Iyamu I, McKee G, Haag D, Gilbert M. Defining the role of digital public health in the evolving digital health landscape: policy and practice implications in Canada. Health Promot Chronic Dis Prev Can 2024 Feb;44(2):66-69 [FREE Full text] [doi: 10.24095/hpcdp.44.2.04] [Medline: 38353941]
38. Khan S, Kandukuri DK, Subramaniyan EU, MohanaSundaram A. Harnessing the untapped potential of digital twin technology in digital public health interventions. Explor Digit Health Technol 2023 Nov 1;5(4):11-16. [doi: 10.37349/edht.2023.00003]
39. Maaß L, Pan CC, Freye M. Mapping digital public health interventions among existing digital technologies and internet-based interventions to maintain and improve population health in practice: protocol for a scoping review. JMIR Res Protoc 2022 Mar 31;11(3):e33404 [FREE Full text] [doi: 10.2196/33404] [Medline: 35357321]
40. Odone A, Buttigieg S, Ricciardi W, Azzopardi-Muscat N, Staines A. Public health digitalization in Europe. Eur J Public Health 2019 Oct 01;29(Supplement_3):28-35 [FREE Full text] [doi: 10.1093/eurpub/ckz161] [Medline: 31738441]
41. Wienert J, Jahnel T, Maaß L. What are digital public health interventions? First steps toward a definition and an intervention classification framework. J Med Internet Res 2022 Jun 28;24(6):e31921 [FREE Full text] [doi: 10.2196/31921] [Medline: 35763320]
42. Burgun A, Bernal-Delgado E, Kuchinke W, van Staa T, Cunningham J, Lettieri E, et al. Health data for public health: towards new ways of combining data sources to support research efforts in Europe. Yearb Med Inform 2017 Aug 19;26(01):235-240. [doi: 10.1055/S-0037-1606509]

43. Canfell OJ, Davidson K, Sullivan C, Eakin E, Burton-Jones A. Data sources for precision public health of obesity: a scoping review, evidence map and use case in Queensland, Australia. BMC Public Health 2022 Mar 24;22(1):584 [FREE Full text] [doi: 10.1186/s12889-022-12939-x] [Medline: 35331189]
44. Canfell OJ, Davidson K, Woods L, Sullivan C, Cocoros NM, Klompas M, et al. Precision public health for non-communicable diseases: an emerging strategic roadmap and multinational use cases. Front. Public Health 2022 Apr 8;10:854525 [FREE Full text] [doi: 10.3389/fpubh.2022.854525] [Medline: 35462850]
45. Sielemann K, Hafner A, Pucker B. The reuse of public datasets in the life sciences: potential risks and rewards. PeerJ 2020;8:e9954 [FREE Full text] [doi: 10.7717/peerj.9954] [Medline: 33024631]
46. Deal JA, Betz J, Lin FR, Reed NS. Interpreting results from epidemiologic studies. Semin Hear 2021 Feb 15;42(1):3-9 [FREE Full text] [doi: 10.1055/s-0041-1725996] [Medline: 33883787]
47. DiPietro NA. Methods in epidemiology: observational study designs. Pharmacotherapy 2010 Oct 06;30(10):973-984. [doi: 10.1592/phco.30.10.973] [Medline: 20874034]
48. Cancer epidemiology: principles and methods. International Agency for Research on Cancer. 1999. URL: https://publications.iarc.fr/_publications/media/download/4106/a72983d2524d630f6ad0539f747e4780447f1ae3.pdf [accessed 2024-04-29]
49. Munnangi S, Boktor SW. Epidemiology of study design. StatPearls. 2014. URL: https://www.ncbi.nlm.nih.gov/books/NBK470342/ [accessed 2024-04-29]
50. Röhrig B, du Prel JB, Wachtlin D, Blettner M. Types of study in medical research: part 3 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2009 Apr;106(15):262-268 [FREE Full text] [doi: 10.3238/arztebl.2009.0262] [Medline: 19547627]
51. Himali JJ, Baril AA, Cavuoto MG, Yiallourou S, Wiedner CD, Himali D, et al. Association between slow-wave sleep loss and incident dementia. JAMA Neurol 2023 Dec 01;80(12):1326-1333. [doi: 10.1001/jamaneurol.2023.3889] [Medline: 37902739]
52. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. Lancet 2014 Mar 15;383(9921):999-1008 [FREE Full text] [doi: 10.1016/S0140-6736(13)61752-3] [Medline: 24084292]
53. Salines G, De LC. Les cohortes : intérêt, role et position de l'InVS. In: Saint-Maurice (Fra) : Institut de veille sanitaire, janvier 2010, 20 p. Princeton, NJ: Princeton University Press; 2018.
54. Zins M, Goldberg M. L'apport des bases de données d'origine administrative aux cohortes épidémiologiques : l'exemple de la cohorte Constances. Statistique et Société 2014;2(2):43-48 [FREE Full text]
55. Chen W, Wang X, Chen J, You C, Ma L, Zhang W, et al. Household air pollution, adherence to a healthy lifestyle, and risk of cardiometabolic multimorbidity: results from the China health and retirement longitudinal study. Sci Total Environ 2023 Jan 10;855:158896. [doi: 10.1016/j.scitotenv.2022.158896] [Medline: 36150596]
56. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb 12;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
57. Kesse-Guyot E, Assmann K, Andreeva V, Castetbon K, Méjean C, Touvier M, et al. Metadata correction of: lessons learned from methodological validation research in e-epidemiology. JMIR Public Health Surveill 2016 Oct 31;2(2):e160 [FREE Full text] [doi: 10.2196/publichealth.6876] [Medline: 27797322]
58. Quantin C. Bases de données médicoadministratives et épidémiologie : intérêt, limites. Revue d'Épidémiologie et de Santé Publique 2008 Sep;56(5):S254. [doi: 10.1016/j.respe.2008.06.011]
59. Babik KR, Shockey TM, Moore LL, Wurzelbacher SJ. Standardizing industrial hygiene data collection forms used by workers' compensation insurers. J Occup Environ Hyg 2018 Sep 04;15(9):676-685 [FREE Full text] [doi: 10.1080/15459624.2018.1490022] [Medline: 29985777]
60. Bosward R, Braunack-Mayer A, Frost E, Carter S. Mapping precision public health definitions, terminology and applications: a scoping review protocol. BMJ Open 2022 Feb 23;12(2):e058069 [FREE Full text] [doi: 10.1136/bmjopen-2021-058069] [Medline: 35197357]
61. Canfell OJ, Kodiyattu Z, Eakin E, Burton-Jones A, Wong I, Macaulay C, et al. Real-world data for precision public health of noncommunicable diseases: a scoping review. BMC Public Health 2022 Nov 24;22(1):2166 [FREE Full text] [doi: 10.1186/s12889-022-14452-7] [Medline: 36434553]
62. Chao K, Sarker MN, Ali I, Firdaus RB, Azman A, Shaed MM. Big data-driven public health policy making: potential for the healthcare industry. Heliyon 2023 Sep;9(9):e19681 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e19681] [Medline: 37809720]
63. Dolley S. Big Data's role in precision public health. Front Public Health 2018 Mar 7;6:68 [FREE Full text] [doi: 10.3389/fpubh.2018.00068] [Medline: 29594091]
64. Doyle CM, Lix LM, Hemmelgarn BR, Paterson JM, Renoux C. Data variability across Canadian administrative health databases: differences in content, coding, and completeness. Pharmacoepidemiol Drug Saf 2020 Jan 10;29 Suppl 1(S1):68-77. [doi: 10.1002/pds.4889] [Medline: 31507029]
65. Gallini A, Moisan F, Maura G, Carcaillon-Bentata L, Leray E, Haesebaert J, et al. [Identification of neurodegenerative diseases in administrative databases in France: a systematic review of the literature]. Rev Epidemiol Sante Publique 2017 Oct;65 Suppl 4:S183-S197. [doi: 10.1016/j.respe.2017.01.115] [Medline: 28341166]

66. Khoury MJ, Iademarco MF, Riley WT. Precision public health for the era of precision medicine. Am J Prev Med 2016 Mar;50(3):398-401 [FREE Full text] [doi: 10.1016/j.amepre.2015.08.031] [Medline: 26547538]
67. McCullough LE, Maliniak ML, Amin AB, Baker JM, Baliashvili D, Barberio J, et al. Epidemiology beyond its limits. Sci Adv 2022 Jun 10;8(23):eabn3328 [FREE Full text] [doi: 10.1126/sciadv.abn3328] [Medline: 35675391]
68. Näher AF, Vorisek CN, Klopfenstein SA, Lehne M, Thun S, Alsalamah S, et al. Secondary data for global health digitalisation. Lancet Digit Health 2023 Feb;5(2):e93-101 [FREE Full text] [doi: 10.1016/S2589-7500(22)00195-9] [Medline: 36707190]
69. Koesten L, Vougiouklis P, Simperl E, Groth P. Dataset reuse: toward translating principles to practice. Patterns (N Y) 2020 Nov 13;1(8):100136 [FREE Full text] [doi: 10.1016/j.patter.2020.100136] [Medline: 33294873]
70. Hu H, Galea S, Rosella L, Henry D. Big data and population health: focusing on the health impacts of the social, physical, and economic environment. Epidemiology 2017 Nov;28(6):759-762. [doi: 10.1097/EDE.0000000000000711] [Medline: 28682850]
71. Shaw RJ, Harron KL, Pescarini JM, Pinto Junior EP, Allik M, Siroky AN, et al. Biases arising from linked administrative data for epidemiological research: a conceptual framework from registration to analyses. Eur J Epidemiol 2022 Dec 05;37(12):1215-1224 [FREE Full text] [doi: 10.1007/s10654-022-00934-w] [Medline: 36333542]
72. Subramanian M, Wojtusciszyn A, Favre L, Boughorbel S, Shan J, Letaief KB, et al. Precision medicine in the era of artificial intelligence: implications in chronic disease management. J Transl Med 2020 Dec 09;18(1):472 [FREE Full text] [doi: 10.1186/s12967-020-02658-5] [Medline: 33298113]
73. Velmovitsky PE, Bevilacqua T, Alencar P, Cowan D, Morita PP. Convergence of precision medicine and public health into precision public health: toward a big data perspective. Front Public Health 2021 Apr 6;9:561873 [FREE Full text] [doi: 10.3389/fpubh.2021.561873] [Medline: 33889555]
74. Zhang H, Hu H, Diller M, Hogan WR, Prosperi M, Guo Y, et al. Semantic standards of external exposome data. Environ Res 2021 Jun;197:111185 [FREE Full text] [doi: 10.1016/j.envres.2021.111185] [Medline: 33901445]
75. Daniels K, Bonnechère B. Harnessing digital health interventions to bridge the gap in prevention for older adults. Front Public Health 2023 Jan 8;11:1281923 [FREE Full text] [doi: 10.3389/fpubh.2023.1281923] [Medline: 38259780]
76. Liu F, Panagiotakos D. Real-world data: a brief review of the methods, applications, challenges and opportunities. BMC Med Res Methodol 2022 Nov 05;22(1):287 [FREE Full text] [doi: 10.1186/s12874-022-01768-6] [Medline: 36335315]
77. Maaß L, Freye M, Pan CC, Dassow HH, Niess J, Jahnel T. The definitions of health apps and medical apps from the perspective of public health and law: qualitative analysis of an interdisciplinary literature overview. JMIR Mhealth Uhealth 2022 Oct 31;10(10):e37980 [FREE Full text] [doi: 10.2196/37980] [Medline: 36315221]
78. Mooney SJ, Garber MD. Sampling and sampling frames in big data epidemiology. Curr Epidemiol Rep 2019 Mar 2;6(1):14-22 [FREE Full text] [doi: 10.1007/s40471-019-0179-y] [Medline: 31360626]
79. Pearson AL, Tribby C, Brown CD, Yang JA, Pfeiffer K, Jankowska MM. Systematic review of best practices for GPS data usage, processing, and linkage in health, exposure science and environmental context research. BMJ Open 2024 Feb 01;14(2):e077036 [FREE Full text] [doi: 10.1136/bmjopen-2023-077036] [Medline: 38307539]
80. Prosperi M, Min JS, Bian J, Modave F. Big data hurdles in precision medicine and precision public health. BMC Med Inform Decis Mak 2018 Dec 29;18(1):139 [FREE Full text] [doi: 10.1186/s12911-018-0719-2] [Medline: 30594159]
81. Safran C. Update on data reuse in health care. Yearb Med Inform 2017 Aug;26(1):24-27 [FREE Full text] [doi: 10.15265/IY-2017-013] [Medline: 29063535]
82. Severo M, Feredj A, Romele A. Soft data and public policy: can social media offer alternatives to official statistics in urban policymaking? Policy Internet 2016 Jul 19;8(3):354-372. [doi: 10.1002/poi3.127]
83. Shakeri Hossein Abad Z, Butler GP, Thompson W, Lee J. Physical activity, sedentary behavior, and sleep on Twitter: multicountry and fully labeled public data set for digital public health surveillance research. JMIR Public Health Surveill 2022 Feb 14;8(2):e32355 [FREE Full text] [doi: 10.2196/32355] [Medline: 35156938]
84. Singh T, Roberts K, Cohen T, Cobb N, Wang J, Fujimoto K, et al. Social media as a research tool (SMaaRT) for risky behavior analytics: methodological review. JMIR Public Health Surveill 2020 Nov 30;6(4):e21660 [FREE Full text] [doi: 10.2196/21660] [Medline: 33252345]
85. Zins M, Cuggia M, Goldberg M. Les données de santé en France - Abondantes mais complexes. Med Sci (Paris) 2021 Feb 16;37(2):179-184 [FREE Full text] [doi: 10.1051/medsci/2021001] [Medline: 33591261]
86. Motahari-Nezhad H, Fgaier M, Mahdi Abid M, Péntek M, Gulácsi L, Zrubka Z. Digital biomarker-based studies: scoping review of systematic reviews. JMIR Mhealth Uhealth 2022 Oct 24;10(10):e35722 [FREE Full text] [doi: 10.2196/35722] [Medline: 36279171]
87. Ayilara OF, Platt RW, Dahl M, Coulombe J, Ginestet PG, Chateau D, et al. Generating synthetic data from administrative health records for drug safety and effectiveness studies. Int J Popul Data Sci 2023 Nov 27;8(1):2176 [FREE Full text] [doi: 10.23889/ijpds.v8i1.2176] [Medline: 38414538]
88. Fadel M, Petot J, Gourraud PA, Descatha A. Flexibility of a large blindly synthetized avatar database for occupational research: example from the CONSTANCES cohort for stroke and knee pain. PLoS One 2024 Jul 31;19(7):e0308063 [FREE Full text] [doi: 10.1371/journal.pone.0308063] [Medline: 39083487]
89. Giuffrè M, Shung DL. Harnessing the power of synthetic data in healthcare: innovation, application, and privacy. NPJ Digit Med 2023 Oct 09;6(1):186 [FREE Full text] [doi: 10.1038/s41746-023-00927-3] [Medline: 37813960]

90. Kamel Boulos MN, Zhang P. Digital twins: from personalised medicine to precision public health. J Pers Med 2021 Jul 29;11(8):745 [FREE Full text] [doi: 10.3390/jpm11080745] [Medline: 34442389]
91. Vallevik VB, Babic A, Marshall SE, Elvatun S, Brøgger HB, Alagaratnam S, et al. Can I trust my fake data - a comprehensive quality assessment framework for synthetic tabular data in healthcare. Int J Med Inform 2024 May;185:105413 [FREE Full text] [doi: 10.1016/j.ijmedinf.2024.105413] [Medline: 38493547]
92. Chamot S, Manaouil C, Fantoni S, Frimat P, Al-Salameh A, Desailloud R, et al. Système national des données de santé (SNDS) : perspectives en santé au travail et environnementale. Arch Mal Prof Environ 2023 Nov;84(6):101875. [doi: 10.1016/j.admp.2023.101875]
93. Cohen S, Gilutz H, Marelli AJ, Iserin L, Benis A, Bonnet D, et al. Administrative health databases for addressing emerging issues in adults with CHD: a systematic review. Cardiol Young 2018 Jun 29;28(6):844-853. [doi: 10.1017/S1047951118000446] [Medline: 29704902]
94. Fraser HS, Mugisha M, Bacher I, Ngenzi JL, Seebregts C, Umubyeyi A, et al. Factors influencing data quality in electronic health record systems in 50 health facilities in Rwanda and the role of clinical alerts: cross-sectional observational study. JMIR Public Health Surveill 2024 Jul 03;10:e49127 [FREE Full text] [doi: 10.2196/49127] [Medline: 38959048]
95. Gavrielov-Yusim N, Friger M. Use of administrative medical databases in population-based research. J Epidemiol Community Health 2014 Mar;68(3):283-287. [doi: 10.1136/jech-2013-202744] [Medline: 24248997]
96. Hinds A, Lix LM, Smith M, Quan H, Sanmartin C. Quality of administrative health databases in Canada: a scoping review. Can J Public Health 2016 Jun 27;107(1):e56-e61 [FREE Full text] [doi: 10.17269/cjph.107.5244] [Medline: 27348111]
97. Howell CR, Zhang L, Clay OJ, Dutton G, Horton T, Mugavero MJ, et al. Social determinants of health phenotypes and cardiometabolic condition prevalence among patients in a large academic health system: latent class analysis. JMIR Public Health Surveill 2024 Aug 07;10:e53371 [FREE Full text] [doi: 10.2196/53371] [Medline: 39113389]
98. Jones KH, Heys S, Tingay KS, Jackson P, Dibben C. The good, the bad, the clunky: improving the use of administrative data for research. Int J Popul Data Sci 2018 Dec 04;4(1):587 [FREE Full text] [doi: 10.23889/ijpds.v4i1.587] [Medline: 32935024]
99. Namulanda G, Qualters J, Vaidyanathan A, Roberts E, Richardson M, Fraser A, et al. Electronic health record case studies to advance environmental public health tracking. J Biomed Inform 2018 Mar;79:98-104 [FREE Full text] [doi: 10.1016/j.jbi.2018.02.012] [Medline: 29476967]
100. Nicholls SG, Quach P, von Elm E, Guttmann A, Moher D, Petersen I, et al. The REporting of studies conducted using observational routinely-collected health data (RECORD) statement: methods for arriving at consensus and developing reporting guidelines. PLoS One 2015;10(5):e0125620 [FREE Full text] [doi: 10.1371/journal.pone.0125620] [Medline: 25965407]
101. Silverwood R, Rajah N, Calderwood L, De Stavola B, Harron K, Ploubidis G. Examining the quality and population representativeness of linked survey and administrative data: guidance and illustration using linked 1958 National Child Development Study and Hospital Episode Statistics data. Int J Popul Data Sci 2024 Jan 09;9(1):2137 [FREE Full text] [doi: 10.23889/ijpds.v9i1.2137] [Medline: 38425790]
102. Zuiderwijk A, Türk BO, Brazier F. Identifying the most important facilitators of open research data sharing and reuse in Epidemiology: a mixed-methods study. PLoS One 2024 Feb 8;19(2):e0297969 [FREE Full text] [doi: 10.1371/journal.pone.0297969] [Medline: 38330007]
103. Hobensack M, Song J, Scharp D, Bowles KH, Topaz M. Machine learning applied to electronic health record data in home healthcare: a scoping review. Int J Med Inform 2023 Feb;170:104978 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104978] [Medline: 36592572]
104. Tassi MF, le Meur N, Stéfic K, Grammatico-Guillon L. Performance of French medico-administrative databases in epidemiology of infectious diseases: a scoping review. Front Public Health 2023 May 12;11:1161550 [FREE Full text] [doi: 10.3389/fpubh.2023.1161550] [Medline: 37250067]
105. Yoshida S, Takeuchi M, Tanaka-Mizuno S, Mizuno K, Nakashima M, Fukasawa T, et al. Clinical epidemiology and pharmacoepidemiology studies with real-world databases. Proc Jpn Acad Ser B Phys Biol Sci 2022;98(10):517-528 [FREE Full text] [doi: 10.2183/pjab.98.026] [Medline: 36504194]
106. Howren A, Aviña-Zubieta JA, Puyat JH, Esdaile JM, Da Costa D, De Vera MA. Defining depression and anxiety in individuals with rheumatic diseases using administrative health databases: a systematic review. Arthritis Care Res (Hoboken) 2020 Feb 09;72(2):243-255. [doi: 10.1002/acr.24048] [Medline: 31421021]
107. Kim MK, Rouphael C, McMichael J, Welch N, Dasarathy S. Challenges in and opportunities for electronic health record-based data analysis and interpretation. Gut Liver 2024 Mar 15;18(2):201-208 [FREE Full text] [doi: 10.5009/gnl230272] [Medline: 37905424]
108. Lee S, Xu Y, D Apos Souza AG, Martin EA, Doktorchik C, Zhang Z, et al. Unlocking the potential of electronic health records for health research. Int J Popul Data Sci 2020 Jan 30;5(1):1123 [FREE Full text] [doi: 10.23889/ijpds.v5i1.1123] [Medline: 32935049]
109. Pletscher F, Mändli Lerch K, Glinz D. Willingness to share anonymised routinely collected clinical health data in Switzerland: a cross-sectional survey. Swiss Med Wkly 2022 Jun 20;152(2324):w30182 [FREE Full text] [doi: 10.4414/smw.2022.w30182] [Medline: 35752970]

110. Yuan H, Ali MS, Brouwer ES, Girman CJ, Guo JJ, Lund JL, ISPE Comparative Effectiveness Research Special Interest Group. Real-world evidence: what it is and what it can tell us according to the International Society for Pharmacoepidemiology (ISPE) Comparative Effectiveness Research (CER) Special Interest Group (SIG). Clin Pharmacol Ther 2018 Aug 07;104(2):239-241. [doi: 10.1002/cpt.1086] [Medline: 29733448]
111. Andrew A, Zhou J, Gui J, Harrison A, Shi X, Li M, et al. Pesticides applied to crops and amyotrophic lateral sclerosis risk in the U.S. Neurotoxicology 2021 Dec;87:128-135 [FREE Full text] [doi: 10.1016/j.neuro.2021.09.004] [Medline: 34562505]
112. Chazard E, Ficheur G, Caron A, Lamer A, Labreuche J, Cuggia M, et al. Secondary use of healthcare structured data: the challenge of domain-knowledge based extraction of features. Stud Health Technol Inform 2018;255:15-19. [Medline: 30306898]
113. Grath-Lone LM, Jay MA, Blackburn R, Gordon E, Zylbersztejn A, Wiljaars L, et al. What makes administrative data "research-ready"? A systematic review and thematic analysis of published literature. Int J Popul Data Sci 2022;7(1):1718 [FREE Full text] [doi: 10.23889/ijpds.v6i1.1718] [Medline: 35520099]
114. Hashimoto RE, Brodt ED, Skelly AC, Dettori JR. Administrative database studies: goldmine or goose chase? Evid Based Spine Care J 2014 Oct 24;5(2):74-76 [FREE Full text] [doi: 10.1055/s-0034-1390027] [Medline: 25278880]
115. Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, et al. The national healthcare system claims databases in France, SNIIRAM and EGB: powerful tools for pharmacoepidemiology. Pharmacoepidemiol Drug Saf 2017 Aug 24;26(8):954-962. [doi: 10.1002/pds.4233] [Medline: 28544284]
116. Al-Sahab B, Leviton A, Loddenkemper T, Paneth N, Zhang B. Biases in electronic health records data for generating real-world evidence: an overview. J Healthc Inform Res 2024 Mar 14;8(1):121-139. [doi: 10.1007/s41666-023-00153-2] [Medline: 38273982]
117. Bastarache L, Brown JS, Cimino JJ, Dorr DA, Embi PJ, Payne PR, et al. Developing real-world evidence from real-world data: transforming raw data into analytical datasets. Learn Health Syst 2022 Jan 14;6(1):e10293 [FREE Full text] [doi: 10.1002/lrh2.10293] [Medline: 35036557]
118. Fang J, Kapral MK, Richards J, Robertson A, Stamplecoski M, Silver FL. The Registry of Canadian Stroke Network : an evolving methodology. Acta Neurol Taiwan 2011 Jun;20(2):77-84. [Medline: 21739386]
119. Fillinger S, de la Garza L, Peltzer A, Kohlbacher O, Nahnsen S. Challenges of big data integration in the life sciences. Anal Bioanal Chem 2019 Oct 28;411(26):6791-6800. [doi: 10.1007/s00216-019-02074-9] [Medline: 31463515]
120. Fonteneau L, Le Meur N, Cohen-Akenine A, Pessel C, Brouard C, Delon F, et al. [The use of administrative health databases in infectious disease epidemiology and public health]. Rev Epidemiol Sante Publique 2017 Oct;65 Suppl 4:S174-S182 [FREE Full text] [doi: 10.1016/j.respe.2017.03.131] [Medline: 28624133]
121. Silenou BC, Avalos M, Helmer C, Berr C, Pariente A, Jacqmin-Gadda H. Health administrative data enrichment using cohort information: comparative evaluation of methods by simulation and application to real data. PLoS One 2019 Jan 31;14(1):e0211118 [FREE Full text] [doi: 10.1371/journal.pone.0211118] [Medline: 30703112]
122. Wang SV, Schneeweiss S. Data checks before registering study protocols for health care database analyses. JAMA 2024 May 07;331(17):1445-1446. [doi: 10.1001/jama.2024.2988] [Medline: 38587830]
123. Yamana H, Yamasaki K, Jo T, Yatsuhashi H, Yasunaga H. A descriptive analysis of acute viral hepatitis using a database with electronic medical records and claims data. Ann Clin Epidemiol 2023;5(4):107-112 [FREE Full text] [doi: 10.37737/ace.23014] [Medline: 38504953]
124. Gianfrancesco MA, Goldstein ND. A narrative review on the validity of electronic health record-based research in epidemiology. BMC Med Res Methodol 2021 Oct 27;21(1):234 [FREE Full text] [doi: 10.1186/s12874-021-01416-5] [Medline: 34706667]
125. Denisov NS, Kamenskikh EM, Fedorova OS. Trends in population-based studies: molecular and digital epidemiology (review). Sovrem Tekhnologii Med 2022 Aug 29;14(4):60-70 [FREE Full text] [doi: 10.17691/stm2022.14.4.07] [Medline: 37179982]
126. Hamm NC, Hamad AF, Wall-Wieler E, Roos LL, Plana-Ripoll O, Lix LM. Multigenerational health research using population-based linked databases: an international review. Int J Popul Data Sci 2021 Oct 07;6(1):1686 [FREE Full text] [doi: 10.23889/ijpds.v6i1.1686] [Medline: 34734126]
127. Poulain R, Tarek MF, Beheshti R. Improving fairness in AI models on electronic health records: the case for federated learning methods. FAccT 23 (2023) 2023 Jun;2023:1599-1608 [FREE Full text] [doi: 10.1145/3593013.3594102] [Medline: 37990734]
128. Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L, et al. Value of a national administrative database to guide public decisions: from the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. Rev Epidemiol Sante Publique 2017 Oct;65 Suppl 4:S149-S167. [doi: 10.1016/j.respe.2017.05.004] [Medline: 28756037]
129. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. Annu Rev Public Health 2016;37:61-81 [FREE Full text] [doi: 10.1146/annurev-publhealth-032315-021353] [Medline: 26667605]

130. Chan JM, Carroll MW, Smyth M, Hamilton Z, Evans D, McGrail K, et al. Comparing health administrative and clinical registry data: trends in incidence and prevalence of pediatric inflammatory bowel disease in British Columbia. Clin Epidemiol 2021;13:81-90 [FREE Full text] [doi: 10.2147/CLEP.S292546] [Medline: 33603489]
131. Chowdhury T, Hemmelgarn B. Evidence-based decision-making 6: utilization of administrative databases for health services research. In: Parfrey PS, Barrett BJ, editors. Clinical Epidemiology: Practice and Methods. Cham, Switzerland: Springer; 2015:469-484.
132. Chishtie J, Sapiro N, Wiebe N, Rabatach L, Lorenzetti D, Leung AA, et al. Use of epic electronic health record system for health care research: scoping review. J Med Internet Res 2023 Dec 15;25:e51003 [FREE Full text] [doi: 10.2196/51003] [Medline: 38100185]
133. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. Annu Rev Public Health 2011;32:91-108. [doi: 10.1146/annurev-publhealth-031210-100700] [Medline: 21219160]
134. Mazzali C, Duca P. Use of administrative data in healthcare research. Intern Emerg Med 2015 Jun 25;10(4):517-524. [doi: 10.1007/s11739-015-1213-9] [Medline: 25711312]
135. McPheeters ML, Sathe NA, Jerome RN, Carnahan RM. Methods for systematic reviews of administrative database studies capturing health outcomes of interest. Vaccine 2013 Dec 30;31 Suppl 10:K2-K6. [doi: 10.1016/j.vaccine.2013.06.048] [Medline: 24331070]
136. Sato S, Yasunaga H. A review of studies using Japanese nationwide administrative claims databases. Ann Clin Epidemiol 2023;5(2):58-64 [FREE Full text] [doi: 10.37737/ace.23008] [Medline: 38505730]
137. Mues KE, Liede A, Liu J, Wetmore J, Zaha R, Bradbury BD, et al. Use of the Medicare database in epidemiologic and health services research: a valuable source of real-world evidence on the older and disabled populations in the US. Clin Epidemiol 2017;9:267-277 [FREE Full text] [doi: 10.2147/CLEP.S105613] [Medline: 28533698]
138. Bernstein CN. Large registry epidemiology in IBD. Inflamm Bowel Dis 2017 Nov;23(11):1941-1949. [doi: 10.1097/MIB.0000000000001279] [Medline: 28991858]
139. Bloom PP, Tapper EB. The use of administrative data to investigate the population burden of hepatic encephalopathy. J Clin Med 2020 Nov 10;9(11):3620 [FREE Full text] [doi: 10.3390/jcm9113620] [Medline: 33182743]
140. Hudson M, Tascilar K, Suissa S. Comparative effectiveness research with administrative health data in rheumatoid arthritis. Nat Rev Rheumatol 2016 Jun 15;12(6):358-366. [doi: 10.1038/nrrheum.2016.34] [Medline: 27080692]
141. Mitchell E, Waring T, Ahern E, O'Donovan D, O'Reilly D, Bradley DT. Predictors and consequences of homelessness in whole-population observational studies that used administrative data: a systematic review. BMC Public Health 2023 Aug 24;23(1):1610 [FREE Full text] [doi: 10.1186/s12889-023-16503-z] [Medline: 37612701]
142. Ulrich EH, So G, Zappitelli M, Chanchlani R. A review on the application and limitations of administrative health care data for the study of acute kidney injury epidemiology and outcomes in children. Front Pediatr 2021 Oct 27;9:742888 [FREE Full text] [doi: 10.3389/fped.2021.742888] [Medline: 34778133]
143. Curl CL, Spivak M, Phinney R, Montrose L. Synthetic pesticides and health in vulnerable populations: agricultural workers. Curr Environ Health Rep 2020 Mar 20;7(1):13-29 [FREE Full text] [doi: 10.1007/s40572-020-00266-5] [Medline: 31960353]
144. Demos K, Sazakli E, Jelastopulu E, Charokopos N, Ellul J, Leotsinidis M. Does farming have an effect on health status? A comparison study in west Greece. Int J Environ Res Public Health 2013 Feb 26;10(3):776-792 [FREE Full text] [doi: 10.3390/ijerph10030776] [Medline: 23442558]
145. Kachuri L, Harris MA, MacLeod JS, Tjepkema M, Peters PA, Demers PA. Cancer risks in a population-based study of 70,570 agricultural workers: results from the Canadian census health and Environment cohort (CanCHEC). BMC Cancer 2017 May 19;17(1):343 [FREE Full text] [doi: 10.1186/s12885-017-3346-x] [Medline: 28525996]
146. Ramos AK, Girdžiūtė L, Starič J, Rautianinen RH. Identifying "vulnerable agricultural populations" at risk for occupational injuries and illnesses: a European perspective. J Agromedicine 2021 Jul 27;26(3):340-345. [doi: 10.1080/1059924X.2020.1771498] [Medline: 32458756]
147. Sturm ET, Castro C, Mendez-Colmenares A, Duffy J, Burzynska AZ, Stallones L, et al. Risk factors for brain health in agricultural work: a systematic review. Int J Environ Res Public Health 2022 Mar 13;19(6):3373 [FREE Full text] [doi: 10.3390/ijerph19063373] [Medline: 35329061]
148. Guindo Y, Parent ME, Richard H, Luce D, Barul C. Expert-based assessment of chemical and physical exposures, and organizational factors, in past agricultural jobs. Environ Res 2024 Dec 15;263(Pt 3):120238 [FREE Full text] [doi: 10.1016/j.envres.2024.120238] [Medline: 39461702]
149. Komarek AM, De Pinto A, Smith VH. A review of types of risks in agriculture: what we know and what we need to know. Agric Syst 2020 Feb;178:102738. [doi: 10.1016/j.agsy.2019.102738]
150. Leon ME, Schinasi LH, Lebailly P, Beane Freeman LE, Nordby KC, Ferro G, et al. Pesticide use and risk of non-Hodgkin lymphoid malignancies in agricultural cohorts from France, Norway and the USA: a pooled analysis from the AGRICOH consortium. Int J Epidemiol 2019 Oct 01;48(5):1519-1535 [FREE Full text] [doi: 10.1093/ije/dyz017] [Medline: 30880337]
151. Pérès K, Matharan F, Allard M, Amieva H, Baldi I, Barberger-Gateau P, et al. Health and aging in elderly farmers: the AMI cohort. BMC Public Health 2012 Jul 27;12(1):558 [FREE Full text] [doi: 10.1186/1471-2458-12-558] [Medline: 22839764]

152. Petit P, Chamot S, Al-Salameh A, Cancé C, Desailloud R, Bonneterre V. Farming activity and risk of treated thyroid disorders: insights from the TRACTOR project, a nationwide cohort study. Environ Res 2024 May 15;249:118458 [FREE Full text] [doi: 10.1016/j.envres.2024.118458] [Medline: 38365059]
153. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
154. The Joanna Briggs Institute Reviewers' Manual 2015 methodology for JBI scoping reviews. The Joanna Briggs Institute. 2015. URL: https://reben.com.br/revista/wp-content/uploads/2020/10/Scoping.pdf [accessed 2024-04-03]
155. Beydon M, Pinto S, De Rycke Y, Fautrel B, Mariette X, Seror R, et al. Risk of cancer for patients with rheumatoid arthritis versus general population: a national claims database cohort study. Lancet Reg Health Eur 2023 Dec;35:100768 [FREE Full text] [doi: 10.1016/j.lanepe.2023.100768] [Medline: 37954220]
156. Moysan M. Des archives à la base de données : valeur(s) des données en épidémiologie. Hist Méd Santé 2022 Dec 15(22):151-164 [FREE Full text] [doi: 10.4000/hms.6274]
157. Syed R, Eden R, Makasi T, Chukwudi I, Mamudu A, Kamalpour M, et al. Digital health data quality issues: systematic review. J Med Internet Res 2023 Mar 31;25:e42615 [FREE Full text] [doi: 10.2196/42615] [Medline: 37000497]
158. Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F. An approach for detecting, quantifying, and visualizing the evolution of a research field: a practical application to the Fuzzy Sets Theory field. J Informetr 2011 Jan;5(1):146-166. [doi: 10.1016/j.joi.2010.10.002]
159. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: an overview and guidelines. J Bus Res 2021 Sep;133:285-296. [doi: 10.1016/j.jbusres.2021.04.070]
160. Moral-Muñoz JA, Herrera-Viedma E, Santisteban-Espejo A, Cobo MJ. Software tools for conducting bibliometric analysis in science: an up-to-date review. Prof Inf J 2020 Jan 19;29(1) [FREE Full text] [doi: 10.3145/epi.2020.ene.03]
161. Bienert IR, Oliveira RC, Andrade PB, Caramori CA. Bibliometric indexes, databases and impact factors in cardiology. Rev Bras Cir Cardiovasc 2015;30(2):254-259 [FREE Full text] [doi: 10.5935/1678-9741.20150019] [Medline: 26107458]
162. Bihari A, Tripathi S, Deepak A. A review on h-index and its alternative indices. J Inf Sci 2021 May 31;49(3):624-665. [doi: 10.1177/01655515211014478]
163. Shao Y, Chien TW, Jang FL. The use of radar plots with the Yk-index to identify which authors contributed the most to the journal of Medicine in 2020 and 2021: a bibliometric analysis. Medicine (Baltimore) 2022 Nov 11;101(45):e31033 [FREE Full text] [doi: 10.1097/MD.0000000000031033] [Medline: 36397440]
164. Kumar S, Kumar S. Trends of collaborative research in journals of oilseeds research (India), 1993-2004. Indian J Agric Library Inf Services 2008;24:90.
165. Wu T, Duan Y, Zhang T, Tian W, Liu H, Deng Y. Research trends in the application of artificial intelligence in oncology: a bibliometric and network visualization study. Front Biosci (Landmark Ed) 2022 Aug 31;27(9):254 [FREE Full text] [doi: 10.31083/j.fbl2709254] [Medline: 36224012]
166. Wang B, Zhao P. Worldwide research productivity in the field of back pain: a bibliometric analysis. Medicine (Baltimore) 2018 Oct;97(40):e11566 [FREE Full text] [doi: 10.1097/MD.0000000000011566] [Medline: 30290588]
167. Aria M, Cuccurullo C, D'Aniello L, Misuraca M, Spano M. Thematic analysis as a new culturomic tool: the social media coverage on COVID-19 pandemic in Italy. Sustainability 2022 Mar 20;14(6):3643. [doi: 10.3390/su14063643]
168. Petit P, Leroyer A, Chamot S, Fumery M, Bonneterre V. Farming activities and risk of inflammatory bowel disease: a French nationwide population-based cohort study. J Crohns Colitis 2024 Sep 03;18(9):1415-1429. [doi: 10.1093/ecco-jcc/jjae050] [Medline: 38605515]
169. Montazeri A, Mohammadi S, M Hesari P, Ghaemi M, Riazi H, Sheikhi-Mobarakeh Z. Preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO): a minimum requirements. Syst Rev 2023 Dec 15;12(1):239 [FREE Full text] [doi: 10.1186/s13643-023-02410-2] [Medline: 38102710]
170. Aria M, Cuccurullo C. bibliometrix : an R-tool for comprehensive science mapping analysis. J Informetr 2017 Nov;11(4):959-975. [doi: 10.1016/j.joi.2017.08.007]
171. Spitzer WO, Hill GB, Chambers LW, Helliwell BE, Murphy HB. The occupation of fishing as a risk factor in cancer of the lip. N Engl J Med 1975 Aug 28;293(9):419-424. [doi: 10.1056/nejm197508282930903]
172. Alavanja MC, Samanic C, Dosemeci M, Lubin J, Tarone R, Lynch CF, et al. Use of agricultural pesticides and prostate cancer risk in the Agricultural Health Study cohort. Am J Epidemiol 2003 May 01;157(9):800-814. [doi: 10.1093/aje/kwg040] [Medline: 12727674]
173. Alavanja MC, Dosemeci M, Samanic C, Lubin J, Lynch CF, Knott C, et al. Pesticides and lung cancer risk in the agricultural health study cohort. Am J Epidemiol 2004 Nov 01;160(9):876-885. [doi: 10.1093/aje/kwh290] [Medline: 15496540]
174. Andersen A, Barlow L, Engeland A, Kjaerheim K, Lynge E, Pukkala E. Work-related cancer in the Nordic countries. Scand J Work Environ Health 1999;25 Suppl 2:1-116 [FREE Full text] [Medline: 10507118]
175. Andreotti G, Koutros S, Hofmann JN, Sandler DP, Lubin JH, Lynch CF, et al. Glyphosate use and cancer incidence in the agricultural health study. J Natl Cancer Inst 2018 May 01;110(5):509-516 [FREE Full text] [doi: 10.1093/jnci/djx233] [Medline: 29136183]

176. Elbaz A, Clavel J, Rathouz PJ, Moisan F, Galanaud J, Delemotte B, et al. Professional exposure to pesticides and Parkinson disease. Ann Neurol 2009 Oct;66(4):494-504. [doi: 10.1002/ana.21717] [Medline: 19847896]
177. Engel LS, Hill DA, Hoppin JA, Lubin JH, Lynch CF, Pierce J, et al. Pesticide use and breast cancer risk among farmers' wives in the agricultural health study. Am J Epidemiol 2005 Jan 15;161(2):121-135. [doi: 10.1093/aje/kwi022] [Medline: 15632262]
178. Eriksson M, Karlsson M. Occupational and other environmental factors and multiple myeloma: a population based case-control study. Br J Ind Med 1992 Feb 01;49(2):95-103 [FREE Full text] [doi: 10.1136/oem.49.2.95] [Medline: 1536825]
179. Hemminki K, Saloniemi I, Salonen T, Partanen T, Vainio H. Childhood cancer and parental occupation in Finland. J Epidemiol Community Health 1981 Mar 01;35(1):11-15 [FREE Full text] [doi: 10.1136/jech.35.1.11] [Medline: 7264527]
180. Karjalainen A, Kurppa K, Martikainen R, Klaukka T, Karjalainen J. Work is related to a substantial portion of adult-onset asthma incidence in the Finnish population. Am J Respir Crit Care Med 2001 Aug 15;164(4):565-568. [doi: 10.1164/ajrccm.164.4.2012146] [Medline: 11520716]
181. Karjalainen A, Martikainen R, Klaukka T, Saarinen K, Uitti J. Risk of asthma among Finnish patients with occupational rhinitis. Chest 2003 Jan;123(1):283-288. [doi: 10.1378/chest.123.1.283] [Medline: 12527633]
182. Kristensen P, Andersen A, Irgens LM, Laake P, Bye AS. Incidence and risk factors of cancer among men and women in Norwegian agriculture. Scand J Work Environ Health 1996 Feb;22(1):14-26 [FREE Full text] [doi: 10.5271/sjweh.104] [Medline: 8685669]
183. Lee WJ, Sandler DP, Blair A, Samanic C, Cross AJ, Alavanja MC. Pesticide use and colorectal cancer risk in the Agricultural Health Study. Int J Cancer 2007 Jul 15;121(2):339-346 [FREE Full text] [doi: 10.1002/ijc.22635] [Medline: 17390374]
184. Levecque C, Elbaz A, Clavel J, Richard F, Vidal JS, Amouyel P, et al. Association between Parkinson's disease and polymorphisms in the nNOS and iNOS genes in a community-based case-control study. Hum Mol Genet 2003 Jan 01;12(1):79-86. [doi: 10.1093/hmg/ddg009] [Medline: 12490535]
185. Mills PK, Yang R. Prostate cancer risk in California farm workers. J Occup Environ Med 2003 Mar;45(3):249-258. [doi: 10.1097/01.jom.0000058339.05741.0c] [Medline: 12661182]
186. Morrison H, Savitz D, Semenciw R, Hulka B, Mao Y, Morison D, et al. Farming and prostate cancer mortality. Am J Epidemiol 1993 Feb 01;137(3):270-280. [doi: 10.1093/oxfordjournals.aje.a116674] [Medline: 8452135]
187. Orton SM, Wald L, Confavreux C, Vukusic S, Krohn J, Ramagopalan S, et al. Association of UV radiation with multiple sclerosis prevalence and sex ratio in France. Neurology 2011 Feb;76(5):425-431. [doi: 10.1212/wnl.0b013e31820a0a9f]
188. Pouchieu C, Piel C, Carles C, Gruber A, Helmer C, Tual S, et al. Pesticide use in agriculture and Parkinson's disease in the AGRICAN cohort study. Int J Epidemiol 2018 Feb 01;47(1):299-310. [doi: 10.1093/ije/dyx225] [Medline: 29136149]
189. Pukkala E, Martinsen JI, Lynge E, Gunnarsdottir HK, Sparén P, Tryggvadottir L, et al. Occupation and cancer - follow-up of 15 million people in five Nordic countries. Acta Oncol 2009 Jul 21;48(5):646-790. [doi: 10.1080/02841860902913546] [Medline: 19925375]
190. Reif J, Pearce N, Fraser J. Cancer risks in New Zealand farmers. Int J Epidemiol 1989 Dec;18(4):768-774. [doi: 10.1093/ije/18.4.768] [Medline: 2621012]
191. Stokes L, Stark A, Marshall E, Narang A. Neurotoxicity among pesticide applicators exposed to organophosphates. Occup Environ Med 1995 Oct 01;52(10):648-653 [FREE Full text] [doi: 10.1136/oem.52.10.648] [Medline: 7489054]
192. Vingård E, Alfredsson L, Goldie I, Hogstedt C. Occupation and osteoarthrosis of the hip and knee: a register-based cohort study. Int J Epidemiol 1991 Dec;20(4):1025-1031. [doi: 10.1093/ije/20.4.1025] [Medline: 1800399]
193. Vukusic S, Van Bockstael V, Gosselin S, Confavreux C. Regional variations in the prevalence of multiple sclerosis in French farmers. J Neurol Neurosurg Psychiatry 2007 Jul 18;78(7):707-709 [FREE Full text] [doi: 10.1136/jnnp.2006.101196] [Medline: 17299020]
194. Weichenthal S, Villeneuve PJ, Burnett RT, van Donkelaar A, Martin RV, Jones RR, et al. Long-term exposure to fine particulate matter: association with nonaccidental and cardiovascular mortality in the agricultural health study cohort. Environ Health Perspect 2014 Jun;122(6):609-615 [FREE Full text] [doi: 10.1289/ehp.1307277] [Medline: 24633320]
195. Wigle DT, Semenciw RM, Wilkins K, Riedel D, Ritter L, Morrison HI, et al. Mortality study of Canadian male farm operators: non-Hodgkin's lymphoma mortality and agricultural practices in Saskatchewan. J Natl Cancer Inst 1990 Apr 04;82(7):575-582. [doi: 10.1093/jnci/82.7.575] [Medline: 2313734]
196. Wiklund K, Dich J, Holm LE, Eklund G. Risk of cancer in pesticide applicators in Swedish agriculture. Br J Ind Med 1989 Nov 01;46(11):809-814 [FREE Full text] [doi: 10.1136/oem.46.11.809] [Medline: 2590647]
197. Horwitz H, Ahlgren B, Nærum E. Effect of occupation on risk of developing MS: an insurance cohort study. BMJ Open 2013 Jun 20;3(6):e002894 [FREE Full text] [doi: 10.1136/bmjopen-2013-002894] [Medline: 23794592]
198. Karttunen JP, Rautiainen RH, Quendler E. Gender division of farm work and occupational injuries. J Agric Saf Health 2019 Jul 16;25(3):117-127. [doi: 10.13031/jash.13177] [Medline: 32429612]
199. Keller JE, Howe HL. Risk factors for lung cancer among nonsmoking Illinois residents. Environ Res 1993 Jan;60(1):1-11. [doi: 10.1006/enrs.1993.1001] [Medline: 8432260]
200. Marant Micallef C, Shield KD, Vignat J, Baldi I, Charbotel B, Fervers B, et al. Cancers in France in 2015 attributable to occupational exposures. Int J Hyg Environ Health 2019 Jan;222(1):22-29. [doi: 10.1016/j.ijheh.2018.07.015] [Medline: 30174219]

201. Moisan F, Spinosi J, Dupupet JL, Delabre L, Mazurie J, Goldberg M, et al. The relation between type of farming and prevalence of Parkinson's disease among agricultural workers in five French districts. Mov Disord 2011 Feb 01;26(2):271-279 [FREE Full text] [doi: 10.1002/mds.23370] [Medline: 21412834]
202. Toch-Marquardt M, Menvielle G, Eikemo TA, Kulhánová I, Kulik MC, Bopp M, Euro-GBD-SE consortium. Occupational class inequalities in all-cause and cause-specific mortality among middle-aged men in 14 European populations during the early 2000s. PLoS One 2014 Sep 30;9(9):e108072 [FREE Full text] [doi: 10.1371/journal.pone.0108072] [Medline: 25268702]
203. Tüchsen F, Andersen O, Costa G, Filakti H, Marmot MG. Occupation and ischemic heart disease in the European Community: a comparative study of occupations at potential high risk. Am J Ind Med 1996 Oct;30(4):407-414. [doi: 10.1002/(SICI)1097-0274(199610)30:4<407::AID-AJIM5>3.0.CO;2-S] [Medline: 8892545]
204. Wang CC, Lin GL, Lin YJ, Chen WL, Wu WT. Occupational health surveillance and detection of emerging occupational diseases among Taiwan farmers, through analysis of national-based farmers' and medico-administrative databases. Am J Ind Med 2023 Jan 17;66(1):85-93. [doi: 10.1002/ajim.23443] [Medline: 36398405]
205. Langholz PL, Sommerseth HL. Old-age mortality and social class in northern Norway in the first half of the twentieth century. Scand Econ Hist Rev 2023 Sep 13;72(2):153-170. [doi: 10.1080/03585522.2023.2255599]
206. Das AV, Bagga B, Joseph J. Clinical profile and demographic distribution of keratitis: an electronic medical record-driven data analytics from an eye care network in India. Ocul Immunol Inflamm 2024 Jul 21;32(5):722-726. [doi: 10.1080/09273948.2023.2199331] [Medline: 37084281]
207. Lemarchand C, Tual S, Levêque-Morlais N, Perrier S, Belot A, Velten M, AGRICAN group. Cancer incidence in the AGRICAN cohort study (2005-2011). Cancer Epidemiol 2017 Aug;49:175-185 [FREE Full text] [doi: 10.1016/j.canep.2017.06.003] [Medline: 28709062]
208. Scott E, Hirabayashi L, Levenstein A, Krupa N, Jenkins P. The development of a machine learning algorithm to identify occupational injuries in agriculture using pre-hospital care reports. Health Inf Sci Syst 2021 Dec 29;9(1):31 [FREE Full text] [doi: 10.1007/s13755-021-00161-9] [Medline: 34422257]
209. Wilke RA, Qamar M, Lupu RA, Gu S, Zhao J. Chronic kidney disease in agricultural communities. Am J Med 2019 Oct;132(10):e727-e732 [FREE Full text] [doi: 10.1016/j.amjmed.2019.03.036] [Medline: 30998912]
210. Engel LS, Werder E, Satagopan J, Blair A, Hoppin JA, Koutros S, et al. Insecticide use and breast cancer risk among farmers' wives in the agricultural health study. Environ Health Perspect 2017 Sep 06;125(9):097002 [FREE Full text] [doi: 10.1289/EHP1295] [Medline: 28934092]
211. Louis LM, Lerro CC, Friesen MC, Andreotti G, Koutros S, Sandler DP, et al. A prospective study of cancer risk among agricultural health study farm spouses associated with personal use of organochlorine insecticides. Environ Health 2017 Sep 06;16(1):95 [FREE Full text] [doi: 10.1186/s12940-017-0298-1] [Medline: 28874165]
212. Elbaek Pedersen J, Hansen J. Risk of breast cancer in daughters of agricultural workers in Denmark. Environ Res 2024 Jan 01;240(Pt 1):117374. [doi: 10.1016/j.envres.2023.117374] [Medline: 37866542]
213. Kasten-Arias C, Hodes T, Marino M, Kaufmann J, Lucas JA, Estela Vasquez Guzman C, et al. Healthcare utilization for asthma exacerbation among children of migrant and seasonal farmworkers. Prev Med Rep 2024 Feb;38:102598 [FREE Full text] [doi: 10.1016/j.pmedr.2024.102598] [Medline: 38283959]
214. Mengist Dessie A, Fenta Feleke S, Getaye Workie S, Getinet Abebe T, Mossu Chanie Y, Kassa Yalew A. Prevalence of skin disease and its associated factors among primary schoolchildren: a cross-sectional study from a Northern Ethiopian town. Clin Cosmet Investig Dermatol 2022;15:791-801 [FREE Full text] [doi: 10.2147/CCID.S361051] [Medline: 35521561]
215. Mills PK, Zahm SH. Organophosphate pesticide residues in urine of farmworkers and their children in Fresno County, California. Am J Ind Med 2001;40(5):571-577. [doi: 10.1002/ajim.10007.abs]
216. Andrew N, Beare R, Ravipati T, Parker E, Snowdon D, Naude K, et al. Developing a linked electronic health record derived data platform to support research into healthy ageing. Int J Popul Data Sci 2023 Jun 12;8(1):2129 [FREE Full text] [doi: 10.23889/ijpds.v8i1.2129] [Medline: 37670961]
217. Contiero P, Borgini A, Bertoldi M, Abita A, Cuffari G, Tomao P, et al. An epidemiological study to investigate links between atmospheric pollution from farming and SARS-CoV-2 mortality. Int J Environ Res Public Health 2022 Apr 12;19(8):4637 [FREE Full text] [doi: 10.3390/ijerph19084637] [Medline: 35457503]
218. Lotterman A, Baliatsas C, de Rooij MM, Huss A, Jacobs J, Dückers M, et al. Increased risk of pneumonia amongst residents living near goat farms in different livestock-dense regions in the Netherlands. PLoS One 2023 Jul 5;18(7):e0286972 [FREE Full text] [doi: 10.1371/journal.pone.0286972] [Medline: 37405987]
219. Nerich V, Monnet E, Weill A, Vallier N, Vanbockstael V, Auleley G, et al. Fine-scale geographic variations of inflammatory bowel disease in France: correlation with socioeconomic and house equipment variables. Inflamm Bowel Dis 2010 May;16(5):813-821. [doi: 10.1002/ibd.21122] [Medline: 19902547]
220. Nilles JD, Lim D, Boyer MP, Wilson BD, Betar RA, Showalter HA, et al. The occurrence of bone and joint cancers and their association with rural living and radon exposure in Iowa. Environ Geochem Health 2023 Mar 05;45(3):925-940 [FREE Full text] [doi: 10.1007/s10653-022-01261-5] [Medline: 35381949]

221. Yzermans CJ, Moleman YP, Spreeuwenberg P, Nielen MM, Dückers ML, Smit LA, et al. Risk of pneumonia in the vicinity of goat farms: a comparative assessment of temporal variation based on longitudinal health data. Pneumonia (Nathan) 2023 Sep 05;15(1):13 [FREE Full text] [doi: 10.1186/s41479-023-00115-7] [Medline: 37667350]
222. Petit P, Bosson-Rieutort D, Maugard C, Gondard E, Ozenfant D, Joubert N, et al. The TRACTOR project: TRACking and MoniToring occupational risks in agriculture using French insurance health data (MSA). Ann Work Expo Health 2022 Mar 15;66(3):402-411. [doi: 10.1093/annweh/wxab083] [Medline: 34562080]
223. Mills PK, Kwong S. Cancer incidence in the United Farmworkers of America (UFW), 1987–1997. Am J Ind Med 2001;40(5):596-603. [doi: 10.1002/ajim.1125.abs]
224. Guillien A, Laurent L, Soumagne T, Puyraveau M, Laplante J, Andujar P, et al. Anxiety and depression among dairy farmers: the impact of COPD. Int J Chron Obstruct Pulmon Dis 2018;13:1-9 [FREE Full text] [doi: 10.2147/COPD.S143883] [Medline: 29296078]
225. A consortium of agricultural cohort studies. AGRICOH. URL: https://agricoh.iarc.fr/ [accessed 2024-05-17]
226. Jouneau S, Pinault M, Gouyet T, Brinchault G, Guillot S, Viel J, GERPRA. [AIRBAg study: preliminary results after one year of screening for COPD in dairy farmers]. Rev Mal Respir 2015 Mar;32(3):240-248. [doi: 10.1016/j.rmr.2014.03.014] [Medline: 25847201]
227. Soto S, Yoder AM, Nuño T, Aceves B, Sepulveda R, Rosales CB. Health conditions among farmworkers in the southwest: an analysis of the National Agricultural Workers survey. Front Public Health 2022 Nov 3;10:962085 [FREE Full text] [doi: 10.3389/fpubh.2022.962085] [Medline: 36407981]
228. Contaxis N, Clark J, Dellureficio A, Gonzales S, Mannheimer S, Oxley PR, et al. Ten simple rules for improving research data discovery. PLoS Comput Biol 2022 Feb 10;18(2):e1009768 [FREE Full text] [doi: 10.1371/journal.pcbi.1009768] [Medline: 35143479]
229. Vesteghem C, Brøndum RF, Sønderkær M, Sommer M, Schmitz A, Bødker JS, et al. Implementing the FAIR Data Principles in precision oncology: review of supporting initiatives. Brief Bioinform 2020 May 21;21(3):936-945 [FREE Full text] [doi: 10.1093/bib/bbz044] [Medline: 31263868]
230. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 2016 Mar 15;3(1):160018 [FREE Full text] [doi: 10.1038/sdata.2016.18] [Medline: 26978244]
231. Roux P, Guillien A, Soumagne T, Ritter O, Laplante J, Travers C, et al. Smoking habits in French farmers: a cross-sectional study. BMC Public Health 2017 Feb 04;17(1):166 [FREE Full text] [doi: 10.1186/s12889-017-4030-4] [Medline: 28160776]
232. Audi C, Baïz N, Maesano CN, Ramousse O, Reboulleau D, Magnan A, et al. Serum cytokine levels related to exposure to volatile organic compounds and PM2.5 in dwellings and workplaces in French farmers - a mechanism to explain nonsmoking COPD. Int J Chron Obstruct Pulmon Dis 2017;12:1363-1374 [FREE Full text] [doi: 10.2147/COPD.S117866] [Medline: 28503065]
233. Baldi I, Gruber A, Rondeau V, Lebailly P, Brochard P, Fabrigoule C. Neurobehavioral effects of long-term exposure to pesticides: results from the 4-year follow-up of the PHYTONER study. Occup Environ Med 2011 Feb 22;68(2):108-115. [doi: 10.1136/oem.2009.047811] [Medline: 21097948]
234. Christensen CH, Platz EA, Andreotti G, Blair A, Hoppin JA, Koutros S, et al. Coumaphos exposure and incident cancer among male participants in the Agricultural Health Study (AHS). Environ Health Perspect 2010 Jan;118(1):92-96 [FREE Full text] [doi: 10.1289/ehp.0800446] [Medline: 20056581]
235. Achard P, Maugard C, Cancé C, Spinosi J, Ozenfant D, Maître A, et al. Medico-administrative data combined with agricultural practices data to retrospectively estimate pesticide use by agricultural workers. J Expo Sci Environ Epidemiol 2020 Jul 04;30(4):743-755. [doi: 10.1038/s41370-019-0166-x] [Medline: 31484997]
236. Hohenadel K, Harris SA, McLaughlin JR, Spinelli JJ, Pahwa P, Dosman JA, et al. Exposure to multiple pesticides and risk of non-Hodgkin lymphoma in men from six Canadian provinces. Int J Environ Res Public Health 2011 Jun 21;8(6):2320-2330 [FREE Full text] [doi: 10.3390/ijerph8062320] [Medline: 21776232]
237. Chung EY, Risi D, Holt JL, Lonergan M, Kotwal S, Yong K, et al. Retrospective study on the epidemiology of antineutrophil cytoplasmic autoantibodies-associated vasculitis in two Australian health districts. Intern Med J 2022 Apr 13;52(4):605-613. [doi: 10.1111/imj.15098] [Medline: 33040456]
238. Li X, Sundquist J, Sundquist K. Socioeconomic and occupational risk factors for rheumatoid arthritis: a nationwide study based on hospitalizations in Sweden. J Rheumatol 2008 Jun;35(6):986-991. [Medline: 18464310]
239. Apalla Z, Lallas A, Sotiriou E, Lazaridou E, Vakirlis E, Trakatelli M, et al. Farmers develop more aggressive histologic subtypes of basal cell carcinoma. Experience from a tertiary hospital in Northern Greece. J Eur Acad Dermatol Venereol 2016 Apr 17;30 Suppl 3(S3):17-20. [doi: 10.1111/jdv.13605] [Medline: 26995018]
240. Lope V, Pollán M, Gustavsson P, Plato N, Pérez-Gómez B, Aragonés N, et al. Occupation and thyroid cancer risk in Sweden. J Occup Environ Med 2005 Sep;47(9):948-957. [doi: 10.1097/01.jom.0000169564.21523.5d] [Medline: 16155480]
241. Nordby KC, Andersen A, Kristensen P. Incidence of lip cancer in the male Norwegian agricultural population. Cancer Causes Control 2004 Aug;15(6):619-626. [doi: 10.1023/b:caco.0000036169.90864.e2]

242. Kristensen P, Andersen A, Irgens LM. Hormone-dependent cancer and adverse reproductive outcomes in farmers' families--effects of climatic conditions favoring fungal growth in grain. Scand J Work Environ Health 2000 Aug;26(4):331-337 [FREE Full text] [doi: 10.5271/sjweh.550] [Medline: 10994799]
243. Døssing M, Petersen KT, Vyberg M, Olsen JH. Liver cancer among employees in Denmark. Am J Ind Med 1997 Sep;32(3):248-254. [doi: 10.1002/(sici)1097-0274(199709)32:3<248::aid-ajim10>3.0.co;2-v] [Medline: 9219654]
244. Klingelschmidt J, Chastang JF, Khireddine-Medouni I, Chérié-Challine L, Niedhammer I. [Occupational factors associated with suicide among French employees from the special agricultural social security scheme (MSA) working between 2007 and 2013]. Rev Epidemiol Sante Publique 2020 Feb;68(1):1-8. [doi: 10.1016/j.respe.2019.09.008] [Medline: 31843361]
245. Manninen P, Heliövaara M, Riihimäki H, Mäkelä P. Does psychological distress predict disability? Int J Epidemiol 1997 Oct;26(5):1063-1070. [doi: 10.1093/ije/26.5.1063] [Medline: 9363529]
246. Zöller B, Li X, Sundquist J, Sundquist K. Socioeconomic and occupational risk factors for venous thromboembolism in Sweden: a nationwide epidemiological study. Thromb Res 2012 May;129(5):577-582. [doi: 10.1016/j.thromres.2011.07.050] [Medline: 21868069]
247. Lucyk K, Lu M, Sajobi T, Quan H. Administrative health data in Canada: lessons from history. BMC Med Inform Decis Mak 2015 Aug 19;15(1):69 [FREE Full text] [doi: 10.1186/s12911-015-0196-9] [Medline: 26286712]
248. Kab S, Spinosi J, Chaperon L, Dugravot A, Singh-Manoux A, Moisan F, et al. Agricultural activities and the incidence of Parkinson's disease in the general French population. Eur J Epidemiol 2017 Mar 9;32(3):203-216. [doi: 10.1007/s10654-017-0229-z] [Medline: 28185034]
249. Perrin L, Spinosi J, Chaperon L, Kab S, Moisan F, Ebaz A. Pesticides expenditures by farming type and incidence of Parkinson disease in farmers: a French nationwide study. Environ Res 2021 Jun;197:111161. [doi: 10.1016/j.envres.2021.111161] [Medline: 33887276]
250. Carles C, Bouvier G, Lebailly P, Baldi I. Use of job-exposure matrices to estimate occupational exposure to pesticides: a review. J Expo Sci Environ Epidemiol 2017 Mar;27(2):125-140. [doi: 10.1038/jes.2016.25] [Medline: 27189257]
251. Tudi M, Li H, Li H, Wang L, Lyu J, Yang L, et al. Exposure routes and health risks associated with pesticide application. Toxics 2022 Jun 19;10(6):335 [FREE Full text] [doi: 10.3390/toxics10060335] [Medline: 35736943]
252. Ahmad MF, Ahmad FA, Alsayegh AA, Zeyaullah M, AlShahrani AM, Muzammil K, et al. Pesticides impacts on human health and the environment with their mechanisms of action and possible countermeasures. Heliyon 2024 Apr 15;10(7):e29128 [FREE Full text] [doi: 10.1016/j.heliyon.2024.e29128] [Medline: 38623208]
253. Cavalier H, Trasande L, Porta M. Exposures to pesticides and risk of cancer: evaluation of recent epidemiological evidence in humans and paths forward. Int J Cancer 2023 Mar 01;152(5):879-912 [FREE Full text] [doi: 10.1002/ijc.34300] [Medline: 36134639]
254. Cresto N, Forner-Piquer I, Baig A, Chatterjee M, Perroy J, Goracci J, et al. Pesticides at brain borders: impact on the blood-brain barrier, neuroinflammation, and neurological risk trajectories. Chemosphere 2023 May;324:138251. [doi: 10.1016/j.chemosphere.2023.138251] [Medline: 36878369]
255. Richardson JR, Fitsanakis V, Westerink RH, Kanthasamy AG. Neurotoxicity of pesticides. Acta Neuropathol 2019 Sep 13;138(3):343-362 [FREE Full text] [doi: 10.1007/s00401-019-02033-9] [Medline: 31197504]
256. Sun J, Fang R, Wang H, Xu D, Yang J, Huang X, et al. A review of environmental metabolism disrupting chemicals and effect biomarkers associating disease risks: where exposomics meets metabolomics. Environ Int 2022 Jan;158:106941 [FREE Full text] [doi: 10.1016/j.envint.2021.106941] [Medline: 34689039]
257. Blair A, Freeman LB. Epidemiologic studies in agricultural populations: observations and future directions. J Agromedicine 2009 May 07;14(2):125-131 [FREE Full text] [doi: 10.1080/10599240902779436] [Medline: 19437268]
258. Carvalho JG, Augusto HC, Ferraz R, Delerue-Matos C, Fernandes VC. Micro(nano)plastic and related chemicals: emerging contaminants in environment, food and health impacts. Toxics 2024 Oct 20;12(10):762 [FREE Full text] [doi: 10.3390/toxics12100762] [Medline: 39453182]
259. Clarke K, Manrique A, Sabo-Attwood T, Coker ES. A narrative review of occupational air pollution and respiratory health in farmworkers. Int J Environ Res Public Health 2021 Apr 13;18(8):4097 [FREE Full text] [doi: 10.3390/ijerph18084097] [Medline: 33924663]
260. Domingo NG, Balasubramanian S, Thakrar SK, Clark MA, Adams PJ, Marshall JD, et al. Air quality-related health damages of food. Proc Natl Acad Sci U S A 2021 May 18;118(20):e2013637118 [FREE Full text] [doi: 10.1073/pnas.2013637118] [Medline: 33972419]
261. Hao Y, Sun H, Zeng X, Dong G, Kronzucker HJ, Min J, et al. Smallholder vegetable farming produces more soil microplastics pollution than large-scale farming. Environ Pollut 2023 Jan 15;317:120805. [doi: 10.1016/j.envpol.2022.120805] [Medline: 36470457]
262. Ng EL, Huerta Lwanga E, Eldridge SM, Johnston P, Hu H, Geissen V, et al. An overview of microplastic and nanoplastic pollution in agroecosystems. Sci Total Environ 2018 Jun 15;627:1377-1388. [doi: 10.1016/j.scitotenv.2018.01.341] [Medline: 30857101]
263. van Kersen W, Bossers A, de Steenhuijsen Piters WA, de Rooij MM, Bonten M, Fluit AC, et al. Air pollution from livestock farms and the oropharyngeal microbiome of COPD patients and controls. Environ Int 2022 Nov;169:107497 [FREE Full text] [doi: 10.1016/j.envint.2022.107497] [Medline: 36088872]

264. Ganter M. Zoonotic risks from small ruminants. Vet Microbiol 2015 Dec 14;181(1-2):53-65. [doi: 10.1016/j.vetmic.2015.07.015] [Medline: 26275853]
265. Golomazou E, Mamedova S, Eslahi AV, Karanis P. Cryptosporidium and agriculture: a review. Sci Total Environ 2024 Mar 15;916:170057. [doi: 10.1016/j.scitotenv.2024.170057] [Medline: 38242460]
266. McDaniel CJ, Cardwell DM, Moeller Jr RB, Gray GC. Humans and cattle: a review of bovine zoonoses. Vector Borne Zoonotic Dis 2014 Jan;14(1):1-19 [FREE Full text] [doi: 10.1089/vbz.2012.1164] [Medline: 24341911]
267. Arora K, Xu L, Bhagianadh D. Dementia and cognitive decline in older adulthood: are agricultural workers at greater risk? J Gerontol B Psychol Sci Soc Sci 2021 Sep 13;76(8):1629-1643 [FREE Full text] [doi: 10.1093/geronb/gbab005] [Medline: 33406265]
268. Douphrate DI, Rosecrance JC, Reynolds SJ, Stallones L, Gilkey DP. Tractor-related injuries: an analysis of workers' compensation data. J Agromedicine 2009 May 07;14(2):198-205. [doi: 10.1080/10599240902773215] [Medline: 19437278]
269. Dos Anjos Magri C, Garófallo Garcia R, Binotto E, Duarte da Silva Lima N, de Alencar Nääs I, Sgavioli S, et al. Occupational risk factors in health of broiler-farm workers: a systematic review. Arch Environ Occup Health 2021 Oct 15;76(8):482-493. [doi: 10.1080/19338244.2020.1832036] [Medline: 33054688]
270. Barbosa Junior M, Sokulski CC, Salvador R, Pinheiro E, de Francisco AC, Trojan F. What kills the agricultural worker? A systematic review on suicide. Rural Remote Health 2021 Aug;21(3):6067 [FREE Full text] [doi: 10.22605/RRH6067] [Medline: 34384223]
271. Garrett-Wright D, Malin C, Jones MS. Mental health in farming communities. J Psychosoc Nurs Ment Health Serv 2023 Oct;61(10):39-43. [doi: 10.3928/02793695-20230424-02] [Medline: 37134283]
272. Lunner Kolstrup C, Kallioniemi M, Lundqvist P, Kymäläinen HR, Stallones L, Brumby S. International perspectives on psychosocial working conditions, mental health, and stress of dairy farm operators. J Agromedicine 2013 Jul 03;18(3):244-255. [doi: 10.1080/1059924X.2013.796903] [Medline: 23844791]
273. Petit P, Gandon G, Dubuc M, Vuillerme N, Bonneterre V. Agricultural activities and risk of treatment for depressive disorders among the entire French agricultural workforce: the TRACTOR project, a nationwide retrospective cohort study. Lancet Reg Health Eur 2023 Aug;31:100674 [FREE Full text] [doi: 10.1016/j.lanepe.2023.100674] [Medline: 37408876]
274. Daghagh Yazd S, Wheeler SA, Zuo A. Key risk factors affecting farmers' mental health: a systematic review. Int J Environ Res Public Health 2019 Dec 02;16(23):4849 [FREE Full text] [doi: 10.3390/ijerph16234849] [Medline: 31810320]
275. Nordby K, Andersen A, Irgens LM, Kristensen P. Indicators of mancozeb exposure in relation to thyroid cancer and neural tube defects in farmers' families. Scand J Work Environ Health 2005 Apr;31(2):89-96 [FREE Full text] [doi: 10.5271/sjweh.855] [Medline: 15864902]
276. Kauppinen T, Uuksulainen S, Saalo A, Mäkinen I, Pukkala E. Use of the Finnish Information System on Occupational Exposure (FINJEM) in epidemiologic, surveillance, and other applications. Ann Occup Hyg 2014 Apr;58(3):380-396. [doi: 10.1093/annhyg/met074] [Medline: 24401793]
277. Shim YK, Mlynarek SP, van Wijngaarden E. Parental exposure to pesticides and childhood brain cancer: U.S. Atlantic coast childhood brain cancer study. Environ Health Perspect 2009 Jun;117(6):1002-1006 [FREE Full text] [doi: 10.1289/ehp.0800209] [Medline: 19590697]
278. Baldi I, Carles C, Blanc-Lapierre A, Fabbro-Peray P, Druet-Cabanac M, Boutet-Robinet E, PESTIMAT Group. A French crop-exposure matrix for use in epidemiological studies on pesticides: PESTIMAT. J Expo Sci Environ Epidemiol 2017 Jan 23;27(1):56-63. [doi: 10.1038/jes.2015.72] [Medline: 26696463]
279. Oltramare C, Mediouni Z, Shoman Y, Hopf NB, Graczyk H, Berthet A. Determinants of pesticide exposure in occupational studies: a meta-analysis. Toxics 2023 Jul 18;11(7):623 [FREE Full text] [doi: 10.3390/toxics11070623] [Medline: 37505588]
280. Burstyn I, Lavoué J, Van Tongeren M. Aggregation of exposure level and probability into a single metric in job-exposure matrices creates bias. Ann Occup Hyg 2012 Nov;56(9):1038-1050. [doi: 10.1093/annhyg/mes031] [Medline: 22986426]
281. Fadel M, Evanoff BA, Andersen JH, d'Errico A, Dale AM, Leclerc A, et al. Not just a research method: if used with caution, can job-exposure matrices be a useful tool in the practice of occupational medicine and public health? Scand J Work Environ Health 2020 Sep 01;46(5):552-553 [FREE Full text] [doi: 10.5271/sjweh.3900] [Medline: 32367143]
282. Peters S, Vienneau D, Sampri A, Turner MC, Castaño-Vinyals G, Bugge M, et al. Occupational exposure assessment tools in Europe: a comprehensive inventory overview. Ann Work Expo Health 2022 Jun 06;66(5):671-686 [FREE Full text] [doi: 10.1093/annweh/wxab110] [Medline: 34935027]
283. Petit P, Gondard E, Gandon G, Moreaud O, Sauvée M, Bonneterre V. Agricultural activities and risk of Alzheimer's disease: the TRACTOR project, a nationwide retrospective cohort study. Eur J Epidemiol 2024 Mar 10;39(3):271-287 [FREE Full text] [doi: 10.1007/s10654-023-01079-0] [Medline: 38195954]
284. List of classifications. International Agency for Research on Cancer. URL: https://monographs.iarc.who.int/list-of-classifications [accessed 2024-10-30]
285. Desai B, Sahni S, Jordan H, Sahni R, Reinbeau R, Nguyen A, et al. Risk of cardiovascular disease in male farmers over the age of 45: a review of literature. Cureus 2022 May;14(5):e24642 [FREE Full text] [doi: 10.7759/cureus.24642] [Medline: 35663675]

286. Kim H, Jung W, Jung S, Cho S, Jung I, Song H, et al. Is farming a risk occupation for cardio-cerebrovascular diseases? A scoping review on cardio-cerebrovascular disease risk in farmers. J Prev Med Public Health 2024 Aug 23;57(6):521-529 [FREE Full text] [doi: 10.3961/jpmph.24.302] [Medline: 39438011]
287. Hulkkonen S, Shiri R, Auvinen J, Miettunen J, Karppinen J, Ryhänen J. Risk factors of hospitalization for carpal tunnel syndrome among the general working population. Scand J Work Environ Health 2020 Jan 01;46(1):43-49 [FREE Full text] [doi: 10.5271/sjweh.3835] [Medline: 31168633]
288. Kim J, Youn K, Park J. Risk factors for musculoskeletal disorders in Korean farmers: survey on occupational diseases in 2020 and 2022. Healthcare (Basel) 2024 Oct 11;12(20):2026 [FREE Full text] [doi: 10.3390/healthcare12202026] [Medline: 39451441]
289. Hemminki K, Niemi ML, Saloniemi I, Vainio H, Hemminki E. Spontaneous abortions by occupation and social class in Finland. Int J Epidemiol 1980 Jun;9(2):149-153. [doi: 10.1093/ije/9.2.149] [Medline: 7409966]
290. Strand K, Wergeland E, Bjerkedal T. Fertility patterns according to occupational grouping in Norway, 1989. Scand J Soc Med 1996 Mar 01;24(1):50-54. [doi: 10.1177/140349489602400108] [Medline: 8740876]
291. Liu Y, Huang L, Hu F, Zhang X. Investigating frailty, polypharmacy, malnutrition, chronic conditions, and quality of life in older adults: large population-based study. JMIR Public Health Surveill 2024 Oct 11;10:e50617 [FREE Full text] [doi: 10.2196/50617] [Medline: 39145920]
292. Lie A, Skogstad M, Johannessen HA, Tynes T, Mehlum IS, Nordby K, et al. Occupational noise exposure and hearing: a systematic review. Int Arch Occup Environ Health 2016 Apr 7;89(3):351-372 [FREE Full text] [doi: 10.1007/s00420-015-1083-5] [Medline: 26249711]
293. Kraemer JG, Aebi S, Oppliger A, Hilty M. The Indoor-Air Microbiota of pig farms drives the composition of the pig farmers' nasal microbiota in a season-dependent and farm-specific manner. Appl Environ Microbiol 2019 May 01;85(9):e03038 [FREE Full text] [doi: 10.1128/AEM.03038-18] [Medline: 30824439]
294. Mucci N, Tommasi E, Chiarelli A, Lulli LG, Traversini V, Galea RP, et al. WORKbiota: a systematic review about the effects of occupational exposure on microbiota and workers' health. Int J Environ Res Public Health 2022 Jan 18;19(3):1043 [FREE Full text] [doi: 10.3390/ijerph19031043] [Medline: 35162072]
295. Shukla SK, Ye Z, Sandberg S, Reyes I, Fritsche TR, Keifer M. The nasal microbiota of dairy farmers is more complex than oral microbiota, reflects occupational exposure, and provides competition for staphylococci. PLoS One 2017 Aug 29;12(8):e0183898 [FREE Full text] [doi: 10.1371/journal.pone.0183898] [Medline: 28850578]
296. Sun J, Liao XP, D'Souza AW, Boolchandani M, Li S, Cheng K, et al. Environmental remodeling of human gut microbiota and antibiotic resistome in livestock farms. Nat Commun 2020 Mar 18;11(1):1427 [FREE Full text] [doi: 10.1038/s41467-020-15222-y] [Medline: 32188862]
297. Vilela C, Araújo B, Soares-Guedes C, Caridade-Silva R, Martins-Macedo J, Teixeira C, et al. From the gut to the brain: is microbiota a new paradigm in Parkinson’s disease treatment? Cells 2024 Apr 30;13(9):770 [FREE Full text] [doi: 10.3390/cells13090770] [Medline: 38727306]
298. Gois MF, Fernández-Pato A, Huss A, Gacesa R, Wijmenga C, Weersma RK, et al. Impact of occupational pesticide exposure on the human gut microbiome. Front Microbiol 2023 Aug 10;14:1223120 [FREE Full text] [doi: 10.3389/fmicb.2023.1223120] [Medline: 37637104]
299. Christen P, Schnell R. Thirty-three myths and misconceptions about population data: from data capture and processing to linkage. Int J Popul Data Sci 2023 Jan 31;8(1):2115 [FREE Full text] [doi: 10.23889/ijpds.v8i1.2115] [Medline: 37636835]
300. Dattani N, Hardelid P, Davey J, Gilbert R. Accessing electronic administrative health data for research takes time. Arch Dis Child 2013 May 28;98(5):391-392. [doi: 10.1136/archdischild-2013-303730] [Medline: 23448863]
301. Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible sources of bias in primary care electronic health record data use and reuse. J Med Internet Res 2018 May 29;20(5):e185 [FREE Full text] [doi: 10.2196/jmir.9134] [Medline: 29844010]
302. Bousquet P, Caillet P, Coeuret-Pellicer M, Goulard H, Kudjawu Y, Le Bihan C, et al. [Using cancer case identification algorithms in medico-administrative databases: literature review and first results from the REDSIAM Tumors group based on breast, colon, and lung cancer]. Rev Epidemiol Sante Publique 2017 Oct;65 Suppl 4:S236-S242. [doi: 10.1016/j.respe.2017.04.057] [Medline: 28576379]
303. Goldberg M, Carton M, Doussin A, Fagot-Campagna A, Heyndrickx E, Lemaitre M, et al. [The REDSIAM network]. Rev Epidemiol Sante Publique 2017 Oct;65 Suppl 4:S144-S148. [doi: 10.1016/j.respe.2017.06.001] [Medline: 28844426]
304. Lipscombe LL, Hwee J, Webster L, Shah BR, Booth GL, Tu K. Identifying diabetes cases from administrative data: a population-based validation study. BMC Health Serv Res 2018 May 02;18(1):316 [FREE Full text] [doi: 10.1186/s12913-018-3148-0] [Medline: 29720153]
305. Caputo M, Pecere A, Sarro A, Mele C, Ucciero A, Pagano L, et al. Incidence and prevalence of hyperthyroidism: a population-based study in the Piedmont Region, Italy. Endocrine 2020 Jul 13;69(1):107-112. [doi: 10.1007/s12020-020-02222-7] [Medline: 32056093]
306. Habib RR, Hojeij S, Elzein K. Gender in occupational health research of farmworkers: a systematic review. Am J Ind Med 2014 Dec 16;57(12):1344-1367 [FREE Full text] [doi: 10.1002/ajim.22375] [Medline: 25227724]

307. Biswas A, Harbin S, Irvin E, Johnston H, Begum M, Tiong M, et al. Sex and gender differences in occupational hazard exposures: a scoping review of the recent literature. Curr Environ Health Rep 2021 Dec 27;8(4):267-280 [FREE Full text] [doi: 10.1007/s40572-021-00330-8] [Medline: 34839446]
308. Biswas A, Tiong M, Irvin E, Zhai G, Sinkins M, Johnston H, et al. Gender and sex differences in occupation-specific infectious diseases: a systematic review. Occup Environ Med 2024 Aug 29;81(8):425-432 [FREE Full text] [doi: 10.1136/oemed-2024-109451] [Medline: 39168602]
309. Comfort N, Re DB. Sex-specific neurotoxic effects of organophosphate pesticides across the life course. Curr Environ Health Rep 2017 Dec 23;4(4):392-404 [FREE Full text] [doi: 10.1007/s40572-017-0171-y] [Medline: 29063415]
310. Wang W, Jin J, He R, Gong H. Gender differences in pesticide use knowledge, risk awareness and practices in Chinese farmers. Sci Total Environ 2017 Jul 15;590-591:22-28. [doi: 10.1016/j.scitotenv.2017.03.053] [Medline: 28285132]
311. Xu L, Xu X, Wu X, Kuang H, Xu C. Sex-dependent environmental health risk analysis of flupyradifurone. Environ Sci Technol 2022 Feb 01;56(3):1841-1853. [doi: 10.1021/acs.est.1c07726] [Medline: 35041393]
312. Ketchen Jr DJ, Roccapriore AY, Connelly BL. Using old data: when is it appropriate? J Manag 2023 Jun 04;49(8):2541-2548. [doi: 10.1177/01492063231177785]
313. Vallot C. Epigenomic tumor evolution under the spotlight: the promises of single-cell approaches. C R Biol 2022 May 11;345(1):11-16 [FREE Full text] [doi: 10.5802/crbiol.75] [Medline: 35787617]
314. Horgan D, Hajduch M, Vrana M, Soderberg J, Hughes N, Omar MI, et al. European health data space-an opportunity now to grasp the future of data-driven healthcare. Healthcare (Basel) 2022 Aug 26;10(9):1629 [FREE Full text] [doi: 10.3390/healthcare10091629] [Medline: 36141241]
315. Martani A, Geneviève LD, Wangmo T, Maurer J, Crameri K, Erard F, et al. Sensing the (digital) pulse. Future steps for improving the secondary use of data for research in Switzerland. Digit Health 2023 Apr 20;9:20552076231169826 [FREE Full text] [doi: 10.1177/20552076231169826] [Medline: 37113255]
316. Raab R, Küderle A, Zakreuskaya A, Stern AD, Klucken J, Kaissis G, et al. Federated electronic health records for the European Health Data Space. Lancet Digit Health 2023 Nov;5(11):e840-e847 [FREE Full text] [doi: 10.1016/S2589-7500(23)00156-5] [Medline: 37741765]
317. Pampel H, Weisweiler NL, Strecker D, Witt M, Vierkant P, Elger K, et al. re3data - indexing the global research data repository landscape since 2012. Sci Data 2023 Aug 29;10(1):571 [FREE Full text] [doi: 10.1038/s41597-023-02462-y] [Medline: 37644024]
318. Suhr B, Dungl J, Stocker A. Search, reuse and sharing of research data in materials science and engineering-a qualitative interview study. PLoS One 2020 Sep 15;15(9):e0239216 [FREE Full text] [doi: 10.1371/journal.pone.0239216] [Medline: 32931508]
319. Noaeen M, Doiron C, Syer J, Brook J. Advancing population health through open environmental data platforms. Curr Top Behav Neurosci 2024;68:297-323. [doi: 10.1007/7854_2024_512] [Medline: 39112811]
320. Dul J, van Raaij E, Caputo A. Advancing scientific inquiry through data reuse: necessary condition analysis with archival data. Strateg Change 2023 Nov 02;33(1):35-40. [doi: 10.1002/jsc.2562]
321. Catalogue des bases de données individuelles en santé en France. INSERM. 2020. URL: https://epidemiologie-france.aviesan.fr/ [accessed 2024-10-30]
322. Plateforme ouverte des données publiques françaises. République Française. URL: https://www.data.gouv.fr/fr/ [accessed 2024-10-30]
323. Dataset search. Google. URL: https://datasetsearch.research.google.com/ [accessed 2024-10-30]
324. Kogevinas M, Schlünssen V, Mehlum IS, Turner MC. The OMEGA-NET international inventory of occupational cohorts. Ann Work Expo Health 2020 Jul 01;64(6):565-568 [FREE Full text] [doi: 10.1093/annweh/wxaa039] [Medline: 32556221]
325. Peters S. Although a valuable method in occupational epidemiology, job-exposure -matrices are no magic fix. Scand J Work Environ Health 2020 May 01;46(3):231-234 [FREE Full text] [doi: 10.5271/sjweh.3894] [Medline: 32356897]
326. Nittas V, Puhan MA, von Wyl V. Toward a working definition of eCohort studies in health research: narrative literature review. JMIR Public Health Surveill 2021 Jan 21;7(1):e24588 [FREE Full text] [doi: 10.2196/24588] [Medline: 33475521]
327. Singh AV, Bansod G, Mahajan M, Dietrich P, Singh SP, Rav K, et al. Digital transformation in toxicology: improving communication and efficiency in risk assessment. ACS Omega 2023 Jun 20;8(24):21377-21390. [doi: 10.1021/acsomega.3c00596] [Medline: 37360489]
328. Soussi A, Zero E, Sacile R, Trinchero D, Fossa M. Smart sensors and smart data for precision agriculture: a review. Sensors (Basel) 2024 Apr 21;24(8):2647 [FREE Full text] [doi: 10.3390/s24082647] [Medline: 38676264]
329. Bukhtiyarova O, Abderrazak A, Chiu Y, Sparano S, Simard M, Sirois C. Major areas of interest of artificial intelligence research applied to health care administrative data: a scoping review. Front Pharmacol 2022 Jul 18;13:944516 [FREE Full text] [doi: 10.3389/fphar.2022.944516] [Medline: 35924057]
330. Caruana A, Bandara M, Musial K, Catchpoole D, Kennedy PJ. Machine learning for administrative health records: a systematic review of techniques and applications. Artif Intell Med 2023 Oct;144:102642. [doi: 10.1016/j.artmed.2023.102642] [Medline: 37783537]

331. Leal Neto O, Von Wyl V. Digital transformation of public health for noncommunicable diseases: narrative viewpoint of challenges and opportunities. JMIR Public Health Surveill 2024 Jan 25;10:e49575 [FREE Full text] [doi: 10.2196/49575] [Medline: 38271097]
332. Wang J, Kharrat FG, Gariépy G, Gagné C, Pelletier J, Massamba VK, et al. Predicting the population risk of suicide using routinely collected health administrative data in Quebec, Canada: model-based synthetic estimation study. JMIR Public Health Surveill 2024 Jun 28;10:e52773 [FREE Full text] [doi: 10.2196/52773] [Medline: 38941610]
333. Ben-Ari T, Boé J, Ciais P, Lecerf R, Van der Velde M, Makowski D. Causes and implications of the unforeseen 2016 extreme yield loss in the breadbasket of France. Nat Commun 2018 Apr 24;9(1):1627 [FREE Full text] [doi: 10.1038/s41467-018-04087-x] [Medline: 29691405]
334. Bhawra J, Elsahli N, Patel J. Applying digital technology to understand human experiences of climate change impacts on food security and mental health: scoping review. JMIR Public Health Surveill 2024 Jul 23;10:e54064 [FREE Full text] [doi: 10.2196/54064] [Medline: 39042453]
335. Comi M, Becot F, Bendixsen C. Automation, climate change, and the future of farm work: cross-disciplinary lessons for studying dynamic changes in agricultural health and safety. Int J Environ Res Public Health 2023 Mar 08;20(6):4778 [FREE Full text] [doi: 10.3390/ijerph20064778] [Medline: 36981685]
336. Ebi KL, Vanos J, Baldwin JW, Bell JE, Hondula DM, Errett NA, et al. Extreme weather and climate change: population health and health system implications. Annu Rev Public Health 2021 Apr 01;42:293-315 [FREE Full text] [doi: 10.1146/annurev-publhealth-012420-105026] [Medline: 33406378]
337. El Khayat M, Halwani DA, Hneiny L, Alameddine I, Haidar MA, Habib RR. Impacts of climate change and heat stress on farmworkers' health: a scoping review. Front Public Health 2022 Feb 8;10:782811 [FREE Full text] [doi: 10.3389/fpubh.2022.782811] [Medline: 35211437]
338. Shahzad A, Ullah S, Dar AA, Sardar MF, Mehmood T, Tufail MA, et al. Nexus on climate change: agriculture and possible solution to cope future climate change stresses. Environ Sci Pollut Res 2021 Jan 29;28(12):14211-14232. [doi: 10.1007/S11356-021-12649-8]
339. Gibb K, Beckman S, Vergara XP, Heinzerling A, Harrison R. Extreme heat and occupational health risks. Annu Rev Public Health 2024 May;45(1):315-335 [FREE Full text] [doi: 10.1146/annurev-publhealth-060222-034715] [Medline: 38166501]
340. Lefèvre-Arbogast S, Chaker J, Mercier F, Barouki R, Coumoul X, Miller GW, et al. Assessing the contribution of the chemical exposome to neurodegenerative disease. Nat Neurosci 2024 May 29;27(5):812-821 [FREE Full text] [doi: 10.1038/s41593-024-01627-1] [Medline: 38684891]
341. Shan L, Heusinkveld HJ, Paul KC, Hughes S, Darweesh SK, Bloem BR, et al. Towards improved screening of toxins for Parkinson's risk. NPJ Parkinsons Dis 2023 Dec 19;9(1):169 [FREE Full text] [doi: 10.1038/s41531-023-00615-9] [Medline: 38114496]
342. Abdelzaher H, Tawfik SM, Nour A, Abdelkader S, Elbalkiny ST, Abdelkader M, et al. Climate change, human health, and the exposome: utilizing OMIC technologies to navigate an era of uncertainty. Front Public Health 2022 Sep 21;10:973000 [FREE Full text] [doi: 10.3389/fpubh.2022.973000] [Medline: 36211706]
343. DeBord DG, Carreón T, Lentz TJ, Middendorf PJ, Hoover MD, Schulte PA. Use of the "exposome" in the practice of epidemiology: a primer on -Omic technologies. Am J Epidemiol 2016 Aug 15;184(4):302-314 [FREE Full text] [doi: 10.1093/aje/kwv325] [Medline: 27519539]
344. Waffenschmidt S, Knelangen M, Sieben W, Bühn S, Pieper D. Single screening versus conventional double screening for study selection in systematic reviews: a methodological systematic review. BMC Med Res Methodol 2019 Jun 28;19(1):132 [FREE Full text] [doi: 10.1186/s12874-019-0782-0] [Medline: 31253092]
345. Waffenschmidt S, Sieben W, Jakubeit T, Knelangen M, Overesch I, Bühn S, et al. Increasing the efficiency of study selection for systematic reviews using prioritization tools and a single-screening approach. Syst Rev 2023 Sep 14;12(1):161 [FREE Full text] [doi: 10.1186/s13643-023-02334-x] [Medline: 37705060]

Abbreviations

AGR: annual growth rate
AGRICAN: Agriculture and Cancer
AGRICOH: agricultural cohort
AHD: administrative health database
AHS: Agricultural Health Study
AI: artificial intelligence
AMI: Aging Multidisciplinary Investigation
ANR: French National Research Agency (Agence Nationale de la Recherche in French)
AS: association strength
BIBLIO: preliminary guideline for reporting bibliometric reviews of the biomedical literature
COPD: chronic obstructive pulmonary disease
EHR: electronic health record

EMR: electronic medical record
FAIR: findable, accessible, interoperable, and reusable
FERMA: risk factors of the rural environment and the allergic and respiratory disease
GDP: gross domestic product
JEM: job-exposure matrix
MIAI: Multidisciplinary Institute in Artificial Intelligence
MSA: Mutualité Sociale Agricole
NOCCA: Nordic Occupational Cancer Study
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
TRACTOR: Tracking and Monitoring Occupational Risks in Agriculture

---

Edited by A Mavragani; submitted 05.06.24; peer-reviewed by G Wu, F Clerc; comments to author 30.09.24; revised version received 08.10.24; accepted 07.11.24; published 09.01.25.

Please cite as:
Petit P, Vuillerme N
Leveraging Administrative Health Databases to Address Health Challenges in Farming Populations: Scoping Review and Bibliometric Analysis (1975-2024)
JMIR Public Health Surveill 2025;11:e62939
URL: https://publichealth.jmir.org/2025/1/e62939
doi:10.2196/62939
PMID:

Review

When Infodemic Meets Epidemic: Systematic Literature Review

Chaimae Asaad[1,2], MSc; Imane Khaouja[1], PhD; Mounir Ghogho[1,3], PhD; Karim Baïna[2], PhD

[1]TICLab, College of Engineering and Architecture, International University of Rabat, Salé, Morocco

[2]ENSIAS, Alqualsadi, Rabat IT Center, Mohammed V University, Rabat, Morocco

[3]University of Leeds, Leeds, United Kingdom

Corresponding Author:
Chaimae Asaad, MSc
TICLab
College of Engineering and Architecture
International University of Rabat
Shore Rocade, Rocade S, Rabat
Salé, 11103
Morocco
Phone: 212 (5)30112063
Email: chaimae.asaad@uir.ac.ma

Abstract

Background: Epidemics and outbreaks present arduous challenges, requiring both individual and communal efforts. The significant medical, emotional, and financial burden associated with epidemics creates feelings of distrust, fear, and loss of control, making vulnerable populations prone to exploitation and manipulation through misinformation, rumors, and conspiracies. The use of social media sites has increased in the last decade. As a result, significant amounts of public data can be leveraged for biosurveillance. Social media sites can also provide a platform to quickly and efficiently reach a sizable percentage of the population; therefore, they have a potential role in various aspects of epidemic mitigation.

Objective: This systematic literature review aimed to provide a methodical overview of the integration of social media in 3 epidemic-related contexts: epidemic monitoring, misinformation detection, and the relationship with mental health. The aim is to understand how social media has been used efficiently in these contexts, and which gaps need further research efforts.

Methods: Three research questions, related to epidemic monitoring, misinformation, and mental health, were conceptualized for this review. In the first PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) stage, 13,522 publications were collected from several digital libraries (PubMed, IEEE Xplore, ScienceDirect, SpringerLink, MDPI, ACM, and ACL) and gray literature sources (arXiv and ProQuest), spanning from 2010 to 2022. A total of 242 (1.79%) papers were selected for inclusion and were synthesized to identify themes, methods, epidemics studied, and social media sites used.

Results: Five main themes were identified in the literature, as follows: epidemic forecasting and surveillance, public opinion understanding, fake news identification and characterization, mental health assessment, and association of social media use with psychological outcomes. Social media data were found to be an efficient tool to gauge public response, monitor discourse, identify misleading and fake news, and estimate the mental health toll of epidemics. Findings uncovered a need for more robust applications of lessons learned from epidemic "postmortem documentation." A vast gap exists between retrospective analysis of epidemic management and result integration in prospective studies.

Conclusions: Harnessing the full potential of social media in epidemic-related tasks requires streamlining the results of epidemic forecasting, public opinion understanding, and misinformation detection, all while keeping abreast of potential mental health implications. Proactive prevention has thus become vital for epidemic curtailment and containment.

(JMIR Public Health Surveill 2025;11:e55642) doi:10.2196/55642

KEYWORDS

epidemics; social media; epidemic surveillance; misinformation; mental health

Introduction

Background

The global community braved the COVID-19 crisis, with multiple emerging variants, more than 6 million deaths, and 764 million cases being registered [1]. COVID-19 was dubbed "an individual and collective traumatic event," and has "directly or indirectly affected every individual in the world" [2]. Four years later, the world is still grappling with the emotional and socioeconomic aftermath of this crisis [3].

However, COVID-19 has not been the first crisis of its kind to affect global public health. Multiple epidemics have spanned the last 2 decades, causing varying degrees of instability and disease burden [4]. An epidemic is defined as "the occurrence in a community or region of cases of an illness, specific health-related behavior, or other health-related events clearly in excess of normal expectancy" [5]. When an epidemic "occurs worldwide or over a very wide area, crosses international boundaries, and affects a large number of people," it qualifies as a pandemic [5].

Epidemics are often linked to major feelings of uncertainty and loss. The 2014 Ebola outbreak caused rampant fear behaviors in West Africa [6]. The SARS outbreak has created a range of psychiatric conditions, including posttraumatic stress disorder, depressive disorders, and other anxiety spectrum disorders, such as panic, agoraphobia, and social phobia [7]. COVID-19 was associated with major stigma and psychological pressure, further aggravating feelings of guilt, shame, regret, sadness, self-pity, anger, internalized emotions, overwhelmed feelings, negative self-talk, unrealistic expectations, and perceived sense of failure [2]. During epidemics and outbreaks, mistrust of governments and health workers, misinformation, rumors, and conspiracies [8] present challenges to containment and can have a negative impact on mitigation efforts [9-11]. The particular vulnerability surrounding epidemics could render social media users highly suggestible and at risk for fake news acceptance and dissemination [12]. The substantial financial and medical burden imposed by outbreaks and epidemics, in addition to the substantial challenges arising in their progression and aftermath, further complicates the mental health toll they take on the affected population and on vulnerable communities [13].

The control strategies put in place in public health crises to contain the spread of infection are highly dependent on the transmission method and rate [14]. For instance, during COVID-19, various containment measures were adopted, including school closures, shut-downs of nonessential businesses, bans on mass gatherings, travel restrictions, border closures, and curfews [14]. These measures, although necessary for mitigation, can worsen emotional states, contribute to the exacerbation of preexisting socioeconomic inequalities in mental health [15], and lead to unhealthy coping mechanisms, such as problematic internet use, social media addiction, and emotional overeating [16-18].

During epidemics, social media platforms fulfill various functions ranging from informational support to emotional and peer support [19]. They are often a solemn companion offering a tool for connection, a space to grieve, and an instrument of outrage [19]. It is not surprising that the use of social media platforms massively increased during the COVID-19 pandemic [20], rendering them almost essential, ubiquitous, and a catalyst for change, for better and for worse [21].

Social media platforms offer significant amounts of data that can be leveraged for biosurveillance and syndromic surveillance of epidemics and outbreaks [22]. Biosurveillance provides early warning and situational awareness of events using diverse data streams [22]. Efforts directed at facilitating both the early detection and forecasting of disease outbreaks have been increasing in the past 2 decades [22]. Through the analysis of a variety of data sources, syndromic surveillance aims to discern individual and population health indicators before confirmed diagnoses are made [23] using trackable or exhibited behavioral patterns, symptoms, signs, or laboratory findings [23].

Understanding how social media shapes our experiences and preparedness during epidemics, and characterizing the roles it can fulfill, could allow for an improved apprehension of how to efficiently harness this resource for prevention efforts or alleviation of burden of disease [24].

Literature reviews have shown interest in understanding the roles social media fulfills during times of crisis, especially in the last decade [12,25-27]. Social media roles related to the facilitation of public health management, prevention of misinformation, and management of public health behavior and response were found to be of utmost priority [24], and social media topics related to surveillance and monitoring of public attitudes and perceptions, as well as mental health, misinformation, and fake news, were found to be the most well-developed research topics [28]. These 3 particular facets of social media's intersection with epidemics have not been approached in existing reviews; therefore, a gap remains for the research questions (RQs) proposed in this systematic literature review.

This Review

This review aimed to examine the "epidemic-social media" relationship and delineate its various aspects, as well as identify the methods used in harnessing social media in epidemics, with a particular focus on monitoring and surveillance, misinformation, and mental health. In light of the current state of global public health, it is vital to understand how a tool as influential as social media can shape the population's response in times of crisis and how it can be leveraged.

This systematic literature review outlines 3 RQs as follows: (1) How is social media harnessed for epidemic monitoring and management? (RQ1); (2) How is social media used for capturing and managing misinformation during epidemics? (RQ2); and (3) How is social media related to mental health during epidemics? (RQ3).

The remainder of this paper is organized as follows. Methods pertaining to the search strategy and extraction process are detailed in the Methods section. Results of the systematic review are synthesized in the Results section. Discussion of the major issues and practical implications as well as identified directions for future research are presented in the Discussion section. Conclusions are summarized in the Conclusion section.

Methods

Overview

This systematic review builds upon the preferred reporting items outlined in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [29] (Multimedia Appendix 1).

Proposed RQs

The RQs proposed in this systematic literature review examine the epidemic-social media relationship from different perspectives. The first RQ aims to identify potential uses of social media in the context of epidemic management and mitigation. The second RQ examines potential methods used in the context of social media misinformation as it relates to epidemics. Furthermore, the third RQ aims to discern potential aspects of the relationship between social media and public mental health during epidemics.

Search Strategy

A systematic literature search was undertaken at the beginning of June 2021. A collaborative planning and task allocation process was developed and updated at each stage of the study. The systematic search was conducted across multiple digital libraries—PubMed, IEEE Xplore, ACM Digital Library, ScienceDirect, MDPI, ACL, SpringerLink, arXiv, and ProQuest. Gray literature sources (arXiv and ProQuest) were used to complement the search and reduce publication bias as they provide a venue for authors to share studies with null or negative results that might otherwise not be disseminated.

The RQs were used as a guideline to identify search keywords. The search terms used included “social media” and “epidemics,” with variations depending on the RQ’s objectives and the database searched. For RQ1, the search results of the query (“social media” AND “epidemics”) were complemented by the results of the query (“social media” AND “epidemics” AND “monitoring” AND “tracking”). The combination of these 2 queries allowed for result-filtering without overlimiting the output. The query (“social media” AND “epidemics” AND “fake news”) was used for RQ2. A combination of the queries (“social media” AND “epidemics” AND “mental health” AND “support system”) and (“social media” AND epidemic AND “mental health” AND addiction) was used for RQ3.

These queries were adapted to each database based on its settings. All searches used the parameters *full-text* or *all metadata* in the queries. All searches covered the time range 2010 to 2022.

Table 1 details the number of publications (without duplication) retrieved for screening from each database for each RQ.

Table 1. Output of search strategy for research questions (RQs) 1, 2, and 3.

Database	RQ1	RQ2	RQ3
IEEE Xplore	259	27	54
ScienceDirect	1180	371	240
SpringerLink	2189	367	2188
ACL	90	20	121
ACM Digital Library	672	923	1795
MDPI	178	113	70
arXiv	1544	5	0
ProQuest	226	26	127
PubMed	54	149	217

Study Selection and Data Extraction Strategy

At the initial screening stage, 3 authors assessed the titles and abstracts against the inclusion criteria. Publications included after this screening stage were then retrieved in full-text version, and subsequently screened in the eligibility stage. Three of the authors read the full-text articles independently to ascertain their relevance with regard to the search terms and the research aims. All disagreements on the included articles were resolved by consensus.

To organize the screening process, Rayyan [30], a web application facilitating the collaborative review process and screening process for systematic literature reviews, was used by the authors to import all articles initially collected and screen them following a “blind on” setting, where decisions and labels of any collaborator were not visible to others. Publications with inclusion disagreements were then identified after dropping the “blind on” setting and resolved among authors.

The inclusion and exclusion criteria specified the aims of the review and were agreed upon by all authors (Textbox 1). For a publication to be selected, it needed to address the RQs and be published within the time range. The publication was excluded if it was not a journal paper, conference proceedings paper, or peer-reviewed workshop or symposium paper. Long abstracts and posters were excluded. Publications related to the HIV or tuberculosis epidemic were excluded to preserve the homogeneity of the review. Tuberculosis is a bacterial infection with a high burden of disease, especially in developing countries, while HIV is the virus responsible for AIDS [1]. Both tuberculosis and HIV or AIDS are classified as ongoing worldwide public health issues by the World Health

Organization (WHO) and the Centers for Disease Control and Prevention [1]. Given the particularities of both tuberculosis and HIV or AIDS and the high volume of literature review publications related to them [31], the authors agreed to consider both beyond the scope of this review.

Textbox 1. Inclusion and exclusion criteria for the study selection process.

Inclusion criteria for studies

- Within the scope of one of the research questions
- Published between 2010 and 2022
- Relates to an epidemic or pandemic within the last 2 decades
- Includes the use of a social media site
- Is a journal, conference, or workshop paper

Exclusion criteria for studies

- Tuberculosis, HIV, or, noninfectious diseases
- Online forums or traditional media
- Book, e-book, letter to editor, magazine, abstracts, case reports, comments, reviews, or poster

In the data extraction stage, the final list of papers was analyzed to answer the RQs and extract pertinent information. The final stage of the PRISMA guidelines [29] was considered in this phase. The following data were extracted from selected papers: authors, publication year, epidemic studied, social media site used, theme, identified method, and key findings. All the related data were extracted independently by 2 investigators. When necessary, differences were resolved by discussing, examining, and negotiating with a third investigator.

Quality Assessment

The quality of the included studies in this review was appraised using a set checklist of quality criteria. Papers that did not fulfill at least 4 out of the 5 quality criteria were excluded. The checklist was defined as follows:

1. Are the study objectives clearly defined?
2. Are the methods clearly defined and applied?
3. Are the methods applied successfully and correctly?
4. Are accuracy values and efficiency and confidence levels reported?
5. Are limitations clearly reported and adequately represented?
6. Do the contributions outweigh the limitations of the study?

The quality criteria were formulated based on our understanding of the current state of research in this field and the research gap this systematic review is attempting to fill. The papers were assessed for their ability to answer the RQs and enrich the literature while fulfilling quality standards.

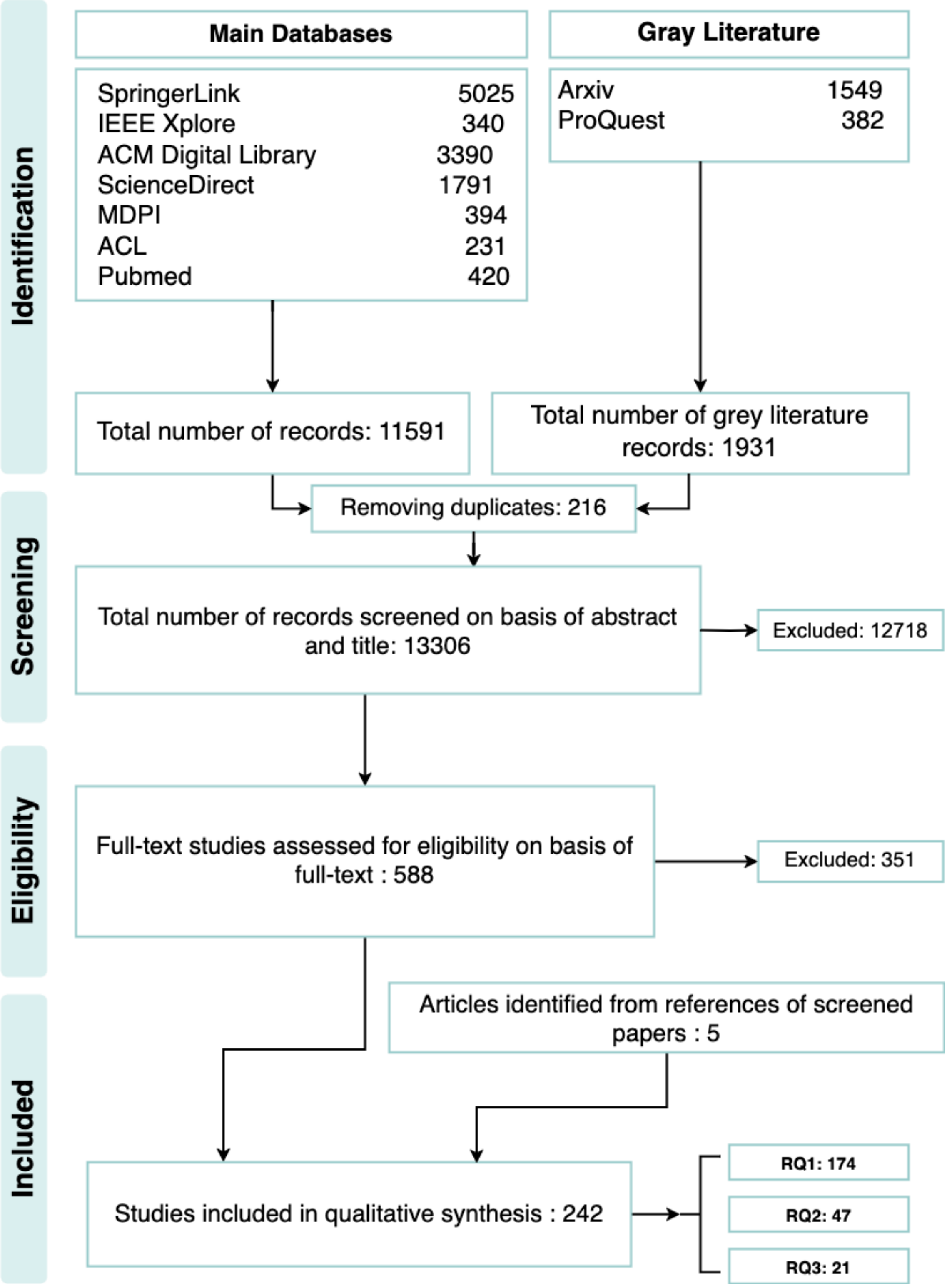
Bias was evaluated in this systematic literature review from 2 aspects. First, the risk of bias based on inclusion was limited through the use of multiple reviewers. Second, publication bias was limited by including gray literature which reports negative and null results. To enhance the quality of this review, the authors monitored the planned review tasks and ensured continuous progress monitoring. Collaborative worksheets were created to keep track of scheduled tasks and deadlines, and to note pertinent observations. Validation of the extracted data from selected papers was conducted by the authors and peer-reviewing was maintained at every stage of the systematic review process.

Results

Characteristics of the Selected Papers

The search process resulted in a total of 13,522 articles distributed over both the main and gray databases used. After the removal of duplicates, 13,306 (98.4%) titles remained. Of these, 12,718 (95.58%) studies were excluded after the title and abstract screening, as they did not fulfill the inclusion criteria. A flow diagram of the results of literature collection, screening, eligibility, and inclusion is presented in Figure 1.

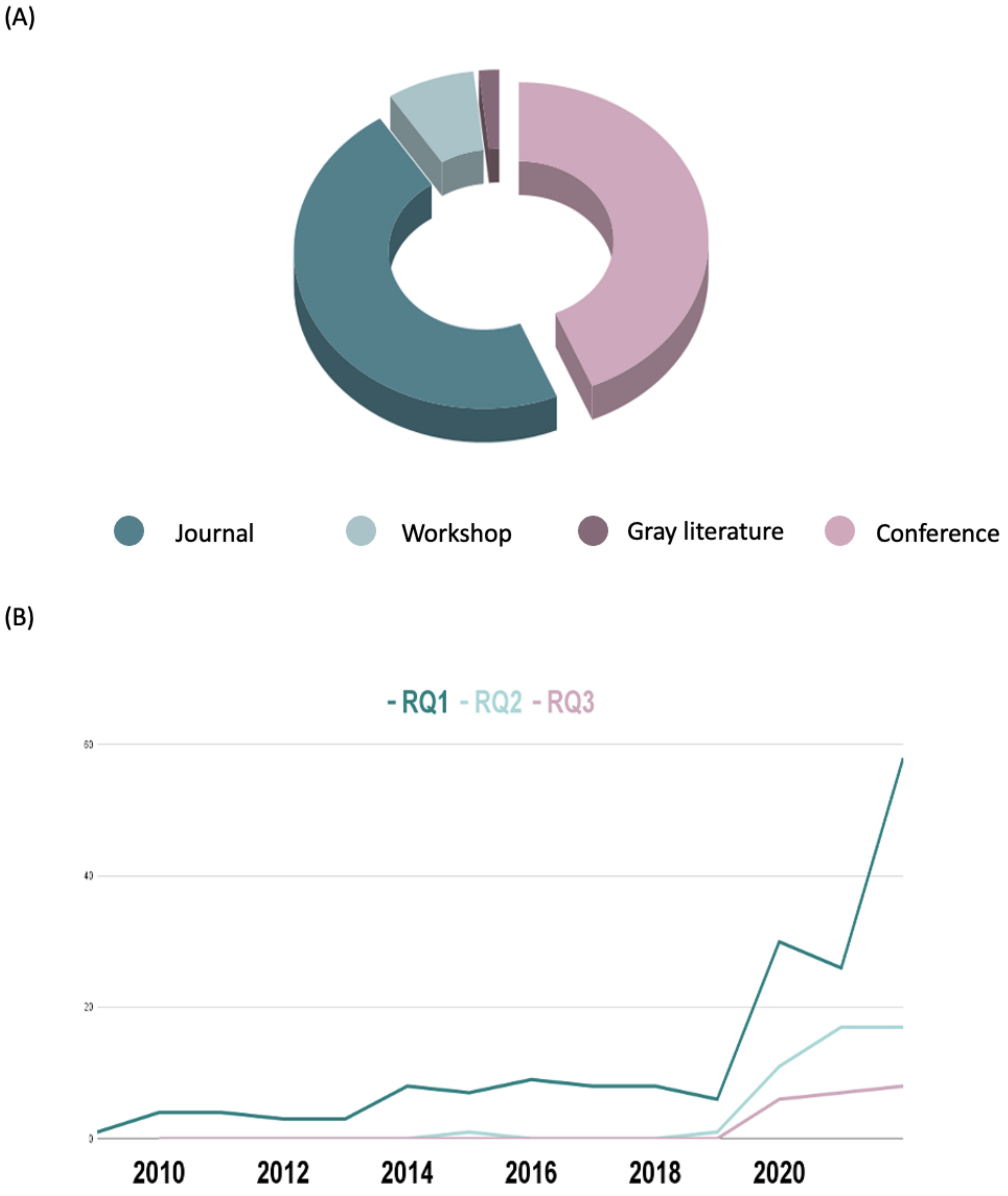
Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for the selection of articles of the literature reviewed. RQ: research question.



Of the 588 studies that were full-text screened, 351 (59.7%) did not meet the inclusion criteria. 5 (1.4%) papers were identified from reference lists of included papers. A total of 242 (67.9%) studies were selected for inclusion in this review as summarized in subsection Answers to RQs.

The papers included in the review were distributed as follows: 47.1% (114/242) were journal papers, 43.8% (106/242) were publications of conference proceedings, 7.4% (18/242) were workshop and symposium publications, while 1.7% (4/242) were gray literature (Figure 2A).

Figure 2. Distribution of selected papers by (A) type and (B) year. RQ: research question.

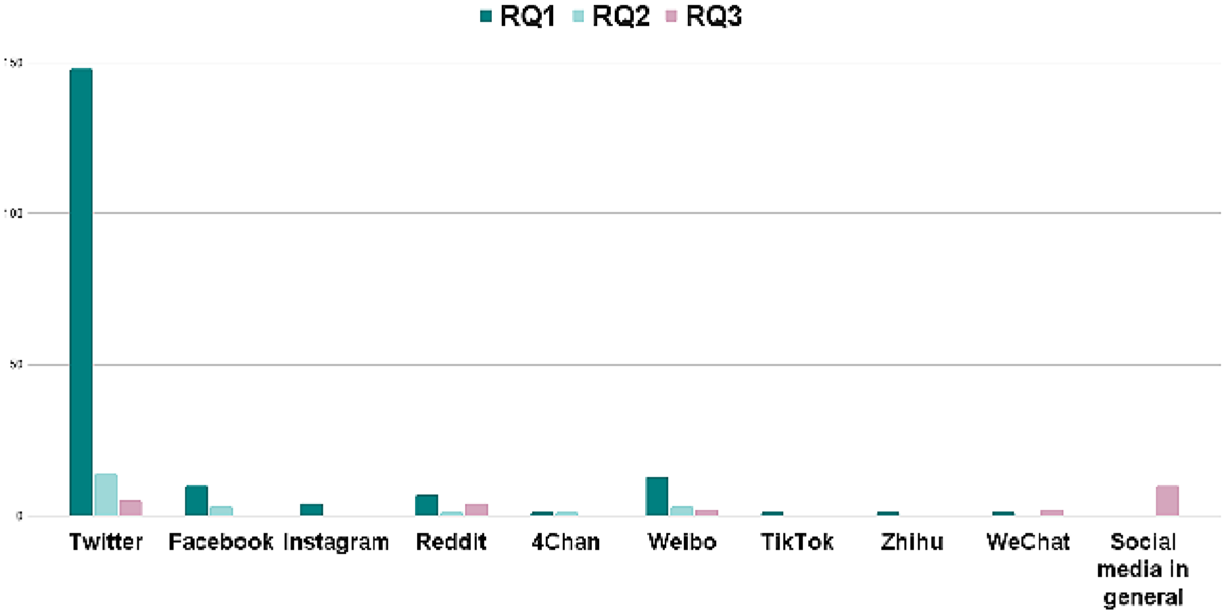


The publications spanned from 2010 to 2022. As can be seen in Figure 2B, the number of publications peaked in 2020 and continued to increase for all RQs. All the selected papers that answeredRQ3 spanned the from 2020 to 2021. A similar distribution was seen in papers that answeredRQ2, where selected papers were from 2015, 2019, 2020, and 2021. RQ1, which studies the aspects of epidemic management and mitigation using social media, included the highest number of papers and spanned the entire decade.

Social Media Platforms Used

Several social media platforms were used in the literature selected for this systematic review. X (formerly Twitter) is one of the most widely used platforms for sharing “microblogs.” These short messages are called tweets and can take up to 280 characters. In contrast, Weibo, is a popular platform to share and discuss individual information and life activities as well as celebrity news in China. As can be seen in Figure 3, X followed by Weibo seems to be the platform of choice for most works aiming to study epidemic monitoring and mitigation through social media (RQ1) and epidemic-related misinformation on social media (RQ2). For epidemic and social media–related mental health aspects, most works seem to take a generalist approach rather than a platform-specific one. Compared with other social media sites, such as Facebook and Instagram, which predominantly include heterogeneous posts, X offers a more concise "microblog" format.

Figure 3. Number of selected publications using each social media platform included in the systematic literature review. RQ: research question.



Epidemics Studied

The selected literature discussed multiple epidemics (Figures 4 and 5), including vital hemorrhagic fevers and influenza-like illness (ILI).

Dengue fever and Zika fever are mosquito-borne diseases caused by the dengue virus and Zika virus, respectively, and spread by several species of female mosquitoes of the Aedes genus [1]. The disease is now endemic in more than 100 countries with potential risk in other areas [1,32]. The WHO declared the Zika outbreak of 2016 and the Ebola outbreak in 2019 as public health emergencies of international concern (PHEICs) [1].

ILI is a nonspecific respiratory illness characterized by fever, fatigue, cough, and other symptoms. Cases of ILI can be caused either by influenza strains or by other viruses, such as coronaviruses. Influenza remains a global and year-round disease burden and causes illnesses that range in severity and sometimes lead to hospitalization and death. Seasonal influenza epidemics are mainly caused by influenza A and B viruses [1]. The influenza A virus subtype strain H1N1, commonly referred to as the swine flu, disproportionately affects children and younger people. H1N1 was declared a PHEIC in 2009 and then designated a pandemic [1]. Coronaviruses include SARS, MERS (Middle East respiratory syndrome), which can be contracted through direct or indirect contact with infected animals [1], as well as COVID-19 caused by the SARS-CoV-2 virus. The latter was designated a PHEIC and a pandemic by the WHO. As of April 26, 2023, the official death toll from COVID-19 reached 6,915,268 [1].

The highest number of selected publications for all RQs related to COVID-19, followed by influenza (Figure 4). This trend is due, in part, to the volume of the COVID-19 research output [33].

Figure 4. Number of selected publications pertaining to each epidemic included in the systematic literature review. MERS: Middle East respiratory virus; RQ: research question.

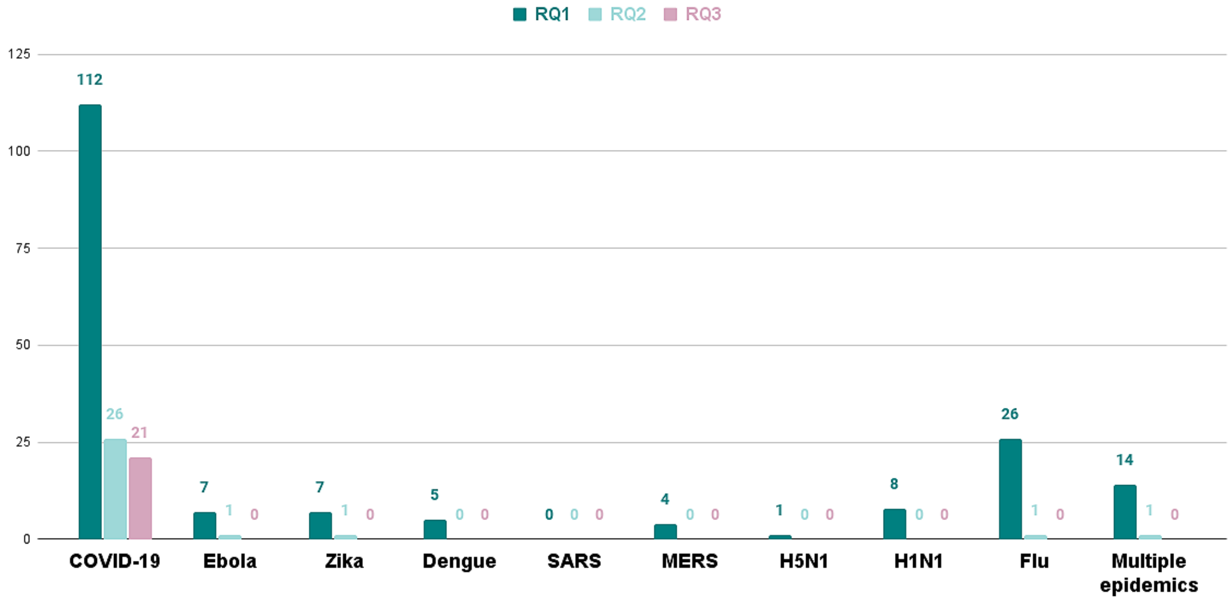
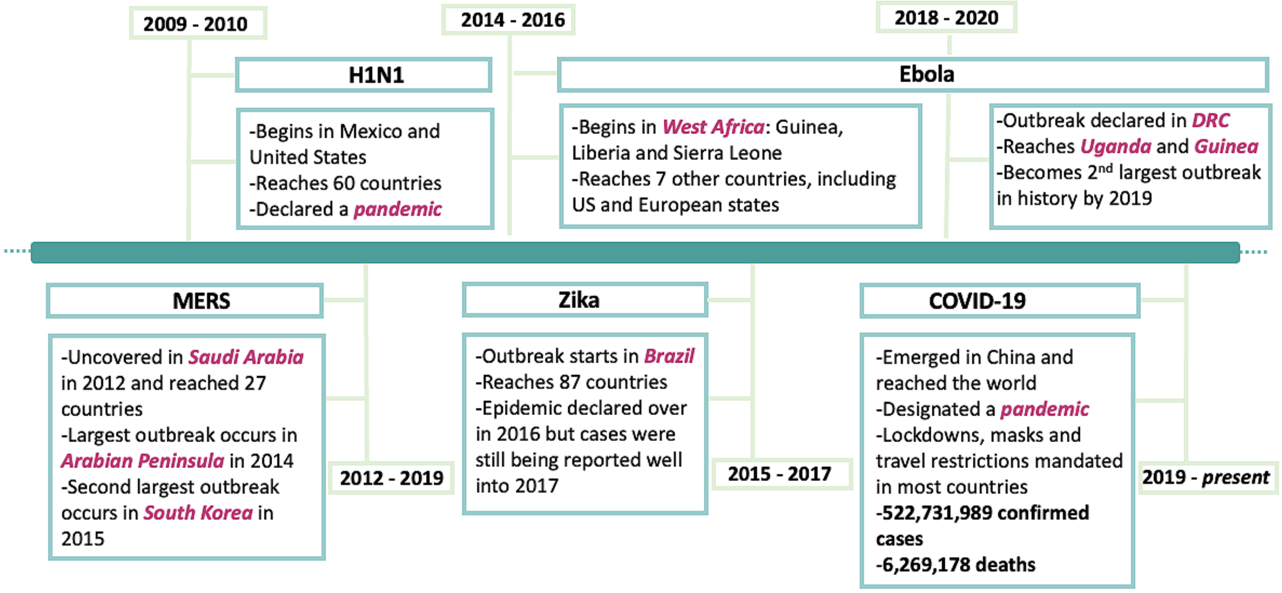


Figure 5. Timeline of the epidemics and pandemics spanning the last decade and included in the systematic literature review. SARS is pre-2009 and dengue fever has caused multiple outbreaks. Both are not illustrated in the timeline but are included in the systematic literature review. DRC: Democratic Republic of the Congo; MERS: Middle East respiratory syndrome.



Answers to RQs

Overview

A thematic analysis of the selected literature was conducted with the aim of identifying the main themes of each RQ. Themes were identified following the objectives of the paper and its results. For each theme, papers were organized by method, social media platform used, and the epidemic studied. Methods were grouped categorically. For instance, content analysis includes automated, linguistic, thematic, qualitative, or quantitative analysis, while dictionary-based classification entails a lexicon-based classification. Machine learning (ML) classification includes conventional ML models, while deep learning (DL) entails methods based on artificial neural networks with representation learning. Although it must be acknowledged that overlaps exist, the categorization used in this paper is based on the most distinctive and predominant use or theoretical approach of each method. This categorization is meant to facilitate a structured analysis and discussion of the literature by grouping papers according to their primary methodological approach, thus allowing a clearer comparison and contrast of their contributions, strengths, and limitations.

RQ1. Social Media for Epidemic Monitoring and Management

Overview

Social media platforms offer significant amounts of data, which can be potentially useful in biosurveillance and syndromic surveillance of epidemics and outbreaks.

Two main themes were identified in the selected papers that addressed how social media could be used in epidemic management, namely, (1) *epidemic surveillance and forecasting*, and (2) *public opinion understanding*.

Epidemic Surveillance and Forecasting

Several works proposed a dictionary-based classification of X for the surveillance of COVID-19 [34,35], dengue fever [36,37], Ebola [38,39], H1N1 [40,41], influenza [42-50], Zika [51,52], MERS [53], and a combination of epidemics [53-55]. Similar epidemic surveillance applications using dictionary-based classification were conducted using data from Weibo for Ebola [56] and influenza [57,58], Reddit for Zika [59], and Facebook for MERS and other epidemics [53].

Reported results indicated that epidemic surveillance can be achieved using varying strategies. For instance, social distancing–related tweets can be grouped into categories, such as implementation, purpose, social disruption, and adaptation, and used to quantify the spatiotemporal prevalence and evolution of COVID-19 social distancing on X [34]. Similarly, official social media channels of information and health organizations, such as the Centers for Disease Control and Prevention, WHO, and National Institutes of Health (NIH) can be monitored, and their X data can be classified to recognize “alarming” news and “concerning” news [35]. Dengue fever reported surveillance strategies to include systems aggregating social media data with weather and flood information [36], and using volume, location, time, and public perception as spatiotemporal dimensions [37].

Results also reported keyword-based data extraction and classification as a strategy for the creation of an Ebola monitoring platforming in China using Weibo data [56] and in Africa using X data [39], and for multiple epidemics [53-55].

Regression analysis was reported to be used for tracking and forecasting influenza [46,47,49,57] and Zika [52], and the Markov switching model was used for real-time early-stage influenza detection with emotion factors for epidemic and nonepidemic segmentation [58]. Statistical analysis was used to study the relationship between human activities collected from Sina Weibo and morbidity patterns and at-risk areas during COVID-19 in China [60].

Correlation analysis reported that X, in addition to other sources, could not provide an Ebola alert more than a week before the WHO and that X’s message volume was correlated more with news article volume than with the number of Ebola cases [38].

Additional dictionary-based surveillance methods include quantitative analysis, filtering, and normalization of X data for H1N1 [40,41] and Zika [51]; mathematical modeling of influenza trends using geo-tagged X streams [42]; time series for X symptom reporting matching ILI [43]; keyword analysis for Zika risk assessment [59,61] as well as influenza risk surveillance [44] and condition aggravation [45], and upcoming influenza spike detection [50], sentiment analysis [48].

Different methodologies using conventional ML were reported to be used for dengue-related event monitoring [62] and lazy associative classification [63]; influenza detection [64-66]; influenza activity monitoring [67]; seasonal influenza trend prediction [68,69]; ILI prevalence prediction [70] and awareness or infection classification [71]; location-specific influenza state detection [58,72,73]; avian influenza outbreak detection [74]; disease-related category classification [75-77] for Ebola, MERS, and dengue fever; guideline-related category classification for ILIs [78]; supervised text classification [79]; topic classification for symptomatic manifestation and prevention of mosquito-borne diseases [80]; infectious disease analytics [81,82] and COVID-19 case forecasting [83], and X-enabled contact tracing [84] and early detection [85].

Conventional ML models used for epidemic surveillance and monitoring include support vector machine (SVM), naive Bayes (NB), and logistic regression (LR).

Several DL techniques were applied for epidemic monitoring [86-90]; fine-tuning of semisupervised model with unlabeled COVID-19 dataset [91]; disease-infected individual detection in tweets using bidirectional encoder representations from transformers (BERT)–based model and disease-infection region identification using spatial analysis [92]; classification of Zika- and Ebola-related tweets [93], COVID-19 related tweets [94], and influenza-related information [95,96]; H1N1 outbreak forecasting and individual-level disease progressing using semisupervised multilayer perceptron and a online stochastic training algorithm [97]; and correlation of X reports on H1N1 infectious disease control using gray wolf optimizer and least square method [98]. Results reported that mathematical modeling can be used to understand the influence of X on the spread of H1N1 [98,99] and information entropy to quantify the impact of social network information [100].

Results showed that social network theory and social network analysis can be used for the prediction of infected groups and early detection of contagious outbreaks in social media [101-104], and that topic modeling techniques, such as latent Dirichlet allocation (LDA) can be used for epidemic intelligence [105-110] to detect major epidemic-related events [111], monitor information spread [112], and rank epidemic-related tweets [113].

Understanding Public Opinion

Several methods were used in the selected literature to extract and analyze public opinions expressed on social media. These methods were based on content analysis of social media data, linguistic analysis, qualitative analysis, lexicon-based analysis, sentiment analysis, valence aware dictionary and sentiment reasoner–based sentiment analysis, topic modeling, conventional ML models, and DL models.

Social media content analysis was used to analyze public discourse around H1N1- [114] and Zika-related risks [61], inspect social media coverage related to influenza vaccinations [115] and COVID-19 vaccinations [116-127], measure public health concerns [128], identify stances toward policies, such as social distancing and face masks [129], identify emotional composition of online discourse before and after COVID-19 [130], and inspect the presence and escalation of negative sentiments toward China [131]. Latent semantic analysis and LDA were used to mine opinions on X related to the hashtag #IndiaFightsCorona [132]. Topic detection and sentiment analysis were performed for opinion mining, concern exploration, and public opinion analysis in the context of epidemics [133-142], and for pattern analysis [143-145]. Social media content analysis was also used for tracking information spread [146], narratives and information voids [147], monitoring engagement [148-152] and emotional response [153-161], requests for medical assistance [162], health behavior changes [163], governmental response [164,165], and physicians' opinions [166].

Public reaction tracking and investigation were performed using SVM and NB for topic and sentiment analysis [127,167-175]; SVM, NB, and random forest (RF) for social media content classification (eg, caution, advice, notifications, donations, etc) [176]; crisis analysis [177,178]; clustering for topic extraction [179]; and LR for prevention category tweet classification [180]. ML was used to analyze public discourse against masks [181], extract insights on policy response [182], and understand expressions of help-seeking during COVID-19 [183].

BERT-based models were used for public sentiment assessment of data related to COVID-19 available in X [184-186]. Multilingual COVID-19 emotion prediction was performed using a fine-tuned BERT "BERTmoticon" [187], while bias and user opinion were identified using a GPT [188]. A language model for Arabic Moroccan dialect was used for topic modeling, emotion recognition, and polarity analysis [189]. long short-term memory (LSTM), BERT, and enhanced language representation with informative entities were used to analyze the evolution of sentiments in the face of the public health crisis due to COVID-19 [190]. Bi-LSTM with an attention mechanism was used for sentiment analysis of COVID-19–related tweets [191]. Term-frequency analysis was adopted to build an emerging topic graph [192], while the k-means algorithm, LR, SVM, and NB were used to identify COVID-19–related topics [193]. An extra tree and convolutional neural network-based ensemble model was reported to have outperformed conventional ML models in a sentiment classification task [194]. French COVID-19 tweet classification was performed using FlauBERT [195], while opinion monitoring was achieved using a combination of LSTM and global vectors for word representations [196]. Convolutional neural network was used for COVID-19 personal health mention detection [197].

Findings of analyses performed in the context of Ebola, Zika, and influenza revealed that social media posts from health organizations were highly effective when incorporating visuals and that public response was more affected by these communications when they acknowledged the concerns and fear of the community [198]. In the context of Ebola, findings highlighted that online blame was directed toward the affected populations as well as figures with whom social media users had preexisting political frustrations [199].

Analysis of X discussions in relation to COVID-19 revealed the presence of negative sentiments and an association between the words "coronavirus" and "China" [200]; a gradual increase in calls for social distancing, quarantining, and working from home among social media users [201]; a growing number of anger expressions directed at individuals refusing sanitary protocols; and the frequent use of the words "family," "life," "health," and "death" [201]. Analysis of X hashtags also revealed categories, such as quarantine, panic buying, school closures, lockdowns, frustration, and hope [201], as well as mentions of mental health issues and gratitude for essential workers [201]. Other categories and themes identified or used for manual annotation of topics discussed on social media include resource provision, employment and strategies [87], statistics, prevention, hygiene, diagnosis, politics, world news [202], conspiracy, economy, mortality, origin, and outbreak [203].

Findings also indicated increased levels of connectivity and agency coordination during the early-stage response to COVID-19 [87]. Disregarding COVID-19–imposed sanitary and government recommendations was potentially linked to uncertainty in times of crisis, overwhelm by "noise" presented on social media, and varying socioeconomic factors [204].

Results revealed that social media analytics were an efficient approach to capture the attitudes and perceptions of the public during COVID-19 as mentioned in studies by Yigitcanlar et al [205] and Xia et al [206]. Fear and collectivism were identified as predictors of people's preventive intention in the context of COVID-19 [207]. "Sadness" appeared to spike after the WHO declared COVID-19 as a pandemic, while "anger" and "disgust" spiked after the death toll surpassed the hundred thousand in the United States [187].

Tables 2 and 3 summarize the methods, epidemics, and social media used in studies pertaining to epidemic forecasting and prediction and understanding of public opinion.

Table 2. Summary of methodologies used in studies addressing the first part of research question 1 (epidemic surveillance and forecasting).

Method, epidemic studied, and social media used	References
Dictionary-based classification	
COVID-19	
X (formerly Twitter)	[34,35]
Sina Weibo	[60]
Dengue fever	
X	[36,37]
Ebola	
X	[38,39]
Weibo	[56]
H1N1 or swine flu	
X	[40,41]
Influenza or flu	
X	[42-50]
Sina Weibo, Tancent Weibo	[57,58]
Zika	
X	[51,52]
Reddit	[59]
MERS[a]	
X	[53]
Facebook	[53]
Multiple epidemics	
X	[53-55]
Facebook	[53]
ML[b] **classification**	
Dengue fever	
X	[62,63]
Influenza or flu	
X	[64-73]
Sina Weibo, Tancent Weibo	[58,73]
Facebook	[68]
H1N1 or swine flu	
X	[78]
H5N1 or avian influenza	
X	[74]
MERS	
X	[75-77]
Ebola	
X	[93]
Zika	
X	[93]
COVID-19	

Method, epidemic studied, and social media used	References
X	[83-85]
Multiple epidemics	
X	[75-77,79-82]
DL[c] classification	
COVID-19	
X	[35,86,88-91,94]
Ebola	
X	[93]
Zika	
X	[93]
Influenza or flu	
X	[95]
Multiple epidemics	
X	[92]
Mathematical modeling	
COVID-19	
WeChat	[100]
H1N1 or swine flu	
X	[97-99]
Social network analysis	
COVID-19	
X	[101,102]
Influenza or flu	
Facebook	[104]
Multiple epidemics	
X	[103]
Topic modeling	
COVID-19	
X	[88,102,105,106]
Dengue fever	
X	[111]
Ebola	
X	[112]
Influenza or flu	
X	[107]
Weibo	[108]
Zika	
X	[109]
Multiple epidemics	
X	[113]

[a] MERS: Middle East respiratory syndrome.

[b] ML: machine learning.

[c] DL: deep learning.

Table 3. Summary of methodologies used in studies addressing the second part of research question 1 (understanding public opinion).

Method, epidemic studied and social media used	References
Content analysis	
COVID-19	
X	[87,123,126,141,144,146,148,151,156,157,159,160,200,201]
Instagram	[202]
Reddit	[155,157,204]
TikTok	[122]
Weibo	[151,162]
Facebook	[126,149]
Ebola	
X	[198,199]
Facebook	[199]
Instagram	[198]
Zika	
Reddit	[61]
Influenza or flu	
X	[115]
Facebook	[115]
H1N1 or swine flu	
X	[115]
Facebook	[115]
Dictionary-based classification	
COVID-19	
X	[128,129,147,153,158,165,166,203,205,206]
Facebook	[147]
Instagram	[147]
Reddit	[147]
Weibo	[129]
H1N1 or swine flu	
X	[114]
ML[a] classification	
COVID-19	
X	[127,132,164-167,177-179,181]
Weibo	[176]
Facebook	[182]
Instagram	[182]
Zika	
X	[180]
DL[b] classification	
COVID-19	
X	[118-120,130,184-189,191,194-197]
Weibo	[190]

Method, epidemic studied and social media used	References
Topic modeling	
COVID-19	
X	[116,133-135,138,143,154,163,189,192,193]
Reddit	[139]
Weibo	[136,137]
Zhihu	[183]
Social network analysis	
COVID-19	
X	[117,121,124,125,140,142,145,150,152]
Facebook	[161]

[a] ML: machine learning.

[b] DL: deep learning.

RQ2. Social Media for Misinformation Management During Epidemics

Overview

Misinformation, or "fake news," has become a social phenomenon and has received increased attention in the past few years. Although the term, "fake news" has been around since the 1890s [208], the emergence and exponential rise in popularity of social media platforms has brought the term to the "front page." Fake news can fall into multiple categories depending on the intent and form it takes [208]. For instance, fake news can be false information and rumor fabrication (eg, celebrity gossip), hoaxes (eg, doomsday 2012), conspiracy theories (Q-Anon), and satire (eg, The Onion). The intent can range from deception for the purposes of monetary or personal gain to satirizing real news.

One main theme was identified in the selected papers that addressed how social media could be used in misinformation management during epidemics, namely, misinformation detection and characterization. Three subsequent subthemes were identified based on the scope of selected literature, namely: fake news identification, fake news characterization, and information distortion and conspiracy theories.

Misinformation Detection and Characterization

Overview

The selected literature focused on the inspection of news or claims shared on social media, with the aim of classifying them based on trustworthiness. Several methods were used to analyze social media content and detect misleading information, such as expert annotation, DL models, and social network analysis. While some papers focused on technical approaches to the detection of fake news, other studies tried to identify various characteristics related to the source or propagation of fake news.

Fake News Identification

Several works performed fake news identification using DL models [209-211] with conventional ML models for comparison or as baselines. A modified 3-layer-each LSTM and gated recurrent unit were used along with 6 conventional ML classification models (decision trees, LR, k-nearest neighbors, RF, SVM, and NB) on a "Covid-19 fake news Twitter dataset" [212] to identify fake news [210]. Findings reported that the best test results were obtained by LSTM (2 layers), with an accuracy of 98.6%, a precision of 98.55%, a recall of 98.6%, and an F_1-score of 98.5% [210]. Similarly, a multilayer perceptron, LR, decision trees, RF, NB, SVM, and gradient boosting were used for COVID-19 fake news detection in tweets and concluded that RF outperformed other models with an accuracy of 78%, a recall of 100%, a precision of 85%, and an F_1-score of 83% [211]. Expert annotated tweets were used to evaluate the performance of a BERT-based misinformation detection system [213]. Findings suggest that knowledge about the domain vocabulary helps domain-adapted models in predicting the correct stance, as it did for retrieval.

Detecting misleading and fake news was also performed by several studies using methods based on pretrained transformer models, bi-LSTM networks, artificial neural networks, convolutional neural networks, deep transfer learning [214-220], and using hybrid methodologies [221-227].

A semisupervised probabilistic graphical model that aimed to jointly learn the interactions between user trustworthiness, content reliability, and post credibility for influenza posts' credibility analysis outperformed baseline models (RF and Bayesian network) with an accuracy of 71.7% on data from Sina Weibo [209]. LR was performed on a small dataset of Facebook comments to detect fake news [228]. Several ML models, including gradient boosting classifier, LR, RF classifier, and decision tree classification were used in multiple works for fake news classification on social media [229-233].

Other works seeking to curtail the misinformation of COVID-19–related news and support reliable information dissemination used manual analysis through fact-checkers as well as consensus to verify the veracity and correctness of selected tweets and social media posts [234,235]. This is illustrated in a use case analyzing Facebook and X content in both English and Amharic [234] and an Ebola study [235].

Fake News Characterization

A manual annotation of tweet sources following 5 categories (academic, government, media, health professional, and public) allowed for the creation of a gold standard dataset for training a LR model based on 6 million Arabic tweets related to infectious viruses, such as MERS and COVID-19 [236]. Rumor detection using a top-down strategy consisting of extracting posts associated with previously identified rumors reported an 84.03% accuracy for the LR classifier [236]. Higher precision was obtained at the expense of higher runtime using ML models [232]. Similarly, topic modeling based on the k-means algorithm was used to identify sources of COVID-19–related rumors [193]. An entropy-based method was used to investigate the potential control of COVID-19 rumors [237] and content analysis was used to evaluate rumor dissemination and official responses during COVID-19 [238].

Semantic correlations between textual content and attached images were mined using a pretrained convolutional neural network to learn image representations and use them to enhance textual representations and train a fake news detector [239].

Content analysis showed that fake news from multiple sources could be classified using a taxonomy of health and non–health-related types and reported that the response of the public health system was debilitated by the propagation of fake news [240]. Roots of misinformation were categorized as politically related, false medical information, celebrity and pop culture related, religious belief related, and fraud and criminality related [241]. A comparison of fake news sources between China, Iran, and the United States showed that fake science is the main "root" of misinformation in China, while counterexpertise, that is, the rejection of mainstream academic expertise, politically motivated and governmentally sourced misinformation is the most prevalent source of fake news in the United States. In Iran, discourse about COVID-19 was found to be politically manipulated by the government, while official religious figures hindered the dissemination of accurate information [241]. Statistical analysis found bias of sentiment in fake news, as well as biases of gender of the user and media use with respect to real news [242].

Bot detection using BERT was performed as a potential strategy to improve fake news detection [243]. Findings imply that the ratio of real news to fake news is very similar between human accounts and bot accounts, and bot detection could not improve the performance of the fake news detection model [243].

Findings of an information mutation study using A Lite BERT reported that misinformation propagation could potentially be exacerbated by user commentary and found a positive association between information mutation and spreading outcome [244].

The findings of a propagation analysis showed that false claims propagate faster than partially false claims and that tweets containing misinformation are more often concerned with discrediting other information on social media [245].

An investigation leveraging neural networks and quantitative content analysis that aimed to reveal the conditions that lead audiences to accept and disseminate a fake claim as it relates to the Zika virus showed that Zika tweets, including threat cues and protection cues, are positively associated with the likelihood of sharing fake news [246]. In addition, findings of a descriptive analysis showed that the quality of news sources varies considerably with regard to information on COVID-19 [247], Results of a computational analysis indicated that the COVID-19 infodemic is highly characteristic of community structure, shaped by ideological orientation, typology of fake news, and geographic areas of reference [248]. Data from X indicated that content could be labeled according to political affiliation, media source, and type of source (political, satire, mainstream media, science, conspiracy or junk science, clickbait, and fake or hoax) [248].

Information Distortion and Conspiracy Theories

Information distortion in X cascades was found to be linked to oversimplification, distortion of logical links, omission of facts, and a shift in the medical topic to political and business disputes [249]. Risk amplification by information dramatization appeared to be linked to controversial topics as well as social and cultural influences [250].

Manual content and semantic analysis and topic modeling (LDA) techniques of tweet content were conducted through an examination of key term distribution, context, and medical terminology verification [249]. In a COVID-19 5G conspiracy use case, LDA and social network analysis were used to identify several topics from dataset of tweets [251] related to "5G conspiracy" and "5G threat" and discuss topics, including 5G towers, radiation effects, network, and radiation [252,253]. Emerging COVID-19–related conspiracy theories were detected by estimating narrative networks with an underlying graphical model and using a collection of data from Reddit subreddits and 4Chan threads related to the pandemic [254]. Findings identified multiple central conspiracy theories illustrated by examples, such as incorporating the COVID-19 conspiracy into Q-Anon conspiracy, #scamdemic and #plandemic [255], 5G as the cause of COVID-19 [252,253], antivax conspiracy, Bill Gates, #filmyourhospital conspiracy [256], and Pizzagate conspiracy [254]. Table 4 summarizes the methods, epidemics, and social media used in studies pertaining to misinformation management and detection.

Table 4. Summary of methods used in papers addressing research question 2–misinformation identification and characterization.

Method, epidemic studied, and social media used	References
ML[a] classification	
COVID-19	
X	[193,210,211,221-227,229-232]
Facebook	[228,230]
Sina Weibo	[233]
Multiple epidemics	
X	[236]
DL[b] classification	
COVID-19	
X	[210,213-218,220-227,239]
Facebook	[156]
Instagram	[217]
Weibo	[219,239]
Topic modeling	
COVID-19	
X	[237,249,252,255]
Social network analysis	
COVID-19	
X	[248,252,253,256]
Reddit, 4Chan	[254]
Probabilistic graph modeling	
Influenza	
Weibo	[209]
Manual content analysis	
COVID-19	
X	[234,241,248,249]
Facebook	[234,241,247]
Weibo	[241,250]
Instagram	[241]
Ebola	
X	[235]
Quantitative content analysis	
COVID-19	
X	[213,242,245]
Weibo	[250]
Zika	
X	[246]

[a] ML: machine learning.

[b] DL: deep learning.

RQ3. Social Media's Relationship With Mental Health During Epidemics

Overview

During the implementation of restrictive measures requiring limited social contact, social media can become one of the few methods to safely engage with others, rendering it the sole support system of vulnerable populations. Mental health deterioration can manifest in expressions shared on the internet and be used to gauge the toll epidemics and subsequent containment strategies could potentially take on individuals.

Two main themes were identified in the selected papers addressing how social media can be integrated in aspects of public mental health management during epidemics, namely, (1) social media as a tool to gauge the mental health toll of epidemics, and (2) impact of social media consumption during epidemics on mental health.

Mental Health Assessment Using Social Media

Assessment of mental health state was performed using conventional ML [257-259], DL [260-262], and topic modeling techniques [263,264]. Psychological profiles of Weibo users were predicted using ML and online ecological recognition with emotional measures and cognitive indicators, such as anxiety, depression, Oxford happiness, social risk judgment, and life satisfaction [257]. LSTM was used to estimate the rate of depression in the population during the COVID-19 pandemic using Reddit data [260]. Topic modeling, expert intervention, and X data were used to evaluate the possible effects of critical factors related to COVID-19 on the mental well-being of the population in a psychological vulnerability study [263].

Findings revealed that negative emotional indicators of psychological traits increased in anxiety and depression after COVID-19 was declared an epidemic or pandemic [257,262], while life satisfaction and happiness decreased [257]. A 53% average increase in depression rate of Reddit users was noted in selected months after the pandemic [260], and negative psychological vulnerability manifested in negative emotions toward social distancing and hospitalization [263]. Financial burden was found to increase the odds of depressive nonsuicidal thoughts for individuals who suffered job loss during COVID-19 [264]. Results indicated the beginning of recovery following the immediate mental health impact of the COVID-19 pandemic [259].

Table 5 summarizes the methods, epidemics, and social media used in studies pertaining to the use of social media as a tool to gauge the mental health toll of epidemics.

Table 5. Summary of methods used in papers addressing the first part of research question 3 (mental health assessment using social media).

Method and epidemic studied	Social media used	References
ML[a] classification		
COVID-19		
	Weibo	[257]
	X	[258,259]
	Reddit	[259]
DL[b] classification		
COVID-19		
	Reddit	[260,261]
	X	[262]
Topic modeling		
COVID-19		
	X	[263,264]
	Reddit	[264]

[a] ML: machine learning.

[b] DL: deep learning.

Association of Social Media Consumption and Mental Health

Multiple papers conducted cross-sectional studies and statistical analysis to investigate the association between social media consumption and mental health complications during epidemics, particularly during COVID-19. Several studies relied on regression analysis, online surveys, the Generalized Anxiety Disorder Scale, and the Patient Health Questionnaire.

Findings revealed that frequent Sina Weibo use was associated with higher anxiety, depression, and a combination of both [265], and compulsive WeChat use was associated with social media fatigue, emotional stress, and social anxiety [266]. Frequent use of WeChat during COVID-19 was also associated with depression and secondary trauma and was found to be a significant predictor of both [19], while close contact with individuals with COVID-19, along with spending ≥2 hours daily on COVID-19–related news on WeChat was associated with probable anxiety and depression in community-based adults [267]. The association between social media consumption and anxiety and depression was found to be statistically significant

[265,268,269] and positively associated with emotional overeating and anxiety in individuals with neuroticism [18].

The association between the mental health of students receiving higher education and social media use during COVID-19 confinement was analyzed, and results indicated that students in the 18 to 24 years age group, who were not in a relationship and who had lower academic results, presented the highest levels of addiction to social media [16]. Significant positive associations were found between relatedness, need, frustration, and social media addiction, as well as between social media addiction, depressive symptoms, and loneliness [17]. Excessive social media use was also found to fully mediate the relationship between COVID-19–related life concerns and schizotypal traits [270].

Appropriate guidance of adolescents in the use of social networking sites was found to have a potential impact on the mitigation of negative emotions during the COVID-19 pandemic [271].

On the positive side, social media use was found to be rewarding for Wuhan’s residents through information sharing and emotional and peer support [19]. Social media breaks were reported to have the potential to promote well-being during the COVID-19 pandemic [19]. In addition, positive mental health and mindfulness appeared to serve as protective factors, and positive mental health was found to be a mediator between the COVID-19 burden and addictive social media use [272].

Table 6 summarizes the methods, epidemics, and social media used in studies pertaining to the association of social media use with mental health issues during epidemics.

Table 6. Summary of methods used in papers addressing the second part of research question 3 (association of social media consumption with mental health).

Method, epidemic studied, and social media used	References
Statistical analysis	
COVID-19	
WeChat	[19,266]
Sina Weibo	[265]
Social media in general	[16-18,267-273]

Discussion

Principal Findings

This systematic literature review conceptualized 3 RQs to investigate if, when, and how social media can be harnessed for successful epidemic management and mitigation, effective curtailment of fake news propagation, and a refined understanding of social media’s relationship with mental health during epidemics. It presented a systematic categorization and summary of methods, social media sites, and epidemics broached in the 242 selected works and identified potential research directions and practical implications related to the RQs.

Papers selected pertaining to RQ1 comprised the highest number of papers and included publications from all years of the decade, illustrating continuous and ongoing efforts by the scientific community to harness social media’s potential for improved containment measures during epidemics.

COVID-19 was found to be the epidemic most studied in selected papers. This is due to the rapid increase of COVID-19–related publications since the first year of the pandemic. The frequency of publication and the volume of the academic output contributed to the creation of the COVID-19 Open Research Dataset [33]. A similar rising trend was seen in RQ2. This can be explained by the emergence of the “fake news” phenomena on social media and its particular increase in times of crisis. The selected publications answering RQ3 were published from 2020 to 2022. Papers that pertained to RQ3 were much lesser in number than those that pertained to RQ1 and RQ2. Given the mental health aspect of this particular RQ, a potential inference can be made suggesting a very recent interest in mental health as it relates to social media and epidemics. X was found to be the most used social media site in the selected literature, potentially suggesting its attractiveness to works conducting linguistic analysis and classification tasks. This can also be due to the differences in the popularity of social media sites by geographic location and key demographics. The availability of application programming interfaces to crawl data is also a major factor in choosing specific social media platforms as data sources.

General Discussion

The systematic literature review presented in this paper differs from existing reviews and aims to cover a different gap in the literature. Existing works have taken an interest in a broader range of crises, including noninfectious diseases and health risk behaviors [12], disasters in general [25], and new and reemerging infectious diseases [26]. Focus was directed toward effectively targeting vulnerable populations to test interventions and improve health outcomes [12], collective behavior [25], and generalized perspectives on emergency situations [27]. Differences other than scope include data sources, time range, and volume of literature. The review presented in this paper covered a broader time range, included gray literature, and reviewed a sizable volume of research papers.

The review’s findings indicated that social media was found to be an effective way to understand the public’s reactions and engagement during epidemics [205]. Monitoring topics of discussion during epidemics allowed for insights on whether aspects of epidemic management needed improvement, whether the public agrees with government decisions, and which emotions are linked to the onset of epidemics and mitigation protocols [198,204-206]. Analysis of opinions related to aspects, such as COVID-19 vaccinations were proposed and could be used to give feedback to governments and health organizations to implement better suited protocols [116,122,124-126] for mitigation, and to identify topics of misinformation, and therefore offer clarifications or conduct further awareness efforts to combat rumors and conspiracies [254]. Results also indicated that social media can be used in case forecasting [83], X-enabled contact tracing [84], early detection [85], tracking adherence to preventive guidelines, such as wearing masks and social distancing [205,206], and monitoring symptomatic self-expressions of infection [80]. Misinformation detection on social media was performed as a classification task, manually using experts and fact checkers, and using artificial intelligence techniques; however, presented several challenges. Misinformation often used language styles of academics and health professionals to deceive the public [236] and propagated faster when it included higher levels of threat due to the collective stress reaction it generated [246]. "Troll" accounts were found to play the second most prominent role is misinformation spread and present a "substantial cause for concern" [248]. Other challenges of misinformation detection related to limitations of studies due to the use of small batches of data [252], false positives [228], and a "politicization" of neutral health emergency crises [235].

Although epidemics were found to cause negative emotions and mental health issues [260,262,263], many expressions of positive emotions were noted [257], reflecting group cohesiveness rather than pure personal emotions. Group threats contributed to the manifestation of more beneficial behaviors and social solidarity [269]. Viewing heroic acts, speeches from experts, and knowledge of the disease and prevention methods were associated with more positive effects and less expressions of depression [269]. Media content, including useful information for self-protection was found to be potentially helpful to people during epidemics and may enhance active coping, prevention behaviors, and instill a sense of control [269]. The use of social media during epidemics, although linked with manifestations of anxiety and depression, appeared to benefit Wuhan residents and was perceived as an important activity during lockdown [19]. Balancing social media use to obtain ample informational as well as emotional and peer support, while avoiding the potential mental health toll, is a difficult task for users, especially without the availability of alternative and easily accessible sources of health information [19].

Using social media data for mental health assessment has its challenges and limitations. It can add a population or demographic bias to results, given that some social media sites are predominantly used by younger people or are more or less popular depending on the country [257,263]. Depending on the social media site (eg, Reddit), the user pool skews younger, and thus could be more prone to depression [260]. Moreover, some analyses are based on a weekly basis, with a relatively large granularity, which has certain influences on reflecting the changing trend of social mentality in a timely manner [257]. The qualitative nature of the results obtained and interpreted by domain experts limits the generalization of the findings and requires more corroborating results. Consequently, findings may need additional data to be strengthened [260,263]. As for works pertaining to the association of social media consumption with psychological outcomes, a causal link has not been established due to the cross-sectional nature of the contributions. Studies reflected a single point in time for participants, therefore, further longitudinal studies are necessary. In addition, the surveys were conducted on the web, and consequently, respondent bias is possible [265]. The recruitment of all participants from the same country and from one social media platform can introduce bias to studies [266,268], in addition to potential gender biases and sample representativeness [18,19], and recall bias related to self-reporting [269]. The results could not exclude the possibility of residual confounding caused by unmeasured factors.

A change can be seen in the evolution of research themes over time and through different epidemics. A sizable number of works focused on the influenza epidemic surveillance using lexicon-based and dictionary-based classifications, as well as classical ML techniques. This volume of literature could potentially be linked to the influenza prediction "wave" that preceded, paralleled, and followed the dereliction of the "Google flu trend" after its failure to predict major outbreaks [274]. Although various methods were used, ML and DL techniques were most frequently used for COVID-19 surveillance. Scientific contributions evolved with the emergence of more epidemics. COVID-19 appeared to have benefited from the digitization of literature as well as the development and improvements taking place in the fields of natural language processing, ML, DL, social network analysis, and topic modeling. The global nature of the COVID-19 crisis generated an influx of publications and contributions. The theme of misinformation management has also evolved with epidemics and with the proliferation of social media fake news, bots, troll accounts, and widely propagated conspiracy theories. COVID-19 has been the subject of multiple controversies and conspiracies, which encouraged scientific efforts to study potential curtailment methods. As for the mental health aspect, all publications pertaining to the scope of RQ3 were related to COVID-19, and it appeared that previous epidemics were not subject to social media association analysis. This could be due to the fear linked to COVID-19 and the challenging nature of sanitary measures such as global lockdowns and social distancing, which led to an increase in social media reliance. It could also be due to the decade's zeitgeist which brought online mental health discussions and awareness front and center.

Identified Issues

One of the major issues identified was the lack of preemptive measures building on the results of previous studies and aiming to implement social media–enabled processes in real time or near real time. Lessons learned are not efficiently integrated in crisis mitigation measures nor used as building blocks for optimized proactive prevention. A synergy between government health agencies, research communities, and the public would allow for the success of social-media public health initiatives. Such collaborative efforts require effective and trustworthy interactions. This highlights an additional issue related to the relative inefficiency of social media campaigns. Populations need to be targeted for both informative purposes and for active emotional support. Understanding public opinion is useful to gauge sentiments and reactions, and therefore it is important to remedy the gap for applications integrating extracted opinions in targeted epidemic management.

Because of the medical and financial burden of epidemics, mental health concerns are often ignored by both governments and the public. As a result, the manifestation of several mental health–related symptoms becomes more prevalent as epidemics progress. In the case of the Ebola outbreak in 2014, symptoms of posttraumatic stress disorder and anxiety-depression were more prevalent even after a year of the Ebola response [199]. When limited resources are geared for epidemic containment, the health care system focuses majorly on emergency services. Therefore, individuals with substance abuse and dependency disorders may see deterioration in their mental health [13]. During community crises, event-related information is often sought in an effort to retain a sense of control in the face of fear and uncertainty and their psychological manifestations. When misleading misinformation is propagated on social media, perceptions of risk are distorted, leading to extreme public panic, stigmatization, and marginalization [13]. Psychological interventions and psychosocial support would have a direct impact on the improvement of public mental health during epidemics.

Directions for Future Research

We identified several issues and gaps in the literature related to the RQs of this systematic literature review and suggest potential paths for future research.

Given the recognized impact of epidemics on mental health and the prevalent use of social media platforms during times of crisis, it is necessary to explore the aspects of social media leading to mental health deterioration during epidemics. Potential factors range from increased consumption levels of social media, social media addiction, emotional fatigue due to overwhelm, and consumption of “sad” content. Investigating which aspects of social media use are responsible for worsening states of mental health and mental health disorders would allow a targeted approach to curbing this negative impact during times of crisis. As for health-related fake news, it is important to understand what makes citizens prone to engaging in fake news sharing. Specifically, features identifying both an individual’s and a group’s susceptibility to believe and share misinformation need to be determined and categorized. Levels of education, geographic and demographic profiles, cultural influences, and psychological vulnerability are potential features requiring further investigation in their association with fake news dissemination on social media and within communities.

Epidemics are rapidly changing phenomena requiring fast interventions and decision-making. Although postcrisis analysis is imperative for an improved understanding of lessons learned, proactive epidemic management is vital and would have the most impact on mitigation efforts. Integrating artificial intelligence techniques into this proactive surveillance could further optimize this process.

In addition, misinformation propagation has a significant impact on the success of interventions given that both the components of exaggerated fear and apathy linked to misinformation can hinder management efforts. However, the investigation of misinformation needs to be extended to include potential links between misinformation and mental health deterioration.

Practical Implications

This work has several potential practical implications pertaining to different entities.

Implications for governing entities include the development of an efficient misinformation correction strategy to fight incorrect information, rumors, and conspiracy theories related to epidemics; the development of clear communication channels for knowledge dissemination to build trust with the public; the development of interventions to limit the impact of epidemics on stress responses (anxiety, depression) due to distorted risk perceptions; the bolstering of public awareness efforts on sanitary measures and proactive protection; and the insurance of the supply of medical staff available to treat patients, as well as psychological support staff to assist patients and their families in navigating the ramifications of infection and loss of loved ones.

Implications for social media platforms include taking a leadership position in the management of epidemic-related fake news by implementing built-in fact-checking processes and assisting health agencies and scientific entities in disseminating factual information about the disease, its symptoms, its potential risk, and efficient sanitary measures for the public to adopt.

Implications for the public include improving community resilience during epidemics using social media groups and assisting in combating misinformation.

Limitations

The results of this review should be considered in light of several limitations. The data sources used in this review did not cover all existing scientific databases, and therefore, cannot generalize findings to the entirety of the literature. The scope of the review focused on specific aspects of the epidemic-social media relationship, and so does not provide a general overview. Although the process of data extraction and analysis was undertaken with extreme diligence, there can be potential for bias. Despite our recognition of the inherent limitations of any search strategy, we have ensured our commitment to the rigor and transparency of the systematic review process.

Conclusions

Given the collective experience of epidemics, responses by communities can often provide insight into the degree of adherence toward preventive measures as well as mitigation protocols. In an effort to control the spread of epidemics, governments, public health institutions, and health care professionals generally issue guidelines for the public through online portals, news sources, and in the past decade, social media. Online “chatter” can indicate the public’s response to these guidelines, and their sentiments toward the epidemic itself or specific topics related to it, such as vaccinations, treatments, mortality rates, etc. Mitigation efforts require collaborative strategies and public involvement; therefore, gaining insight into public opinion and response can prove vital in the success or failure of such efforts.

It is evident that epidemic preparedness and mitigation protocols need to be adjusted to deal with the special challenges that accompany the technological revolution taking place, especially in light of the considerable impact of the ongoing infodemic. In addition, it is vital to have effective ways to exploit the full potential of social media without risking the toll it could potentially take on users' mental health. The systematic literature review presented in this paper covers several key aspects of the relationship between epidemics and social media, especially with respect to fake news and mental health. Methods used to answer RQs are categorized. The findings of this review could shed light on broader implications related to data quality concerns and privacy considerations in epidemic surveillance, thus highlighting the lack of works proposing ethical, legal, and technical frameworks to accompany scientific efforts. Learning from past crises and integrating a digital and social media-enabled infrastructure into public health protocols could make a difference in future preparedness levels.

Acknowledgments

The authors thank the anonymous reviewers for their remarks and improvements to the paper. This work was carried out as part of the eCovid and CovInov projects which are funded by the Centre National de la Recherche Scientifique et Technique and the Rabat-Salé-Kénitra region.All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Data Availability

The datasets generated during and analyzed during this study are available in the Github repository [275].

Authors' Contributions

CA and MG conceived the study. CA and IK designed the experiments. CA, IK, and MG carried out the research. CA and IK prepared the first draft of the manuscript. MG and KB contributed to the experimental design and preparation of the manuscript. All authors were involved in the revision of the draft manuscript and have agreed to the final content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[DOCX File , 33 KB - publichealth_v11i1e55642_app1.docx]

References

1. Health topics. World Health Organization. URL: https://www.who.int/health-topics [accessed 2024-12-24]
2. Mukhtar S. Psychological health during the coronavirus disease 2019 pandemic outbreak. Int J Soc Psychiatry 2020 Aug;66(5):512-516 [FREE Full text] [doi: 10.1177/0020764020925835] [Medline: 32434402]
3. Duarte Alonso A, Kok SK, Bressan A, O'Shea M, Sakellarios N, Koresis A, et al. COVID-19, aftermath, impacts, and hospitality firms: an international perspective. Int J Hosp Manag 2020 Oct;91:102654 [FREE Full text] [doi: 10.1016/j.ijhm.2020.102654] [Medline: 32863526]
4. Major epidemics of the modern era. Council on Foreign Relations. URL: https://www.cfr.org/timeline/major-epidemics-modern-era [accessed 2024-12-24]
5. Porta M. A Dictionary of Epidemiology. Oxford, UK: Oxford University Press; 2008.
6. Shultz JM, Baingana F, Neria Y. The 2014 Ebola outbreak and mental health: current status and recommended response. JAMA 2015 Feb 10;313(6):567-568. [doi: 10.1001/jama.2014.17934] [Medline: 25532102]
7. Mak IW, Chu CM, Pan PC, Yiu MG, Chan VL. Long-term psychiatric morbidities among SARS survivors. Gen Hosp Psychiatry 2009 Jul;31(4):318-326 [FREE Full text] [doi: 10.1016/j.genhosppsych.2009.03.001] [Medline: 19555791]
8. Saroj A, Pal S. Use of social media in crisis management: a survey. Int J Disaster Risk Reduct 2020 Sep;48:101584. [doi: 10.1016/j.ijdrr.2020.101584]
9. Lin TH, Chang MC, Chang CC, Chou YH. Government-sponsored disinformation and the severity of respiratory infection epidemics including COVID-19: a global analysis, 2001-2020. Soc Sci Med 2022 Mar;296:114744 [FREE Full text] [doi: 10.1016/j.socscimed.2022.114744] [Medline: 35124544]
10. Pavela Banai I, Banai B, Mikloušić I. Beliefs in COVID-19 conspiracy theories, compliance with the preventive measures, and trust in government medical officials. Curr Psychol 2022 May 26;41(10):7448-7458 [FREE Full text] [doi: 10.1007/s12144-021-01898-y] [Medline: 34075284]

11. Pfeffer B, Goreis A, Reichmann A, Bauda I, Klinger D, Bock MM, et al. Coping styles mediating the relationship between perceived chronic stress and conspiracy beliefs about COVID-19. Curr Psychol 2022 Aug 16;43(15):1-9 [FREE Full text] [doi: 10.1007/s12144-022-03625-7] [Medline: 35990199]
12. Charles-Smith LE, Reynolds TL, Cameron MA, Conway M, Lau EH, Olsen JM, et al. Using social media for actionable disease surveillance and outbreak management: a systematic literature review. PLoS One 2015;10(10):e0139701 [FREE Full text] [doi: 10.1371/journal.pone.0139701] [Medline: 26437454]
13. Roy A, Singh AK, Mishra S, Chinnadurai A, Mitra A, Bakshi O. Mental health implications of COVID-19 pandemic and its response in India. Int J Soc Psychiatry 2021 Aug;67(5):587-600 [FREE Full text] [doi: 10.1177/0020764020950769] [Medline: 32873106]
14. Kaimann D, Tanneberg I. What containment strategy leads us through the pandemic crisis? An empirical analysis of the measures against the COVID-19 pandemic. PLoS One 2021 Jun 21;16(6):e0253237 [FREE Full text] [doi: 10.1371/journal.pone.0253237] [Medline: 34153058]
15. Serrano-Alarcón M, Kentikelenis A, Mckee M, Stuckler D. Impact of COVID-19 lockdowns on mental health: evidence from a quasi-natural experiment in England and Scotland. Health Econ 2022 Feb 12;31(2):284-296 [FREE Full text] [doi: 10.1002/hec.4453] [Medline: 34773325]
16. Oliveira AP, Nobre JR, Luis H, Luis LS, Pinho LG, Albacar-Riobóo N, et al. Social media use and its association with mental health and internet addiction among Portuguese higher education students during COVID-19 confinement. Int J Environ Res Public Health 2022 Dec 30;20(1):664 [FREE Full text] [doi: 10.3390/ijerph20010664] [Medline: 36612991]
17. Cheng C, Lau YC. Social media addiction during COVID-19-mandated physical distancing: relatedness needs as motives. Int J Environ Res Public Health 2022 Apr 12;19(8):4621 [FREE Full text] [doi: 10.3390/ijerph19084621] [Medline: 35457489]
18. Gao Y, Ao H, Hu X, Wang X, Huang D, Huang W, et al. Social media exposure during COVID-19 lockdowns could lead to emotional overeating via anxiety: the moderating role of neuroticism. Appl Psychol Health Well Being 2022 Feb;14(1):64-80 [FREE Full text] [doi: 10.1111/aphw.12291] [Medline: 34265167]
19. Zhong B, Huang Y, Liu Q. Mental health toll from the coronavirus: social media usage reveals Wuhan residents' depression and secondary trauma in the COVID-19 outbreak. Comput Human Behav 2021 Jan;114:106524 [FREE Full text] [doi: 10.1016/j.chb.2020.106524] [Medline: 32836728]
20. Social media and COVID-19: a global study of digital crisis interaction among Gen Z and Millennials. World Health Organization. 2021 Dec 1. URL: https://www.who.int/news-room/feature-stories/detail/social-media-covid-19-a-global-study-of-digital-crisis-interaction-among-gen-z-and-millennials [accessed 2024-12-24]
21. Kapoor KK, Tamilmani K, Rana NP, Patil P, Dwivedi YK, Nerur S. Advances in social media research: past, present and future. Inf Syst Front 2017 Nov 6;20(3):531-558. [doi: 10.1007/s10796-017-9810-y]
22. Walters RA, Harlan PA, Nelson NP, Hartley DM, Voeller JG. Data sources for biosurveillance. In: Wiley Handbook of Science and Technology for Homeland Security. Hoboken, NJ: John Wiley & Sons; 2010.
23. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, Mostashari F, et al. Implementing syndromic surveillance: a practical guide informed by the early experience. J Am Med Inform Assoc 2004 Nov 21;11(2):141-150 [FREE Full text] [doi: 10.1197/jamia.M1356] [Medline: 14633933]
24. Selerio EJ, Caladcad JA, Catamco MR, Capinpin EM, Ocampo L. Emergency preparedness during the COVID-19 pandemic: modelling the roles of social media with fuzzy DEMATEL and analytic network process. Socioecon Plann Sci 2022 Aug;82:101217 [FREE Full text] [doi: 10.1016/j.seps.2021.101217] [Medline: 35001981]
25. Eismann K, Posegga O, Fischbach K. Collective behaviour, social media, and disasters: a systematic literature review. In: Proceedings of the European Conference on Information Systems. 2016 Presented at: ECIS 2016; June 12-15, 2016; Istanbul, Turkey.
26. Tang L, Bie B, Park SE, Zhi D. Social media and outbreaks of emerging infectious diseases: a systematic review of literature. Am J Infect Control 2018 Sep;46(9):962-972 [FREE Full text] [doi: 10.1016/j.ajic.2018.02.010] [Medline: 29628293]
27. Abdulhamid NG, Ayoung DA, Kashefi A, Sigweni B. A survey of social media use in emergency situations: a literature review. Inf Dev 2020 Mar 26;37(2):274-291. [doi: 10.1177/0266666920913894]
28. Michailidis PD. Visualizing social media research in the age of COVID-19. Information 2022 Aug 03;13(8):372. [doi: 10.3390/info13080372]
29. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. [The PRISMA 2020 statement: an updated guideline for reporting systematic reviewsDeclaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas]. Rev Panam Salud Publica 2022;46:e112 [FREE Full text] [doi: 10.26633/RPSP.2022.112] [Medline: 36601438]
30. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
31. Taggart T, Grewe ME, Conserve DF, Gliwa C, Roman Isler M. Social media and HIV: a systematic review of uses of social media in HIV communication. J Med Internet Res 2015 Nov 02;17(11):e248 [FREE Full text] [doi: 10.2196/jmir.4387] [Medline: 26525289]
32. Guzman MG, Harris E. Dengue. The Lancet 2015 Jan;385(9966):453-465. [doi: 10.1016/s0140-6736(14)60572-9]

33. Wang LL, Lo K, Chandrasekhar Y, Reas R, Yang J, Eide D, et al. CORD-19: the COVID-19 open research dataset. arXiv Preprint posted online on April 22, 2020 [FREE Full text]
34. Kwon J, Grady C, Feliciano JT, Fodeh SJ. Defining facets of social distancing during the COVID-19 pandemic: Twitter analysis. J Biomed Inform 2020 Nov;111:103601 [FREE Full text] [doi: 10.1016/j.jbi.2020.103601] [Medline: 33065264]
35. Rao HR, Vemprala N, Akello P, Valecha R. Retweets of officials' alarming vs reassuring messages during the COVID-19 pandemic: implications for crisis management. Int J Inf Manage 2020 Dec;55:102187 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102187] [Medline: 32836644]
36. Othman MK, Danuri MS. Proposed conceptual framework of Dengue Active Surveillance System (DASS) in Malaysia. In: Proceedings of the International Conference on Information and Communication Technology. 2016 Presented at: ICICTM 2016; May 16-17, 2016; Kuala Lumpur, Malaysia. [doi: 10.1109/icictm.2016.7890783]
37. Gomide J, Veloso A, Meira WJ, Almeida V, Benevenuto F, Ferraz F, et al. Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. In: Proceedings of the 3rd International Web Science Conference. 2011 Presented at: WebSci '11; June 15-17, 2011; Koblenz, Germany. [doi: 10.1145/2527031.2527049]
38. Yom-Tov E. Ebola data from the internet: an opportunity for syndromic surveillance or a news event? In: Proceedings of the 5th International Conference on Digital Health 2015. 2015 Presented at: DH '15; May 18-20, 2015; Florence, Italy. [doi: 10.1145/2750511.2750512]
39. Marc T, Kodzo A. How useful are social networks for analyzing epidemics? The example of Twitter for monitoring the 2018–2019 Ebola epidemic in Africa. In: Proceedings of the International Multi-Conference on: “Organization of Knowledge and Advanced Technologies”. 2020 Presented at: OCTA 2020; February 6-8, 2020; Tunis, Tunisia. [doi: 10.1109/octa49274.2020.9151663]
40. Szomszor M, Kostkova P, de Quincey E. #Swineflu: Twitter predicts Swine Flu outbreak in 2009. In: Proceedings of the Third International Conference on Electronic Healthcare. 2010 Presented at: eHealth 2010; December 13-15, 2010; Casablanca, Morocco. [doi: 10.1007/978-3-642-23635-8_3]
41. Quincey D, Kostkova P. Early warning and outbreak detection using social networking websites: the potential of Twitter. In: Proceedings of the Second International ICST Conference on Electronic Healthcare. 2009 Presented at: eHealth 2009; September 23-25, 2009; Istanbul, Turkey. [doi: 10.1007/978-3-642-11745-9_4]
42. Chon J, Raymond R, Wang H, Wang F. Modeling flu trends with real-time geo-tagged Twitter data streams. In: Proceedings of the 10th International Conference on Wireless Algorithms, Systems, and Applications. 2015 Presented at: WASA 2015; August 10-12, 2015; Qufu, China. [doi: 10.1007/978-3-319-21837-3_7]
43. Stilo G, Velardi P, Tozzi AE, Gesualdo F. Predicting flu epidemics using Twitter and historical data. In: Proceedings of the International Conference on Brain Informatics and Health. 2014 Presented at: BIH 2014; August 11-14, 2014; Warsaw, Poland. [doi: 10.1007/978-3-319-09891-3_16]
44. Hwang MH, Wang S, Cao G, Padmanabhan A, Zhang Z. Spatiotemporal transformation of social media geostreams: a case study of Twitter for flu risk analysis. In: Proceedings of the 4th ACM SIGSPATIAL International Workshop on GeoStreaming. 2013 Presented at: IWGS '13; November 5, 2013; Orlando, FL. [doi: 10.1145/2534303.2534310]
45. Talvis K, Chorianopoulos K, Kermanidis KL. Real-time monitoring of flu epidemics through linguistic and statistical analysis of Twitter messages. In: Proceedings of the 9th International Workshop on Semantic and Social Media Adaptation and Personalization. 2014 Presented at: SMAP 2014; November 6-7, 2014; Corfu, Greece. [doi: 10.1109/smap.2014.38]
46. Lampos V, Zou B, Cox IJ. Enhancing feature selection using word embeddings: the case of flu surveillance. In: Proceedings of the 26th International Conference on World Wide Web. 2017 Presented at: WWW '17; April 3-7, 2017; Perth, Australia. [doi: 10.1145/3038912.3052622]
47. Lampos V, Bie TD, Cristianini N. Flu detector-tracking epidemics on Twitter. In: Proceedings of the Machine Learning and Knowledge Discovery in Databases. 2010 Presented at: ECML PKDD 2010; September 20-24, 2010; Barcelona, Spain. [doi: 10.1007/978-3-642-15939-8_42]
48. Byrd K, Mansurov A, Baysal O. Mining Twitter data for influenza detection and surveillance. In: Proceedings of the International Workshop on Software Engineering in Healthcare Systems. 2016 Presented at: SEHS '16; May 14-22, 2016; Austin, Texas. [doi: 10.1145/2897683.2897693]
49. Comito C, Forestiero A, Pizzuti C. Improving influenza forecasting with web-based social data. In: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2018 Presented at: ASONAM 2018; August 28-31, 2018; Barcelona, Spain. [doi: 10.1109/asonam.2018.8508563]
50. Kostkova P, Szomszor M, St. Louis C. #swineflu: the use of Twitter as an early warning and risk communication tool in the 2009 swine flu pandemic. ACM Trans Manage Inf Syst 2014 Jul;5(2):1-25. [doi: 10.1145/2597892]
51. Abouzahra M, Tan J. Twitter vs. Zika—the role of social media in epidemic outbreaks surveillance. Health Policy Technol 2021 Mar;10(1):174-181. [doi: 10.1016/j.hlpt.2020.10.014]
52. Masri S, Jia J, Li C, Zhou G, Lee MC, Yan G, et al. Use of Twitter data to improve Zika virus surveillance in the United States during the 2016 epidemic. BMC Public Health 2019 Jun 14;19(1):761 [FREE Full text] [doi: 10.1186/s12889-019-7103-8] [Medline: 31200692]

53. Chaudhary S, Naaz S. Use of big data in computational epidemiology for public health surveillance. In: Proceedings of the International Conference on Computing and Communication Technologies for Smart Nation. 2017 Presented at: IC3TSN 2017; October 12-14, 2017; Gurgaon, India. [doi: 10.1109/ic3tsn.2017.8284467]
54. Romano S, Di Martino S, Kanhabua N, Mazzeo A, Nejdl W. Challenges in detecting epidemic outbreaks from social networks. In: Proceedings of the 30th International Conference on Advanced Information Networking and Applications Workshops. 2016 Presented at: WAINA 2016; March 23-25, 2016; Crans-Montana, Switzerland. [doi: 10.1109/waina.2016.111]
55. Ji X, Chun SA, Geller J. Epidemic outbreak and spread detection system based on Twitter data. In: Proceedings of the First International Conference on Health Information Science. 2012 Presented at: HIS 2012; April 8-10, 2012; Beijing, China. [doi: 10.1007/978-3-642-29361-0_19]
56. Feng S, Hossain L. Weibo surveillance of public awareness to Ebola disaster in China. In: Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016. 2016 Presented at: IntelliSys 2016; September 21-22, 2016; London, UK. [doi: 10.1007/978-3-319-56991-8_58]
57. Zhang Q, Gioannini C, Paolotti D, Perra N, Perrotta D, Quaggiotto M, et al. Social data mining and seasonal influenza forecasts: the FluOutlook platform. In: Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases. 2015 Presented at: ECML PKDD 2015; September 7-11, 2015; Porto, Portugal. [doi: 10.1007/978-3-319-23461-8_21]
58. Sun X, Ye J, Ren F. Real time early-stage influenza detection with emotion factors from Sina Microblog. In: Proceedings of the Fifth Workshop on South and Southeast Asian Natural Language Processing. 2014 Presented at: WSSANLP@COLING 2014; August 23, 2014; Dublin, Ireland. [doi: 10.3115/v1/w14-5511]
59. Gui X, Kou Y, Pine KH, Chen Y. Managing uncertainty: using social media for risk assessment during a public health crisis. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17; May 6-11, 2017; Denver, CO. [doi: 10.1145/3025453.3025891]
60. Yuan M, Liu T, Yang C. Exploring the relationship among human activities, COVID-19 morbidity, and at-risk areas using location-based social media data: knowledge about the early pandemic stage in Wuhan. Int J Environ Res Public Health 2022 May 27;19(11):6523 [FREE Full text] [doi: 10.3390/ijerph19116523] [Medline: 35682104]
61. Gui X, Kou Y, Pine K, Ladaw E, Kim H, Suzuki-Gill E, et al. Multidimensional risk communication: public discourse on risks during an emerging epidemic. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC URL: https://doi.org/10.1145/3173574.3173788 [doi: 10.1145/3173574.3173788]
62. Souza RC, de Brito DE, Cardoso RL, de Oliveira DM, Meira WJ, Pappa GL. An evolutionary methodology for handling data scarcity and noise in monitoring real events from social media data. In: Proceedings of the 14th Ibero-American Conference on AI. 2014 Presented at: IBERAMIA 2014; November 24-27, 2014; Santiago de Chile, Chile. [doi: 10.1007/978-3-319-12027-0_24]
63. Souza RC, Assunção RM, de Oliveira DM, de Brito DE, Meira WJ. Infection hot spot mining from social media trajectories. In: Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases. 2016 Presented at: ECML PKDD 2016; September 19-23, 2016; Riva del Garda, Italy. [doi: 10.1007/978-3-319-46227-1_46]
64. Zhang F, Luo J, Li C, Wang X, Zhao Z. Detecting and analyzing influenza epidemics with social media in China. In: Proceedings of the 18th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining. 2014 Presented at: PAKDD 2014; May 13-16, 2014; Tainan, Taiwan. [doi: 10.1007/978-3-319-06608-0_8]
65. Culotta A. Towards detecting influenza epidemics by analyzing Twitter messages. In: Proceedings of the First Workshop on Social Media Analytics. 2010 Presented at: SOMA '10; July 25-28, 2010; Washington, DC. [doi: 10.1145/1964858.1964874]
66. Wakamiya S, Kawai Y, Aramaki E. After the boom no one tweets: microblog-based influenza detection incorporating indirect information. In: Proceedings of the Sixth International Conference on Emerging Databases: Technologies, Applications, and Theory. 2016 Presented at: EDB '16; October 17-19, 2016; Jeju, Republic of Korea. [doi: 10.1145/3007818.3007822]
67. Hassan Zadeh A, Zolbanin HM, Sharda R, Delen D. Social media for nowcasting flu activity: spatio-temporal big data analysis. Inf Syst Front 2019 Jan 5;21(4):743-760. [doi: 10.1007/s10796-018-9893-0]
68. Achrekar H, Gandhe A, Lazarus R, Yu SH, Liu B. Online social networks flu trend tracker: a novel sensory approach to predict flu trends. In: Proceedings of the 5th International Joint Conference on Biomedical Engineering Systems and Technologies. 2012 Presented at: BIOSTEC 2012; February 1-4, 2012; Vilamoura, Portugal. [doi: 10.1007/978-3-642-38256-7_24]
69. Aramaki E, Maskawa S, Morita M. Twitter catches the flu: detecting influenza epidemics using Twitter. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2011 Presented at: EMNLP '11; July 27-31, 2011; Edinburgh, UK.
70. Zhang K, Arablouei R, Jurdak R. Predicting prevalence of influenza-like illness from geo-tagged tweets. In: Proceedings of the 26th International Conference on World Wide Web Companion. 2017 Presented at: WWW '17 Companion; April 3-7, 2017; Perth, Australia. [doi: 10.1145/3041021.3051150]

71. Lamb A, Paul MJ, Dredze M. Separating fact from fear: tracking flu infections on Twitter. In: Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2013 Presented at: NAACL HLT 2013; June 9-15, 2013; Atlanta, GA.
72. Sun X, Ye J, Ren F. Detecting influenza states based on hybrid model with personal emotional factors from social networks. Neurocomputing 2016 Oct;210:257-268. [doi: 10.1016/j.neucom.2016.01.107]
73. Sun X, Ye J, Ren F. Hybrid model based influenza detection with sentiment analysis from social networks. In: Proceedings of the 4th National Conference on Social Media Processing. 2015 Presented at: SMP 2015; November 16-17, 2015; Guangzhou, China. [doi: 10.1007/978-981-10-0080-5_5]
74. Yousefinaghani S, Dara R, Poljak Z, Bernardo TM, Sharif S. The assessment of Twitter's potential for outbreak detection: avian influenza case study. Sci Rep 2019 Dec 03;9(1):18147 [FREE Full text] [doi: 10.1038/s41598-019-54388-4] [Medline: 31796768]
75. Rudra K, Sharma A, Ganguly N, Imran M. Classifying and summarizing information from microblogs during epidemics. Inf Syst Front 2018 Mar 20;20(5):933-948 [FREE Full text] [doi: 10.1007/s10796-018-9844-9] [Medline: 32214879]
76. Rudra K, Sharma A, Ganguly N, Imran M. Classifying information from microblogs during epidemics. In: Proceedings of the 2017 International Conference on Digital Health. 2017 Presented at: DH '17; July 2-5, 2017; London, UK. [doi: 10.1145/3079452.3079491]
77. Ghosh S, Rudra K, Ghosh S, Ganguly N, Podder S, Balani N, et al. Identifying multi-dimensional information from microblogs during epidemics. In: Proceedings of the ACM India Joint International Conference on Data Science and Management of Data. 2019 Presented at: CODS-COMAD '19; January 3-5, 2019; Kolkata, India. [doi: 10.1145/3297001.3297030]
78. Collier N, Son NT, Nguyen NM. OMG U got flu? Analysis of shared health messages for bio-surveillance. J Biomed Sem 2011;2(Suppl 5):S9. [doi: 10.1186/2041-1480-2-s5-s9]
79. Sousa L, de Mello R, Cedrim D, Garcia A, Missier P, Uchôa A, et al. VazaDengue: an information system for preventing and combating mosquito-borne diseases with social networks. Inf Syst 2018 Jun;75:26-42. [doi: 10.1016/j.is.2018.02.003]
80. Jain VK, Kumar S. Effective surveillance and predictive mapping of mosquito-borne diseases using social media. J Comput Sci 2018 Mar;25:406-415. [doi: 10.1016/j.jocs.2017.07.003]
81. Hong Y, Sinnott RO. A social media platform for infectious disease analytics. In: Proceedings of the 18th International Conference on Computational Science and Its Applications. 2018 Presented at: ICCSA 2018; July 2-5, 2018; Melbourne, Australia. [doi: 10.1007/978-3-319-95162-1_36]
82. Swain S, Seeja KR. Analysis of epidemic outbreak in Delhi using social media data. In: Proceedings of the Second International Conference on Information, Communication and Computing Technology. 2017 Presented at: ICICCT 2017; May 13, 2017; New Delhi, India. [doi: 10.1007/978-981-10-6544-6_3]
83. Comito C. Sensing social media to forecast COVID-19 cases. In: Proceedings of the IEEE Symposium on Computers and Communications. 2022 Presented at: ISCC 2022; June 30-July 3, 2022; Rhodes, Greece. [doi: 10.1109/iscc55528.2022.9913033]
84. Perera D, Bamunusinghe J. Contact tracing of covid-19 patients using tweets. In: Proceedings of the 2nd International Conference on Advanced Research in Computing. 2022 Presented at: ICARC 2022; February 23-24, 2022; Belihuloya, Sri Lanka. [doi: 10.1109/icarc54489.2022.9753859]
85. Lazebnik T, Bunimovich-Mendrazitsky S, Ashkenazi S, Levner E, Benis A. Early detection and control of the next epidemic wave using health communications: development of an artificial intelligence-based tool and its validation on COVID-19 data from the US. Int J Environ Res Public Health 2022 Nov 30;19(23):16023 [FREE Full text] [doi: 10.3390/ijerph192316023] [Medline: 36498096]
86. Mundotiya RK, Baruah R, Srivastava B, Singh AK. NLPRL at WNUT-2020 task 2: ELMo-based system for identification of COVID-19 tweets. In: Proceedings of the Sixth Workshop on Noisy User-Generated Text. 2020 Presented at: W-NUT 2020; November 19, 2020; Online. [doi: 10.18653/v1/2020.wnut-1.60]
87. Wang Y, Hao H, Platt LS. Examining risk and crisis communications of government agencies and stakeholders during early-stages of COVID-19 on Twitter. Comput Human Behav 2021 Jan;114:106568 [FREE Full text] [doi: 10.1016/j.chb.2020.106568] [Medline: 32982038]
88. Cornelius J, Ellendorff T, Furrer L, Rinaldi F. COVID-19 twitter monitor: aggregating and visualizing COVID-19 related trends in social media. In: Proceedings of the Fifth Social Media Mining for Health Applications Workshop & Shared Task. 2020 Presented at: SMM4H@COLING 2020; December 12, 2020; Barcelona, Spain. [doi: 10.18653/v1/w19-3208]
89. Klein AZ, Magge A, O'Connor K, Flores Amaro JI, Weissenbacher D, Gonzalez Hernandez G. Toward using Twitter for tracking COVID-19: a natural language processing pipeline and exploratory data set. J Med Internet Res 2021 Jan 22;23(1):e25314 [FREE Full text] [doi: 10.2196/25314] [Medline: 33449904]
90. Zhou Y, Jiang JY, Chen X, Wang W. #StayHome or #Marathon?: social media enhanced pandemic surveillance on spatial-temporal dynamic graphs. In: Proceedings of the 30th ACM International Conference on Information & Knowledge Management. 2021 Presented at: CIKM '21; November 1-5, 2021; Virtual Event. [doi: 10.1145/3459637.3482222]

91. Lwowski B, Najafirad P. COVID-19 surveillance through Twitter using self-supervised and few shot learning. In: Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. 2020 Presented at: COVID-19@EMNLP 2020; November 19, 2020; Virtual Event. [doi: 10.18653/v1/2020.nlpcovid19-2.9]
92. Amin S, Uddin MI, Zeb MA, Alarood AA, Mahmoud M, Alkinani MH. Detecting dengue/flu infections based on tweets using LSTM and word embedding. IEEE Access 2020;8:189054-189068. [doi: 10.1109/access.2020.3031174]
93. Khatua A, Khatua A, Cambria E. A tale of two epidemics: contextual Word2Vec for classifying Twitter streams during outbreaks. Inf Process Manage 2019 Jan;56(1):247-257. [doi: 10.1016/j.ipm.2018.10.010]
94. Yu X, Xie Z, Mashhadi A, Hong L. Multi-task models for multi-faceted classification of pandemic information on social media. In: Proceedings of the 14th ACM Web Science Conference 2022. 2022 Presented at: WebSci '22; June 26-29, 2022; Barcelona, Spain. [doi: 10.1145/3501247.3531552]
95. Wang CK, Singh O, Tang ZL, Dai HJ. Using a recurrent neural network model for classification of tweets conveyed influenza-related information. In: Proceedings of the International Workshop on Digital Disease Detection using Social Media 2017. 2017 Presented at: DDDSM-2017; November 27, 2017; Taipei, Taiwan.
96. Kia MA, Khaksefidi FE. Twitter flu trend: a hybrid deep neural network for tweet analysis. In: Proceedings of the 42nd SGAI International Conference on Artificial Intelligence. 2022 Presented at: AI 2022; December 13-15, 2022; Cambridge, UK. [doi: 10.1007/978-3-031-21441-7_3]
97. Zhao L, Chen J, Chen F, Jin F, Wang W, Lu CT, et al. Online flu epidemiological deep modeling on disease contact network. Geoinformatica 2019 Jul 25;24(2):443-475. [doi: 10.1007/s10707-019-00376-9]
98. Huo HF, Jing SL, Wang XY, Xiang H. Modeling and analysis of a H1N1 model with relapse and effect of Twitter. Physica A 2020 Dec;560:125136. [doi: 10.1016/j.physa.2020.125136]
99. Huo HF, Zhang XM. Modeling the influence of Twitter in reducing and increasing the spread of influenza epidemics. Springerplus 2016 Jan 27;5(1):88 [FREE Full text] [doi: 10.1186/s40064-016-1689-4] [Medline: 26848428]
100. Nie Q, Liu Y, Zhang D, Jiang H. Dynamical SEIR model with information entropy using COVID-19 as a case study. IEEE Trans Comput Soc Syst 2021 Aug;8(4):946-954. [doi: 10.1109/tcss.2020.3046712]
101. Spurlock K, Elgazzar H. Predicting COVID-19 infection groups using social networks and machine learning algorithms. In: Proceedings of the 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference. 2020 Presented at: UEMCO; October 28-31, 2020; New York, NY. [doi: 10.1109/uemcon51285.2020.9298093]
102. Haupt MR, Jinich-Diamant A, Li J, Nali M, Mackey TK. Characterizing Twitter user topics and communication network dynamics of the “liberate” movement during COVID-19 using unsupervised machine learning and social network analysis. Online Soc Netw Media 2021 Jan;21:100114. [doi: 10.1016/j.osnem.2020.100114]
103. Bagavathi A, Krishnan S. Social sensors early detection of contagious outbreaks in social media. In: Ahram T, editor. Advances in Artificial Intelligence, Software and Systems Engineering. Cham, Switzerland: Springer; 2019.
104. Martín G, Marinescu MC, Singh DE, Carretero J. Leveraging social networks for understanding the evolution of epidemics. BMC Syst Biol 2011 Dec 23;5(S3):S14. [doi: 10.1186/1752-0509-5-s3-s14]
105. Zheng H, Goh DH, Lee EW, Lee CS, Theng YL. Uncovering topics related to COVID-19 pandemic on Twitter. In: Proceedings of the 22nd International Conference on Asia-Pacific Digital Libraries. 2020 Presented at: ICADL 2020; November 30-December 1, 2020; Kyoto, Japan. [doi: 10.1007/978-3-030-64452-9_28]
106. Shin M, Han S, Park S, Cha M. A risk communication event detection model via contrastive learning. In: Proceedings of the 3rd NLP4IF Workshop on NLP for Internet Freedom. 2020 Presented at: NLP4IF 2020; December 12, 2020; Barcelona, Spain.
107. Chen L, Hossain KS, Butler P, Ramakrishnan N, Prakash BA. Flu gone viral: syndromic surveillance of flu on Twitter using temporal topic models. In: Proceedings of the IEEE International Conference on Data Mining. 2014 Presented at: ICDM 2014; December 14-17, 2014; Shenzhen, China. [doi: 10.1109/icdm.2014.137]
108. Fu Q, Hu C, Xu W, He X, Zhang T. Detect and analyze flu outlier events via social network. In: Proceedings of the Web Technologies and Applications. 2014 Presented at: APWeb 2014; September 5, 2014; Changsha, China. [doi: 10.1007/978-3-319-11119-3_13]
109. Nolasco D, Oliveira J. Subevents detection through topic modeling in social media posts. Future Gener Comput Syst 2019 Apr;93:290-303. [doi: 10.1016/j.future.2018.09.008]
110. Zheng H, Goh DH, Lee CS, Lee EW, Theng YL. Uncovering temporal differences in COVID-19 tweets. Proc Assoc Inf Sci Technol 2020 Oct 22;57(1):e233 [FREE Full text] [doi: 10.1002/pra2.233] [Medline: 33173810]
111. Missier P, Romanovsky A, Miu T, Pal A, Daniilakis M, Garcia A, et al. Tracking dengue epidemics using Twitter content classification and topic modelling. In: Proceedings of the Current Trends in Web Engineering. 2016 Presented at: ICWE 2016; June 6-9, 2016; Lugano, Switzerland. [doi: 10.1007/978-3-319-46963-8_7]
112. Odlum M, Yoon S. What can we learn about the Ebola outbreak from tweets? Am J Infect Control 2015 Jun;43(6):563-571 [FREE Full text] [doi: 10.1016/j.ajic.2015.02.023] [Medline: 26042846]
113. Diaz-Aviles E, Stewart A. Tracking Twitter for epidemic intelligence: case study: EHEC/HUS outbreak in Germany, 2011. In: Proceedings of the 4th Annual ACM Web Science Conference. 2012 Presented at: WebSci '12; June 22-24, 2012; Evanston, Illinois. [doi: 10.1145/2380718.2380730]

114. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov 29;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
115. Lehmann BA, Ruiter RA, Kok G. A qualitative study of the coverage of influenza vaccination on Dutch news sites and social media websites. BMC Public Health 2013 Jun 05;13:547 [FREE Full text] [doi: 10.1186/1471-2458-13-547] [Medline: 23738769]
116. Kobayashi R, Takedomi Y, Nakayama Y, Suda T, Uno T, Hashimoto T, et al. Evolution of public opinion on COVID-19 vaccination in Japan: large-scale Twitter data analysis. J Med Internet Res 2022 Dec 22;24(12):e41928 [FREE Full text] [doi: 10.2196/41928] [Medline: 36343186]
117. Martin S, Vanderslott S. "Any idea how fast 'It's just a mask!' can turn into 'It's just a vaccine!'": from mask mandates to vaccine mandates during the COVID-19 pandemic. Vaccine 2022 Dec 05;40(51):7488-7499 [FREE Full text] [doi: 10.1016/j.vaccine.2021.10.031] [Medline: 34823912]
118. Zhang C, Xu S, Li Z, Liu G, Dai D, Dong C. The evolution and disparities of online attitudes toward COVID-19 vaccines: year-long longitudinal and cross-sectional study. J Med Internet Res 2022 Jan 21;24(1):e32394 [FREE Full text] [doi: 10.2196/32394] [Medline: 34878410]
119. Argyris YA, Zhang N, Bashyal B, Tan PN. Using deep learning to identify linguistic features that facilitate or inhibit the propagation of anti- and pro-vaccine content on social media. In: Proceedings of the IEEE International Conference on Digital Health. 2022 Presented at: ICDH 2022; July 10-16, 2022; Barcelona, Spain. [doi: 10.1109/icdh55609.2022.00025]
120. Islam T, Goldwasser D. Understanding COVID-19 vaccine campaign on Facebook using minimal supervision. In: Proceedings of the IEEE International Conference on Big Data. 2022 Presented at: BigData 2022; December 17-20, 2022; Osaka, Japan. [doi: 10.1109/bigdata55660.2022.10021123]
121. Durmaz N, Hengirmen E. The dramatic increase in anti-vaccine discourses during the COVID-19 pandemic: a social network analysis of Twitter. Hum Vaccin Immunother 2022 Dec 31;18(1):2025008 [FREE Full text] [doi: 10.1080/21645515.2021.2025008] [Medline: 35113767]
122. Basch CH, Meleo-Erwin Z, Fera J, Jaime C, Basch CE. A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok. Hum Vaccin Immunother 2021 Aug 03;17(8):2373-2377 [FREE Full text] [doi: 10.1080/21645515.2021.1894896] [Medline: 33764283]
123. Kariyapperuma KR, Banujan K, Wijeratna PM, Kumara BT. Classification of covid19 vaccine-related tweets using deep learning. In: Proceedings of the International Conference on Data Analytics for Business and Industry. 2022 Presented at: ICDABI 2022; October 25-26, 2022; Sakhir, Bahrain. [doi: 10.1109/icdabi56818.2022.10041615]
124. Stella M, Vitevitch MS, Botta F. Cognitive networks extract insights on COVID-19 vaccines from English and Italian popular tweets: anticipation, logistics, conspiracy and loss of trust. Big Data Cogn Comput 2022 May 12;6(2):52. [doi: 10.3390/bdcc6020052]
125. Jabalameli S, Xu Y, Shetty S. Spatial and sentiment analysis of public opinion toward COVID-19 pandemic using twitter data: at the early stage of vaccination. Int J Disaster Risk Reduct 2022 Oct 01;80:103204 [FREE Full text] [doi: 10.1016/j.ijdrr.2022.103204] [Medline: 35935613]
126. Feizollah A, Anuar NB, Mehdi R, Firdaus A, Sulaiman A. Understanding COVID-19 halal vaccination discourse on Facebook and Twitter using aspect-based sentiment analysis and text emotion analysis. Int J Environ Res Public Health 2022 May 21;19(10):6269 [FREE Full text] [doi: 10.3390/ijerph19106269] [Medline: 35627806]
127. Khan K, Yadav S. Sentiment analysis on covid-19 vaccine using Twitter data: a NLP approach. In: Proceedings of the IEEE 9th Region 10 Humanitarian Technology Conference. 2021 Presented at: R10-HTC 2021; September 30-October 2, 2021; Bangalore, India. [doi: 10.1109/r10-htc53172.2021.9641515]
128. Ji X, Chun SA, Wei Z, Geller J. Twitter sentiment classification for measuring public health concerns. Soc Netw Anal Min 2015;5(1):13 [FREE Full text] [doi: 10.1007/s13278-015-0253-5] [Medline: 32226558]
129. Wang S, Schraagen M, Sang ET, Dastani M. Public sentiment on governmental COVID-19 measures in Dutch social media. In: Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. 2020 Presented at: EMNLP 2020; November 16-20, 2020; Online. [doi: 10.18653/v1/2020.nlpcovid19-2.17]
130. Zhunis A, Lima G, Song H, Han J, Cha M. Emotion bubbles: emotional composition of online discourse before and after the COVID-19 outbreak. In: Proceedings of the ACM Web Conference 2022. 2022 Presented at: WWW '22; April 25-29, 2022; Virtual event. [doi: 10.1145/3485447.3512132]
131. Muñoz LM, Ramirez MF, Camargo JE. A data-driven method for measuring the negative impact of sentiment towards China in the context of COVID-19. In: Proceedings of the Third International Conference on Applied Informatics. 2020 Presented at: ICAI 2020; October 29-31, 2020; Ota, Nigeria. [doi: 10.1007/978-3-030-61702-8_15]
132. Shah CS, Sebastian MP. Sentiment analysis and topic modelling of Indian government’s Twitter handle #IndiaFightsCorona. In: Proceedings of the International Conference on Transfer and Diffusion of IT. 2020 Presented at: TDIT 2020; December 18-19, 2020; Tiruchirappalli, India. [doi: 10.1007/978-3-030-64861-9_30]
133. Garcia K, Berton L. Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA. Appl Soft Comput 2021 Mar;101:107057 [FREE Full text] [doi: 10.1016/j.asoc.2020.107057] [Medline: 33519326]

134. Massaro M, Tamburro P, La Torre M, Dal Mas F, Thomas R, Cobianchi L, et al. Non-pharmaceutical interventions and the infodemic on Twitter: lessons learned from Italy during the covid-19 pandemic. J Med Syst 2021 Mar 06;45(4):50 [FREE Full text] [doi: 10.1007/s10916-021-01726-7] [Medline: 33675427]
135. Shi J, Li W, Yang Y, Yao N, Bai Q, Yongchareon S, et al. Automated concern exploration in pandemic situations - COVID-19 as a use case. In: Proceedings of the 17th Pacific Rim Knowledge Acquisition Workshop. 2020 Presented at: PKAW 2020; January 7-8, 2021; Yokohama, Japan. [doi: 10.1007/978-3-030-69886-7_15]
136. Chen L, Huang X, Zhang H, Niu B. Covid-19 public opinion analysis based on LDA topic modeling and data visualization. In: Proceedings of the Third International Conference on Machine Learning for Cyber Security. 2020 Presented at: ML4CS 2020; October 8-10, 2020; Guangzhou, China. [doi: 10.1007/978-3-030-62460-6_20]
137. Zhu B, Zheng X, Liu H, Li J, Wang P. Analysis of spatiotemporal characteristics of big data on social media sentiment with COVID-19 epidemic topics. Chaos Solitons Fractals 2020 Nov;140:110123 [FREE Full text] [doi: 10.1016/j.chaos.2020.110123] [Medline: 32834635]
138. Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z. Top concerns of tweeters during the COVID-19 pandemic: infoveillance study. J Med Internet Res 2020 Apr 21;22(4):e19016 [FREE Full text] [doi: 10.2196/19016] [Medline: 32287039]
139. Whitfield C, Liu Y, Anwar M. Surveillance of COVID-19 pandemic using social media: a reddit study in North Carolina. In: Proceedings of the 12th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics. 2021 Presented at: BCB '21; August 1-4, 2021; Gainesville, FL. [doi: 10.1145/3459930.3469550]
140. Kirabo L, Namara M, Mcneese N. The power of the blue tick (): Ugandans' experiences and engagement on Twitter at the onset of the COVID-19 pandemic. In: Proceedings of the 3rd African Human-Computer Interaction Conference: Inclusiveness and Empowerment. 2021 Presented at: AfriCHI '21; March 8-12, 2021; Maputo, Mozambique. [doi: 10.1145/3448696.3448746]
141. Thakur N, Han CY. An exploratory study of tweets about the SARS-CoV-2 Omicron variant: insights from sentiment analysis, language interpretation, source tracking, type classification, and embedded URL detection. COVID 2022 Jul 25;2(8):1026-1049. [doi: 10.3390/covid2080076]
142. Avasthi S, Chauhan R, Acharjya DP. Information extraction and sentiment analysis to gain insight into the COVID-19 crisis. In: Proceedings of the International Conference on Innovative Computing and Communications. 2021 Presented at: ICICC 2021; February 20-21, 2021; Delhi, India. [doi: 10.1007/978-981-16-2594-7_28]
143. Sabareesha SS, Bhattacharjee S, Shetty RD. Pattern analysis of COVID-19 based on geotagged social media data with sociodemographic factors. In: Proceedings of the 27th International Conference on Automation and Computing. 2022 Presented at: ICAC 2022; September 1-3, 2022; Bristol, UK. [doi: 10.1109/icac55051.2022.9911119]
144. Hoque MU, Lee K, Beyer JL, Curran SR, Gonser KS, Lam NS, et al. Analyzing tweeting patterns and public engagement on Twitter during the recognition period of the COVID-19 pandemic: a study of two U.S. States. IEEE Access 2022;10:72879-72894. [doi: 10.1109/access.2022.3189670]
145. Koukaras P, Tjortjis C, Rousidis D. Mining association rules from COVID-19 related Twitter data to discover word patterns, topics and inferences. Inf Syst 2022 Nov;109:102054 [FREE Full text] [doi: 10.1016/j.is.2022.102054] [Medline: 36569358]
146. Comito C. How COVID-19 information spread in U.S.? The role of Twitter as early indicator of epidemics. IEEE Trans Serv Comput 2022 May 1;15(3):1193-1205. [doi: 10.1109/tsc.2021.3091281]
147. Purnat TD, Vacca P, Burzo S, Zecchin T, Wright A, Briand S, et al. WHO digital intelligence analysis for tracking narratives and information voids in the COVID-19 infodemic. Stud Health Technol Inform 2021 May 27;281:989-993. [doi: 10.3233/SHTI210326] [Medline: 34042821]
148. Lorenzoni V, Andreozzi G, Bazzani A, Casigliani V, Pirri S, Tavoschi L, et al. How Italy tweeted about COVID-19: detecting reactions to the pandemic from social media. Int J Environ Res Public Health 2022 Jun 24;19(13):7785 [FREE Full text] [doi: 10.3390/ijerph19137785] [Medline: 35805444]
149. Pang PC, Jiang W, Pu G, Chan KS, Lau Y. Social media engagement in two governmental schemes during the COVID-19 pandemic in Macao. Int J Environ Res Public Health 2022 Jul 23;19(15):8976 [FREE Full text] [doi: 10.3390/ijerph19158976] [Medline: 35897346]
150. Abramova O, Batzel K, Modesti D. Collective response to the health crisis among German Twitter users: a structural topic modeling approach. Int J Inf Manage Data Insights 2022 Nov;2(2):100126. [doi: 10.1016/j.jjimei.2022.100126]
151. Wang T, Brooks I, Bashir M. Public reaction on social media during COVID-19: a comparison between Twitter and Weibo. In: Proceedings of the 2021 Intelligent Computing Conference. 2021 Presented at: IC 2021; July 15-16, 2021; Online. [doi: 10.1007/978-3-030-80119-9_38]
152. Bogović PK, Meštrović A, Martinčić-Ipšić S. Topic modeling for tracking COVID-19 communication on Twitter. In: Proceedings of the 28th International Conference on Information and Software Technologies. 2022 Presented at: ICIST 2022; October 13-15, 2022; Kaunas, Lithuania.
153. León-Sandoval E, Zareei M, Barbosa-Santillán LI, Falcón Morales LE, Pareja Lora A, Ochoa Ruiz G. Monitoring the emotional response to the COVID-19 pandemic using sentiment analysis: a case study in Mexico. Comput Intell Neurosci 2022;2022:4914665 [FREE Full text] [doi: 10.1155/2022/4914665] [Medline: 35634092]

154. Perera S, Perera I, Ahangama S. Exploring Twitter messages during the COVID-19 pandemic in Sri Lanka: topic modelling and emotion analysis. In: Proceedings of the 2nd International Conference on Advanced Research in Computing. 2022 Presented at: ICARC 2022; February 23-24, 2022; Belihuloya, Sri Lanka. [doi: 10.1109/icarc54489.2022.9753866]
155. Abro HU, Shah ZS, Abbasi H. Analysis of COVID-19 effects on wellbeing - study of Reddit posts using natural language processing techniques. In: Proceedings of the International Conference on Emerging Trends in Electrical, Control, and Telecommunication Engineering. 2022 Presented at: ETECTE 2022; December 2-4, 2022; Lahore, Pakistan. [doi: 10.1109/etecte55893.2022.10007300]
156. Sreeraag G, Shynu PG. A comparative analysis of tweets from the South Indian states based on COVID-19 Omicron wave. In: Proceedings of the International Conference on Computing, Communication, Security and Intelligent Systems. 2022 Presented at: IC3SIS 2022; June 23-25, 2022; Kochi, India. [doi: 10.1109/ic3sis54991.2022.9885333]
157. Fukuda S, Nanba H, Shoji H. Changes in interests and emotional responses to news coverage of coronavirus disease 2019 case numbers over time. In: Proceedings of the Joint 12th International Conference on Soft Computing and Intelligent Systems and 23rd International Symposium on Advanced Intelligent Systems. 2022 Presented at: SCIS&ISIS 2022; November 29-December 2, 2022; Ise, Japan. [doi: 10.1109/scisisis55246.2022.10001983]
158. Nandan M, Mitra S, Dey S. Tweet classification and sentiment analysis of Covid 19 epidemic by applying hybrid based techniques. In: Proceedings of the IEEE VLSI Device Circuit and System. 2022 Presented at: VLSI DCS 2022; February 26-27, 2022; Kolkata, India. [doi: 10.1109/vlsidcs53788.2022.9811438]
159. Qazi A, Qazi J, Naseer K, Zeeshan M, Khan B, Dey S. Sentiment analysis of nationwide lockdown amid COVID 19: evidence from Pakistan. In: Proceedings of the IEEE 7th International Conference on Information Technology and Digital Applications. 2022 Presented at: ICITDA 2022; November 4-5, 2022; Yogyakarta, Indonesia. [doi: 10.1109/icitda55840.2022.9971397]
160. Pitroda H. Opinion mining approach: to understand public reactions towards reopening of schools and colleges/universities in India. In: Proceedings of the 2nd International Conference on Innovative Practices in Technology and Management. 2022 Presented at: ICIPTM 2022; February 23-25, 2022; Gautam Buddha Nagar, India. [doi: 10.1109/iciptm54933.2022.9754208]
161. Chang A, Xian X, Liu MT, Zhao X. Health communication through positive and solidarity messages amid the COVID-19 pandemic: automated content analysis of Facebook uses. Int J Environ Res Public Health 2022 May 19;19(10):6159 [FREE Full text] [doi: 10.3390/ijerph19106159] [Medline: 35627696]
162. Yang W, Wu Z, Mok NY, Ma X. How to save lives with microblogs? Lessons from the usage of Weibo for requests for medical assistance during COVID-19. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; April 29-May 5, 2022; New Orleans, LA. [doi: 10.1145/3491102.3517591]
163. Maylawati DS, Ramdhani MA. Indonesian citizens’ health behavior in a pandemics: Twitter conversation analysis using latent Dirichlet allocation. In: Proceedings of the 8th International Conference on Wireless and Telematics. 2022 Presented at: ICWT 2022; July 21-22, 2022; Yogyakarta, Indonesia. [doi: 10.1109/icwt55831.2022.9935502]
164. Li L, Aldosery A, Vitiugin F, Nathan N, Novillo-Ortiz D, Castillo C, et al. The response of governments and public health agencies to COVID-19 pandemics on social media: a multi-country analysis of Twitter discourse. Front Public Health 2021 Sep 28;9:716333 [FREE Full text] [doi: 10.3389/fpubh.2021.716333] [Medline: 34650948]
165. Neves JC, de França TC, Bastos MP, de Carvalho PV, Gomes JO. Analysis of government agencies and stakeholders' Twitter communications during the first surge of COVID-19 in Brazil. Work 2022 Oct 19;73(s1):S81-S93. [doi: 10.3233/WOR-211213] [Medline: 36214022]
166. Wahbeh A, Nasralah T, Al-Ramahi M, El-Gayar O. Mining physicians' opinions on social media to obtain insights into COVID-19: mixed methods analysis. JMIR Public Health Surveill 2020 Jun 18;6(2):e19276 [FREE Full text] [doi: 10.2196/19276] [Medline: 32421686]
167. Addawood A, Alsuwailem A, Alohali A, Alajaji D, Alturki M, Alsuhaibani J, et al. Tracking and understanding public reaction during COVID-19: Saudi Arabia as a use case. In: Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. 2020 Presented at: COVID-19@EMNLP 2020; November 16-20, 2020; Online. [doi: 10.18653/v1/2020.nlpcovid19-2.24]
168. Comito C. COVID-19 concerns in US: topic detection in Twitter. In: Proceedings of the 25th International Database Engineering & Applications Symposium. 2021 Presented at: IDEAS '21; July 14-16, 2021; Montreal, QC. [doi: 10.1145/3472163.3472169]
169. Kumari KR, Ratna T, Gayathri T. Machine learning technique with spider monkey optimization for COVID-19 sentiment analysis. In: Proceedings of the International Conference on Computing, Communication and Power Technology. 2022 Presented at: IC3P 2022; January 7-8, 2022; Visakhapatnam, India. [doi: 10.1109/ic3p52835.2022.00070]
170. Paramarta V, Mailangkay A, Amalia H, Chrismas D. Public sentiment analysis of Indonesian tweets about COVID-19 vaccination using different machine learning approaches. In: Proceedings of the Seventh International Conference on Informatics and Computing. 2022 Presented at: ICIC 2022; December 8-9, 2022; Bali, Indonesia. [doi: 10.1109/icic56845.2022.10006896]

171. Bhatia S, Alhaider M, Alarjani M. Sentiment analysis for Arabic tweets on covid-19 using computational techniques. In: Proceedings of the 12th International Conference on Cloud Computing, Data Science & Engineering. 2022 Presented at: Confluence 2022; January 27-28, 2022; Noida, India. [doi: 10.1109/confluence52989.2022.9734188]
172. Arias F, Guerra-Adames A, Zambrano ME, Quintero-Guerra E, Tejedor-Flores N. Analyzing Spanish-language public sentiment in the context of a pandemic and social unrest: the Panama case. Int J Environ Res Public Health 2022 Aug 19;19(16):10328 [FREE Full text] [doi: 10.3390/ijerph191610328] [Medline: 36011965]
173. Elsaka T, Afyouni I, Hashem I, Al Aghbari Z. Spatio-temporal sentiment mining of COVID-19 Arabic social media. ISPRS Int J Geo Inf 2022 Sep 02;11(9):476. [doi: 10.3390/ijgi11090476]
174. Rahmanti AR, Chien CH, Nursetyo AA, Husnayain A, Wiratama BS, Fuad A, et al. Social media sentiment analysis to monitor the performance of vaccination coverage during the early phase of the national COVID-19 vaccine rollout. Comput Methods Programs Biomed 2022 Jun;221:106838 [FREE Full text] [doi: 10.1016/j.cmpb.2022.106838] [Medline: 35567863]
175. Kanakaraddi S, Chikaraddi AK, Aivalli N, Maniyar J, Singh N. Sentiment analysis of Covid-19 tweets using machine learning and natural language processing. In: Pandian AP, Narayanan M, Palanisamy R, Senjyu T, editors. Proceedings of Third International Conference on Intelligent Computing, Information and Control Systems. Singapore, Singapore: Springer; 2022.
176. Li L, Zhang Q, Wang X, Zhang J, Wang T, Gao TL, et al. Characterizing the propagation of situational information in social media during COVID-19 epidemic: a case study on Weibo. IEEE Trans Comput Soc Syst 2020 Apr;7(2):556-562. [doi: 10.1109/tcss.2020.2980007]
177. Guo R, Xu K. A large-scale analysis of COVID-19 Twitter dataset in a new phase of the pandemic. In: Proceedings of the IEEE 12th International Conference on Electronics Information and Emergency Communication. 2022 Presented at: ICEIEC 2022; July 15-17, 2022; Beijing, China. [doi: 10.1109/iceiec54567.2022.9835047]
178. Buvanasri AK, Meenakshi R, Karthika S. Applications of open source intelligence in crisis analysis—a COVID-19 case study. In: Proceedings of the ICT Systems and Sustainability. 2021 Presented at: ICT4SD 2021; July 22-23, 2021; Goa, India. [doi: 10.1007/978-981-16-5987-4_32]
179. Satu MS, Khan MI, Mahmud M, Uddin S, Summers MA, Quinn JM, et al. TClustVID: a novel machine learning classification model to investigate topics and sentiment in COVID-19 tweets. Knowl Based Syst 2021 Aug 17;226:107126 [FREE Full text] [doi: 10.1016/j.knosys.2021.107126] [Medline: 33972817]
180. Mandal S, Rath M, Wang Y, Patra BG. Predicting Zika prevention techniques discussed on Twitter: an exploratory study. In: Proceedings of the 2018 Conference on Human Information Interaction & Retrieval. 2018 Presented at: CHIIR '18; March 11-15, 2018; New Brunswick, NJ. [doi: 10.1145/3176349.3176874]
181. Al-Ramahi M, Elnoshokaty A, El-Gayar O, Nasralah T, Wahbeh A. Public discourse against masks in the COVID-19 era: infodemiology study of Twitter data. JMIR Public Health Surveill 2021 Apr 05;7(4):e26780 [FREE Full text] [doi: 10.2196/26780] [Medline: 33720841]
182. Liang G, Zhao J, Lau HY, Leung CW. Using social media to analyze public concerns and policy responses to COVID-19 in Hong Kong. ACM Trans Manag Inf Syst 2021 Sep 08;12(4):1-20. [doi: 10.1145/3460124]
183. Li L, Hua L, Gao F. Int J Environ Res Public Health 2022 Dec 31;20(1):780 [FREE Full text] [doi: 10.3390/ijerph20010780] [Medline: 36613100]
184. Mahor K, Manjhvar AK. Public sentiment assessment of coronavirus-specific tweets using a transformer-based BERT classifier. In: Proceedings of the International Conference on Edge Computing and Applications. 2022 Presented at: ICECAA 2022; October 13-15, 2022; Tamilnadu, India. [doi: 10.1109/icecaa55415.2022.9936448]
185. Rifat N, Ahsan M, Gomes R, Chowdhury M. COVID-19 sentiment analysis applying BERT. In: Proceedings of the IEEE International Conference on Electro Information Technology. 2022 Presented at: eIT 2022; May 19-21, 2022; Mankato, MN. [doi: 10.1109/eit53891.2022.9813777]
186. Dewangan A, Sharma A, Badholia A. COVID-19 tweet analysis using deep convolutional neural network (DCNN)-BERT. In: Proceedings of the 4th International Conference on Inventive Research in Computing Applications. 2022 Presented at: ICIRCA 2022; September 21-23, 2022; Coimbatore, India. [doi: 10.1109/icirca54612.2022.9985742]
187. Stoikos S, Izbicki M. Multilingual emoticon prediction of tweets about COVID-19. In: Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotion's in Social Media. 2020 Presented at: PEOPLES 2020; December 13, 2020; Online.
188. Feldman P, Tiwari S, Cheah CS, Foulds JR, Pan S. Analyzing COVID-19 tweets with transformer-based language models. arXiv Preprint posted online on April 20, 2021 [FREE Full text]
189. Ghanem A, Asaad C, Hafidi H, Moukafih Y, Guermah B, Sbihi N, et al. Real-time infoveillance of moroccan social media users' sentiments towards the COVID-19 pandemic and its management. Int J Environ Res Public Health 2021 Nov 19;18(22):12172 [FREE Full text] [doi: 10.3390/ijerph182212172] [Medline: 34831927]
190. Lyu X, Chen Z, Wu D, Wang W. Sentiment analysis on Chinese Weibo regarding COVID-19. In: Proceedings of the 9th CCF International Conference on Natural Language Processing and Chinese Computing. 2020 Presented at: NLPCC 2020; October 14-18, 2020; Zhengzhou, China. [doi: 10.1007/978-3-030-60450-9_56]

191. Pan Z, Zhou S, Sun N. Twitter sentiment analysis of COVID-19 vaccine based on BiLSTM with attention mechanism. In: Proceedings of the 4th International Conference on Advances in Computer Technology, Information Science and Communications. 2022 Presented at: CTISC 2022; April 22-24, 2022; Suzhou, China. [doi: 10.1109/ctisc54888.2022.9849814]
192. De Santis E, Martino A, Rizzi A. An infoveillance system for detecting and tracking relevant topics from Italian tweets during the COVID-19 event. IEEE Access 2020;8:132527-132538. [doi: 10.1109/access.2020.3010033]
193. Alsudias L, Rayson P. COVID-19 and Arabic Twitter: how can Arab world governments and public health organizations learn from social media? In: Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020. 2020 Presented at: NLP COVID-19 ACL 2020; July 5-10, 2020; Online.
194. Umer M, Sadiq S, Karamti H, Abdulmajid Eshmawi A, Nappi M, Usman Sana M, et al. ETCNN: extra tree and convolutional neural network-based ensemble model for COVID-19 tweets sentiment classification. Pattern Recognit Lett 2022 Dec;164:224-231 [FREE Full text] [doi: 10.1016/j.patrec.2022.11.012] [Medline: 36407854]
195. Malo S, Bayala TR, Kinda Z. French COVID-19 tweets classification using FlauBERT layers. In: Proceedings of the Congress on Intelligent Systems. 2021 Presented at: CIS 2021; September 4-5, 2021; Bengaluru, India. [doi: 10.1007/978-981-16-9416-5_59]
196. Miao L, Last M, Litvak M. Tracking social media during the COVID-19 pandemic: the case study of lockdown in New York State. Expert Syst Appl 2022 Jan;187:115797 [FREE Full text] [doi: 10.1016/j.eswa.2021.115797] [Medline: 34566273]
197. Luo L, Wang Y, Liu H. COVID-19 personal health mention detection from tweets using dual convolutional neural network. Expert Syst Appl 2022 Aug 15;200:117139 [FREE Full text] [doi: 10.1016/j.eswa.2022.117139] [Medline: 35399189]
198. Guidry JP, Jin Y, Orr CA, Messner M, Meganck S. Ebola on Instagram and Twitter: how health organizations address the health crisis in their social media engagement. Public Relat Rev 2017 Sep;43(3):477-486. [doi: 10.1016/j.pubrev.2017.04.009]
199. Roy M, Moreau N, Rousseau C, Mercier A, Wilson A, Atlani-Duault L. Ebola and localized blame on social media: analysis of Twitter and Facebook conversations during the 2014-2015 Ebola epidemic. Cult Med Psychiatry 2020 Mar 18;44(1):56-79 [FREE Full text] [doi: 10.1007/s11013-019-09635-8] [Medline: 31214902]
200. Chetty N, Alathur S, Kumar V. 2019-nCoV disease control and rehabilitation: insights from Twitter analytics. In: Proceedings of the 5th International Conference on Computing, Communication and Security. 2020 Presented at: ICCCS 2020; October 14-16, 2020; Patna, India. [doi: 10.1109/icccs49678.2020.9277142]
201. Shanthakumar SG, Seetharam A, Ramesh A. Understanding the socio-economic disruption in the United States during COVID-19's early days. arXiv Preprint posted online on April 11, 2020 [FREE Full text]
202. Niknam F, Samadbeik M, Fatehi F, Shirdel M, Rezazadeh M, Bastani P. COVID-19 on Instagram: a content analysis of selected accounts. Health Policy Technol 2021 Mar;10(1):165-173 [FREE Full text] [doi: 10.1016/j.hlpt.2020.10.016] [Medline: 33723502]
203. Essam BA, Abdo MS. How do Arab tweeters perceive the COVID-19 pandemic? J Psycholinguist Res 2021 Jun 14;50(3):507-521 [FREE Full text] [doi: 10.1007/s10936-020-09715-6] [Medline: 32797330]
204. Ölcer S, Yilmaz-Aslan Y, Brzoska P. Lay perspectives on social distancing and other official recommendations and regulations in the time of COVID-19: a qualitative study of social media posts. BMC Public Health 2020 Jun 19;20(1):963 [FREE Full text] [doi: 10.1186/s12889-020-09079-5] [Medline: 32560716]
205. Yigitcanlar T, Kankanamge N, Preston A, Gill PS, Rezayee M, Ostadnia M, et al. How can social media analytics assist authorities in pandemic-related policy decisions? Insights from Australian states and territories. Health Inf Sci Syst 2020 Dec 15;8(1):37 [FREE Full text] [doi: 10.1007/s13755-020-00121-9] [Medline: 33078073]
206. Xia H, An W, Li J, Zhang ZJ. Outlier knowledge management for extreme public health events: understanding public opinions about COVID-19 based on microblog data. Socioecon Plann Sci 2022 Mar;80:100941 [FREE Full text] [doi: 10.1016/j.seps.2020.100941] [Medline: 32921839]
207. Huang F, Ding H, Liu Z, Wu P, Zhu M, Li A, et al. How fear and collectivism influence public's preventive intention towards COVID-19 infection: a study based on big data from the social media. BMC Public Health 2020 Nov 16;20(1):1707 [FREE Full text] [doi: 10.1186/s12889-020-09674-6] [Medline: 33198699]
208. Machete P, Turpin M. The use of critical thinking to identify fake news: a systematic literature review. In: Responsible Design, Implementation and Use of Information and Communication Technology. Cham, Switzerland: Springer International Publishing; 2020.
209. Guo Q, Huang W, Huang K, Liu X. Information credibility: a probabilistic graphical model for identifying credible influenza posts on social media. In: Proceedings of the International Conference on Smart Health. 2015 Presented at: ICSH 2015; November 17-18, 2015; Phoenix, AZ. [doi: 10.1007/978-3-319-29175-8_12]
210. Abdelminaam D, Ismail F, Taha M, Taha A, Houssein E, Nabil A. CoAID-DEEP: an optimized intelligent framework for automated detecting COVID-19 misleading information on Twitter. IEEE Access 2021;9:27840-27867. [doi: 10.1109/access.2021.3058066]
211. Madani Y, Erritali M, Bouikhalene B. Using artificial intelligence techniques for detecting Covid-19 epidemic fake news in Moroccan tweets. Results Phys 2021 Jun;25:104266. [doi: 10.1016/j.rinp.2021.104266]
212. Cui L, Lee D. Coaid: Covid-19 healthcare misinformation dataset. arXiv Preprint posted online on May 22, 2020 [FREE Full text]

213. Hossain T, Logan RLIV, Ugarte A, Matsubara Y, Young S, Singh S. COVIDLies: detecting COVID-19 misinformation on social media. In: Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. 2020 Presented at: COVID-19@EMNLP 2020; November 16-20, 2020; Online. [doi: 10.18653/v1/2020.nlpcovid19-2.11]
214. Malla S, Alphonse PJ. Fake or real news about COVID-19? Pretrained transformer model to detect potential misleading news. Eur Phys J Spec Top 2022 Jan 13;231(18-20):3347-3356 [FREE Full text] [doi: 10.1140/epjs/s11734-022-00436-6] [Medline: 35039760]
215. Alajramy L, Jarrar R. Using artificial neural networks to identify COVID-19 misinformation. In: Proceedings of the 4th Multidisciplinary International Symposium on Disinformation in Open Online Media. 2022 Presented at: MISDOOM 2022; October 11-12, 2022; Boise, ID. [doi: 10.1007/978-3-031-18253-2_2]
216. Samantaray S, Kumar A. Bi-directional long short-term memory network for fake news detection from social media. In: Proceedings of the Intelligent and Cloud Computing. 2021 Presented at: ICICC 2021; February 20-21, 2021; Delhi, India. [doi: 10.1007/978-981-16-9873-6_42]
217. Ghayoomi M, Mousavian M. Deep transfer learning for COVID-19 fake news detection in Persian. Expert Syst 2022 Apr 03;39(8):e13008 [FREE Full text] [doi: 10.1111/exsy.13008] [Medline: 35599852]
218. Raj C, Meel P. ARCNN framework for multimodal infodemic detection. Neural Netw 2022 Feb;146:36-68 [FREE Full text] [doi: 10.1016/j.neunet.2021.11.006] [Medline: 34839091]
219. Tang C, Ma K, Cui B, Ji K, Abraham A. Long text feature extraction network with data augmentation. Appl Intell (Dordr) 2022 Apr 04;52(15):17652-17667 [FREE Full text] [doi: 10.1007/s10489-022-03185-0] [Medline: 35400845]
220. Zuo C, Banerjee R, Chaleshtori FH, Shirazi H, Ray I. Seeing should probably not be believing: the role of deceptive support in COVID-19 misinformation on Twitter. J Data Inf Quality 2022 Dec 28;15(1):1-26. [doi: 10.1145/3546914]
221. Biradar S, Saumya S, Chauhan A. Combating the infodemic: COVID-19 induced fake news recognition in social media networks. Complex Intell Systems 2023 Feb 18;9(3):2879-2891 [FREE Full text] [doi: 10.1007/s40747-022-00672-2] [Medline: 35194546]
222. Jeyasudha J, Seth P, Usha G, Tanna P. Fake information analysis and detection on pandemic in Twitter. SN Comput Sci 2022 Aug 24;3(6):456 [FREE Full text] [doi: 10.1007/s42979-022-01363-y] [Medline: 36035506]
223. Alouffi B, Alharbi A, Sahal R, Saleh H. An optimized hybrid deep learning model to detect COVID-19 misleading information. Comput Intell Neurosci 2021 Nov 15;2021(1):9615034 [FREE Full text] [doi: 10.1155/2021/9615034] [Medline: 34790233]
224. Hussna A, Trisha II, Karim MS, Alam MG. COVID-19 fake news prediction on social media data. In: Proceedings of the IEEE Region 10 Symposium. 2021 Presented at: TENSYMP 2021; August 23-25, 2021; Jeju, Republic of Korea. [doi: 10.1109/tensymp52854.2021.9550957]
225. Ajao O, Garg A, Da Costa-Abreu M. Exploring content-based and meta-data analysis for detecting fake news infodemic: a case study on COVID-19. In: Proceedings of the 12th International Conference on Pattern Recognition Systems. 2022 Presented at: ICPRS 2022; June 7-10, 2022; Saint-Etienne, France. [doi: 10.1109/icprs54038.2022.9854058]
226. Alhakami H, Alhakami W, Baz A, Faizan M, Khan MW, Agrawal A. Evaluating intelligent methods for detecting COVID-19 fake news on social media platforms. Electronics 2022 Aug 03;11(15):2417. [doi: 10.3390/electronics11152417]
227. Ali M, Murtza I, Ejaz A. Pandemic rumor identification on social networking sites: a case study of COVID-19. In: Proceedings of the 2021 5th International Conference on Natural Language Processing and Information Retrieval. 2021 Presented at: NLPIR '21; December 17-20, 2021; Sanya, China. [doi: 10.1145/3508230.3508246]
228. Maakoul O, Boucht S, Hachimi KE, Azzouzi S. Towards evaluating the COVID’19 related fake news problem: case of Morocco. In: Proceedings of the IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science. 2020 Presented at: ICECOCS 2020; December 2-3, 2020; Kenitra, Morocco. [doi: 10.1109/icecocs50124.2020.9314517]
229. Mehta V, Mishra RK. Machine learning based fake news detection on Covid-19 tweets data. In: Proceedings of International Conference on Computational Intelligence and Data Engineering. 2021 Presented at: ICCIDE 2021; August 13-14, 2021; Andhra Pradesh, India. [doi: 10.1007/978-981-16-7182-1_8]
230. Khan S, Hakak S, Deepa N, Prabadevi B, Dev K, Trelova S. Detecting COVID-19-related fake news using feature extraction. Front Public Health 2021 Jan 4;9:788074 [FREE Full text] [doi: 10.3389/fpubh.2021.788074] [Medline: 35059379]
231. Jouyandeh F, Rahmatikargar S, Rahmatikargar B, Zadeh PM. Fake news and COVID-19 vaccination: a comparative study. In: Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2021 Presented at: ASONAM '21; November 8-11, 2021; Virtual Event, Netherlands. [doi: 10.1145/3487351.3490960]
232. Bamiro B, Assayad I. Data-based automatic Covid-19 rumors detection in social networks. In: Proceedings of the Networking, Intelligent Systems and Security. 2021 Presented at: NISS 2021; April 1-2, 2021; Kenitra, Morocco. [doi: 10.1007/978-981-16-3637-0_57]
233. Zhang Y, Guo B, Ding Y, Liu J, Qiu C, Liu S, et al. Investigation of the determinants for misinformation correction effectiveness on social media during COVID-19 pandemic. Inf Process Manag 2022 May;59(3):102935 [FREE Full text] [doi: 10.1016/j.ipm.2022.102935] [Medline: 35400028]
234. Belay EG, Beyene M, Alemu T, Negash A, Tesema T, Mohammed A, et al. Towards curtailing infodemic in the era of COVID-19: a contextualized solution for Ethiopia. In: Proceedings of the HCI International 2020 – Late Breaking Papers:

Interaction, Knowledge and Social Media. 2020 Presented at: HCII 2020; July 19-24, 2020; Copenhagen, Denmark. [doi: 10.1007/978-3-030-60152-2_17]
235. Sell TK, Hosangadi D, Trotochaud M. Misinformation and the US Ebola communication crisis: analyzing the veracity and content of social media messages related to a fear-inducing infectious disease outbreak. BMC Public Health 2020 May 07;20(1):550 [FREE Full text] [doi: 10.1186/s12889-020-08697-3] [Medline: 32375715]
236. Alsudias L, Rayson P. Classifying information sources in Arabic Twitter to support online monitoring of infectious diseases. In: Proceedings of the 3rd Workshop on Arabic Corpus Linguistics. 2019 Presented at: WACL-3 2019; July 22, 2019; Cardiff, UK.
237. Jain L. An entropy-based method to control COVID-19 rumors in online social networks using opinion leaders. Technol Soc 2022 Aug;70:102048 [FREE Full text] [doi: 10.1016/j.techsoc.2022.102048] [Medline: 35765463]
238. Chen B, Chen X, Pan J, Liu K, Xie B, Wang W, et al. Dissemination and refutation of rumors during the COVID-19 outbreak in China: infodemiology study. J Med Internet Res 2021 Feb 15;23(2):e22427 [FREE Full text] [doi: 10.2196/22427] [Medline: 33493124]
239. Zeng J, Zhang Y, Ma X. Fake news detection for epidemic emergencies via deep correlations between text and images. Sustain Cities Soc 2021 Mar;66:102652 [FREE Full text] [doi: 10.1016/j.scs.2020.102652] [Medline: 36570569]
240. Atehortua NA, Patino S. COVID-19, a tale of two pandemics: novel coronavirus and fake news messaging. Health Promot Int 2021 Apr 15;36(2):524-534 [FREE Full text] [doi: 10.1093/heapro/daaa140] [Medline: 33450022]
241. Madraki G, Grasso I, Otala JM, Liu Y, Matthews J. Characterizing and comparing COVID-19 misinformation across languages, countries and platforms. In: Proceedings of the Companion Proceedings of the Web Conference 2021. 2021 Presented at: WWW '21; April 19-23, 2021; Ljubljana, Slovenia. [doi: 10.1145/3442442.3452304]
242. Raj C, Meel P. People lie, actions Don't! Modeling infodemic proliferation predictors among social media users. Technol Soc 2022 Feb;68:101930. [doi: 10.1016/j.techsoc.2022.101930]
243. Heidari M, Zad S, Hajibabaee P, Malekzadeh M, HekmatiAthar S, Uzuner O. BERT model for fake news detection based on social bot activities in the COVID-19 pandemic. In: Proceedings of the IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference. 2021 Presented at: UEMCON 2021; December 1-4, 2021; New York, NY. [doi: 10.1109/uemcon53757.2021.9666618]
244. Yan M, Lin YR, Chung WT. Are mutated misinformation more contagious? A case study of COVID-19 misinformation on Twitter. In: Proceedings of the 14th ACM Web Science Conference 2022. 2022 Presented at: WebSci '22; June 26-29, 2022; Barcelona, Spain. [doi: 10.1145/3501247.3531562]
245. Shahi GK, Dirkson A, Majchrzak TA. An exploratory study of COVID-19 misinformation on Twitter. Online Soc Netw Media 2021 Mar;22:100104 [FREE Full text] [doi: 10.1016/j.osnem.2020.100104] [Medline: 33623836]
246. Valecha R, Srinivasan SK, Volety T, Kwon KH, Agrawal M, Rao HR. Fake news sharing: an investigation of threat and coping cues in the context of the Zika virus. Digital Threats 2021 Apr 15;2(2):1-16. [doi: 10.1145/3410025]
247. Singh L, Bode L, Budak C, Kawintiranon K, Padden C, Vraga E. Understanding high- and low-quality URL sharing on COVID-19 Twitter streams. J Comput Soc Sci 2020 Nov 27;3(2):343-366 [FREE Full text] [doi: 10.1007/s42001-020-00093-6] [Medline: 33263092]
248. Sacco PL, Gallotti R, Pilati F, Castaldo N, De Domenico M. Emergence of knowledge communities and information centralization during the COVID-19 pandemic. Soc Sci Med 2021 Sep;285:114215 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114215] [Medline: 34364158]
249. Ermakova L, Nurbakova D, Ovchinnikova I. Covid or not Covid? Topic shift in information cascades on Twitter. In: Proceedings of the 3rd International Workshop on Rumours and Deception in Social Media. 2020 Presented at: RDSM 2020; December 12, 2020; Barcelona, Spain.
250. Fan Y. The social amplification of risk on Weibo-take the COVID-19 epidemic in China as an example. In: Proceedings of the 2020 Conference on Artificial Intelligence and Healthcare. 2020 Presented at: CAIH2020; October 23-25, 2020; Taiyuan, China. [doi: 10.1145/3433996.3434013]
251. Chen E, Lerman K, Ferrara E. Tracking social media discourse about the COVID-19 pandemic: development of a public coronavirus Twitter data set. JMIR Public Health Surveill 2020 May 29;6(2):e19273 [FREE Full text] [doi: 10.2196/19273] [Medline: 32427106]
252. Bahja M, Safdar GA. Unlink the link between COVID-19 and 5G networks: an NLP and SNA based approach. IEEE Access 2020;8:209127-209137. [doi: 10.1109/access.2020.3039168]
253. Ahmed W, Vidal-Alaball J, Downing J, López Seguí F. COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data. J Med Internet Res 2020 May 06;22(5):e19458 [FREE Full text] [doi: 10.2196/19458] [Medline: 32352383]
254. Shahsavari S, Holur P, Wang T, Tangherlini TR, Roychowdhury V. Conspiracy in the time of corona: automatic detection of emerging COVID-19 conspiracy theories in social media and the news. J Comput Soc Sci 2020;3(2):279-317 [FREE Full text] [doi: 10.1007/s42001-020-00086-5] [Medline: 33134595]
255. Lanier HD, Diaz MI, Saleh SN, Lehmann CU, Medford RJ. Analyzing COVID-19 disinformation on Twitter using the hashtags #scamdemic and #plandemic: retrospective study. PLoS One 2022 Jun 22;17(6):e0268409 [FREE Full text] [doi: 10.1371/journal.pone.0268409] [Medline: 35731785]

256. Ahmed W, López Seguí F, Vidal-Alaball J, Katz MS. COVID-19 and the "film your hospital" conspiracy theory: social network analysis of Twitter data. J Med Internet Res 2020 Oct 05;22(10):e22374 [FREE Full text] [doi: 10.2196/22374] [Medline: 32936771]
257. Li S, Wang Y, Xue J, Zhao N, Zhu T. The impact of COVID-19 epidemic declaration on psychological consequences: a study on active Weibo users. Int J Environ Res Public Health 2020 Mar 19;17(6):2032 [FREE Full text] [doi: 10.3390/ijerph17062032] [Medline: 32204411]
258. Fatimah N, Budi I, Santoso AB, Putra PK. Analysis of mental health during the Covid-19 pandemic in Indonesia using Twitter data. In: Proceedings of the 8th International Conference on Advanced Informatics: Concepts, Theory and Applications. 2021 Presented at: ICAICTA 2021; September 29-30, 2021; Bandung, Indonesia. [doi: 10.1109/icaicta53211.2021.9640265]
259. Davis BD, McKnight DE, Teodorescu D, Quan-Haase A, Chunara R, Fyshe A, et al. Quantifying depression-related language on social media during the COVID-19 pandemic. Int J Popul Data Sci 2020 Mar 30;5(4):1716 [FREE Full text] [doi: 10.23889/ijpds.v5i4.1716] [Medline: 35516163]
260. Wolohan JT. Estimating the effect of COVID-19 on mental health: linguistic indicators of depression during a global pandemic. In: Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020. 2020 Presented at: NLP COVID-19 ACL 2020; July 5-10, 2020; Online.
261. Chen JQ, Qi K, Zhang A, Shalaginov M, Zeng TH. COVID-19 impact on mental health analysis based on Reddit comments. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine. 2022 Presented at: BIBM 2022; December 6-8, 2022; Las Vegas, NV. [doi: 10.1109/bibm55620.2022.9995512]
262. Nandy S, Kumar V. My mind is a prison: a boosted deep learning approach to detect the rise in depression since COVID-19 using a stacked bi-LSTM CatBoost model. In: Proceedings IEEE International Conference on Big Data. 2021 Presented at: BigData 2021; December 15-18, 2021; Orlando, FL. [doi: 10.1109/bigdata52589.2021.9671482]
263. Viviani M, Crocamo C, Mazzola M, Bartoli F, Carrà G, Pasi G. Assessing vulnerability to psychological distress during the COVID-19 pandemic through the analysis of microblogging content. Future Gener Comput Syst 2021 Dec;125:446-459 [FREE Full text] [doi: 10.1016/j.future.2021.06.044] [Medline: 34934256]
264. Kumar R, Mukherjee S, Choi TM, Dhamotharan L. Mining voices from self-expressed messages on social-media: diagnostics of mental distress during COVID-19. Decis Support Syst 2022 Nov;162:113792 [FREE Full text] [doi: 10.1016/j.dss.2022.113792] [Medline: 35542965]
265. Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, et al. Mental health problems and social media exposure during COVID-19 outbreak. PLoS One 2020 Apr 16;15(4):e0231924 [FREE Full text] [doi: 10.1371/journal.pone.0231924] [Medline: 32298385]
266. Pang H. How compulsive WeChat use and information overload affect social media fatigue and well-being during the COVID-19 pandemic? A stressor-strain-outcome perspective. Telemat Inform 2021 Nov;64:101690 [FREE Full text] [doi: 10.1016/j.tele.2021.101690] [Medline: 36567817]
267. Ni MY, Yang L, Leung CM, Li N, Yao XI, Wang Y, et al. Mental health, risk factors, and social media use during the COVID-19 epidemic and cordon sanitaire among the community and health professionals in Wuhan, China: cross-sectional survey. JMIR Ment Health 2020 May 12;7(5):e19009 [FREE Full text] [doi: 10.2196/19009] [Medline: 32365044]
268. Neill R, Blair C, Best P, McGlinchey E, Armour C. Media consumption and mental health during COVID-19 lockdown: a UK cross-sectional study across England, Wales, Scotland and Northern. PsyArXiv Preprints Preprint posted online July 1, 2020 [FREE Full text] [doi: 10.31234/osf.io/5d9fc]
269. Chao M, Xue D, Liu T, Yang H, Hall BJ. Media use and acute psychological outcomes during COVID-19 outbreak in China. J Anxiety Disord 2020 Aug;74:102248 [FREE Full text] [doi: 10.1016/j.janxdis.2020.102248] [Medline: 32505918]
270. Daimer S, Mihatsch LL, Neufeld SA, Murray GK, Knolle F. Investigating the relationship of COVID-19 related stress and media consumption with schizotypy, depression, and anxiety in cross-sectional surveys repeated throughout the pandemic in Germany and the UK. Elife 2022 Jul 04;11:e75893 [FREE Full text] [doi: 10.7554/eLife.75893] [Medline: 35781372]
271. Yu L, Du M. Social networking use, mental health, and quality of life of Hong Kong adolescents during the COVID-19 pandemic. Front Public Health 2022 Oct 31;10:1040169 [FREE Full text] [doi: 10.3389/fpubh.2022.1040169] [Medline: 36388293]
272. Brailovskaia J, Margraf J. Positive mental health and mindfulness as protective factors against addictive social media use during the COVID-19 outbreak. PLoS One 2022 Nov 30;17(11):e0277631 [FREE Full text] [doi: 10.1371/journal.pone.0277631] [Medline: 36449483]
273. Zhao N, Zhou G. Social media use and mental health during the COVID-19 pandemic: moderator role of disaster stressor and mediator role of negative affect. Appl Psychol Health Well Being 2020 Dec;12(4):1019-1038 [FREE Full text] [doi: 10.1111/aphw.12226] [Medline: 32945123]
274. Lazer D, Kennedy R, King G, Vespignani A. Big data. The parable of Google Flu: traps in big data analysis. Science 2014 Mar 14;343(6176):1203-1205. [doi: 10.1126/science.1248506] [Medline: 24626916]
275. ChaiAsaad / SLR_EpidemicsAndSocialMedia. GitHub. URL: https://github.com/ChaiAsaad/SLR_EpidemicsAndSocialMedia [accessed 2024-12-24]

Abbreviations

BERT: bidirectional encoder representations from transformers
DL: deep learning
ILI: influenza-like illness
LDA: latent Dirichlet allocation
LR: logistic regression
LSTM: long short-term memory
ML: machine learning
NB: naive Bayes
PHEIC: public health emergency of international concern
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RF: random forest
RQ: research question
SVM: support vector machine
WHO: World Health Organization

Edited by A Mavragani; submitted 19.12.23; peer-reviewed by Y Wei, I Gryech, S Gordon; comments to author 06.02.24; revised version received 25.03.24; accepted 22.05.24; published 03.02.25.

Please cite as:
Asaad C, Khaouja I, Ghogho M, Baïna K
When Infodemic Meets Epidemic: Systematic Literature Review
JMIR Public Health Surveill 2025;11:e55642
URL: https://publichealth.jmir.org/2025/1/e55642
doi:10.2196/55642
PMID:

Review

Exploring the Use of Digital Educational Tools for Sexual and Reproductive Health in Sub-Saharan Africa: Systematic Review

Ramatu Hajia Abdul Hamid Alhassan[1], MPH; Catherine L Haggerty[1], MPH, PhD; Abimbola Fapohunda[2], MS, MPH, DrPH; Nabeeha Jabir Affan[1], MPH; Martina Anto-Ocrah[1,3], MPH, PhD

[1]Department of Epidemiology, University of Pittsburgh School of Public Health, Pittsburgh, PA, United States

[2]Department of Behavioral and Community Health Sciences, University of Pittsburgh School of Public Health, Pittsburgh, PA, United States

[3]Department of Medicine, Division of General Internal Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

Corresponding Author:
Martina Anto-Ocrah, MPH, PhD
Department of Medicine
Division of General Internal Medicine
University of Pittsburgh School of Medicine
230 McKee Place
Pittsburgh, PA, 15213
United States
Phone: 1 000 000 0000
Email: maa509@pitt.edu

Abstract

Background: Adolescents, particularly those in Sub-Saharan Africa, experience major challenges in getting accurate and comprehensive sexual and reproductive health (SRH) information because of sociocultural norms, stigma, and limited SRH educational resources. Digital educational tools, leveraging the widespread use of mobile phones and internet connectivity, present a promising avenue to overcome these barriers and enhance SRH education among adolescents in Sub-Saharan Africa.

Objective: We conducted a systematic review to describe (1) the geographic and demographic distributions (designated objectives 1a and 1b, respectively, given their interrelatedness) and (2) the types and relevant impacts of digital educational tools (objective 2).

Methods: We followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, using databases, such as Ovid-MEDLINE, Google Scholar, PubMed, and ERIC, to conduct literature searches. The selection criteria focused on studies that specifically addressed digital educational tools used to assess or deliver SRH education, their implementation, and their effectiveness among the adolescent population in Sub-Saharan Africa. We used the JBI critical appraisal tools for the quality assessment of papers included in the review.

Results: The review identified 22 studies across Sub-Saharan Africa that met the inclusion criteria. The 22 studies spanned populations in West, Central, East, and South Africa, with an emphasis on youth and adolescents aged 10-24 years, reflecting the critical importance of reaching these age groups with effective, accessible, and engaging health education (objectives 1a and 1b). There was a diverse range of digital tools used, including social media platforms, mobile apps, and gamified learning experiences, for a broad age range of adolescent youth. These methods were generally successful in engaging adolescents by providing them with accessible and relevant SRH information (objective 2). However, challenges, such as the digital divide, the cultural sensitivity of the material, and the necessity for a thorough examination of the long-term influence of these tools on behavior modification, were noted.

Conclusions: Digital educational tools provide great potential to improve SRH education among adolescents in Sub-Saharan Africa. These technologies can help enhance relevant health outcomes and accessibility by delivering information that is easy to understand, interesting, and tailored to their needs. Future research should focus on addressing the identified challenges, including bridging the digital divide, ensuring cultural and contextual relevance of content, and assessing the long-term impact of digital SRH education on adolescent behavior and health outcomes. Policymakers and educators are encouraged to integrate digital tools into SRH educational strategies that target adolescents in order to improve the SRH of this age group and contribute to improving public health in Sub-Saharan Africa.

(JMIR Public Health Surveill 2025;11:e63309) doi:10.2196/63309

KEYWORDS

digital health; adolescents; Africa; sexual health; reproductive health; human-centered design

Introduction

According to the World Health Organization, there are more than 1.2 billion adolescents aged 10 to 19 years with sexual and reproductive health (SRH) problems, such as sexually transmitted infections (STIs) and unplanned pregnancies, many of whom face significant barriers to accessing SRH-related education. With the increasing rates of SRH problems and mental health issues among adolescents, there is a pressing global need for effective educational tools that address the SRH needs of adolescents [1,2]. There are global disparities in adolescents' access to SRH education, with those in low- and middle-income countries (LMICs) facing the greatest barriers to accessing knowledge and resources due to sociocultural norms, beliefs, practices, and stigma, and yet, these parts of the world have the greatest burden of poor adolescent SRH outcomes [3]. For example, over 80% of all adolescents living with HIV are in Sub-Saharan Africa [4], which also has the highest prevalence of adolescent pregnancy globally [5]. Melesse et al [6] conducted a review of demographic and health survey data from 33 African countries and reported that despite improvements in SRH among adolescents in Sub-Saharan Africa, young people in this part of the world continue to have disproportionately high rates of unwanted pregnancies and STIs (including HIV/AIDS). These persistently high rates of poor health outcomes have been attributed, at least in part, to limited access to SRH education, more formally referred to as comprehensive sexuality education, in many Sub-Saharan African countries [7]. Challenges impacting access to sex education in Sub-Saharan Africa include cultural and societal norms that serve as barriers to youths' access to information, especially taboos regarding the open discussion of topics such as family planning, menstruation, and contraception [8]. Additionally, comprehensive sexuality education programs have been constrained by inconsistent funding and poor accountability [9]. It is crucial to promote SRH education to combat the challenges faced by adolescents and foster a more inclusive and equitable society as these young people transition into adulthood. Additionally, universal access to SRH is an essential component of sustainable development goals [10].

Digital educational tools have the potential to effectively provide SRH education for adolescents in Sub-Saharan Africa. Such technologies use the widespread availability of mobile phones and internet infrastructure to innovatively address the unique challenges faced by the youth in accessing current information and resources [11,12]. The use of social media platforms, websites, games, and mobile apps has become prevalent in digital education. These tools can provide instructional content with the potential to stimulate teenagers' interest in learning and promote favorable attitudes and knowledge that may lead to desirable SRH outcomes. The integration of digital tools into SRH education offers a promising avenue to overcome health disparities [13,14]. Mobile health (mHealth), which is the practice of medicine and public health that uses mobile devices, has exploded over the last few years, with mobile phone penetration reaching 90% in some LMICs [15]. The increased usage of mobile phones among younger populations in LMICs presents a valuable opportunity to use mHealth, for example, as a tool to overcome barriers to accessing SRH information and service [16]. Digital technology has the potential to close the information gap for adolescents in LMICs, ensuring equitable access to essential health knowledge for all adolescents irrespective of their geographical location or socioeconomic status.

We have conducted a systematic review to describe: (1) the geographic and demographic distribution (designated objectives 1a and 1b, respectively, given their interrelatedness) and (2) the types and formats of digital educational tools available for delivering SRH education to adolescents in Sub-Saharan Africa and, as appropriate, the impacts of these digital innovations on target populations (objective 2). A systematic review is appropriate for this work as it allows us to comprehensively synthesize the current state of the literature on such an underresearched topic, identify gaps in knowledge, and provide insights on moving the field forward [17]. The public health significance of this review lies in its potential to inform policy and practice by highlighting successful digital interventions that can enhance SRH education for adolescents. Given the increasing prevalence of internet and mobile device usage in Sub-Saharan Africa, identifying effective educational strategies is crucial for addressing the high rates of unintended pregnancies, STIs, and other reproductive health challenges faced by young people in the region. Ultimately, our findings can be used to support the development of targeted evidence-based digital interventions that can significantly improve health outcomes for adolescents in this part of the world.

Methods

Definitions of Digital Tools and Adolescents

We have defined the terms "digital tools" and "adolescents" as presented below.

Digital Tools

Digital tools are technological resources or applications used in the provision of information, support, and resources related to SRH. These tools use digital technology to enhance access, outreach, and effectiveness in the delivery of SRH education and services. They are available through a variety of digital platforms, such as websites, mobile apps, and social media platforms [11]. The use of such tools has become increasingly prevalent in recent years as they offer a convenient and efficient means of accessing information and resources related to SRH. These tools have gained considerable popularity owing to their ease of use and convenience in terms of obtaining information related to SRH.

Adolescents

Adolescents, who are individuals in the intermediate stage of development between childhood and adulthood, commonly referred to as teenagers or young adults, experience substantial transformations in physical structure, thinking, interpersonal relationships, and emotions [18]. As adolescents face major educational and health challenges, especially in the context of SRH-related information, our review is focused on this age group. We used the World Health Organization's definition of "adolescents" as individuals aged 10 to 19 years, "youth" as those aged 15 to 24 years, and "young people" as those aged 10 to 24 years [19]. Therefore, if any studies included these age groups, we erred on the side of caution and included them to ensure we were being as comprehensive as possible in capturing the breadth and scope of the research in this field.

Database and Search Dates

This systematic review adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) criteria [20] to ensure that all relevant primary materials match the review requirements [21]. Additionally, we used the JBI critical appraisal tools for the quality assessment of all papers included in the review [22] (Multimedia Appendix 1). A systematic search of peer-reviewed articles was conducted in November 2023 using Ovid-MEDLINE, with the assistance of a University of Pittsburgh Health Sciences Library System (HSLS) professional, to identify original research articles published between 2013 and 2024. The first Ovid-MEDLINE search was developed using a mix of Medical Subject Heading (MeSH) phrases, keywords, and Boolean operators (eg, "adolescent* or teen* or youth"; "digital health education/ or health promotion/ or sex education/"; "Africa or Angola or Benin or Botswana or Burkina Faso or Burundi or Cabo Verde or Cameroon or Central African Republic or Chad or Comoros or Congo or Cote d'Ivoire or Djibouti or Eritrea or Eswatini or Ethiopia or Gabon or Gambia or Ghana or Accra or Kumasi or Tamale or Sekondi-Takoradi or Guinea or Kenya or Lesotho or Liberia or Madagascar or Malawi or Mali or Mauritania or Mozambique or Namibia or Niger or Nigeria or Rwanda or Sao Tome or Principe or Senegal or Seychelles or Sierra Leone or Somalia or Sudan or Tanzania or Togo or Uganda or Zambia or Zimbabwe.ti,ab,kf.") (Multimedia Appendix 1). The search results were managed and tracked in SciWheel.

Study Selection

We restricted the search to studies conducted in humans and written in the English language. Additional articles were identified by cross-referencing bibliography lists. In alignment with the PRISMA standards, the methodology encompasses 4 phases: identification of sources, assessment of eligibility, detailed screening, and final selection for inclusion [21]. Initially, potential sources were selected by examining their titles and abstracts, ensuring relevance to the research question. This meticulous process facilitated the systematic evaluation and filtering of records, governed by rigorously defined inclusion and exclusion criteria. Articles that advanced beyond the preliminary screening phase were subjected to a comprehensive review of their full texts to determine their eligibility for inclusion in the systematic review. Articles were first reviewed by the first author (RHAHA) in collaboration with the co-author NJA and under the supervision of the second author (CLH) to ensure they met the eligibility criteria. The JBI critical appraisal tools for the quality assessment of all papers included in the review were further used to ensure the papers met the necessary standards for research and publication (internally valid, appropriately randomized, participants appropriately consented, etc). The eligibility criteria are presented in Textbox 1.

Textbox 1. Eligibility criteria.

Inclusion criteria

- Region: Focuses on the Sub-Saharan African region or has significant relevance to Sub-Saharan African populations
- Target population: Adolescents and youth
- Area of focus: Addresses sexual and reproductive health, including sexually transmitted infection education, mental health, sexuality, and positive sexuality
- Mode of delivery: Digital tools, including apps, websites, gaming, e-learning modules, social media, etc
- Timeline: Published within the last 10 years (2013-2024) to ensure relevance
- Language: In English to ensure comprehension and interpretation by research teams who are primarily English speaking

Exclusion criteria

- Region: Excluded Sub-Saharan African populations or did not have a Sub-Saharan Africa focus
- Target population: Outside of the age range of adolescents and young people
- Area of focus: Does not specifically address sexual and reproductive health
- Mode of delivery: Excludes digital tools
- Timeline: Outside of the specified dates
- Language: Not in English

Results

Included Studies

Figure 1 outlines the systematic and comprehensive search and selection process employed in this systematic review, focusing on digital educational tools for adolescent SRH in Sub-Saharan Africa. Initially, a total of 1877 records were identified through meticulous searches across major databases, including OVID-MEDLINE, Google Scholar, PubMed, and ERIC. An additional 4 records were identified through citation searching, augmenting the pool of potential studies. Before screening, 203 records were removed due to duplication, leaving 1674 records for detailed screening. The screening process, grounded in rigorously defined inclusion and exclusion criteria, led to the exclusion of 1656 records for reasons such as not directly addressing SRH digital educational tools, not focusing on the Sub-Saharan African region or LMIC context, targeting populations outside the adolescent age range and young people, and duplications. After the exclusion process, 18 reports were retrieved and assessed for eligibility, and 4 additional studies were identified and included in the review alongside the initially identified studies, bringing the total number of studies included in the review to 22 (Figure 1; Table 1) [22].

Figure 1. Flowchart for our systematic review of digital educational tools for adolescent sexual and reproductive health (SRH) in Sub-Saharan Africa. LMIC: low- and middle-income country.

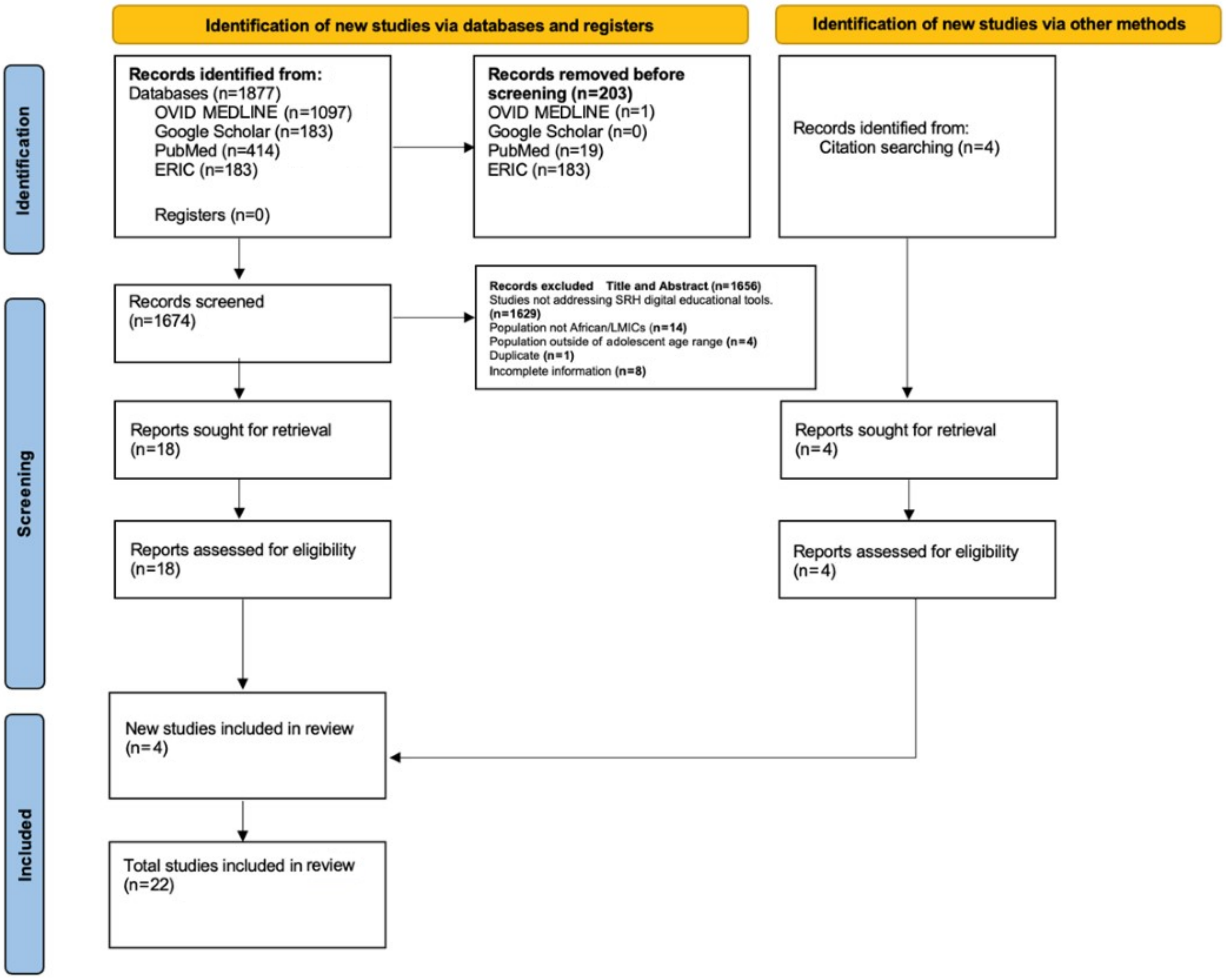


Table 1. Summary of studies on the availability of digital sexual and reproductive health education for adolescents in Sub-Saharan Africa.

Study (authors and year)	Study population/region/country	Study design	Digital tool type	Findings
Dulli et al [23], 2020	349 youth aged 15-24 years from Nigeria	Randomized controlled trial. Participants were randomized into 2 groups: those receiving the SMART Connections intervention and those receiving standard care services.	Social media	Primary results: The intervention did not significantly improve retention (50% vs 45% retention in treatment vs control groups) or social support but did significantly improve HIV-related knowledge (t=–2.96; P=.003).
Fakoya et al [24], 2022	111 adolescents aged 15-19 years from Nigeria, Tanzania, and Ethiopia	Application of youth-led participatory action research (PAR) approaches within human-centered design (HCD)	Youth-engaged version of HCD as part of the Adolescents 360 initiative	Primary results: Successful engagement of youth as project partners and action researchers. Identified opportunities to improve program empathy and responsiveness. Challenges in recruiting “extreme users” due to high competencies needed in HCD. Empathy and design standards during prototyping helped in decision-making. Real-world testing of services and products emphasized the importance of continuing youth-adult partnership.
Feroz et al [16], 2021	Young people (adolescents and youth) aged 10-24 years in low- and middle-income countries (LMICs), specifically Sub-Saharan Africa and Asia	Systematic review	Mobile phones used for mHealth interventions aimed at improving sexual and reproductive health (SRH) outcomes	Primarily used for client education and behavior change communication (n=14, 93%), followed by financial transactions and incentives, and data collection and reporting. Studies evaluated the effect mHealth interventions had on access to SRH services (n=9) and SRH outcomes (n=6). mHealth interventions improved access to SRH services and outcomes by addressing barriers such as provider prejudice, stigmatization, discrimination, and issues of privacy and confidentiality. Included decreased technological literacy, inferior network coverage, and lower linguistic competency.
Gonsalves et al [25], 2018	705 young people aged 13-24 years from Kenya and Peru	Protocol of an open, 3-arm, individually randomized trial. Participants will be randomized into one of three arms: (1) intervention arm receiving SRH information (using ARMADILLO) via mobile phones, (2) control arm receiving no intervention, and (3) a third arm with variations by site (Kenya and Peru).	Mobile phone sexual health information	If proven to be effective, interventions like ARMADILLO can bridge an important gap in achieving universal access to SRH information and education for an otherwise difficult-to-reach group. Outcomes will be assessed via questionnaires administered at baseline, intervention end, and 8 weeks after intervention end. Statistical analysis will include comparisons of proportions (chi-square tests) and means (t-tests) between arms and difference-in-difference techniques.
Haruna et al [26], 2019	348 secondary school participants (students) aged 11-15 years from 3 schools were recruited to participate in this study from Tanzania	Design-based research (DBR)	Digital gamified learning platform named “My Future Begins Today”	A paired t-test revealed a statistically significant improvement in sexual health literacy scores from pretest (mean 26.40, SD 7.29) to posttest (mean 74.12, SD 16.21) (t_{347}=52.230; P<.000), with an average increase of 47.72 points. A 1-way ANOVA showed no significant differences among the 3 groups’ pretest scores, but significant differences were found in posttest scores ($F_{2,345}$=210.43; P<.001), with follow-up Tukey post hoc tests indicating better performance by the gamified learning platform group compared to the traditional teaching group. Positive stakeholder and participant feedback on the gamified approach. Effective use of participatory design in developing educational tools.

Study (authors and year)	Study population/region/country	Study design	Digital tool type	Findings
Nalwanga et al [27], 2021	1086 Kyambogo University students aged 18-30 years from Uganda	Cross-sectional analysis of data from an endline survey of a randomized controlled trial (RCT) and data from use of a mobile phone app over a 6-month period.	Mobile phone app	The mobile phone app demonstrated predominantly positive (responsiveness, nondistracting in-app advertisements, and ease of use) attributes. 86% (n=464) of students who received a recommendation to download the app accepted and downloaded it. Of these students, 81% (n=374) used the app to access SRH information, goods, and services over 6 months. In terms of responsiveness, 55.1% (n=206) of students stated that clinics were responsive (*P*=.11). 75.4% (n=282) said the in-app advertisements were not distracting (*P*=.23). 44.4% (n=166) found that in-app instructions were very easy to use (*P*=.50).
Ippoliti et al [28], 2017	Adolescents and youth aged 10-24 years from Africa (67%), Eurasia (26%), and Latin America (13%)	Global landscape analysis. The authors issued a global call for project resources in 2014 and reviewed the submissions to confirm they met specific inclusion criteria.	Mobile phones (including SMS text messaging and mobile phone apps)	Most projects (n=12, 70%) relied on text messaging to transmit SRH information. The majority of projects were based in Africa (67%), followed by Eurasia (26%) and Latin America (13%). Mobile phones were effectively used to increase the reach of SRH information and services, especially in conservative societies where SRH topics are stigmatized. 70% of the projects relied on text messaging to transmit SRH information. Several projects have been adapted and scaled to other countries, demonstrating the scalability of mHealth interventions.
L'Engle et al [29], 2016	Review targeted adolescents aged 10-24 years and included studies that provided results from mobile phone interventions designed to improve adolescent SRH	Systematic review	Mobile phone interventions, primarily text messaging	Most programs (n=23, 82%) used text messaging. An average of 41% of essential mHealth criteria were met. An average of 82% of methodological reporting criteria were met. Evidence suggests improvements in health promotion campaigns, sexually transmitted infection (STI) screening and follow-up, and medication adherence.
Nolan et al [30], 2020	6000 youth aged 12-19 years across 8 districts in Rwanda	3-arm cluster-randomized. Arms: (1) CyberRwanda self-service, (2) CyberRwanda facilitated, and (3) control schools, which will receive the standard services that are available in the community.	CyberRwanda, a digital health intervention consisting of interactive stories, questions and answers, videos, and an online shop for health products	The study protocol describes the design and intended impact evaluation of the CyberRwanda program using a 3-arm cluster randomized noninferiority trial, but it does not report on the outcomes or effectiveness of the intervention. The primary outcomes are measurements of uptake of a modern method of contraception, initiation of childbearing, and HIV testing, all measured at the participant level. Data analysis will consist of a generalized linear mixed model.
Onukwugha et al [31], 2022	Review includes individuals aged 10-19 years, with consideration for interventions focusing on young people aged 10-24 years in Sub-Saharan Africa	Systematic review	mHealth interventions	mHealth interventions were effective in improving adolescents' uptake of SRH services across a wide range of services, with the strongest evidence for contraceptive use. Interventions with 2-way interactive functions and more behavior change techniques embedded were more effective. Limited data suggested that interventions were inexpensive, but cost-effectiveness was not evaluated.
Patel et al [32], 2022	Adolescents in LMICs (countries not specified by the authors)	Environmental scan is used to identify current SRH mobile apps available in the iOS App Store and Android Play Store.	Mobile apps	2165 mobile apps were initially identified, with only 8 apps meeting the inclusion criteria and assessed using the Mobile App Rating Scale (MARS) tool. The functionality subdomain scored the highest at 4.6. The information subdomain scored the lowest at 2.5. None of the assessed apps contained information on the MARS items related to the evidence base and goals. "Too Shy to Ask" had the highest individual app mean score of 4.1, while "e-SRHR" scored the lowest at 2.3.

Study (authors and year)	Study population/region/country	Study design	Digital tool type	Findings
Pfeiffer et al [33], 2014	60 adolescents aged 15-19 years in Dar es Salaam and Mtwara, Southern Tanzania	Mixed methods study	Social media, with a particular focus on Facebook	Adolescents access the internet mainly through mobile phones (68%). Facebook is the most popular internet site among the youth (73%). Adolescents expressed interest in receiving reproductive and sexual health messages through social media, with 92% of respondents saying Facebook should be used. Youth role models, such as music stars and actors, are influential in delivering health messages.
Rogers et al [34], 2019	Adolescents aged 10-19 years living in Zambia	Content analysis of the TuneMe website	Mobile-optimized website (TuneMe)	The TuneMe platform provided extensive information on SRH and HIV. Topics on relationships (21%) as well as sexual rights and sexual citizenship (11%) were also covered, albeit to a lesser extent. Areas, such as pleasure, violence, diversity, and gender, received significantly less attention. Content was presented in culturally relevant contexts but sometimes portrayed mixed or problematic views on gender norms. 89% of the TuneMe content falls within the IPPF Framework components for CSE, with 37% falling in the SRH and HIV primary category. The Gunning Fog Index score indicated that a minimum of an 8th grade reading level would be required to fully understand TuneMe content.
Rokicki et al [35], 2017	756 female adolescents aged 14-24 years enrolled in secondary schools in Accra, Ghana	Cluster randomized controlled trial. Unidirectional intervention (n=12 schools): participants received text messages with reproductive health information; interactive intervention (n=12 schools): participants engaged in text messaging reproductive health quizzes; control (n=14 schools): did not receive reproductive health–related text messages.	Text messaging program	The unidirectional intervention increased reproductive health knowledge by 11 percentage points (95% CI 7-15). The interactive intervention increased knowledge by 24 percentage points (95% CI 19-28) from a control baseline of 26%. No significant changes in reproductive health outcomes overall. Both interventions lowered the odds of self-reported pregnancy for sexually active participants (unidirectional: OR 0.14, 95% CI 0.03-0.71; interactive: OR 0.15, 95% CI 0.03-0.86). Results showed that text-messaging programs may help improve reproductive health knowledge among adolescents.
Soehnchen et al [13], 2023	77 young adults aged 18-35 years in the resource-poor regions of Kenya	Survey based on the Unified Theory of Acceptance and Use of Technology (UTAUT)	Web-based prototype designed to provide essential sexual health information	77 responses were included in the analysis. High acceptance and usability of the digital tool among the target population. Perceived usefulness, attitude toward health care–integrated evidence technology, and usability had significant positive impacts on acceptance and intention to use. Having the resources and knowledge necessary for the usage of a digital tool had a significant negative impact. The chi-square test statistic was 532 ($P<.001$), meaning it was highly significant compared to the original model. A System Usability Scale (SUS) score of 67.3 indicated the tool's usability as "okay."
Soehnchen et al [8], 2023	17 young female individuals aged 15-25 years from Kenya	Qualitative research design, with a sample size of 12 pilot phase interviews and 5 expert interviews.	Web-based sexual health education app	Stigmatization around sexual health in Kenya leads to myths and a lack of information. Sexual health education is not part of the Kenyan school curriculum, leading to insufficient knowledge about safe contraception, menstruation, and female genital mutilation. A digital app could support and provide education and information for universal equal access. Barriers to using a digital sexual health education app include conservative cultural background, classic text communication, and social affiliation influence.

Study (authors and year)	Study population/region/country	Study design	Digital tool type	Findings
Frankline et al [36], 2020	398 female adolescents aged 10-19 years from Cameroon	Single-centered randomized controlled single-blinded trial. Intervention group: received weekly educational 1-way text messages on sexoreproductive health (199 participants); control group: did not receive the text messages (199 participants).	Mobile phone text messaging (SMS)	Significant increase in mean knowledge, attitude, and practice scores from baseline to the end of the study (from 6.03, 4.01, and 3.45 to 7.99, 5.83, and 4.99, respectively). Statistical significance in the overall improvement in adolescents' perception of sexoreproductive health as a result of the intervention (F=15.12; P=.02 for knowledge; F=60.21; P=.001 for attitudes; F=57; P=.01 for practices). Majority (65.3%) of participants were satisfied with the SMS text messaging service.
Akande et al [37], 2024	1280 in-school adolescents aged 15-17 years from Nigeria	Cluster randomized controlled trial. Intervention group: received the mHealth-based intervention; control group: did not receive the mHealth intervention.	mHealth-based sex education program	Significant improvements in SRH knowledge (F=2117.252; P<.001) and attitudes (F=148.493; P<.001) in the intervention group. No significant change in risky sexual behavior scores in the intervention group (P=.57). Females had higher odds of having good SRH knowledge (aOR 2.5, 95% CI 1.04-6.13). Males had lower odds of practicing protective sexual behavior (aOR 0.3, 95% CI 0.15-0.55), and higher class levels (SS2: aOR 5.2; SS3: aOR 6.2) were associated with more protective behavior.
Rokicki and Fink [15], 2017	756 female adolescents and young women aged 14-24 years from Ghana	Randomized controlled trial. Unidirectional intervention (text messages with reproductive health information), interactive intervention (engaged adolescents in text messaging reproductive health quizzes), and control (no intervention).	Interactive mobile phone quiz (mHealth intervention)	The mHealth intervention was effective in engaging adolescent girls across sociodemographic strata. Participants showed increased health knowledge regarding SRH. Knowledge scores increased from 26% to 32% in the control group, 30% to 45% in the unidirectional group, and 31% to 60% in the interactive group when looking from baseline to the 3-month follow-up. The interactive intervention was more effective than the unidirectional intervention, increasing knowledge scores by 13 percentage points (95% CI 8-18). The intervention reached adolescents who may be at higher risk of poor SRH outcomes.
Haruna et al [38], 2018	120 lower secondary school students aged 11-15 years from Tanzania	Randomized controlled trial. Groups: (1) game-based learning (GBL), (2) gamification, (3) control group (traditional teaching).	GBL and gamification platforms	The Adolescent Sexual Health Literacy Test (ASHLT) showed a significant increase in scores from pretest (mean 29.26, SD 8.689) to posttest (mean 70.36, SD 18.201) (t_{119}=–23.787; P=.001). GBL and gamification methods significantly improved students' motivation, attitude, knowledge, and engagement (MAKE) compared to traditional teaching methods, as evidenced by higher scores in these areas and supported by positive feedback from focus group discussions. Pairwise comparisons revealed significant differences between traditional and experimental methods, with no substantial differences between GBL and gamification, indicating that both experimental approaches were more effective in enhancing sexual health education.
Alhassan et al [39], 2019	250 adolescents and young people aged 18-24 years at the University of Ghana, Legon in Accra, Ghana	Cross-sectional analytical study	Mobile phones (including smartphones)	99% of participants owned mobile phones, with 58% being smartphone users. Male young adults and young adults who owned a smartphone were more likely to use mobile phones for STI education and prevention (P=.000 and P=.01, respectively). The study suggests high mobile phone penetration among young adults, with a belief in the efficacy of mobile phone programs for STI education and prevention.
Ippoliti et al [40], 2021	Adolescents aged 12-19 years from Rwanda	HCD approach	CyberRwanda, a web-based digital health platform	Over 1000 Rwandan youth, caregivers, teachers, health care providers, and government stakeholders engaged. Revealed participants' beliefs, behavioral preferences, and experiences related to family planning and reproductive health. Designed for urban and periurban young people and pharmacy staff. Evaluations will be conducted across 60 schools and 9 youth centers in 8 districts in Rwanda (n=6082 at baseline).

Objectives 1a and 1b: Geographic and Demographic Focus

As shown in Figure 2, the geographical scope of research on digital interventions for SRH education spans diverse settings in Sub-Saharan Africa, each offering unique insights into the challenges and opportunities inherent in deploying technology-based solutions within various cultural and infrastructural contexts. The studies span populations in West, Central, East, and South Africa. This geographical diversity is not merely a backdrop for these studies but a critical factor that shapes the design, implementation, and impact of digital health interventions across the continent; it underscores the importance of contextually adapted solutions. As depicted in Table 1, the studies included youth and adolescents of various ages, integrating variations of human-centered designs (HCDs) that resonated with their lived experiences and preferences. In Nigeria, Dulli et al [23], Fakoya et al [24], and Akande et al [37] highlighted the country's engagement with digital platforms to address SRH among youth and adolescents. Feroz et al [16] and Ippoliti and L'Engle [28] also included studies from Nigeria. Nigeria, with its vast population and significant digital penetration, represents a critical context for understanding how social media and participatory design can be harnessed to meet young people's health education needs. The focus on Nigeria underscores the potential of digital tools to transcend traditional barriers to health education, leveraging the widespread use of social media and mobile phones among the youth. Besides Nigeria, this study included some other countries in West Africa, such as Ghana [16,31,35,39], Senegal, and Mali [28].

Figure 2. Map of Africa showing the countries included in the reviewed studies (green).

The reviewed studies extend beyond West Africa to include Kenya, as illustrated by Gonsalves et al [25], who explored the impact of mobile phone–based sexual health information among young people, as well as Soehnchen et al [13], who assessed acceptance of a digital tool for delivering sexual health education. Some other authors, namely Feroz et al [16], Ippoliti and L'Engle [28], Onukwugha et al [31], and Soehnchen et al [8], also included studies that were performed in Kenya in their reviews. Kenya's inclusion in this body of research is indicative of its role as a hub for technological innovation and digital health initiatives in East Africa. Nalwanga et al [27] performed a cross-sectional analysis to assess the use of a mobile phone app by university students in Uganda, another East African country. Haruna et al [26] focused on secondary school participants in Tanzania, using a gamified learning platform to engage students in SRH education. This emphasis on educational settings highlights the potential of integrating digital tools into formal education systems, showing how gamification can enhance learning outcomes and engagement among school-aged youth. Data from Tanzania, Rwanda, and Ethiopia, which are also East African countries, were included in studies by Fakoya et al [24], Feroz et al [16], Haruna et al [26,38], Ippoliti and L'Engle [28], Ippoliti et al [40], Onukwugha et al [31], Nolan et al [30], and Pfeiffer et al [33]. The cross-regional approach of Gonsalves et al [25], encompassing participants from both Kenya and Peru, further emphasizes the global relevance of digital health interventions and the importance of considering diverse cultural and infrastructural landscapes in their design and deployment.

This review also included some countries in southern Africa, including South Africa, Mozambique, and Zambia. For example, Rogers et al [34] performed a content analysis of TuneMe, a mobile-optimized website that offers extensive information on SRH and HIV. Importantly, the content was presented in culturally relevant contexts. In addition to that study, Feroz et al [16], Ippoliti and L'Engle [28], and Onukwugha et al [31] included data from southern African countries in their systematic reviews.

Democratic Republic of Congo [16] and Cameroon [36] were the 2 Central African countries included in this review. Collectively, these studies paint a picture of a continent actively engaging with digital technologies to overcome the challenges in delivering SRH education. The wide geographical reach of this research, ranging from the populous country of Nigeria to the tech-savvy environments of Kenya, reflects the varied landscapes in which these digital interventions are implemented. Each setting provides unique insights into the interplay between cultural norms, technological accessibility, and health education needs, contributing to a richer understanding of how digital tools can be adapted and scaled across the Sub-Saharan African context.

Target Demographics

Among the studies reviewed, there was a notable emphasis on youth and adolescents, spanning ages 10 to 24 years, which reflected the critical importance of reaching these age groups with effective, accessible, and engaging health education. Although the focus of this study is the adolescent age group, they are often included in broader age groups, such as young people and youth, as shown in Table 1. For instance, Dulli et al [23] concentrated on youth aged 15 to 24 years in Nigeria. The authors performed a randomized controlled trial using social media as a platform to deliver sexual health interventions (SMART Connections). Similar populations were targeted by Gonsalves et al [25] in Kenya and Rokicki et al [35] in Ghana. Similarly, Fakoya et al [24] engaged adolescents aged 15 to 19 years from Nigeria, Tanzania, and Ethiopia, incorporating youth-led participatory action research with an HCD to create interventions that resonate with the lived experiences and preferences of young individuals. The authors directly engaged the target demographic in the study design process. In Rwanda, Nolan et al [30] analyzed data from 6000 youth aged 12 to 19 years in a cluster-randomized trial and Rogers et al [34] studied a similar age group (10 to 19 years) in Zambia.

Further broadening the demographic spectrum, Feroz et al [16] extended their review to young people aged 10 to 24 years in Sub-Saharan Africa and Asia. The Sub-Saharan African countries in that study included Tanzania, Nigeria, Ghana, Uganda, and Kenya. The authors performed a comprehensive review of mHealth interventions, identified a wide range of options, and elucidated their applications as well as barriers and facilitators for adoption.

The inclusive age range spanning from early adolescence to young adulthood acknowledges the diverse needs and challenges faced by individuals as they navigate through different stages of the early life course, emphasizing the importance of tailoring interventions to meet these varying requirements. The more expansive approach of Gonsalves et al [25], by including young people aged 13 to 24 years in Kenya and Peru, illustrates the global relevance of digital sexual health interventions and the need to consider cross-cultural applicability and customization in their design and implementation. This cross-regional study underscores the universal challenges related to sexual health education, while acknowledging the specific nuances that different cultural contexts bring to the fore. Finally, Haruna et al [26] involved secondary school participants from Dar es Salaam, Tanzania in the development of a digital gamified learning platform. This engagement not only emphasizes the educational context as a critical setting for sexual health education but also showcases the potential of gamification as a strategy to enhance learning outcomes and retention among school-aged youth.

Collectively, the studies reflect a strategic and inclusive approach, which is aimed at bridging gaps in access and engagement. By examining the implementation and impact of these digital tools across various African settings, researchers and practitioners can gain valuable insights into the factors that drive the success and scalability of health education interventions, thus paving the way for more inclusive and effective public health strategies.

Objective 2: Digital Tool Types and Effectiveness

The landscape of digital tools deployed for SRH education in various African contexts not only signifies a shift toward more accessible and engaging formats for health education but also reflects the adaptability of health practitioners and researchers to leverage technology to meet the needs of target populations.

With regard to design, Fakoya et al [24] introduced a youth-engaged version of HCD, a problem-solving participatory process that centers the needs, perspectives, and experiences of target populations, when developing solutions to complex challenges. The authors used HCD as part of Adolescents 360, a transdisciplinary initiative to increase the use of modern contraception among 15- to 19-year-old girls in Nigeria, Ethiopia, and Tanzania. The youth-led participatory approach involved 111 “youth designers” trained in HCD methods to inform the design and implementation of the interventions. The authors had successfully engaged youth as project partners and action researchers involved in an iterative cycle of program research, design, and evaluation. However, they noted challenges in recruiting "extreme users" due to the high competencies needed in HCD and emphasized the need for real-world testing (not just prototyping to enhance adolescents' experiences) and the importance of planned and dedicated financial and human resources. An HCD approach was also deployed in Rwanda by Ippoliti et al [40] for their web-based digital health platform CyberRwanda. The involvement of not only youth but also caregivers, teachers, health care providers, and government stakeholders revealed participants' beliefs, behavioral preferences, and experiences related to family planning and reproductive health. The authors noted that the HCD process, although iterative, resulted in significant pivots to the design and implementation of the digital platform and resulted in a superior digital health intervention with and for Rwandan youth.

For deployment and implementation, the ubiquity and accessibility of mobile phones make them an ideal platform for reaching a wide audience, particularly in regions where mobile technology may be more readily available than traditional health care services. In fact, Alhassan et al [39] report that 99% of 250 surveyed adolescents and young people aged 18 to 24 years owned mobile phones and 58% were smartphone users. Young male adults and those who owned a smartphone were more likely to use mobile phones for STI education and prevention. In their systematic review of the use of mobile phone technology in LMICs for mHealth SRH interventions, Feroz et al [16] noted that most of the included mHealth intervention studies (n=14, 93%) focused on behavioral change and patient education using a variety of modalities, including SMS text messaging, video clips, images, and voice communication. mHealth tools improved young people's SRH knowledge, ensured safer sexual behavior, maximized reach and access to family planning information, and improved several HIV outcomes. Barriers to mHealth uptake for SRH included high cost of service, infrastructural or network quality, request for sociodemographic information that could break anonymity, and sociocultural beliefs and norms. Ippoliti and L'Engle [28] also echoed these findings in their global landscape analysis of the use of mobile phones for SRH content among adolescents and youth aged 10 to 24 years in Africa, Eurasia, and Latin America. The authors reported that mobile phones (particularly text messaging) were effectively used to increase the reach of SRH information and services, especially in conservative societies where SRH topics are stigmatized. Nalwanga et al [27] assessed the use of a mobile phone app to increase access to SRH information, goods, and services among university students in Uganda and reported high use of the frequently asked questions portal (71%) and high product use by both sexes (condoms for males [77% use]; sanitary pads for females [94% use]), and the most popularly accessed service was HIV testing and counseling (60% use). Participants appreciated the responsiveness, nondistracting in-app advertisements, and "ease of use" attributes of the mobile phone app. One interesting aspect of this study was the inclusion of a modest co-payment (paid by users), which, interestingly, was not a barrier to usage. Contrary to what has been reported by others, the youth in this study did not find the co-pay to be cost prohibitive. In fact, its incorporation ensured the utility of client payments in future iterations of the app after the pilot period. In a cluster randomized controlled trial of 756 adolescent female individuals aged 14 to 24 years enrolled in secondary schools in Accra, Ghana, Rokicki et al [35] compared the following 3 arms: unidirectional intervention (n=12 schools), where participants received text messages with reproductive health information; interactive intervention (n=12 schools), where participants engaged in text messaging reproductive health quizzes; and control (n=14 schools), where students did not receive reproductive health messaging. They found that the unidirectional intervention increased reproductive health knowledge by 11 percentage points (95% CI 7-15); however, the inclusion of quizzes in the interactive intervention increased knowledge by 24 percentage points (95% CI 19-28) from a control baseline of 26%. Both interventions lowered the odds of self-reported pregnancy for sexually active participants (unidirectional: OR 0.14, 95% CI 0.03-0.71; interactive: OR 0.15, 95% CI 0.03-0.86). These studies highlight the versatility of mobile phones in delivering tailored health education and interventions ranging from SMS text messaging–based information dissemination to more interactive app-based learning experiences and quizzes to reinforce knowledge retention and behavioral change. They can act as a one-stop shop, offering not only educational and clinical counseling services but also condoms, contraceptives, and menstrual products, which give users a holistic experience that can be anonymized to ensure privacy and protection.

The digital gamified learning platform developed in Tanzania by Haruna et al [26,38] represents an innovative leap in engaging and educating secondary school participants in the African setting. The authors used game-based and participatory designs to evaluate the ability of the game to motivate students, improve their attitudes, increase their acquisition of knowledge, and engage them in learning. A paired *t*-test revealed a statistically significant improvement in sexual health literacy scores from pretest (mean 26.40, SD 7.29) to posttest (mean 74.12, SD 16.21) (t_{347}=52.230; *P*<.000), with an average increase of 47.72 points. Additional analyses with ANOVA testing showed no significant differences in pretest scores among the 3 groups, but significant differences were found in posttest scores ($F_{2,345}$=210.43; *P*<.001), with follow-up Tukey post-hoc tests indicating better performance in the gamified learning platform groups than in the traditional teaching group. These findings suggest that gamified learning platforms can enhance motivation, attitude, knowledge acquisition, and engagement in sexual health education more effectively than traditional methods [26].

The research team took the intervention further in a 3-armed randomized clinical trial to compare the effectiveness of game-based learning and gamification to a control approach (traditional teaching method) [38]. The results showed that the average posttest scores were significantly higher for game-based learning (mean 79.94, SD 11.169) and gamification (mean 79.23, SD 9.186) than for the control approach (mean 51.93, SD 18.705) ($F_{2,117}$=54.75; *P*=.001). Additionally, statistically significant differences (*P*≤.05) were found for the constructs of motivation, attitude, knowledge, and engagement (MAKE) for the 2 intervention groups compared to the control arm. Gamification not only enhances learning outcomes but also introduces an element of fun and interaction that can significantly increase participant engagement and knowledge retention [38]. Such gamified platforms offer promising avenues for making health education more appealing and effective for younger people.

Web-based platforms may eliminate the costs associated with the implementation of gamified learning platforms. In a content analysis of the TuneMe website, designed to deliver SRH information to Zambian adolescents aged 10 to 19 years, Rogers et al [34] reported that the website comprehensively covered topics of reproductive anatomy and STIs, including HIV, as well as relationships and sexual rights. However, the challenge of presenting SRH education in culturally relevant contexts was apparent when topics of sexual pleasure and gendered norms, behaviors, expectations, and violence were introduced, demonstrating nuanced beliefs around sexual rights.

Finally, social media emerged as a significant platform in the digital toolbox, with Dulli et al [23] exploring its utility among youth aged 15 to 24 years in Nigeria. The authors performed a randomized controlled trial using social media to deliver sexual health interventions for youth living with HIV (SMART Connections). The design and content of SMART Connections were as informed through workshops conducted with stakeholders and youth living with HIV. The content was delivered anonymously through Facebook over approximately 22 weeks to the intervention and control groups, with daily activities to promote engagement. The primary outcome was retention in HIV treatment, while secondary outcomes included antiretroviral therapy (ART) adherence, HIV knowledge, and social support. Although the intervention did not significantly improve the primary outcome of retention (50% and 45% retention in the treatment and control groups, respectively), there was a significant improvement in HIV-related knowledge for the SMART Connections group compared to controls (t=–2.96; P=.003). Intervention group participants overwhelmingly reported that the intervention was useful, that they enjoyed taking part, and that they would recommend it to peers. Similarly, Pfeiffer et al [33] found that Facebook was the most used social media platform for sexual health promotion in Southern Tanzania for youth aged 15 to 19 years. Adolescents were most interested in SRH messages through humorous posts, links, and clips, as well as through youth role models like music stars and actors in a manner that reflected up-to-date trends of modern youth culture. Both studies highlighted the integral role of social media in the daily lives of youth, positioning it as a critical channel for delivering sexual health education.

Collectively, these digital tools, ranging from social media and participatory design to mobile phone apps and gamification, reflect a dynamic and multifaceted digital ecosystem for SRH education. Each tool, with its unique features and applications, contributes to a broader strategy aimed at enhancing the accessibility, engagement, and effectiveness of health interventions in different African settings. These diversified digital tools not only address the varied needs and preferences of target demographics but also showcase the potential of technology in transforming health education and empowerment.

Discussion

In this research endeavor, we conducted a systematic literature review to achieve the following two objectives: (1) describe the geographic and demographic distributions of digital educational tools for SRH education for adolescents in Sub-Saharan Africa (objectives 1a and 1b) and (2) examine the types, formats, and impacts of these digital tools on target populations (objective 2). For the first objective, we identified 22 studies conducted across West, Central, East, and South Africa that focused on digital educational tools for SRH education targeting youth and adolescents aged 10 to 24 years. These studies offer valuable insights into the types of digital tools currently being employed in SRH education and highlight the regions where such tools are being most actively implemented. For the second objective, we found that a variety of digital tools, including social media, mobile apps, and gamified learning, were used to engage adolescents with accessible SRH information. These methods were generally effective, although challenges, such as the digital divide, cultural sensitivity, and the need for assessing long-term behavior change, were noted.

The 22 studies included in this review covered a broad spectrum of digital interventions, ranging from mobile apps to social media campaigns, and reflected a growing interest in digital platforms as a means of delivering SRH education to young people in Sub-Saharan Africa. The geographical distribution of the studies indicates that digital tools are being used across various cultural contexts, with notable contributions from West (eg, Nigeria and Mali), East (eg, Kenya and Tanzania), Central (eg, Cameroon and Democratic Republic of Congo), and Southern Africa (eg, Mozambique and South Africa). This regional diversity is significant as it shows that digital SRH interventions are becoming increasingly accessible across different subregions of Africa, each with unique sociocultural challenges and opportunities. Our findings suggest that digital tools can overcome geographic barriers, offering a means to engage with youth who may otherwise be excluded from formal SRH education programs due to geographical isolation, economic constraints, or cultural taboos around discussing SRH topics [41,42]. These findings are consistent with the broader literature on digital health interventions in the African region, which has highlighted the promising role of digital tools in expanding access to SRH education [31,43]. For example, in Nigeria, mobile apps focused on HIV prevention and sexual health have been shown to improve knowledge and self-efficacy among adolescents [44]. In Kenya, WhatsApp-based health campaigns have successfully increased awareness of sexual health issues among young people [45]. Conducting cross-country comparisons through systematic reviews, as performed in this work, provides valuable insights into the best practices that can be adapted and scaled across diverse African contexts. For example, WhatsApp is widely used around the world, and rather than reinventing the wheel, could developers in Nigeria adapt and implement the WhatsApp-based intervention designed for Kenyan youth by Chory et al [45] for a similar population in Nigeria? What would be the short- and long-term impacts of such an approach? Moreover, would policymakers be more inclined to support a “repackaged” intervention that is cost-effective, evidence-based, and already proven in another African context? Future research should explore these cross-country comparisons in greater depth, examining the potential for scaling successful digital interventions across African nations. By redesigning these tools to reflect the cultural, linguistic, and societal contexts of each country, researchers can assess the feasibility and cost-effectiveness of adapting existing interventions, potentially saving time and resources compared to developing entirely new solutions.

Our second objective revealed that while a variety of digital tools were used to engage adolescents with accessible SRH information, mobile apps and gamified learning emerged as particularly popular and effective approaches. Mobile apps allow adolescents to access information in a flexible, anonymous, and user-friendly format, which helps overcome the barriers posed by traditional methods of SRH education [46]. This flexibility is particularly important in contexts where discussing sexual health is stigmatized, as it empowers youth to seek information discreetly and at their own pace. Additionally, gamified learning, by integrating interactive elements, not only captures adolescents' attention but also reinforces learning through engaging experiential methods, which may result in better knowledge retention and a deeper understanding of SRH topics [31]. Gamified learning, in particular, aligns with the preferences of a digitally native generation, where interactive and visually stimulating content is highly favored. Social media, including platforms like Facebook, WhatsApp, and Instagram, were also used to reach adolescents, leveraging peer influence and interactive communication strategies. This social aspect of learning enhances engagement by making the information feel more relevant and personal for youth. However, a major issue noted across the studies was the "digital divide," which is used to describe the unequal access to technology and internet connectivity that can limit the effectiveness of digital SRH education programs [47]. While mobile phones are widespread in Sub-Saharan Africa, reliable internet access remains a significant barrier, particularly in rural and underdeveloped regions [45]. This digital divide may result in inequitable access to SRH resources, especially for marginalized youth [45]. Innovative solutions, such as an artificial intelligence–powered phone, have the potential to bridge the digital divide [48]. This system delivers information and advice through voice calls, eliminating the need for a smartphone or internet connection and making information accessible to individuals without reliable internet or those who cannot afford smartphones and data plans [48]. By using artificial intelligence to understand natural language queries and provide tailored information from a knowledge base, these technologies can significantly improve access to SRH education in underserved areas with limited technological infrastructure [48]. Future research should explore effective solutions to overcome the digital divide, such as offline digital tools as described, low-cost data plans, and internet infrastructure development in rural and underserved areas.

Cultural sensitivity also emerged as a critical issue. In many African contexts, cultural norms and values significantly shape how SRH topics are perceived and discussed, and failure to account for these factors can limit the effectiveness of digital SRH interventions. For instance, discussions around issues like sexual consent, gender equality, and contraception can be seen as a taboo in some regions, which may lead to resistance or disengagement from the target audience [41]. Thus, tailoring content to reflect local values, languages, and cultural norms around sexuality is critical, as failing to consider these factors could undermine the effectiveness of SRH interventions. Involving adolescents in the creation and assessment of these tools can provide essential insights into their preferences, needs, and challenges, which can significantly enhance the acceptability and effectiveness of digital interventions [49]. Moreover, including parents could create opportunities to address cultural taboos and foster more open discussions between parents and adolescents on SRH [42]. Such involvement could lead to long-term benefits, as parent-adolescent communication is a key factor in reducing adverse sexual behaviors and improving overall sexual health outcomes [42]. According to Agbeve et al [42], African parents who hold sex-positive attitudes are more inclined to engage in discussions with their adolescents that help them make informed decisions regarding their sexual health and overall well-being. When parents are actively involved in the design and implementation of digital interventions, they can provide support, guidance, and reinforcement of the messages delivered through digital platforms [50,51]. Parental involvement can help create a more comprehensive approach to sexual education, ensuring that teens not only receive the digital information but also have a trusted adult to turn to for advice and clarification [50,51]. By fostering open communication between parents and teens, digital interventions can be better tailored to the specific needs of the family, ultimately promoting healthier decision-making and outcomes for young people.

Future studies could also explore avenues for incorporating comprehensive SRH education into the formal education system, expanding the scope of topics to include healthy relationships, positive sexuality, mental health, gender equality, and consent, as was done by Fakoya et al [24]. Integrating these topics into the curriculum would not only equip adolescents with essential knowledge about SRH but also foster a more holistic understanding of well-being and personal development. By addressing issues like mental health and gender equality alongside traditional SRH topics, this approach could help challenge harmful stereotypes, promote healthier attitudes toward sexuality, and empower young people to make informed decisions, leading to positive long-term social and health outcomes.

Finally, there was a noted gap regarding the impact of these digital tools on long-term adolescent sexual behavioral change. The majority of studies measured short-term outcomes, such as knowledge acquisition, while long-term studies assessing outcomes, such as improved contraceptive use and reductions in risky sexual behavior, remained limited. The anonymity and privacy afforded by digital platforms can empower adolescents to seek information and services without fear of stigma, promoting more proactive and informed health behaviors that can be tracked over time [46]. Tracking tools can be incorporated in the design phase of the intervention development and can have input from adolescent end users. Long-term data support the sustainability of digital interventions, providing valuable information that can aid in the buy-in of investors, software developers, public health enthusiasts, and policymakers. Regardless, the implications of our findings are multifaceted and show that digital SRH tools are a promising avenue for enhancing adolescent sexual health education across the African region.

The use of a systematic literature review approach in this work has some limitations that are worth mentioning. The main limitations of our work are related to the methodology. There is a risk of selection bias, meaning that some relevant studies may have been omitted or that the included studies may not fully represent all available research on the topic. However, we believe that our extensive use of multiple databases (Ovid-MEDLINE, Google Scholar, PubMed, and ERIC) mitigated the risk of selection bias inherent in systematic reviews. A meta-analysis could further inform the selection bias risk and the heterogeneity of the included studies, and quantify the associations reported.

In conclusion, although digital SRH tools are a promising avenue for enhancing adolescent sexual health education in Sub-Saharan Africa, there is a need for infrastructure improvements to ensure the tools are equitably accessible and a need for understanding long-term impacts and sustainability. Policymakers, educators, and public health researchers must collaborate to ensure that digital innovations in SRH education are grounded in evidence, culturally sensitive, and widely accessible to all adolescents, regardless of socioeconomic status or geographical location. This will address the sustainable development goals of reducing the global disparities in SRH health (goal 10) and achieving good health and well-being across nations (goal 3).

Acknowledgments

MAO was funded by National Institutes of Health (NIH) National Institute of Neurological Disorders and Stroke (NINDS) K01 Award# K01NS121199.

Data Availability

All processes are outlined in the Methods section. Additional information beyond what is reported in the manuscript (eg, template data collection forms, data extracted from the included studies, data used for all analyses, analytic codes, etc) is available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Line-by-line search to retrieve titles and abstracts.
[DOCX File , 17 KB - publichealth_v11i1e63309_app1.docx]

Multimedia Appendix 2
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File (Adobe PDF File), 173 KB - publichealth_v11i1e63309_app2.pdf]

References

1. Melcher J, Lavoie J, Hays R, D'Mello R, Rauseo-Ricupero N, Camacho E, et al. Digital phenotyping of student mental health during COVID-19: an observational study of 100 college students. J Am Coll Health 2023 Apr;71(3):736-748. [doi: 10.1080/07448481.2021.1905650] [Medline: 33769927]
2. Wadham E, Green C, Debattista J, Somerset S, Sav A. New digital media interventions for sexual health promotion among young people: a systematic review. Sex Health 2019 Apr;16(2):101-123. [doi: 10.1071/SH18127] [Medline: 30819326]
3. Chandra-Mouli V, Neal S, Moller A. Adolescent sexual and reproductive health for all in sub-Saharan Africa: a spotlight on inequalities. Reprod Health 2021 Jun 17;18(Suppl 1):118 [FREE Full text] [doi: 10.1186/s12978-021-01145-4] [Medline: 34134737]
4. Armstrong A, Nagata JM, Vicari M, Irvine C, Cluver L, Sohn AH, et al. A Global Research Agenda for Adolescents Living With HIV. J Acquir Immune Defic Syndr 2018 Aug 15;78 Suppl 1(1):S16-S21 [FREE Full text] [doi: 10.1097/QAI.0000000000001744] [Medline: 29994915]
5. Kassa GM, Arowojolu AO, Odukogbe AA, Yalew AW. Prevalence and determinants of adolescent pregnancy in Africa: a systematic review and Meta-analysis. Reprod Health 2018 Nov 29;15(1):195 [FREE Full text] [doi: 10.1186/s12978-018-0640-2] [Medline: 30497509]

6. Melesse DY, Mutua MK, Choudhury A, Wado YD, Faye CM, Neal S, et al. Adolescent sexual and reproductive health in sub-Saharan Africa: who is left behind? BMJ Glob Health 2020;5(1):e002231 [FREE Full text] [doi: 10.1136/bmjgh-2019-002231] [Medline: 32133182]
7. Wangamati CK. Comprehensive sexuality education in sub-Saharan Africa: adaptation and implementation challenges in universal access for children and adolescents. Sex Reprod Health Matters 2020 Dec;28(2):1851346 [FREE Full text] [doi: 10.1080/26410397.2020.1851346] [Medline: 33295853]
8. Soehnchen C, Rietz A, Weirauch V, Meister S, Henningsen M. Creating an Intercultural User-Centric Design for a Digital Sexual Health Education App for Young Women in Resource-Poor Regions of Kenya: Qualitative Self-Extended Double Diamond Model for Requirements Engineering Analysis. JMIR Form Res 2023 Nov 03;7:e50304 [FREE Full text] [doi: 10.2196/50304] [Medline: 37921860]
9. Achen D, Fernandes D, Kemigisha E, Rukundo GZ, Nyakato VN, Coene G. Trends and Challenges in Comprehensive Sex Education (CSE) Research in Sub-Saharan Africa: a Narrative Review. Curr Sex Health Rep 2023 May 06:1-9 [FREE Full text] [doi: 10.1007/s11930-023-00362-1] [Medline: 37362203]
10. Otu A, Ukpeh I, Okuzu O, Yaya S. Leveraging mobile health applications to improve sexual and reproductive health services in Nigeria: implications for practice and policy. Reprod Health 2021 Jan 23;18(1):21 [FREE Full text] [doi: 10.1186/s12978-021-01069-z] [Medline: 33485356]
11. Hubert C, Estrada F, Campero L, Heredia-Pi IB, Villalobos A, Suárez-López L, et al. Designing Digital Tools Capable of Reaching Disadvantaged Adolescents and Improving their Sexual Health: A Mexican Experience. Journal of Health Care for the Poor and Underserved 2021;32(2S):62-84. [doi: 10.1353/hpu.2021.0051]
12. Simuyaba M, Hensen B, Phiri M, Mwansa C, Mwenge L, Kabumbu M, et al. Engaging young people in the design of a sexual reproductive health intervention: Lessons learnt from the Yathu Yathu ("For us, by us") formative study in Zambia. BMC Health Serv Res 2021 Jul 29;21(1):753 [FREE Full text] [doi: 10.1186/s12913-021-06696-7] [Medline: 34325696]
13. Soehnchen C, Weirauch V, Schmook R, Henningsen M, Meister S. An acceptance analysis of a sexual health education digital tool in resource-poor regions of Kenya: an UTAUT based survey study. BMC Womens Health 2023 Dec 19;23(1):676 [FREE Full text] [doi: 10.1186/s12905-023-02839-6] [Medline: 38114976]
14. Kaitano B. Technology and sexual health education in sub-Saharan Africa. Council on Business & Society Insights. 2023 Feb 02. URL: https://cobsinsights.org/2023/02/02/technology-and-sexual-health-education-in-sub-saharan-africa/ [accessed 2024-03-18]
15. Rokicki S, Fink G. Assessing the reach and effectiveness of mHealth: evidence from a reproductive health program for adolescent girls in Ghana. BMC Public Health 2017 Dec 20;17(1):969 [FREE Full text] [doi: 10.1186/s12889-017-4939-7] [Medline: 29262823]
16. Feroz AS, Ali NA, Khoja A, Asad A, Saleem S. Using mobile phones to improve young people sexual and reproductive health in low and middle-income countries: a systematic review to identify barriers, facilitators, and range of mHealth solutions. Reprod Health 2021 Jan 16;18(1):9 [FREE Full text] [doi: 10.1186/s12978-020-01059-7] [Medline: 33453723]
17. Sukhera J. Narrative reviews: Flexible, rigorous, and practical. J Grad Med Educ 2022 Aug;14(4):414-417 [FREE Full text] [doi: 10.4300/JGME-D-22-00480.1] [Medline: 35991099]
18. Van Ryn RM, Nyrongo L. STI education for adolescents. SPARK (Scholarly Publications and Repository of Knowledge). 2023. URL: https://spark.siue.edu/dnpprojects/263 [accessed 2024-03-18]
19. Youth-centred digital health interventions: a framework for planning, developing and implementing solutions with and for young people. World Health Organization. 2020 Oct 26. URL: https://iris.who.int/handle/10665/336223 [accessed 2024-03-18]
20. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009 Oct;62(10):1006-1012 [FREE Full text] [doi: 10.1016/j.jclinepi.2009.06.005] [Medline: 19631508]
21. Salameh J, Bossuyt PM, McGrath TA, Thombs BD, Hyde CJ, Macaskill P, et al. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist. BMJ 2020 Aug 14;370:m2632 [FREE Full text] [doi: 10.1136/bmj.m2632] [Medline: 32816740]
22. Critical appraisal tools. JBI. URL: https://jbi.global/critical-appraisal-tools [accessed 2024-06-04]
23. Dulli L, Ridgeway K, Packer C, Murray KR, Mumuni T, Plourde KF, et al. A Social Media-Based Support Group for Youth Living With HIV in Nigeria (SMART Connections): Randomized Controlled Trial. J Med Internet Res 2020 Jun 02;22(6):e18343 [FREE Full text] [doi: 10.2196/18343] [Medline: 32484444]
24. Fakoya I, Cole C, Larkin C, Punton M, Brown E, Ballonoff Suleiman A. Enhancing Human-Centered Design With Youth-Led Participatory Action Research Approaches for Adolescent Sexual and Reproductive Health Programming. Health Promot Pract 2022 Jan;23(1):25-31. [doi: 10.1177/15248399211003544] [Medline: 33858267]
25. Gonsalves L, Hindin MJ, Bayer A, Carcamo CP, Gichangi P, Habib N, et al. Protocol of an open, three-arm, individually randomized trial assessing the effect of delivering sexual and reproductive health information to young people (aged 13-24) in Kenya and Peru via mobile phones: adolescent/youth reproductive mobile access and delivery initiative for love and life outcomes (ARMADILLO) study stage 2. Reprod Health 2018 Jul 11;15(1):126 [FREE Full text] [doi: 10.1186/s12978-018-0568-6] [Medline: 29996854]

26. Haruna H, Zainuddin Z, Mellecker RR, Chu SK, Hu X. An iterative process for developing digital gamified sexual health education for adolescent students in low-tech settings. ILS 2019 Nov 11;120(11/12):723-742. [doi: 10.1108/ils-07-2019-0066]
27. Nalwanga R, Nuwamanya E, Nuwasiima A, Babigumira JU, Asiimwe FT, Babigumira JB. Utilization of a mobile phone application to increase access to sexual and reproductive health information, goods, and services among university students in Uganda. Reprod Health 2021 May 17;18(1):95 [FREE Full text] [doi: 10.1186/s12978-020-01037-z] [Medline: 34001178]
28. Ippoliti NB, L'Engle K. Meet us on the phone: mobile phone programs for adolescent sexual and reproductive health in low-to-middle income countries. Reprod Health 2017 Jan 17;14(1):11 [FREE Full text] [doi: 10.1186/s12978-016-0276-z] [Medline: 28095855]
29. L'Engle KL, Mangone ER, Parcesepe AM, Agarwal S, Ippoliti NB. Mobile Phone Interventions for Adolescent Sexual and Reproductive Health: A Systematic Review. Pediatrics 2016 Sep 01;138(3):49. [doi: 10.1542/peds.2016-0884] [Medline: 27553221]
30. Nolan C, Packel L, Hope R, Levine J, Baringer L, Gatare E, et al. Design and impact evaluation of a digital reproductive health program in Rwanda using a cluster randomized design: study protocol. BMC Public Health 2020 Nov 13;20(1):1701 [FREE Full text] [doi: 10.1186/s12889-020-09746-7] [Medline: 33187485]
31. Onukwugha FI, Smith L, Kaseje D, Wafula C, Kaseje M, Orton B, et al. The effectiveness and characteristics of mHealth interventions to increase adolescent's use of Sexual and Reproductive Health services in Sub-Saharan Africa: A systematic review. PLoS One 2022;17(1):e0261973 [FREE Full text] [doi: 10.1371/journal.pone.0261973] [Medline: 35061757]
32. Patel A, Louie-Poon S, Kauser S, Lassi Z, Meherali S. Environmental scan of mobile apps for promoting sexual and reproductive health of adolescents in low- and middle-income countries. Front Public Health 2022;10:993795 [FREE Full text] [doi: 10.3389/fpubh.2022.993795] [Medline: 36504952]
33. Pfeiffer C, Kleeb M, Mbelwa A, Ahorlu C. The use of social media among adolescents in Dar es Salaam and Mtwara, Tanzania. Reprod Health Matters 2014 May;22(43):178-186 [FREE Full text] [doi: 10.1016/S0968-8080(14)43756-X] [Medline: 24908469]
34. Rogers E, Hemal K, Tembo Z, Mukanu M, Mbizvo M, Bellows B. Comprehensive Sexuality Education for Adolescents in Zambia via the Mobile-Optimized Website TuneMe: A Content Analysis. American Journal of Sexuality Education 2019 Jul 08;15(1):82-98. [doi: 10.1080/15546128.2019.1635546]
35. Rokicki S, Cohen J, Salomon JA, Fink G. Impact of a text-messaging program on adolescent reproductive health: a cluster-randomized trial in Ghana. Am J Public Health 2017 Feb;107(2):298-305. [doi: 10.2105/AJPH.2016.303562] [Medline: 27997236]
36. Frankline S, Dickson S, Omer T, Joseph B. A protocol for a randomized controlled trial on mobile phone text messaging to improve sexo-reproductive health in Cameroon. J Public Health Epidemiol 2020 Jul 31;12(3):179-185 [FREE Full text] [doi: 10.5897/jphe2019.1141]
37. Akande OW, Muzigaba M, Igumbor EU, Elimian K, Bolarinwa OA, Musa OI, et al. The effectiveness of an m-Health intervention on the sexual and reproductive health of in-school adolescents: a cluster randomized controlled trial in Nigeria. Reprod Health 2024 Jan 13;21(1):6 [FREE Full text] [doi: 10.1186/s12978-023-01735-4] [Medline: 38218840]
38. Haruna H, Hu X, Chu SKW, Mellecker RR, Gabriel G, Ndekao PS. Improving sexual health education programs for adolescent students through game-based learning and gamification. Int J Environ Res Public Health 2018 Sep 17;15(9):2027 [FREE Full text] [doi: 10.3390/ijerph15092027] [Medline: 30227642]
39. Alhassan RK, Abdul-Fatawu A, Adzimah-Yeboah B, Nyaledzigbor W, Agana S, Mwini-Nyaledzigbor PP. Determinants of use of mobile phones for sexually transmitted infections (STIs) education and prevention among adolescents and young adult population in Ghana: implications of public health policy and interventions design. Reprod Health 2019 Aug 09;16(1):120 [FREE Full text] [doi: 10.1186/s12978-019-0763-0] [Medline: 31399123]
40. Ippoliti N, Sekamana M, Baringer L, Hope R. Using human-centered design to develop, launch, and evaluate a national digital health platform to improve reproductive health for Rwandan youth. Glob Health Sci Pract 2021 Nov 29;9(Suppl 2):S244-S260 [FREE Full text] [doi: 10.9745/GHSP-D-21-00220] [Medline: 34845048]
41. Ackah J, Esia-Donkoh K, Amponsah A, Agbemavi W, Bediako V, Tettey G. Use of digital media by adolescents for sexual and reproductive health and rights communication in sub-Saharan Africa: a protocol for systematic review. Syst Rev 2024 May 14;13(1):130-167 [FREE Full text] [doi: 10.1186/s13643-024-02534-z] [Medline: 38745293]
42. Agbeve AS, Fiaveh DY, Anto-Ocrah M. Parent-adolescent sexuality communication in the African context: a scoping review of the literature. Sex Med Rev 2022 Oct;10(4):554-566 [FREE Full text] [doi: 10.1016/j.sxmr.2022.04.001] [Medline: 36210094]
43. Chipako I, Singhal S, Hollingsworth B. Impact of sexual and reproductive health interventions among young people in sub-Saharan Africa: a scoping review. Front Glob Womens Health 2024;5(2):1344135-1344218 [FREE Full text] [doi: 10.3389/fgwh.2024.1344135] [Medline: 38699461]
44. Oladele D, Iwelunmor J, Gbajabiamila T, Obiezu-Umeh C, Okwuzu J, Nwaozuru U, et al. The 4 youth by youth mHealth photo verification app for HIV self-testing in Nigeria: Qualitative analysis of user experiences. JMIR Form Res 2021 Nov 17;5(11):e25824-e25831 [FREE Full text] [doi: 10.2196/25824] [Medline: 34787579]

45. Chory A, Callen G, Nyandiko W, Njoroge T, Ashimosi C, Aluoch J, et al. A pilot study of a mobile intervention to support mental health and adherence among adolescents living with HIV in Western Kenya. AIDS Behav 2022 Jan;26(1):232-242 [FREE Full text] [doi: 10.1007/s10461-021-03376-9] [Medline: 34292429]
46. Haruna H, Okoye K, Zainuddin Z, Hu X, Chu S, Hosseini S. Gamifying sexual education for adolescents in a low-tech setting: Quasi-experimental design study. JMIR Serious Games 2021 Oct 12;9(4):e19614 [FREE Full text] [doi: 10.2196/19614] [Medline: 34636739]
47. Dowling R, Howell E, Dasco M, Schwartzman J. Digital adolescent sexual and reproductive health in low- and middle-income countries: A scoping review. Youth 2025 Feb 06;5(1):15-60. [doi: 10.3390/youth5010015]
48. No smartphone or internet? No problem; AI-backed phone has the answers. Reuters. 2024. URL: https://www.reuters.com/technology/no-smartphone-or-internet-no-problem-ai-backed-phone-has-answers-2024-03-15/ [accessed 2024-04-24]
49. Decker MJ, Harrison S, Price M, Gutmann-Gonzalez A, Yarger J, Tenney R. Educators' perspectives on integrating technology into sexual health education: implementation study. JMIR Hum Factors 2022 Jan 12;9(1):e31381 [FREE Full text] [doi: 10.2196/31381] [Medline: 35019842]
50. Aventin Á, Gough A, McShane T, Gillespie K, O'Hare L, Young H, et al. Engaging parents in digital sexual and reproductive health education: evidence from the JACK trial. Reprod Health 2020 Aug 27;17(1):132-765 [FREE Full text] [doi: 10.1186/s12978-020-00975-y] [Medline: 32854734]
51. Santa Maria D, Markham C, Bluethmann S, Mullen P. Parent-based adolescent sexual health interventions and effect on communication outcomes: a systematic review and meta-analyses. Perspect Sex Reprod Health 2015 Mar;47(1):37-50 [FREE Full text] [doi: 10.1363/47e2415] [Medline: 25639664]

Abbreviations

HCD: human-centered design
LMIC: low- and middle-income country
SRH: sexual and reproductive health
STI: sexually transmitted infection

Edited by A Mavragani; submitted 16.06.24; peer-reviewed by N Levitz, MDM Fernández Álvarez; comments to author 06.09.24; revised version received 09.10.24; accepted 24.01.25; published 26.02.25.

Please cite as:
Abdul Hamid Alhassan RH, Haggerty CL, Fapohunda A, Affan NJ, Anto-Ocrah M
Exploring the Use of Digital Educational Tools for Sexual and Reproductive Health in Sub-Saharan Africa: Systematic Review
JMIR Public Health Surveill 2025;11:e63309
URL: https://publichealth.jmir.org/2025/1/e63309
doi:10.2196/63309
PMID:40009849

Accessibility of eHealth Before and During the COVID-19 Pandemic Among People With and People Without Impairment: Repeated Cross-Sectional Survey

Linda Pettersson[1,2,3], MD; Stefan Johansson[4,5], PhD; Ingrid Demmelmaier[2], PhD; Lena von Koch[6,7], PhD; Jan Gulliksen[4], Prof Dr; Per-Olof Hedvall[5], PhD; Karl Gummesson[1], PhD; Catharina Gustavsson[1,2,8], PhD

[1]Center for Clinical Research Dalarna, Uppsala University, Nissers väg 3, Falun, Sweden
[2]Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden
[3]Primary Healthcare Center Mora, Region Dalarna, Mora, Sweden
[4]School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden
[5]Certec, Department of Design Sciences, Lund University, Lund, Sweden
[6]Division of Family Medicine and Primary Care, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden
[7]Theme Heart & Vascular and Neuro, Karolinska University Hospital, Stockholm, Sweden
[8]School of Health and Welfare, Dalarna University, Falun, Sweden

Corresponding Author:
Linda Pettersson, MD
Center for Clinical Research Dalarna, Uppsala University, Nissers väg 3, Falun, Sweden

Abstract

Background: The adoption of eHealth accelerated during the COVID-19 pandemic. Inequalities in the adoption of eHealth during the COVID-19 pandemic have been reported, but there are few such studies among people with impairment.

Objectives: This study aimed to investigate self-reported use and difficulty in the use of eHealth before the COVID-19 pandemic compared to during late social distancing restrictions in Sweden, among people with and without impairment, as well as between different types of impairment.

Methods: A cross-sectional survey was distributed twice by snowball sampling to people with self-reported impairment and a general population matched by age, gender, and county. Use and difficulty in the use of six eHealth services were compared between groups using chi-square test and logistic regression with year interaction terms, reported as odds ratio adjusted (aOR) for gender and age with 95% CI.

Results: The surveys included 1631 (in 2019) and 1410 (in 2021) participants with impairment, and 1084 (in 2019) and 1223 (in 2021) participants without. Participants with impairment, compared to those without impairment, reported less use and more difficulty in booking health care appointments online, digital identification, and the Swedish national web portal for health information and eHealth services (1177.se), both before and during the pandemic (*P*=.003 or lower). Video health care appointments were the exception to this disability digital divide in eHealth as video appointment adoption was the most likely among participants with attention, executive, and memory impairments (interaction term aOR 2.10, 95% CI 1.30 - 3.39). Nonuse and difficulty in the use of eHealth were consistently associated with language impairments and intellectual impairments. For example, language impairments were inversely associated with use of the logged-in eHealth services in 1177.se in 2021 (aOR 0.49, 95% CI 0.36 - 0.67) and were associated with difficulty in the use of 1177.se in 2019 (aOR 2.24, 95% CI 1.50 - 3.36) and the logged-in eHealth services in 1177.se in 2021 (aOR 1.89, 95% CI 1.32 - 2.70). Intellectual impairments were inversely associated with the use of the logged-in eHealth services in 1177.se in 2021 (aOR 0.19, 95% CI 0.13 - 0.27).

Conclusions: This repeated cross-sectional survey study, including participants with diverse types of impairment and a control group without impairment, reveals persisting disability digital divides, despite an accelerated adoption of eHealth across the pandemic. eHealth services were not accessible to some groups of people who were identified as being at risk of severe disease during the COVID-19 pandemic. This implies that all people could not use eHealth as a measure of infection protection.

(JMIR Public Health Surveill 2025;11:e64707) doi:10.2196/64707

KEYWORDS

eHealth; impairment; accessibility; digital inclusion; universal design; disability; digital divide; electronic health; COVID-19; pandemic; cross-sectional study; Sweden; online booking; digital identification; web portal; health information; control group; public health; digital health; digital literacy; health informatics; mobile phone

Introduction

Adoption of eHealth accelerated during the COVID-19 pandemic [1]. Before the pandemic, we identified a disability digital divide in eHealth, that is, people with impairment reported less use and more difficulty in the use of eHealth, particularly with some types of impairment (language impairments and intellectual impairments) [2]. Sustainable development [3], legislation, and human rights principles [4] demand accessible eHealth for all. During the pandemic, people with impairment have reported less use of COVID-19 digital services than people without impairment [5], and have reported difficulty in the use of eHealth [6]. Likewise, eHealth websites have failed to pass accessibility assessments [7]. Even though there are multiple studies and even reviews of associations between eHealth and other sociodemographic factors or somatic diseases, there are very few studies of how people with impairment perceive using eHealth. The three survey studies conducted prior to ours did not sample psychiatric or intellectual impairments, and they had no comparison to before the pandemic [5,6,8]. Further, the two of them that investigated the eHealth modality telehealth reported aggregated results of video health care appointments with telephone contacts.

The emergency phase of the COVID-19 pandemic in the spring of 2020 was characterized by uncertainty [9], and exceptional episodes of social distancing restrictions, including rapid digitalization [1]. Reprioritizing was a necessity, in order to provide acute health care in severe cases of COVID-19, triage with health care guidance, and public health assignments. Like the rest of the general population, health care workers were instructed to self-quarantine when having respiratory tract symptoms, which also limited health care availability [10]. To provide health care services in a social distancing manner, health care facilities were reorganized, separating dedicated areas for suspected COVID-19 cases [11] and the population was encouraged to redirect contacts to eHealth instead of face-to-face contacts [1]. Groups at risk for severe COVID-19 were identified and were recommended intensified social distancing restrictions [12], which included people with intellectual impairments, Down’s syndrome, schizophrenia, bipolar disorder, and stroke. The Swedish restrictions were reimposed in a second wave from autumn 2020 to summer 2021 and a third wave from January 2022 to September 2022. This was a time of realization and concern about prolonged postpandemic scenarios [9].

This study aimed to investigate self-reported use and difficulty in the use of eHealth before the COVID-19 pandemic compared to during late social distancing restrictions in Sweden among people with and without impairment, as well as between different types of impairment.

Methods

Study Design

This study had a repeated cross-sectional comparative design, using data from the “Swedes with impairment and the internet 2019” (SMFOI19) survey and the “Swedes with impairment and the internet 2021” (SMFOI21) survey, to people with diverse types of impairment and matched controls.

Study Setting

The publicly financed Swedish national web portal for health information and eHealth services (1177.se) is open to the whole population and is the most commonly used eHealth service in Sweden [13]. To access the personal eHealth services in 1177.se, it is required that the user have digital identification. Digital identification is also needed to access other private eHealth services. In the past decade, private eHealth services that provide video health care appointments have been launched [14]. Subsequently, this has encouraged the public health care regions to also provide video health care appointments.

Participants

Participants were people with self-reported impairment and people from the general population matched to the sample of people with impairment by gender, age, and county of residence. Impairment status was self-reported by one questionnaire item with 43 checkboxes of diagnoses and activity limitations and a free-text response option for reporting “Other impairment” (Multimedia Appendix 1). Participants who received the survey as matched controls, but responded that they had impairment, were reallocated to be analyzed as participants with impairment.

Procedures

The first version of this survey was developed in 2017 to mirror two nationwide Swedish surveys [15,16], out of which the latter is connected to Eurostat [17]. In the SMFOI19 survey, questions on eHealth services were added [2], and SMFOI21 also included questions on video health care appointments. SMFOI19 comprised 47 questions and SMFOI21 comprised 43 questions, on information and communication technology, impairments, and background characteristics. The wording of the questionnaires was developed in collaboration with the Begripsam group, all of whom have lived experience of impairment [18]. Several response modalities were provided (web, pen, phone, or on-site interview) and adaptations were provided upon request (sign language, reading support, speech therapist, or pictograms). Snowball sampling was undertaken from June to October 2019 (SMFOI19) and from May to August 2021 (SMFOI21), respectively [19]. The snowball sampling was distributed by disability networks, as well as by the participants themselves. Then, in February 2020 (SMFOI19) and in March 2022 (SMFOI21), the surveys were posted to six randomly selected matched controls per participant with impairment. Addresses were provided by the Swedish state personal address register [20].

Data Collection

This study used SMFOI19 and SMFOI21 questionnaire items on eHealth, impairments, and background characteristics.

Outcomes

Table 1 provides an overview of outcome variables. In summary, outcomes were measured on the use of eHealth services (ie, whether the services had been used and whether avoiding or preferring use) and difficulty in the use of eHealth services (ie, perceived difficulty or ease of use).

Table . Outcome variables measuring “use of eHealth” and “difficulty in the use of eHealth.”

Outcome variable followed by response options	In the SMFOI19[a] survey	In the SMFOI21[b] survey
Use of booking health care appointments online		
Checkbox items		
“I use booking medical appointments online” or “I use booking dental appointments online”	✓	✓
Use of video health care appointments		
Multiple-choice item		
“Yes, I used video health care appointments prior to the COVID-19 pandemic” or “Yes, I used video health care appointments during the COVID-19 pandemic”		✓
Comparisons were made respectively with the response option “No, I have not used video health care appointments,” whereas “Do not know” was treated as missing value		
Use of private eHealth services (which are services that primarily provide video health care appointments, chat, and drug prescriptions)		
Checkbox items		
Use of five Swedish private eHealth services or “Other private eHealth service”		✓
Use of the Swedish national web portal for health information and eHealth services (1177.se)		
Checkbox items		
“I use the health information website in the 1177.se”		✓
“I use the logged-in eHealth services in the 1177.se”		✓
Use of digital identification		
Checkbox items		
“I use the digital identification app BankID” or “I use other digital identification than BankID”	✓	✓
Avoid booking health care appointments online		
Multiple-choice item		
“If possible, I avoid booking medical appointments online” or “If possible, I avoid booking dental appointments online”	✓	✓
Comparison was made with the response option “I try to book all appointments online,” whereas “Not applicable” was treated as missing value.		
Avoid video health care appointments		
Multiple-choice item		
“If possible, I avoid video health care appointments”		✓
Comparison was made with the response option “I try to get all my appointments as video health care appointments,” whereas “Not applicable” was treated as missing value.		
Difficulty in the use of the Swedish national web portal for health information and eHealth services (1177.se)		
Multiple-choice items		

Outcome variable followed by response options	In the SMFOI19[a] survey	In the SMFOI21[b] survey
Two items in SMFOI21 “The health information website in 1177.se is difficult to use” or		✓
“The logged-in eHealth services in 1177.se is difficult to use.” For comparison between SMFOI19 and SMFOI21, they were combined.		✓
One single item in SMFOI19 “The Swedish national web portal for health information and eHealth services (1177.se) is difficult to use”	✓	
Comparison was made with the response option “It is easy to use,” whereas “Not applicable” was treated as missing value.		
Difficulty in the use of digital identification		
Multiple-choice item		
“It is difficult to use the digital identification app Mobile BankID” or “It is difficult to use other digital identification than BankID”	✓	✓
Comparison was made with the response option “It is easy to use,” whereas “Not applicable” was treated as missing value.		

[a] SMFOI19: Swedes with impairment and the internet 2019.

[b] SMFOI21: Swedes with impairment and the internet 2021.

Data on the use of eHealth were collected in both SMFOI19 and SMFOI21 on booking health care appointments online and digital identification. In SMFOI21, use of eHealth was also collected on video health care appointments “before the pandemic” and “during the pandemic,” private eHealth services, and the Swedish national web portal for health information and eHealth services (1177.se).

Data on the use of eHealth were also collected by questions on whether the respondents “Avoided” or “Preferred” booking health care appointments online in both SMFOI19 and SMFOI21, and video health care appointments in SMFOI21.

Data on difficulties in the use of eHealth were collected in both SMFOI19 and SMFOI21, on the use of the 1177.se and use of digital identification.

Independent Variables

Age and gender were selected to be confounding variables based on empirical studies of known demographic associations with the study outcomes [21,22]. Gender was analyzed as “female” and “male” whereas the response options “Other gender” and “Prefer not to answer” were handled as missing values in the analysis. Age was categorized for analysis into four categories: 16 - 29, 30 - 44, 45 - 69, and 70 years of age and older.

Data on impairment were collected by one question with 43 checkbox multiple response options on activity limitations (eg, difficulties in understanding) and diagnoses (eg, intellectual disability), and a free-text option for “Other impairment” (Multimedia Appendix 1). For the analysis, we applied a conceptual model of purposeful subgrouping of impairments, which we had developed in a previous study [2]. The categorization of types of impairment was based on our competence in human-computer interaction, digital accessibility, and medical science, and with the International Classification of Functioning, Disability, and Health (ICF)-classification as a framework [23], as well as empirical knowledge of comorbidity and functioning [24]. The grouping of impairments was done independently by three of the researchers (LP, SJ, and CG). There was almost complete interrater agreement and consensus in a conceptual model. Our conceptual model comprises 11 types of impairments by merging ICF-Body functions and ICF-Activities, as well as the 7 global and 11 specific ICF-Cognitive functions. The robustness of the conceptual model was assessed in a cluster analysis [25] using SPSS (version 26.0; IBM Corp).

Background Characteristics

Data on participant characteristics were collected to describe the study samples in regard to educational level, native language, occupation, income, urban or rural status, accommodation, cohabitant status, professional support in everyday life, and access to digital devices.

Data Analysis

We compared 11 outcomes between participants with and without impairment, as well as between different types of impairment. Six of the outcomes were compared between the survey years (2021 and 2019; Table 1). The statistical software R (version 4.0.3; R Core Team) was used for the analyses. Pair-wise deletion was applied. The significance level was set to ≤.05. Chi-square test and logistic regression with robust standard errors (Huber-White) adjusted for gender and age were used. Multicollinearity was assessed in relation to a predetermined cut-off. Sensitivity analyses indicated the robustness of the estimates, that is, models were fitted with and without independent variables with wide CIs. Results from the following analyses are visualized in figures and presented as numeric values in Multimedia Appendix 2. Among all participants overall, SMFOI21 was compared to SMFOI19, by logistic regression models, adjusted for types of impairment, in addition to adjusting for confounders. Comparisons between participants with impairment to those without were made by chi-square test. To identify types of impairment associated with nonuse and difficulty in the use of eHealth, year-stratified confounder-adjusted logistic regression was used. Comparisons between types of impairment in SMFOI21 to SMFOI19 were performed by the input of multiplicative interaction terms of types of impairment with survey year [26].

Ethical Considerations

This study was approved by the Swedish Ethical Review Authority (registration 2022-00184-01). The participants provided informed consent prior to participation. No compensation was provided for participation. This study used anonymized data.

Results

The flowchart of participation (Figure 1) includes 1631 (in SMFOI19) and 1410 (in SMFOI21) participants with impairment, along with 1084 (in SMFOI19) and 1223 (in SMFOI21) participants without impairment.

Compared to participants without impairment, more of the participants with impairment were educated in a special school, were outside the labor market or not working, had lower income, lived in rental apartments or supported accommodation, were living alone, had professional support in everyday life, and lacked access to digital devices (Table 2).

In both survey years, participants with impairment had a median of 2 (IQR 1-6 in SMFOI19 and 1-5 in SMFOI21) impairments. The types of impairment were distributed in approximately similar proportions in the two survey years (Figure 2).

Figure 1. Flowchart of participation in the SMFOI19 and SMFOI21 surveys. SMFOI19: Swedes with impairment and the internet 2019; SMFOI21: Swedes with impairment and the internet 2021.

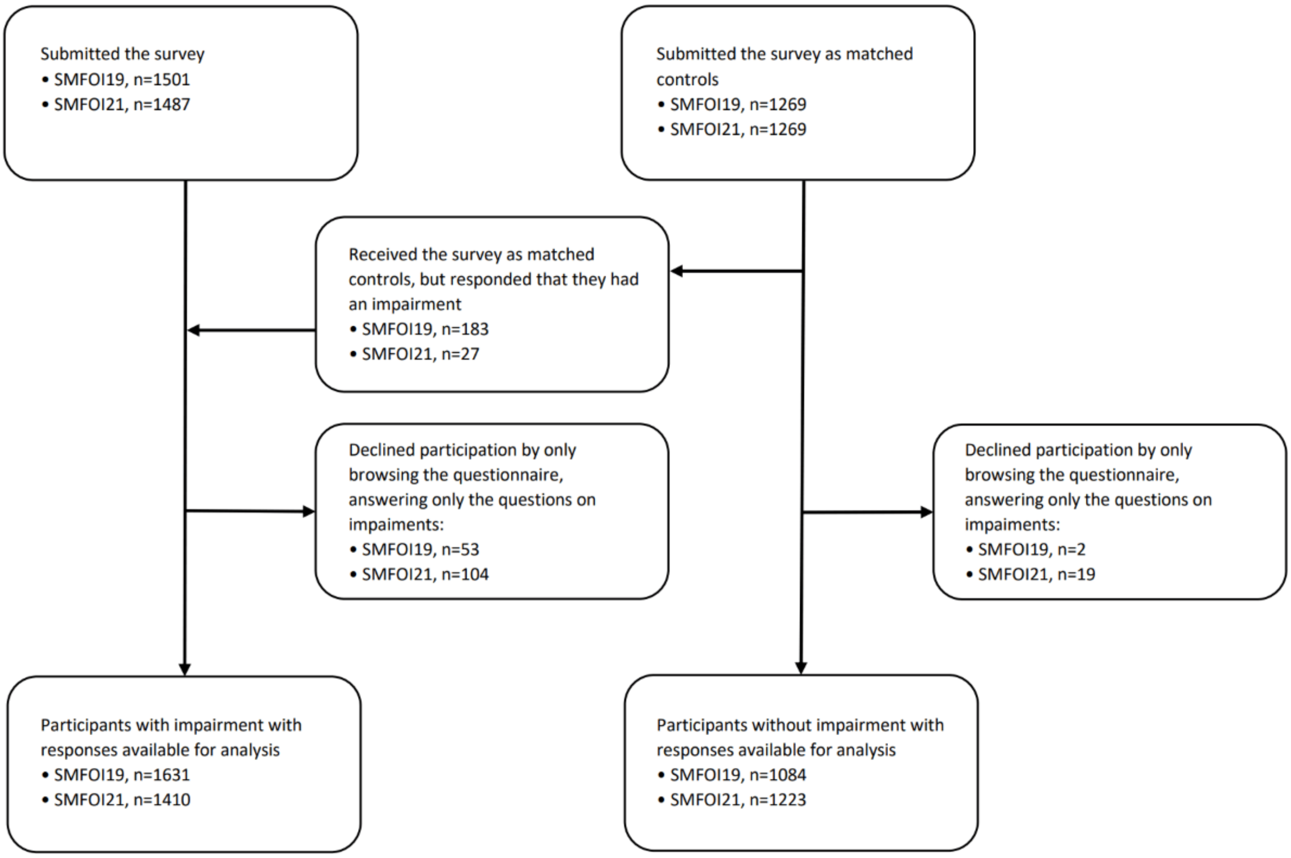


Table . Background characteristics of participants with and without impairment.

	Participants with impairment in the SMFOI19[a] survey (n=1631)	Participants without impairment in the SMFOI19 survey (n=1084)	Participants with impairment in the SMFOI21[b] survey (n=1410)	Participants without impairment in the SMFOI21 survey (n=1223)
Gender, n (%)	n=1356	n=1060	n=1213	n=1193
Women	937 (69)	780 (74)	797 (66)	824 (69)
Men	419 (31)	280 (26)	416 (34)	369 (31)
Age (years), n (%)	n=1388	n=1069	n=1257	n=1200
16 - 29	176 (13)	58 (5)	186 (15)	48 (4)
30 - 44	304 (22)	156 (15)	269 (21)	173 (14)
45 - 69	773 (56)	716 (67)	599 (48)	681 (57)
≥70	135 (10)	139 (13)	203 (16)	298 (25)
Education, n (%)	n=1339	n=1040	n=1222	n=1190
Compulsory school	134 (10)	74 (7)	118 (10)	79 (7)
Special education school	152 (11)	2 (0)	113 (9)	10 (1)
Upper secondary school, Vocational school, or Folk high school (adult education)	567 (42)	442 (43)	488 (40)	447 (38)
University	486 (36)	522 (50)	503 (41)	654 (55)
Native language, n (%)	n=1361	n=1067	n=1253	n=1196
Swedish	1252 (92)	994 (93)	1136 (91)	1065 (89)
Other than Swedish	109 (8)	73 (7)	117 (9)	131 (11)
Occupation, n (%)	n=1367	n=1060	n=1244	n=1194
Retired	273 (17)	304 (28)	327 (26)	478 (40)
Outside of the labor market (disability-related early retirement, daily activity center, or temporary disability allowance)	519 (38)	18 (2)	344 (28)	14 (1)
Working	492 (30)	745 (69)	342 (27)	647 (54)
Student	136 (8)	36 (3)	111 (9)	26 (2)
On the labor market, but not working (unemployed, parental-leave, or sick-leave)	132 (10)	22 (2)	120 (10)	29 (2)
Monthly income (in Swedish Krona, SEK), n (%)	n=1157	n=866	n=1058	n=931
<5000	53 (5)	10 (1)	53 (5)	8 (1)
5000 - 24,999	714 (62)	228 (26)	632 (60)	278 (30)
≥25,000	390 (34)	628 (73)	373 (35)	645 (69)
Urban or rural status, n (%)	n=1368	n=1068	n=1239	n=1188
Rural	191 (14)	176 (16)	169 (14)	168 (14)
Suburban or town	325 (24)	252 (24)	284 (23)	290 (24)
Urban	852 (62)	640 (60)	786 (63)	730 (61)
Accommodation, n (%)	n=1365	n=1059	n=1233	n=1194
Supported accommodation (group living, service apartment, or other supported accommodation)	104 (8)	1 (0)	70 (6)	1 (0)

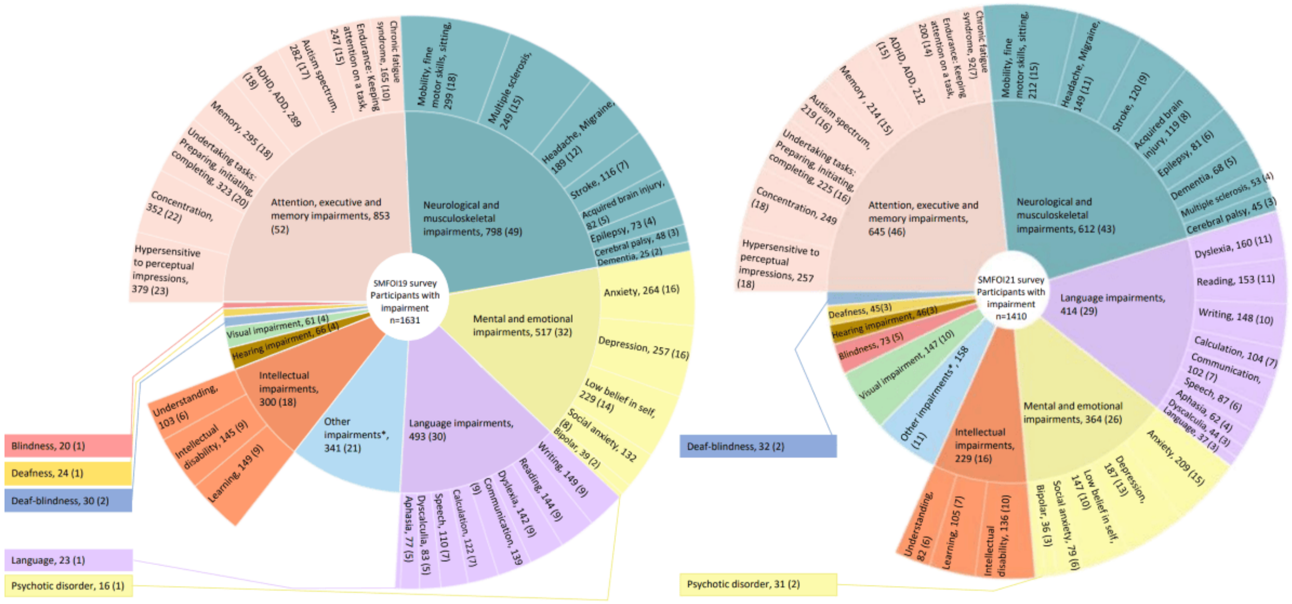
	Participants with impairment in the SMFOI19[a] survey (n=1631)	Participants without impairment in the SMFOI19 survey (n=1084)	Participants with impairment in the SMFOI21[b] survey (n=1410)	Participants without impairment in the SMFOI21 survey (n=1223)
Group living	52 (4)	1 (0)	38 (3)	0 (0)
Service apartment	43 (3)	0 (0)	23 (2)	1 (0)
Other supported accommodation	9 (1)	0 (0)	9 (1)	0 (0)
Apartment, condominium, house, or homeless	1261 (92)	1058 (100)	1163 (94)	1193 (100)
Apartment	470 (34)	168 (16)	407 (33)	225 (19)
Condominium	285 (21)	239 (22)	282 (23)	267 (22)
House	506 (37)	651 (61)	468 (38)	701 (59)
Homeless	0 (0)	0 (0)	6 (0)	0 (0)
Cohabitant status, n (%)	n=1365	n=1066	n=1249	n=1197
Living alone	549 (40)	183 (17)	487 (39)	231 (19)
Cohabiting	816 (60)	883 (83)	762 (61)	966 (81)
Professional support in everyday life, n (%)	n=1381	n=1063	n=1236	n=1190
Have professional support in everyday life	497 (36)	6 (1)	398 (32)	19 (2)
Home based support by municipal care services	112 (8)	2 (0)	91 (7)	4 (0)
Personal assistants	90 (7)	0 (0)	61 (5)	0 (0)
Supported-living staff, support persons or similar	214 (15)	0 (0)	166 (13)	0 (0)
Appointed guardian	111 (8)	0 (0)	71 (6)	0 (0)
Relative	81 (6)	4 (0)	152 (12)	6 (1)
Other support	54 (4)	4 (0)	55 (4)	7 (1)
No support	884 (64)	1057 (99)	838 (68)	1171 (98)
Access to digital devices, n (%)	n=1456	n=1067	n=1316	n=1213
Lack access to at least computer and one portable device	241 (17)	93 (9)	205 (16)	117 (10)
No device	29 (2)	7 (1)	21 (2)	19 (2)
Only computer at home	75 (5)	24 (2)	66 (5)	22 (2)
Only smartphone	78 (5)	19 (2)	44 (3)	30 (2)
Only tablet	17 (1)	1 (0)	13 (1)	1 (0)
Smartphone and tablet	42 (3)	39 (4)	61 (5)	45 (4)
Have access to at least computer and one portable device	1215 (83)	974 (91)	1111 (84)	1096 (90)
Computer and smartphone	418 (29)	264 (25)	395 (30)	362 (30)
Computer and tablet	65 (4)	25 (2)	60 (5)	19 (2)
Computer, smartphone, and tablet	732 (51)	688 (64)	656 (50)	715 (59)
Number of reported impairments	n=1631	n=1084	n=1410	n=1223

	Participants with impairment in the SMFOI19[a] survey (n=1631)	Participants without impairment in the SMFOI19 survey (n=1084)	Participants with impairment in the SMFOI21[b] survey (n=1410)	Participants without impairment in the SMFOI21 survey (n=1223)
Median (IQR), Maximum value	2 (1-6), 20	0 (0-0), 0	2 (1-5), 35	0 (0-0), 0

[a] SMFOI19: Swedes with impairment and the internet 2019.

[b] SMFOI21: Swedes with impairment and the internet 2021.

Figure 2. Description of participants' types of impairment (subgrouped in the inner circle and disaggregated to the level of the questionnaire in the outer circle). The labels are followed by number and proportion of all participants with impairment, n (%). Adds up to over 100% as multiple responses were allowed. The impairment question construction covered both activity limitations (eg, reading difficulties) and diagnoses (eg, dyslexia; Multimedia Appendix 1). *Free-text "other impairments" were in falling frequency: pain, genetic syndromes, paresis, and posttraumatic stress disorder, followed by infrequent responses. ADD: attention deficit disorder; ADHD: attention-deficit/hyperactivity disorder; SMFOI19: Swedes with impairment and the internet 2019; SMFOI21: Swedes with impairment and the internet 2021.



Comparisons of use between the survey years, as well as between participants with and without impairment, are reported in Figures 3a and 4a. Among all participants, all eHealth services were used by more participants in SMFOI21 compared to SMFOI19: booking health care appointments online (adjusted odds ratio [aOR] 2.73, 95% CI 2.39 - 3.10), video health care appointments (aOR 3.34, 95% CI 2.81 - 3.96), and digital identification (aOR 1.74, 95% CI 1.43 - 2.10; Figures 3a and 4a). In both survey years, all but two of the eHealth services were used by fewer participants with impairment compared to participants without impairment: booking health care appointments online ($P<.001$), the information pages in 1177.se ($P<.001$), the logged-in eHealth services in 1177.se ($P<.001$), and digital identification ($P<.001$; Figures 3a and 4a). The exceptions were video health care appointments, which were used by 9% in both groups ($P=.93$) before the pandemic, and 29% with impairment versus 19% without impairment ($P<.001$) during the pandemic, and private eHealth services ($P=.10$; Figure 3a).

Figure 3. Use of booking health care appointments online, video health care appointments, and private eHealth services. The group with impairment (a) is disaggregated into the subgrouped types of impairment from Figure 2 (b-d). The numeric values visualized in this figure are presented in Tables S1 and S3-S5 in Multimedia Appendix 2. *aOR: adjusted odds ratio from logistic regression comparing use of eHealth among all participants overall between before and during the pandemic (reference group is before the pandemic), adjusted for type of impairment (reference is participants without impairment), gender (reference female), and age (reference <30 years of age); †2-sided *P* value from chi-square test comparing participants with and without impairment. ‡Intercept from year-stratified logistic regression (reference group is participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age). §*P* value from logistic regression interaction terms of year with types of impairment (reference is before the pandemic and participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age).

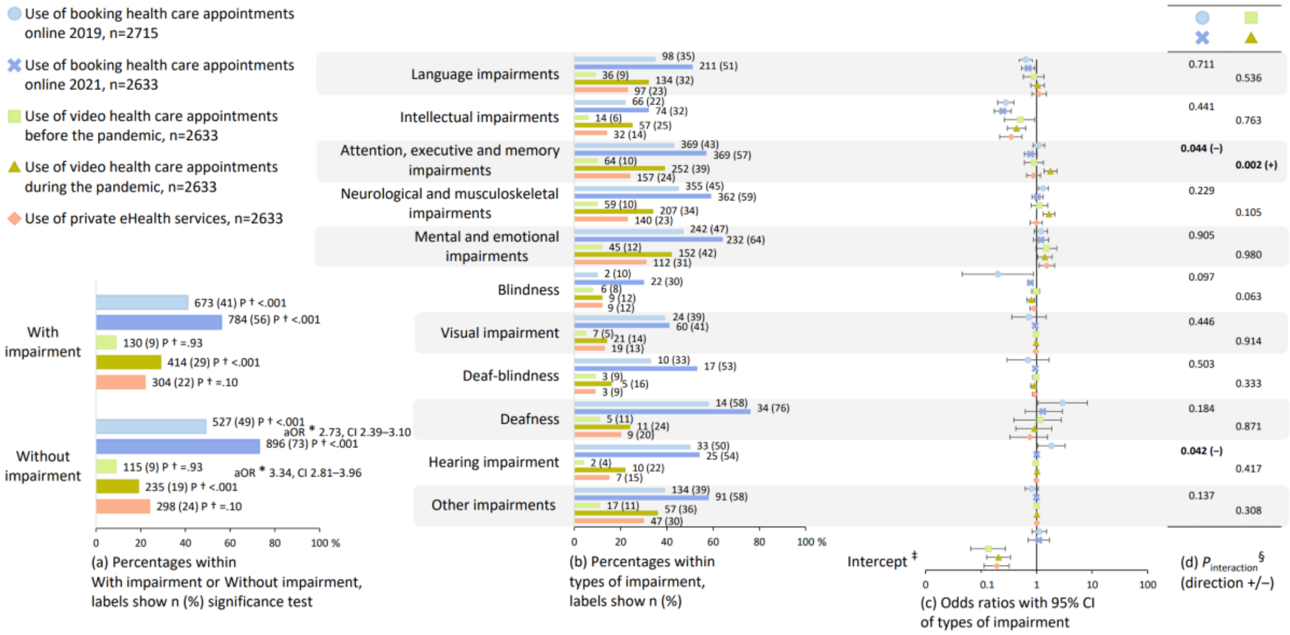
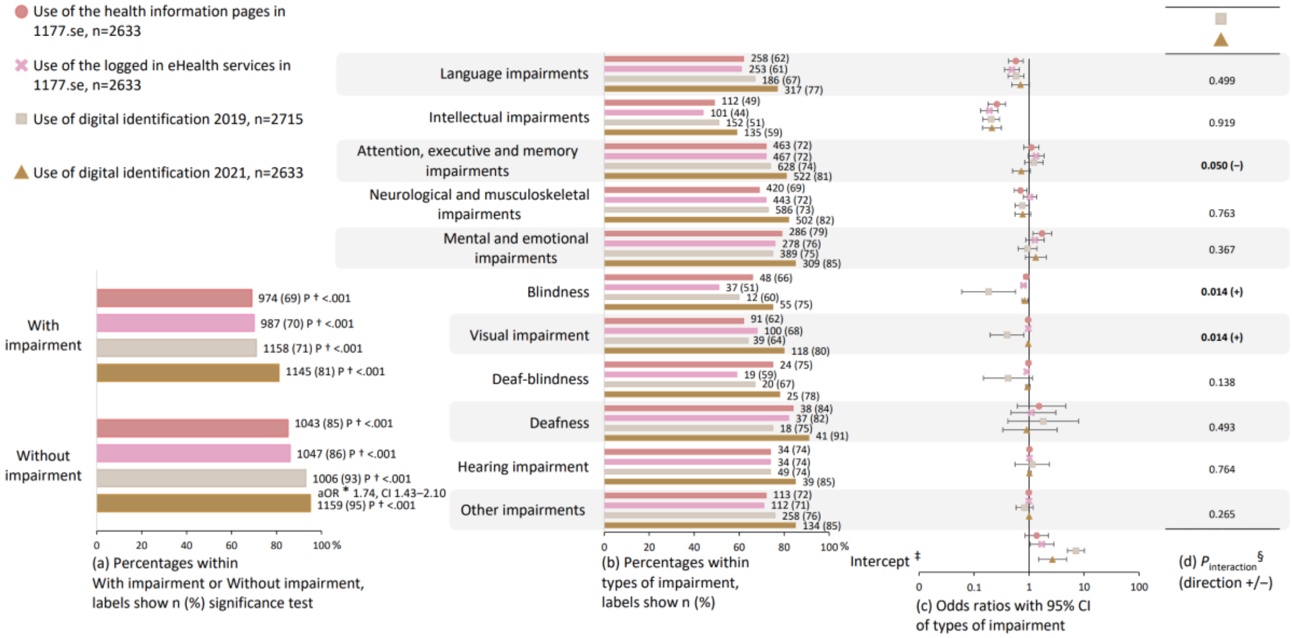


Figure 4. Use of the Swedish national web portal for health information and eHealth services (1177.se) and digital identification. The group with impairment (a) is disaggregated into the subgrouped types of impairment from Figure 2 (b–d). The numeric values visualized in this figure are presented in Tables S1, S6, and S7 in Multimedia Appendix 2. *aOR: adjusted odds ratio from logistic regression comparing use of eHealth among all participants overall between the SMFOI19 and SMFOI21 surveys (reference group is SMFOI19), adjusted for type of impairment (reference is participants without impairment), gender (reference female), and age (reference <30 years of age); †2-sided *P* value from chi-square test comparing participants with and without impairment. ‡Intercept from year-stratified logistic regression (reference group is participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age). §*P* value from logistic regression interaction terms of year with types of impairment (reference is SMFOI19 and participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age). SMFOI19: Swedes with impairment and the internet 2019; SMFOI21: Swedes with impairment and the internet 2021.



Associations between use and types of impairment are reported in Figures 3b, c, 4b, c, and further, changes in the associations between the survey years are reported in Figures 3d and 4d.

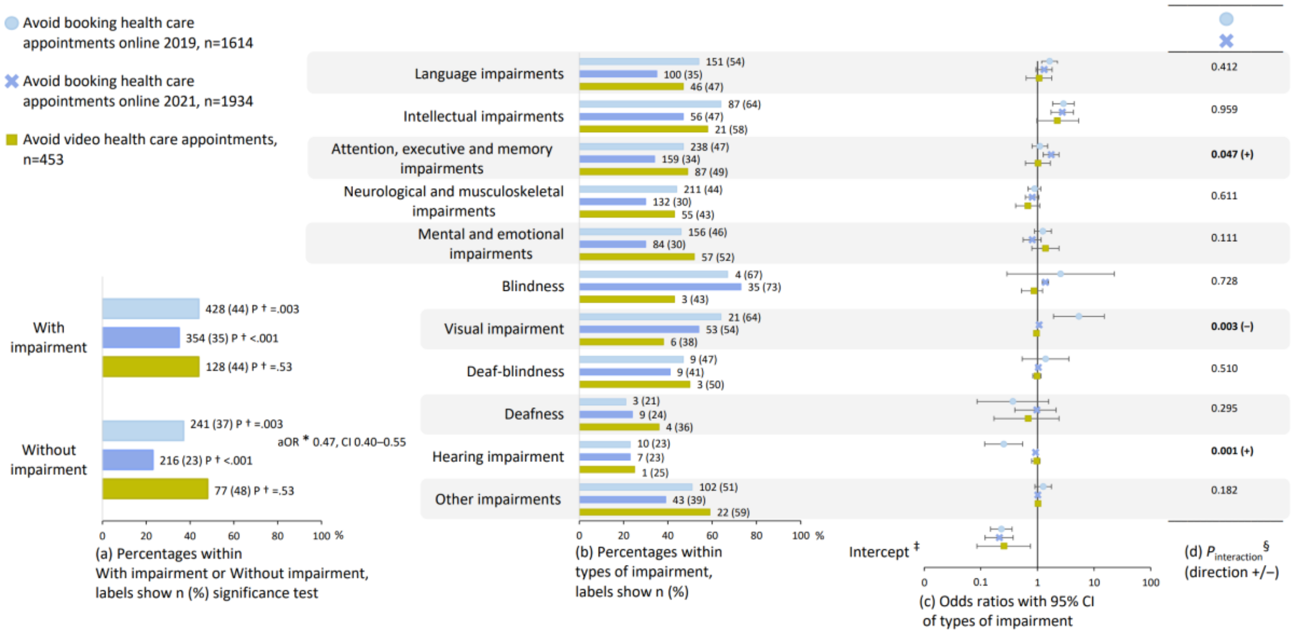
Booking health care appointments online was inversely associated with language impairments in SMFOI19 (aOR 0.64, 95% CI 0.49‐0.83) and in SMFOI21 (aOR 0.70, 95% CI

0.54 - 0.93), with intellectual impairments in SMFOI19 (aOR 0.28, 95% CI 0.20 - 0.39) and in SMFOI21 (aOR 0.25, 95% CI 0.17 - 0.35), and with blindness in SMFOI19 (aOR 0.20, 95% CI 0.05 - 0.88) and in SMFOI21 (aOR 0.78, 95% CI 0.70 - 0.86; Figure 3c). The use of video health care appointments was inversely associated with intellectual impairments before the pandemic (aOR 0.51, 95% CI 0.26 - 0.93) and during the pandemic (aOR 0.44, 95% CI 0.30 - 0.64; Figure 3c). Comparing during the pandemic to before, the use of video health care appointments was the most likely to increase among participants with attention, executive, and memory impairments (interaction term aOR 2.10, 95% CI 1.30 - 3.39; Figure 3d). The use of private eHealth services was associated with mental and emotional impairments (aOR 1.53, 95% CI 1.11 - 2.11) and inversely associated with intellectual impairments (aOR 0.35, 95% CI 0.22 - 0.55; Figure 3c). Use of the 1177.se was inversely associated with three types of impairments: language impairments (aOR 0.58, 95% CI 0.43 - 0.78 and aOR 0.49, 95% CI 0.36 - 0.67), intellectual impairments (aOR 0.26, 95% CI 0.18 - 0.38 and aOR 0.19, 95% CI 0.13 - 0.27), and blindness (aOR 0.89, 95% CI 0.80 - 1.00 and aOR 0.80, 95% 0.72 - 0.88; Figure 4c). The use of digital identification was inversely associated with intellectual impairments in SMFOI19 (aOR 0.21, 95% CI 0.15 - 0.29), and in SMFOI21 (aOR 0.21, 95% CI 0.14 - 0.32), and blindness in SMFOI19 (aOR 0.19, 95% CI 0.06 - 0.57) and SMFOI21 (aOR 0.84, 95% CI 0.75 - 0.96; Figure 4c).

Comparisons of avoiding use between the survey years, as well as between participants with and without impairment, are reported in Figure 5a. Among all participants, avoiding booking health care appointments online was less frequent in SMFOI21 compared to SMFOI19 (aOR 0.47, 95% CI 0.40 - 0.55). Booking health care appointments online was avoided by more participants with impairment compared to participants without impairment in SMFOI19 (44% vs 37%; *P*=.003) and SMFOI21 (35% vs 23%; *P*<.001; Figure 5a). Regarding avoiding video health care appointments, no difference was found between participants with and without impairment (44% vs 48%; *P*=.53; Figure 5).

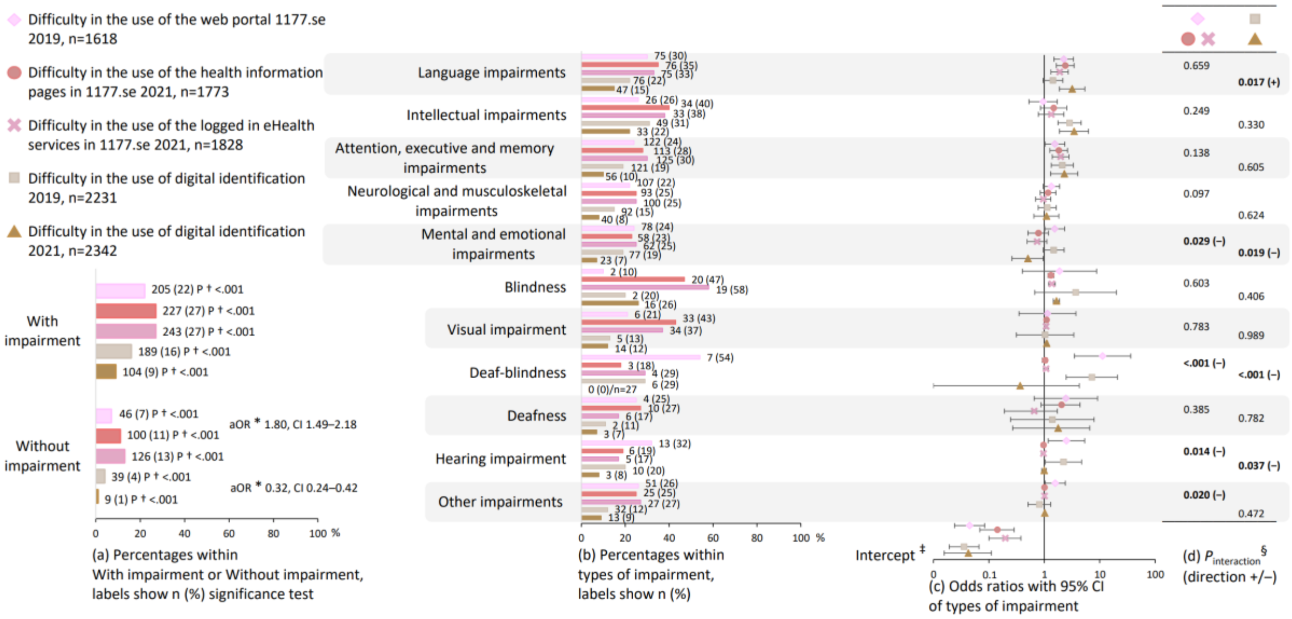
Associations between avoiding use and types of impairment are reported in Figure 5b,c, and further, changes in the associations between the survey years are reported in Figure 5d. Avoiding booking health care appointments online was associated with intellectual impairments in SMFOI19 (aOR 2.88, 95% CI 1.86 - 4.45) and SMFOI21 (aOR 2.74, 95% CI 1.74 - 4.33), and visual impairment in SMFOI19 (aOR 5.40, 95% CI 1.92 - 15.18) and SMFOI21 (aOR 1.07, 95% CI 1.04 - 1.10; Figure 5c). Regarding avoiding video health care appointments, no associations were found with types of impairment (Figure 5c).

Figure 5. Avoid booking health care appointments online and video health care appointments. The group with impairment (a) is disaggregated into the subgrouped types of impairment from Figure 2 (b-d). The numeric values visualized in this figure are presented in Tables S2, S8, and S9 in Multimedia Appendix 2. *aOR: adjusted odds ratio from logistic regression comparing avoidance of eHealth among all participants overall between the SMFOI19 and SMFOI21 surveys (reference group is 2019), adjusted for type of impairment (reference is participants without impairment), gender (reference female), and age (reference <30 years of age); †2-sided *P* value from chi-square test comparing participants with and without impairment. ‡Intercept from year-stratified logistic regression (reference group is participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age). §*P* value from logistic regression interaction terms of year with types of impairment (reference is SMFOI19 and participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age). SMFOI19: Swedes with impairment and the internet 2019; SMFOI21: Swedes with impairment and the internet 2021.



Comparisons of difficulty in the use of eHealth services between the survey years, as well as between participants with and without impairment, are reported in Figure 6a. Among all participants, difficulty in the use of the 1177.se was more frequent in SMFOI21 compared to SMFOI19 (aOR 1.80, 95% CI 1.49 - 2.18; Figure 6a). Difficulty in the use of digital identification was less frequent in SMFOI21 compared to SMFOI19 (aOR 0.32, 95% CI 0.24 - 0.42; Figure 6a). More participants with impairment than participants without impairment reported difficulty in the use of the 1177.se in SMFOI19 (22% vs 7%; *P*<.001), the information pages in 1177.se in SMFOI21 (27% vs 11%; *P*<.001), the logged-in eHealth services in 1177.se in SMFOI21 (27% vs 13%; *P*<.001), and digital identification in SMFOI19 (16% vs 4%; *P*<.001), and in SMFOI21 (9% vs 1%; *P*<.001; Figure 6a).

Figure 6. Difficulty in the use of the Swedish national web portal for health information and eHealth services (1177.se) and digital identification. The group with impairment (a) is disaggregated into the subgrouped types of impairment from Figure 2 (b–d). The numeric values visualized in this figure are presented in Tables S2, S10, and S11 in Multimedia Appendix 2. *aOR: adjusted odds ratio from logistic regression comparing difficulty in the use of eHealth among all participants overall between the SMFOI19 and SMFOI21 surveys (reference group is 2019), adjusted for type of impairment (reference is participants without impairment), gender (reference female) and age (reference <30 years of age); †2-sided *P* value from chi-square test comparing participants with and without impairment. ‡Intercept from year-stratified logistic regression (reference group is participants without impairment), adjusted for gender (reference female) and age (reference <30 years of age). §*P* value from logistic regression interaction terms of year with types of impairment (reference is SMFOI19 and participants without impairment), adjusted for gender (reference female) and age (reference <30 years of age). SMFOI19: Swedes with impairment and the internet 2019; SMFOI21: Swedes with impairment and the internet 2021.



Associations between difficulty in the use of eHealth services and types of impairment are reported in Figure 6b,c, and further, changes in the associations between the survey years are reported in Figure 6d. Difficulty in the use of the 1177.se was associated with two types of impairments: language impairments and attention, executive, and memory impairments (Figure 6c). Language impairments were associated with difficulty in the use of the 1177.se in SMFOI19 (aOR 2.24, 95% CI 1.50 - 3.36), the 1177.se information pages (aOR 2.39, 95% CI 1.65 - 3.44), and the logged-in eHealth services (aOR 1.89, 95% CI 1.32 - 2.70) in SMFOI21. Attention, executive, and memory impairments were associated with difficulty in the use of the 1177.se in SMFOI19 (aOR 1.55, 95% CI 1.03 - 2.34), the 1177.se information pages (aOR 1.82, 95% CI 1.26 - 2.62), and the logged-in eHealth services (aOR 1.97, 95% CI 1.39 - 2.78) in SMFOI21. Difficulty in the use of digital identification was associated with intellectual impairments in SMFOI19 (aOR 2.86, 95% CI 1.77 - 4.62) and in SMFOI21 (aOR 3.44, 95% CI 1.87 - 6.26), and attention, executive, and memory impairments in SMFOI19 (aOR 2.11, 95% CI 1.33 - 3.34) and in SMFOI21 (aOR 2.30, 95% CI 1.30 - 4.02; Figure 6c).

Discussion

Principal Findings

In a rapid digitalization in health care, this first comparison across the COVID-19 pandemic revealed disability digital divides in eHealth that remained proportional. In this comparison of people with and without impairment, the former were disfavored in regard to eHealth. Namely, they reported less use of booking health care appointments online, the 1177.se and digital identification, which is in line with most studies before the pandemic as discussed in our previous article [2] and one during the pandemic [5]. Further, they reported substantial difficulty in the use of eHealth, which is in agreement with a survey conducted during the pandemic [6]. Nonuse and difficulty in the use of eHealth were consistently associated with language impairments and intellectual impairments. Our findings show that the highest adoption of video health care appointments was among participants with attention, executive, and memory impairments, which is in line with other surveys during the pandemic [8,27]. The latter of the cited studies concluded that video health care appointments increased among mental health service users when contacts were redirected to eHealth instead of face-to-face [27]. Accordingly, it is plausible that the adoption of video health care appointments among participants with attention, executive, and memory impairments in this study was due to disrupted face-to-face health care provision.

The Swedish national web portal for health information and eHealth services, 1177.se, was more frequently reported as difficult to use in SMFOI21 than in SMFOI19, both among participants with and without impairment, despite an accelerated adoption of eHealth across the pandemic. This may be interpreted as a sign of forced digitalization, which implies that eHealth followed a previously shown societal pattern of nonaccessible social distancing restrictions [28]. People with intellectual impairments were afflicted on two fronts. They were distinguished as a risk group for severe COVID-19 [12] and thereby recommended intensified social distancing restrictions, which meant that they were encouraged to redirect contacts to eHealth instead of face-to-face contacts with health care if possible, while at the same time, they were facing substantial difficulties in the use of eHealth. Cognitive accessibility is insufficiently covered in eHealth studies [29,30] and guidelines [30]. Known adaptations for cognitive accessibility are intuitive structure, important elements before page scroll, large icons, visualizations, and third-party participation [31,32]. However, there is a lack of conformance to what is known about accessibility [7]. Moreover, in this study, people with language impairments were particularly afflicted by difficulties in the use of digital services, which has previously been shown by multiple studies [10,33], but not all [5]. Text and numerical tasks can be adapted by standards for understandable text, autocorrect, data import, speech recognition, audio, and enabling enough duration [31,32]. Barriers to the use of video appointments can be due to reduced nonverbal communication [10] and natural pauses [33].

Implications

Until this disability digital divide in eHealth is bridged, practitioners should provide eHealth integrated with staff support, or maintain conventional services as alternatives to digitalized options. In line with policy [34], we underscore that research and development must conform to legislated accessibility standards [4,31,32,35]. But in addition, we suggest that user participation in the design of eHealth [36] and adherence to principles for universal design [37] can bring additional benefits, beyond the accessibility standards. Such benefits exceed accessibility, by, for instance, readability, clarity, interoperability, and assistive functions, which improve usability for all users [32,35].

Strengths and Limitations

This is, to our knowledge, the first study comparing the accessibility of eHealth across the COVID-19 pandemic among people with diverse types of impairment relevant to the risk of disablement by poorly designed digital services. The pandemic provoked exceptional conditions, which can provide insights for developing resilient health care systems, in which people with high health care needs are included. This comparison between two cross-sectional surveys was deemed a suitable study design to study an unanticipated historical event, in the absence of an available cohort study.

Snowball sampling [19] is a major reason for achieving representation of seldom-heard populations. However, it is plausible that this sampling method engaged more people with higher digital literacy to respond to the survey. Further, self-assessment of digital literacy can be associated with overestimation of skills [38]. Thus, in summation, while this study did demonstrate the disability digital divide in eHealth, it may have underestimated its magnitude. In the two study samples, participants' characteristics and types of impairment demonstrate similar proportions, which makes it plausible that misclassification did not obstruct drawing inferences on comparisons between samples. To allow representation of individuals with multimorbidity in the analysis, we assessed that multicollinearity did not exceed the predetermined cut-off, and that outcomes were not associated with the number of reported impairments. The SMFOI21 survey sampled more participants in the oldest group (70 years or older) and less participants of 45‐69 years of age, as compared to the SMFOI19 survey. This was handled as we adjusted for age in the statistical analysis. While we adjusted for gender and age, there were socioeconomic differences between people with and without impairment, for which we did not control. The dependent variables were investigated by different types of response options “I use”; “I avoid” or “I prefer”; as well as “is difficult to use” or “is easy to use” in order to capture different aspects of the phenomenon. We consider that avoiding is a dimension of using, in that it measures a type of “not using.” However, it is possible that avoidance also can be associated with difficulty of use, but may also be influenced by other mechanisms, for example, self-efficacy, normative beliefs, or social support. This study focused on the association between types of impairment and eHealth services for patient-provider contact, meaning that other types of eHealth and other factors plausibly mediating the use of eHealth were not investigated, that is, eHealth literacy, socioeconomic factors, or lacking access to digital devices.

Conclusions

Despite an accelerated adoption of eHealth across the COVID-19 pandemic, our results show no relevant difference in the disability digital divide in eHealth between the survey years. In both before and during the pandemic, there were substantial nonuse and difficulty in the use of eHealth among people with impairment. Nonuse and difficulty in the use of eHealth were consistently associated with language impairments and intellectual impairments. This implies that groups recognized as being at risk of severe COVID-19 could not use eHealth, while emphasis was put on eHealth as a measure of infection protection in the social distancing policy. Accessibility standards can improve services. Until then, our results indicate that eHealth in its current state should be provided with optional support, or only be provided as a complement to conventional contacts.

Acknowledgments

Our research partners, the Begripsam group, all of whom have lived experience of impairment, contributed to the study design and development of the questionnaire of the first SMFOI survey (2017) and participated in further elaboration and refinement of the following surveys. Riccardo Lo Martire, statistician at the Center for Clinical Research Dalarna Sweden, provided consultation.

The study was funded by the Center for Clinical Research Dalarna, Sweden, and FORTE - The Swedish Research Council for Health, Working Life and Welfare (grant 2021-01875). The funders had no part in the undertaking of the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Question about impairments.
[DOCX File, 30 KB - publichealth_v11i1e64707_app1.docx]

Multimedia Appendix 2
Tables of all logistic regression models of this study. Unadjusted and adjusted model per outcome.
[DOCX File, 91 KB - publichealth_v11i1e64707_app2.docx]

References

1. Tu K, Sarkadi Kristiansson R, Gronsbell J, et al. Changes in primary care visits arising from the COVID-19 pandemic: an international comparative study by the International Consortium of Primary Care Big Data Researchers (INTRePID). BMJ Open 2022 May 9;12(5):e059130. [doi: 10.1136/bmjopen-2021-059130] [Medline: 35534063]
2. Pettersson L, Johansson S, Demmelmaier I, Gustavsson C. Disability digital divide: survey of accessibility of eHealth services as perceived by people with and without impairment. BMC Public Health 2023 Jan 27;23(1):181. [doi: 10.1186/s12889-023-15094-z] [Medline: 36707791]
3. UN General Assembly. Transforming our world: the 2030 agenda for sustainable development. A/RES/70/1. 2015 URL: http://digitallibrary.un.org/record/1654217 [accessed 2022-04-08]
4. Directive (EU) 2016/2102 of the European Parliament and of the Council of 26 October 2016 on the accessibility of the websites and mobile applications of public sector bodies. 2016. URL: https://eur-lex.europa.eu/eli/dir/2016/2102/oj [accessed 2024-01-31]
5. Park EY. Effect of COVID-19 on internet usage of people with disabilities: a secondary data analysis. Int J Environ Res Public Health 2022 Jun 25;19(13):7813. [doi: 10.3390/ijerph19137813] [Medline: 35805471]
6. Lawford BJ, Hinman RS, Morello R, Oliver K, Spittle A, Bennell KL. Perceptions about the efficacy and acceptability of telephone and video-delivered allied health care for adults with disabilities during the COVID-19 pandemic: a cross-sectional national survey. Arch Phys Med Rehabil 2022 Jul;103(7):1368-1378. [doi: 10.1016/j.apmr.2021.12.017] [Medline: 35041838]
7. Jonsson M, Gustavsson C, Gulliksen J, Johansson S. How have public healthcare providers in Sweden conformed to the European Union’s Web Accessibility Directive regarding accessibility statements on their websites? Univ Access Inf Soc 2025 Mar;24(1):449-462. [doi: 10.1007/s10209-023-01063-1]
8. Xie Z, Hong YR, Jo A, Marlow NM. Telehealth utilization during the covid-19 pandemic among people with functional disabilities in the U.S.: a preliminary analysis of national health interview survey 2020 data. AJPM Focus 2023 Dec;2(4):100149. [doi: 10.1016/j.focus.2023.100149] [Medline: 37920402]
9. Gulliksen J, Lilliesköld J, Stenbom S. The ‘New’ new normal—digitalization and hybridization of work and education before, during and after the COVID-19 pandemic. Interact Comput 2023 Dec 31;35(5):691-706. [doi: 10.1093/iwc/iwac034]
10. Matenge S, Sturgiss E, Desborough J, Hall Dykgraaf S, Dut G, Kidd M. Ensuring the continuation of routine primary care during the COVID-19 pandemic: a review of the international literature. Fam Pract 2022 Jul 19;39(4):747-761. [doi: 10.1093/fampra/cmab115] [Medline: 34611708]
11. Fernemark H, Skagerström J, Seing I, Hårdstedt M, Schildmeijer K, Nilsen P. Working conditions in primary healthcare during the COVID-19 pandemic: an interview study with physicians in Sweden. BMJ Open 2022 Feb 8;12(2):e055035. [doi: 10.1136/bmjopen-2021-055035] [Medline: 35135771]
12. The National Board of Health and Welfare. Final report on risk for severe COVID-19 (S2021/00825) [in swedish: slutredovisning av regeringsuppdrag om risk för särskilt allvarlig sjukdomsutveckling vid COVID-19]. 2021 URL: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/dokument-webb/ovrigt/slutredovisning-riskgrupper-covid-19.pdf [accessed 2023-06-15]
13. Inera Company. 2024. URL: https://www.inera.se/tjanster/statistik-for-ineras-tjanster/statistik-for-1177.se/ [accessed 2024-05-23]
14. Ekman B, Ellegård L. Digitaliseringen av svensk vård och omsorg. In: The Digitalization of Swedish Healthcare: SNS Förlag; 2023:120.
15. Andersson J, Blomdahl F. The swedes and the internet 2022. The Swedish Internet Foundation (IIS). 2022 Oct 26 URL: https://svenskarnaochinternet.se/english/earlier-years-reports/ [accessed 2023-04-14]
16. Statistics Sweden (SCB). Use of computers and the internet by private persons 2016. 2016 Nov 24 URL: https://www.scb.se/hitta-statistik/statistik-efter-amne/forskning-och-det-digitala-samhallet/det-digitala-samhallet/

befolkningens-it-anvandning/pong/publikationer/privatpersoners-anvandning-av-datorer-och-internet-2016/ [accessed 2019-01-05]
17. Eurostat. Digital economy and society—database ICT usage in households and by individuals. URL: http://ec.europa.eu/eurostat/web/digital-economy-and-society/database [accessed 2024-01-31]
18. Begripsam. The Begripsam Company and the Begripsam Group Sweden. URL: https://www.begripsam.se/english [accessed 2024-01-31]
19. Lavrakas P. Encyclopedia of Survey Research Methods: SAGE Publications Ltd; 2008.
20. The Swedish state personal address register (SPAR). The Swedish Tax Agency. URL: https://www.statenspersonadressregister.se/master/start/english-summary [accessed 2024-01-31]
21. Reiners F, Sturm J, Bouw LJW, Wouters EJM. Sociodemographic factors influencing the use of eHealth in people with chronic diseases. Int J Environ Res Public Health 2019 Feb 21;16(4):645. [doi: 10.3390/ijerph16040645] [Medline: 30795623]
22. Mangin D, Parascandalo J, Khudoyarova O, Agarwal G, Bismah V, Orr S. Multimorbidity, eHealth and implications for equity: a cross-sectional survey of patient perspectives on eHealth. BMJ Open 2019 Feb 12;9(2):e023731. [doi: 10.1136/bmjopen-2018-023731] [Medline: 30760515]
23. Ustün TB, Chatterji S, Bickenbach J, Kostanjsek N, Schneider M. The International Classification of Functioning, Disability and Health: a new tool for understanding disability and health. Disabil Rehabil 2003;25(11-12):565-571. [doi: 10.1080/0963828031000137063] [Medline: 12959329]
24. Grigorenko EL, Compton DL, Fuchs LS, Wagner RK, Willcutt EG, Fletcher JM. Understanding, educating, and supporting children with specific learning disabilities: 50 years of science and practice. Am Psychol 2020 Jan;75(1):37-51. [doi: 10.1037/amp0000452] [Medline: 31081650]
25. MacQueen J, editor. Some methods for classification and analysis of multivariate observations. 1967 Presented at: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability; Oakland, CA, USA URL: file://C:/Users/gokul/Downloads/1200512992%20(1).pdf
26. VanderWeele TJ, Knol MJ. A tutorial on interaction. Epidemiol Methods 2014 May 27;3(1):33-72. [doi: 10.1515/em-2013-0005]
27. Ivancic L, Bond DM, Nassar N. Impact of the COVID-19 pandemic on access and use of health services by middle-aged and older Australians. Aust Health Rev 2023 Feb;47(1):100-109. [doi: 10.1071/AH22183] [Medline: 36395815]
28. Chadwick D, Ågren KA, Caton S, et al. Digital inclusion and participation of people with intellectual disabilities during COVID‐19: a rapid review and international bricolage. Policy Practice Intel Disabi 2022 Sep;19(3):242-256. [doi: 10.1111/jppi.12410]
29. Jonsson M, Johansson S, Hussain D, Gulliksen J, Gustavsson C. Development and evaluation of eHealth services regarding accessibility: scoping literature review. J Med Internet Res 2023 Aug 17;25:e45118. [doi: 10.2196/45118] [Medline: 37590050]
30. European Commission: Directorate-General for Communications Networks, Content and Technology, Funka ICF, Kjellstrand S, Laurin S, et al. Inclusive web-accessibility for persons with cognitive disabilities—pilot project study—final report. European Commission. 2022 URL: https://data.europa.eu/doi/10.2759/3048 [accessed 2022-04-06]
31. ISO/TS 82304-2:2021. Health software—part 2: health and wellness apps—quality and reliability. URL: https://www.iso.org/standard/78182.html [accessed 2022-03-07]
32. WHO-ITU. Global standard for accessibility of telehealth services. Geneva: World Health Organization and International Telecommunication Union Licence: CC BY-NC-SA 3.0 IGO. 2022 URL: https://www.who.int/publications/i/item/9789240050464 [accessed 2022-12-01]
33. Ward K, Vagholkar S, Sakur F, Khatri NN, Lau AYS. Visit types in primary care with telehealth use during the COVID-19 pandemic: systematic review. JMIR Med Inform 2022 Nov 28;10(11):e40469. [doi: 10.2196/40469] [Medline: 36265039]
34. World Health Organization. Global strategy on digital health 2020-2025. Geneva: World Health Organization. Licence: CC BY-NC-SA 3.0 IGO. 2021 URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2022-04-07]
35. ISO 21801-1:2020. Cognitive accessibility—part 1: general guidelines. URL: https://www.iso.org/standard/71711.html [accessed 2022-09-27]
36. World Wide Web Consortium (W3C). Involving users in evaluating web accessibility. URL: https://www.w3.org/WAI/test-evaluate/involving-users [accessed 2024-06-15]
37. Persson H, Åhman H, Yngling AA, Gulliksen J. Universal design, inclusive design, accessible design, design for all: different concepts—one goal? On the concept of accessibility—historical, methodological and philosophical aspects. Univers Access Inf Soc 2015 Nov;14(4):505-526. [doi: 10.1007/s10209-014-0358-z]
38. Merritt K, Smith D, Di Renzo Jr JC. An investigation of self-reported computer literacy: is it reliable? Issues Inf Syst 2005;6(1):289-295. [doi: 10.48009/1_iis_2005_289-295]

Abbreviations

aOR: adjusted odds ratio

ICF: International Classification of Functioning, Disability, and Health
SMFOI19: Swedes with impairment and the internet 2019
SMFOI21: Swedes with impairment and the internet 2021

Edited by AS Bhagavathula; submitted 30.07.24; peer-reviewed by M Koks-Leensen; revised version received 09.01.25; accepted 18.02.25; published 28.03.25.

Please cite as:
Pettersson L, Johansson S, Demmelmaier I, von Koch L, Gulliksen J, Hedvall PO, Gummesson K, Gustavsson C
Accessibility of eHealth Before and During the COVID-19 Pandemic Among People With and People Without Impairment: Repeated Cross-Sectional Survey
JMIR Public Health Surveill 2025;11:e64707
URL: https://publichealth.jmir.org/2025/1/e64707
doi:10.2196/64707

Health Equity of Hypertension Management Between Local Residents and Internal Migrants in Shenzhen, China: Cross-Sectional Study

Jinyu He[1,2], MD; Yanjun Li[1], BA; Huatang Zeng[1,3], MSc; Haoran Sun[1,2], MD; Liqun Wu[3], MSc; Zhenzhen Zhu[3], MSc; Ning Zhang[1,2], PhD; Wannian Liang[1,2], PhD

[1]Vanke School of Public Health, Tsinghua University, Beijing, China

[2]Institute of Healthy China, Tsinghua University, Haidian District, Beijing, China

[3]Public Hygiene and Health Commission of Shenzhen Municipality, Shenzhen Health Development Research and Data Management Center, Shenzhen, China

Corresponding Author:
Ning Zhang, PhD
Vanke School of Public Health, Tsinghua University, Beijing, China

Abstract

Background: With hypertension emerging as a global public health concern, taking measures to alleviate its burden is urgently needed. The hypertension management program (HMP) in China is a standardized policy to help people with hypertension to improve their health levels and reduce health inequalities. However, studies focusing on details regarding participation in this program remain scarce.

Objective: This study aims to investigate the participation rates in HMPs in China and examine the health disparities in hypertension management between local residents and internal migrants in Shenzhen.

Methods: This study used the medical record of Shenzhen, Guangdong, China. We included adults with new-onset diagnosis of hypertension after 2017 and focused on patients who have a record in the community health center. We described the basic characteristics of people diagnosed with hypertension, including age, gender, marital status, occupation, education level, and health insurance type. Enrolled rate, follow-up rate, and adherence rate were used to measure the engagement with this program at the city, district, and community levels.

Results: Of the 1,160,214 hypertensive patients, 29.70% (344,526/1,160,214) were local residents, while 70.3% (815,688/1,160,214) were internal migrants. In addition, 55.36% (642,250/1,160,214) were enrolled in the HMP. Of those, 57.52% (369,439/642,25) were followed up. In addition, 62.59% (231,217/369,439) of followed up individuals were adherents. Internal migrants demonstrated a significantly higher participation in the HMP, especially for the enrolled rate (local residents: 156,085/344,526, 45.30% vs internal migrants: 486,165/815,688, 59.60%) and adherence rate (local residents: 44,044/84,635, 52.04% vs internal migrants: 187,173/284,804, 65.72%). Apart from that, local, single, and younger individuals had lower rates compared to their counterparts. There also existed within-district and within-community variation among enrolled, follow-up, and adherence rates.

Conclusions: Based on our research, individuals with different demographic and socioeconomic characteristics and in different regions had different enrolled, follow-up, and adherence rates. Internal migrants in Shenzhen showed a significantly higher participation in the HMP. Given these findings, there exists the potential to enhance the outreach and engagement of local, single, and younger populations through targeted promotional strategies.

(JMIR Public Health Surveill 2025;11:e65548) doi:10.2196/65548

KEYWORDS

health equity; hypertension management; immigrant population; hypertension; China; global health; public health; health disparities; medical record; community health care; native population; immigrant; socioeconomic; disease burden; well-being; chronic disease; community health center

Introduction

Background

The United Nations' Sustainable Development Goals are important global issues. The goals focus on promoting health and well-being, as well as reducing inequalities, highlighting the need for all individuals, regardless of their circumstances, to be able to access quality health care and essential services for optimal health outcomes.

Hypertension presents a significant global health challenge, emerging as a primary contributor to the global disease burden. Analysis from the Global Burden of Disease Study reveals that high systolic blood pressure stands as the foremost risk factor for mortality worldwide, accounting for 10.8 million cardiovascular disease deaths and 11.3 million deaths overall in 2021 [1]. The World Health Organization's latest report highlights a surge in hypertension prevalence, with figures doubling from 650 million individuals in 1990 to an alarming 1.3 billion in 2019 [2]. In China, about 270 million people have hypertension, while research showed that, among the 256 million individuals aged 30-79 years with hypertension, only 52% were aware of their condition, 39% were receiving treatment, and a scant 16% had their blood pressure under control [2,3]. In China, hypertension is one of the earliest chronic diseases undergoing widespread public health intervention. The government has bolstered efforts to combat hypertension through initiatives such as the Basic Public Health Services policy, China Medium- and Long-term Plan for the Prevention and Treatment of Chronic Diseases, National Demonstration Areas for Comprehensive Chronic Disease Prevention and Control, and the National Healthy Lifestyle Action. Chinese strategies underscore the inclusion of hypertension management within basic public health services at the grassroots level, ensuring regular dissemination of national guidelines on hypertension management and treatment. These guidelines provide comprehensive details on the management processes and protocols for diagnosed patients and vulnerable individuals [4]. According to the National Standard for Basic Public Health Services [5], the hypertension management program (HMP) is one of most crucial programs within the Basic Public Health Services. It requires community health centers (CHCs) to establish dedicated medical records for hypertensive patients and provide regular follow-up services without additional fees beyond outpatient costs. According to performance assessments, there must be at least 1 follow-up visit record every 3 months.

Significant health inequities exist in the health policy context; health-related behaviors and beliefs; quality of care; utilization of health care services; and health status [6]. Research has highlighted health inequities between local residents and internal migrants in China. Many internal migrants still face significant barriers in accessing local public health services, particularly in establishing health records and receiving health education. This is more common in economically developed eastern regions, where internal migrants use public health services less often than those in central regions, indicating regional disparities [7]. Furthermore, internal migrants often have poorer health compared to local residents due to socioeconomic and cultural factors, with social integration and income inequality being key determinants. In 2017, internal migrants in Qingdao had relatively better health, while those in Shenzhen faced the greatest health challenges [8].

In terms of hypertension management, treatment, and control within the population [9], previous research has explored the potential impacts of multidimensional factors on the health inequity of hypertension and cardiovascular disease [10]; the factors can be summarized as individual, socioeconomic [11,12], and environmental. A nationally representative cross-sectional study in China revealed that age, BMI, and economic status contributed to the inequitable situation of hypertension prevalence [13]. Another study showed that economic status and educational level are associated with socioeconomic inequalities of health services utilization among patients with hypertension in the Pearl River Delta of China [14]. In addition, due to economic status, the hypertension burden appeared higher in rural areas compared to urban regions in China [15].

As China's first Special Economic Zone, Shenzhen has undergone sweeping changes, including a significant demographic shift due to a large influx of internal migrants from other parts of Mainland China. According to The Shenzhen Statistical Yearbook 2022, the total population of Shenzhen was 17,681,600, of which the registered household population was 5,563,900 (31.5%) and the nonregistered population was 12,117,700 (68.5%). Despite this, Shenzhen, similar to many other regions in China, prioritizes the provision of health care services to its local residents [16]. For example, older adults with local household registrations or local medical insurance are prioritized for free influenza vaccination, while internal migrants without local insurance have to pay out of pocket [17]. Local people aged ≥60 years are eligible for free health checks, whereas internal migrants are required to be ≥65 years old [18].

As a survey indicated that the prevalence rate of hypertension among the residents of Shenzhen is 20.74% [5], and the incidence of hypertension in Shenzhen is increasing, the HMP is accessible to all patients regardless of household registration status, with the aim of reducing health disparities; however, as of yet, there has been no research comparing the health inequity in hypertension management between internal migrants and local residents. In the context of an internal migrant–dense city, understanding the unique challenges and needs of both local residents and internal migrants is essential.

This Study

The primary objective of this baseline report was to describe the characteristics of hypertensive patients and those who have received HMP services at a CHC, as well as the current status of the specific services they receive. The secondary objective was to reveal the health inequity for utilization of hypertension management between local residents and internal migrants in Shenzhen.

Methods

Study Design and Data Source

This study used a cross-sectional design. The data of the study were obtained from CHCs and the government-owned hospitals in Shenzhen, Guangdong, China. All data were extracted with SQL queries from original datasets at the Shenzhen Health Development Research and Data Management Center. The data included the patients' basic information as well as the actual hypertension record and follow-up service record from January 1, 2017, to September 30, 2023. We defined participants without Shenzhen household registration who had been living in Shenzhen for 6 months or more as internal migrants. The HMP is a key component of primary care services for hypertensive patients. Therefore, this study uses the enrollment rate, follow-up rate, and adherence rate of this program to assess the health equity in utilization of primary care services between local residents and internal migrants.

Eligibility Criteria of Resident Records

We included the records of the following residents: (1) those who were documented in either the health information system of a CHC or medical record of government-owned hospitals in Shenzhen, Guangdong, China from January 1, 2017, to September 30 2023; (2) those who were diagnosed with hypertension according to these information systems (patients with gestational hypertension, based on the International Classification of Diseases Tenth Revision code, were excluded); (3) those whose earliest diagnosis date was after January 1, 2017, and (4) those who were aged ≥18 years, which is the age threshold for blood pressure measurements when people visit medical institutions as per the hypertension management policy in Shenzhen.

Sociodemographic Variables

A total of 6 individual-level sociodemographic factors were included in this study: gender (male or female), age (<45 years, 45 - 65 years, and >65 years), marital status (single, married, widowed, or divorced), occupation (technical personnel/office staff, business service industry, agriculture and forestry, industry, other, unemployed), education level (Bachelor's degree or above, college diploma, high school or technical school, junior high school, elementary school or below), and health insurance type (urban employees, urban residents, others, fully self-funded).

Key Variables

The participation in the HMP in this research was assessed using 3 measures. First, "enrollment" in the program was defined as having a hypertensive record, which includes the diagnosis date, blood pressure at the time of diagnosis, lifestyle, specific information about the physicians and institutions providing the services, and patient contact information. This indicates that the patient has visited the CHC and the doctor had the opportunity to interact with them. Enrollment in the HMP is the initial step for doctors when encountering a hypertensive patient for the first time according to the policy. Second, follow-up in this study was defined as patients having at least 2 follow-up records at different times. Third, regular follow-up was defined as patients having an average of more than 4 follow-up visits per year, in accordance with the national policy, which mandates that hypertensive patients should be followed up at least 4 times per year. We used 3 corresponding ratios to evaluate engagement. The enrolled rate refers to the proportion of individuals enrolled out of all patients who have ever been diagnosed with hypertension. The follow-up rate indicates the number of individuals who are followed up out of those enrolled. The adherence rate represents the proportion of individuals who undergo regular follow-up out of the total number of those who are followed up.

Statistical Analysis

Individuals were assessed by household register (internal migrants or local residents). Indicators were estimated for participation in the HMP; for the whole city, this was stratified by sociodemographic factors (resident status, gender, age category, occupation, education level, and insurance type). The city was also stratified by district and community. Categorical variables and prevalence rates were expressed as absolute frequencies and percentages. Missing values were categorized as 1 group for each variable. Logistic regression was used to evaluate the main characteristics associated with participation in the HMP among all patients and internal migrants and local residents as a sensitivity analysis. A 2-sample *z*-test for proportions was used to evaluate whether the observed differences in proportions between internal migrants and local residents were statistically significant. The map figure was drawn using ArcMap 10.8 (Esri) and the other figures were drawn using OriginPro (version 2020b; OriginLab). Statistical analyses were conducted using R Studio (version 4.2.2; Posit).

Ethical Considerations

The data used in this study were anonymized and encrypted, ensuring no access to personally identifiable information. Our data access was authorized through an agreement with government departments. In addition, our study is 1 component of a larger study that was approved by The Tsinghua University Science and Technology Ethics Committee (20230065).

Results

Demographic Characteristics

The total number of diagnosed individuals recorded in the electronic health record was 1,160,214, with 344,526 local residents (29.7%) and 815,688 internal migrants (70.3%). In the overall population, 58.24% (n=675,692) were male and 41.76% (n=484,522) were female. More demographic characteristics of the diagnosed population are detailed in Table 1. Males constituted a higher proportion across all age groups and resident categories, with a noticeable decline in older age groups. Higher educational attainment was more prevalent in younger age groups. Internal migrants had higher proportions of individuals with lower educational levels (elementary school or below). Insurance coverage varied, with urban employee insurance being more common in younger and middle-aged groups, while urban resident insurance was more prevalent in the oldest age group.

Table . Characteristics of residents with hypertension from the electronic health record in Shenzhen by age group (January 1, 2017, to September 30, 2023).

	Total (n=1,160,214), n (%)				Local population (n=344,526), n (%)				Migrant population (n=815,688), n (%)			
Characteristics	18 - 45 years (n=279,179)	46 - 65 years (n=609,238)	>65 years (n=271,797)	Subtotal	18 - 45 years (n=72,361)	46 - 65 years (n=175,887)	>65 years (n=96,278)	Subtotal	18 - 45 years (n=206,818)	46 - 65 years (n=433,351)	>65 years (n=175,519)	Subtotal
Sex												
Male	195,225 (69.9)	356,461 (58.5)	124,006 (45.6)	675,692 (58.2)	47,125 (65.1)	102,826 (58.5)	46,050 (47.8)	196,001 (56.9)	148,100 (71.6)	253,635 (58.5)	77,956 (44.4)	479,691 (58.8)
Female	83,954 (30.1)	252,777 (41.5)	147,791 (54.4)	484,522 (41.8)	25,236 (34.9)	73,061 (41.5)	50,228 (52.2)	148,525 (43.1)	587,18 (28.4)	179,716 (41.5)	97,563 (55.6)	335,997 (41.2)
Marital status												
Single	62,445 (22.4)	35,202 (5.8)	11,072 (4.1)	108,719 (9.4)	18,165 (25.1)	16,762 (9.5)	7141 (7.4)	42,068 (12.2)	44280 (21.4)	18,440 (4.3)	3931 (2.2)	66,651 (8.2)
Married	213,769 (76.6)	567,735 (93.2)	249,868 (91.9)	1,031,372 (88.9)	53,405 (73.8)	156,995 (89.3)	84,909 (88.2)	295,309 (85.7)	160,364 (77.5)	410,740 (94.8)	164,959 (94.0)	736,063 (90.2)
Widowed or divorced	1860 (0.7)	4738 (0.8)	10,293 (3.8)	16,891 (1.5)	358 (0.5)	1553 (0.9)	3935 (4.1)	5846 (1.7)	1502 (0.7)	3185 (0.7)	6358 (3.6)	11,045 (1.4)
Missing	1105 (0.4)	1563 (0.3)	564 (0.2)	3,232 (0.3)	433 (0.6)	577(0.3)	293 (0.3)	1303 (0.4)	672 (0.3)	986 (0.2)	271 (0.2)	1929 (0.2)
Occupation												
Technical personnel or office staff	58,423 (20.9)	86,191 (14.1)	34,917 (12.8)	179,531 (15.5)	25,601 (35.4)	49,862 (28.3)	21,889 (22.7)	97,352 (28.3)	32,822 (15.9)	36,329 (8.4)	13,028 (7.4)	82,179 (10.1)
Business service industry	68,022 (24.4)	134,895 (22.1)	20,419 (7.5)	223,336 (19.2)	18,089 (25.0)	41,752 (23.7)	11,077 (11.5)	70,918 (20.6)	49,933 (24.1)	93,143 (21.5)	9342 (5.3)	152,418 (18.7)
Agriculture and forestry	2617 (0.9)	13,573 (2.2)	23,599 (8.7)	39,789 (3.4)	432 (0.6)	2035 (1.2)	4213 (4.4)	6680 (1.9)	2185 (1.1)	11,538 (2.7)	19,386 (11.0)	33,109 (4.1)
Industry	70,811 (25.4)	114,431 (18.8)	9548 (3.5)	194,790 (16.8)	3024 (4.2)	6877 (3.9)	3988 (4.1)	13,889 (4.0)	67,787 (32.8)	107,554 (24.8)	5560 (3.2)	180,901 (22.2)
Other	37,624 (13.5)	78,876 (12.9)	26,787 (9.9)	143,287 (12.4)	16,709 (23.1)	38,393 (21.8)	16,096 (16.7)	71,198 (20.7)	20,915 (10.1)	40,483 (9.3)	10,691 (6.1)	72,089 (8.8)
Unemployed	41,271 (14.8)	180,768 (29.7)	156,430 (57.6)	378,469 (32.6)	8463 (11.7)	36,932 (21.0)	38,987 (40.5)	84,382 (24.5)	32,808 (15.9)	143,836 (33.2)	117,443 (66.9)	294,087 (36.1)
Missing	411 (0.1)	504 (0.1)	97 (0.0)	1012 (0.1)	43 (0.1)	36 (0.0)	28 (0.0)	107 (0.0)	368 (0.2)	468 (0.1)	69 (0.0)	905 (0.1)
Education level												
Bachelor’s degree or above	42,576 (15.3)	43,995 (7.2)	10,930 (4.0)	97,501 (8.4)	25,497 (35.2)	33,534 (19.1)	7483 (7.8)	66,514 (19.3)	17,079 (8.3)	10,461 (2.4)	3447 (2.0)	30,987 (3.8)
College diploma	46,343 (16.6)	51,669 (8.5)	15,667 (5.8)	113,679 (9.8)	17,230 (23.8)	31,359 (17.8)	9634 (10.0)	58,223 (16.9)	29,113 (14.1)	20,310 (4.7)	6033 (3.4)	55,456 (6.8)

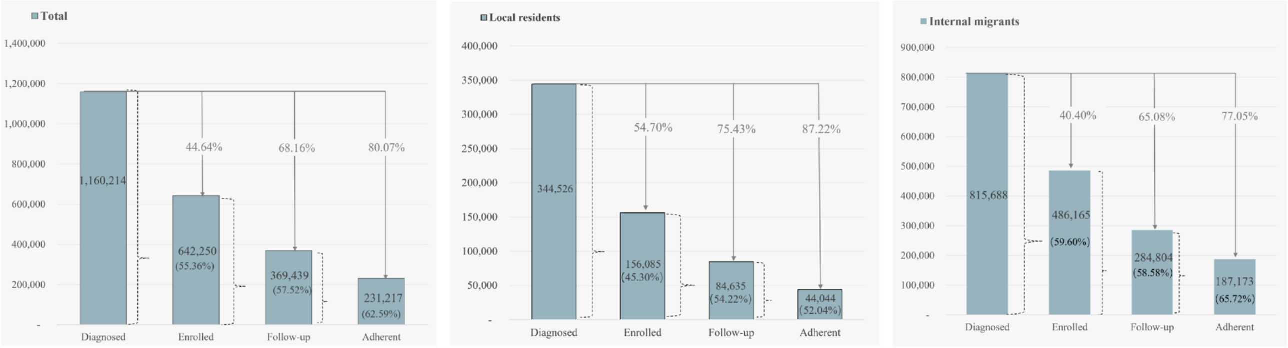
	Total (n=1,160,214), n (%)				Local population (n=344,526), n (%)				Migrant population (n=815,688), n (%)			
High school or technical school	84,781 (30.4)	164,361 (27.0)	52,107 (19.2)	301,249 (26.0)	13,239 (18.3)	53,872 (30.6)	24,805 (25.8)	91,916 (26.7)	71,542 (34.6)	110,489 (25.5)	27,302 (15.6)	209,333 (25.7)
Junior high school	64,310 (23.0)	194,959 (32.0)	65,548 (24.1)	324,817 (28.0)	3196 (4.4)	23,497 (13.4)	20,121 (20.9)	46,814 (13.6)	61,114 (29.5)	171,462 (39.6)	45,427 (25.9)	278,003 (34.1)
Elementary school or below	5205 (1.9)	91,713 (15.1)	108,233 (39.8)	205,151 (17.7)	372 (0.5)	7699 (4.4)	22,928 (23.8)	30,999 (9.0)	4833 (2.3)	84014 (19.4)	85,305 (48.6)	174,152 (21.4)
Missing	35,964 (12.9)	62,541 (10.3)	19,312 (7.1)	117,817 (10.2)	12,827 (17.7)	25,926 (14.7)	11,307 (11.7)	50,060 (14.5)	23,137 (11.2)	36,615 (8.4)	8005 (4.6)	67,757 (8.3)
Insurance type												
Urban employees	167,374 (60.0)	277,068 (45.5)	33,538 (12.3)	477,980 (41.2)	43,049 (59.5)	96,562 (54.9)	27,879 (29.0)	167,490 (48.6)	124,325 (60.1)	180,506 (41.7)	5659 (3.2)	310,490 (38.1)
Urban residents	27,268 (9.8)	74,175 (12.2)	40,742 (15.0)	142,185 (12.3)	11,510 (15.9)	44,050 (25.0)	35,570 (36.9)	91,130 (26.5)	15,758 (7.6)	30,125 (7.0)	5172 (2.9)	51,055 (6.3)
Other	65,806 (23.6)	222,946 (36.6)	184,410 (67.8)	473,162 (40.8)	8336 (11.5)	17,126 (9.7)	24,302 (25.2)	49,764 (14.4)	57,470 (27.8)	205,820 (47.5)	160,108 (91,.2)	423,398 (51.9)
Fully self-funded	1085 (0.4)	2681 (0.4)	1406 (0.5)	5172 (0.4)	458 (0.6)	749 (0.4)	308 (0.3)	1515 (0.4)	627 (0.3)	1932 (0.4)	1098 (0.6)	3657 (0.5)
Missing	17,646 (6.3)	32,368 (5.3)	11,701 (4.3)	61,715 (5.3)	9008 (12.4)	17,400 (9.9)	8219 (8.5)	34,627 (10.1)	8638 (4.2)	14,968 (3.5)	3482 (2.0)	27,088 (3.3)

City-Level HMP Participants

In Total

Among all diagnosed individuals, 55.36% (n=642,250) were enrolled in the HMP. We found that 57.52% (n=369,429) of those followed up. Furthermore, 62.59% (n=231,217) of the individuals who were followed up demonstrated adherence. Compared to local residents, internal migrants have a significantly higher participation rate (*P*<.001) in the HMP, as shown in Figure 1 (Table S1 in Multimedia Appendix 1).

Figure 1. City-level hypertension management program participants by household registration.



Enrollment Rate

Among all diagnosed individuals (n=1,160,214), 58.1% (n=642,250) were enrolled in the HMP. As is shown in Table 2, the enrollment rates were similar between different sexes, with males at 55.27% (n=373,452) and females at 55.48% (n=268,798). For internal residents, the enrollment rate was higher for females, with a rate of 60.15% (202,115/335,997) for females and 59.22% (284,050/479,691) for males. The opposite was seen for local migrants. Industrial workers had the highest enrollment rate, followed by those in agriculture and forestry, then the unemployed, with the lowest rate among those in the “other” category, and technical and administrative personnel just above that. For the local population, the business service industry was above the “other” group. The lower the educational level, the higher the enrollment rate. Older individuals had higher enrollment rates; the enrollment rate was only 39.27% (n=109,623) among those aged 18 - 44 years, but it increased to 64.14% (n=174,335) among those older than 65 years. Divorced and widowed individuals represented the smallest proportion of the diagnosed population, but they had the highest enrollment rate at 67.52% (n=11,405), followed by married individuals at 59.16% (n=610,208), with single individuals having the lowest rate at 18.41% (n=20,018). The trends in enrolled rates across different age groups, marital statuses, and educational levels were similar between local residents and internal migrants. The logistic regression results indicated that all the research variables examined were significant factors influencing the enrollment rate (Table S2 in Multimedia Appendix 2).

Table . Participation rates in hypertension management programs by household registration in Shenzhen (January 1, 2017, to September 30, 2023).

Indicator	Enrolled individuals, n (%)[a]			Follow-up individuals, n (%)[a]			Adherent individuals, n (%)[a]		
Characteristics	Total (n=642,250)	Local residents (n=156,085)	Internal migrants (n=486,165)	Total (n=369,439)	Local residents (n=84,635)	Internal migrants (n=284,804)	Total (n=231,217)	Local residents (n=44,044)	Internal migrants (n=187,173)
Gender									
Male	373,452 (55.27)	89,402 (45.61)	284,050 (59.22)	220,066 (58.93)	48,848 (54.64)	171,218 (60.28)	140,515 (63.85)	26,207 (53.65)	114,308 (66.76)
Female	268,798 (55.48)	66,683 (44.9)	202,115 (60.15)	149,373 (55.57)	35,787 (53.67)	113,586 (56.2)	90,702 (60.72)	17,837 (49.84)	72,865 (64.15)
Age group (years)									
18 - 44	109,623 (39.27)	18,782 (25.96)	90,841 (43.92)	65,685 (59.92)	9565 (50.93)	56,120 (61.78)	45,436 (69.17)	5950 (62.21)	39,486 (70.36)
45 - 64	358,292 (58.81)	83,580 (47.52)	274,712 (63.39)	206,746 (57.7)	45,131 (54)	161,615 (58.83)	127,871 (61.85)	23,454 (51.97)	104,417 (64.61)
≥65	174,335 (64.14)	53,723 (55.8)	120,612 (68.72)	97,008 (55.64)	29,939 (55.73)	67,069 (55.61)	57,910 (59.7)	14,640 (48.9)	43,270 (64.52)
Marital status									
Single	20,018 (18.41)	5240 (12.46)	14,778 (22.17)	10,057 (50.24)	2428 (46.34)	7629 (51.62)	7166 (71.25)	1451 (59.76)	5715 (74.91)
Married	610,208 (59.16)	147,202 (49.85)	463,006 (62.90)	352,829 (57.82)	80,168 (54.46)	272,661 (58.89)	220,041 (62.36)	41,532 (51.81)	178,509 (65.47)
Widowed or divorced	11,405 (67.52)	3412 (58.36)	7993 (72.37)	6426 (56.34)	2001 (58.65)	4425 (55.36)	3923 (61.05)	1039 (51.92)	2884 (65.18)
Missing	619 (19.15)	231 (17.73)	388 (20.11)	127 (20.52)	38 (16.45)	89 (22.94)	87 (68.5)	22 (57.89)	65 (73.03)
Occupation									
Technical personnel or office staff	91,870 (51.17)	47,737 (49.04)	44,133 (53.7)	52,918 (57.6)	26,918 (56.39)	26,000 (58.91)	29,624 (55.98)	13,365 (49.65)	16,259 (62.53)
Business service industry	122,120 (54.68)	34,758 (49.01)	87,362 (57.32)	70,627 (57.83)	18,956 (54.54)	51,671 (59.15)	43,445 (61.51)	9998 (52.74)	33,447 (64.73)
Agriculture and forestry	24,092 (60.55)	3577 (53.55)	20,515 (61.96)	13,895 (57.67)	2015 (56.33)	11,880 (57.91)	8382 (60.32)	1043 (51.76)	7339 (61.78)
Industry	134,668 (69.13)	7855 (56.56)	126,813 (70.1)	88,760 (65.91)	4568 (58.15)	84,192 (66.39)	58,049 (65.4)	2405 (52.65)	55,644 (66.09)
Other	44,554 (31.09)	18,754 (26.34)	25,800 (35.79)	21,457 (48.16)	9216 (49.14)	12,241 (47.45)	12,915 (60.19)	4779 (51.86)	8136 (66.47)
Unemployed	224,840 (59.41)	43,397 (51.43)	181,443 (61.7)	121,778 (54.16)	22,962 (52.91)	98,816 (54.46)	78,798 (64.71)	12,454 (54.24)	66,344 (67.14)
Missing	106 (10.47)	7 (6.54)	99 (10.94)	4 (3.77)	0 (0)	4 (4.04)	4 (100)	0 (0)	4 (100)
Education level									
Bachelor's degree or above	40,803 (41.85)	28,613 (43.02)	12,190 (39.34)	21,729 (53.25)	15,248 (53.29)	6481 (53.17)	11,926 (54.89)	7798 (51.14)	4128 (63.69)
College diploma	53,614 (47.16)	27,111 (46.56)	26,503 (47.79)	29,547 (55.11)	14,640 (54)	14,907 (56.25)	17,369 (58.78)	7610 (51.98)	9759 (65.47)
High school or technical school	169,218 (56.17)	47,939 (52.16)	121,279 (57.94)	98,505 (58.21)	26,429 (55.13)	72,076 (59.43)	59,936 (60.85)	13,514 (51.13)	46,422 (64.41)

Indicator	Enrolled individuals, n (%)[a]			Follow-up individuals, n (%)[a]			Adherent individuals, n (%)[a]		
Characteristics	Total (n=642,250)	Local residents (n=156,085)	Internal migrants (n=486,165)	Total (n=369,439)	Local residents (n=84,635)	Internal migrants (n=284,804)	Total (n=231,217)	Local residents (n=44,044)	Internal migrants (n=187,173)
Junior high school	212,807 (65.52)	26,355 (56.3)	186,452 (67.07)	127,916 (60.11)	14,595 (55.38)	113,321 (60.78)	82,575 (64.55)	7776 (53.28)	74,799 (66.01)
Elementary school or below	144,562 (70.47)	18,481 (59.62)	126,081 (72.4)	83,165 (57.53)	10,490 (56.76)	72,675 (57.64)	54,244 (65.22)	5547 (52.88)	48,697 (67.01)
Missing	21,246 (18.03)	7586 (15.15)	13,660 (20.16)	8577 (40.37)	3233 (42.62)	5344 (39.12)	5167 (60.24)	1799 (55.64)	3368 (63.02)
Insurance type									
Urban employees	270,312 (56.55)	80,523 (48.08)	189,789 (61.13)	166,048 (61.43)	44,582 (55.37)	121,466 (64)	99,579 (59.97)	22,840 (51.23)	76,739 (63.18)
Urban residents	81,890 (57.59)	51,233 (56.22)	30,657 (60.05)	46,235 (56.46)	28,688 (56)	17,547 (57.24)	25,561 (55.28)	14,726 (51.33)	10,835 (61.75)
Fully self-funded	280,172 (59.21)	20,301 (40.79)	259,871 (61.38)	153,701 (54.86)	9797 (48.26)	143,904 (55.38)	103,928 (67.62)	5586 (57.02)	98,342 (68.34)
Other	2519 (48.70)	456 (30.1)	2063 (56.41)	1352 (53.67)	217 (47.59)	1135 (55.02)	878 (64.94)	130 (59.91)	748 (65.9)
Missing	7357 (11.92)	3572 (10.32)	3785 (13.97)	2103 (28.59)	1351 (37.82)	752 (19.87)	1271 (60.44)	762 (56.4)	509 (67.69)

[a] All percentages are calculated in relation to a broader population. For instance, the proportion of males within the enrolled individuals stands at 55.27% and is calculated as n/N × 100. Here, "n" represents the number of males within the enrolled individuals subset, while "N" denotes the number of males within the entire population. For N values, please refer to the corresponding "subtotal" columns in Table 1.

Follow-Up Rate

Of patients who have been enrolled in the hypertension program, about 57.52% (369,439/642,250) were followed up. As is shown in Table 2, the follow-up rate was higher among internal migrants (284,804/486,165, 58.58%), while it was 54.22% (84,635/156,085) among the local population. The proportions of those who follow up were higher among men (male: 220,066/373,452, 58.93%; female: 149,373/268,798, 55.57%) and industrial workers (88,760/134,668, 65.91%). For other characteristics, there were significant differences in follow-up rates between local residents and internal migrants. For the local population, follow-up rates increased with age: 50.93% (9565/18,782) for those aged 18-44 years and 55.73% (29,939/53,723) for those over 65 years. However, for internal migrants, the trend was the opposite: the follow-up rate for those aged 18-44 years was 61.78% (56,120/90,841), while it was 55.61% (67,069/120,612) for those over 65 years. Among local residents, the highest follow-up rate was observed among divorced and widowed individuals, at 58.65% (2001/3412). Conversely, for internal migrants, married individuals had the highest follow-up rate, at 58.89% (272,661/463,006). Single individuals exhibited the lowest follow-up rates in both populations, with 46.34% (2428/5240) of local residents and 51.62% (7629/14,778) of internal migrants following up. Among local residents, the highest follow-up rate was observed in individuals with an educational level of elementary school or below and those with urban resident insurance. Conversely, among internal migrants, the highest follow-up rate was found in individuals with a junior high school education and those with urban employee insurance. The logistic regression results indicated that all the research variables examined were significant factors influencing the follow-up rate (Table S2 in Multimedia Appendix 2).

Adherence Rate

A total of 62.59% (231,217/369,439) of patients who followed up continued to do so at least 4 times per year. As is shown in Table 2, similar to the enrolled rate and follow-up rate, the adherence rate was higher among internal migrants (187,173/284,804, 65.72%) compared to local residents (44,044/84,635, 52.04%). Among the followed-up patients, men (140,515/220,066, 63.85%), single individuals (7166/10,057, 71.25%), and people aged 18-44 years (45,436/65,685, 69.17%) had the highest adherence rates. Individuals with urban employee and urban resident insurance did not have the highest adherence rates. These trends were consistent between both local and internal migrants. However, among local residents, those engaged in commercial and industrial occupations had the highest adherence rates, at 52.74% (9998/18,956) and 52.65% (2405/4568), respectively. Conversely, in the internal migrants group, the unemployed and those classified under other categories had the highest adherence rates, at 67.14% (66,344/98,816) and 66.47% (8136/12,241), respectively. Regarding educational level, among local residents, those with a junior high school level of education had the highest adherence rate at 53.28% (7776/14,595), whereas among internal migrants, those with an educational level of elementary school or below had the highest adherence rate. The logistic regression results indicated that all the research variables examined were significant factors influencing the adherence rate (Table S2 in Multimedia Appendix 2).

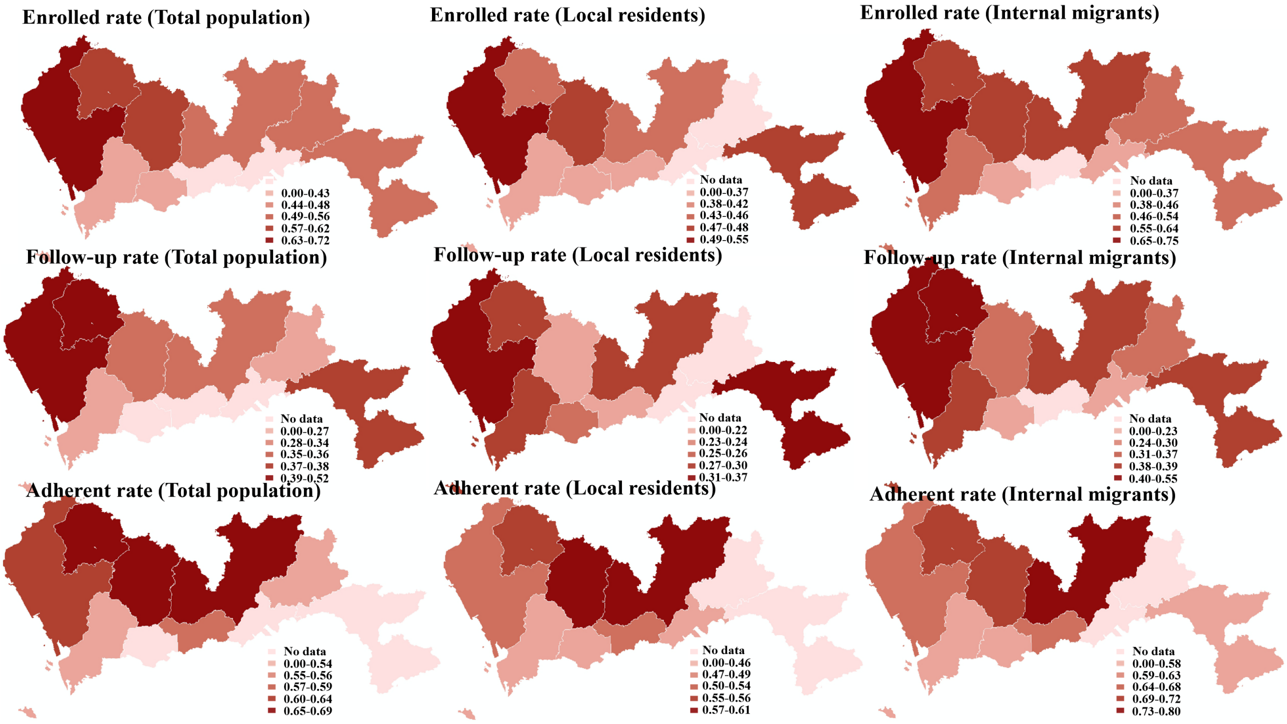
District-Level Hypertension Management

As shown in Figure 2, the highest enrolled rate was detected in Baoan district (180,301/250,036, 72.11%), Guangming district (33,308/54,063, 61.61%), and Longhua district (74,115/124,878, 59.35%), while the lowest enrolled rate was found in Luohu district (52,413/132,023, 39.70%), Yantian district (8609/19,924, 43.21%), and Futian district (59,962/132,337, 45.31%). It is worth noting that, apart from Luohu district, the enrolled rate of internal migrants was higher than that of local residents.

Regarding follow up rate, Baoan district, Guangming district, and Dapeng New Area showed the highest rates, which were 52.21% (131,171/251,237), 45.86% (24,915/54,328), and 36.72% (4301/11,713), while the lowest rate was found in the same places as the enrolled rate. The enrolled rate of internal migrants was higher than that of local residents in all districts.

Considering the adherence rate, Longhua district (31,779/45,917, 69.21%), Longgang district (54,088/79,965, 67.64%), and Guangming district (17,949/26,658, 67.33%) were ranked as the first 3. Similarly, the adherence rate of internal migrants was higher than that of local residents in all districts.

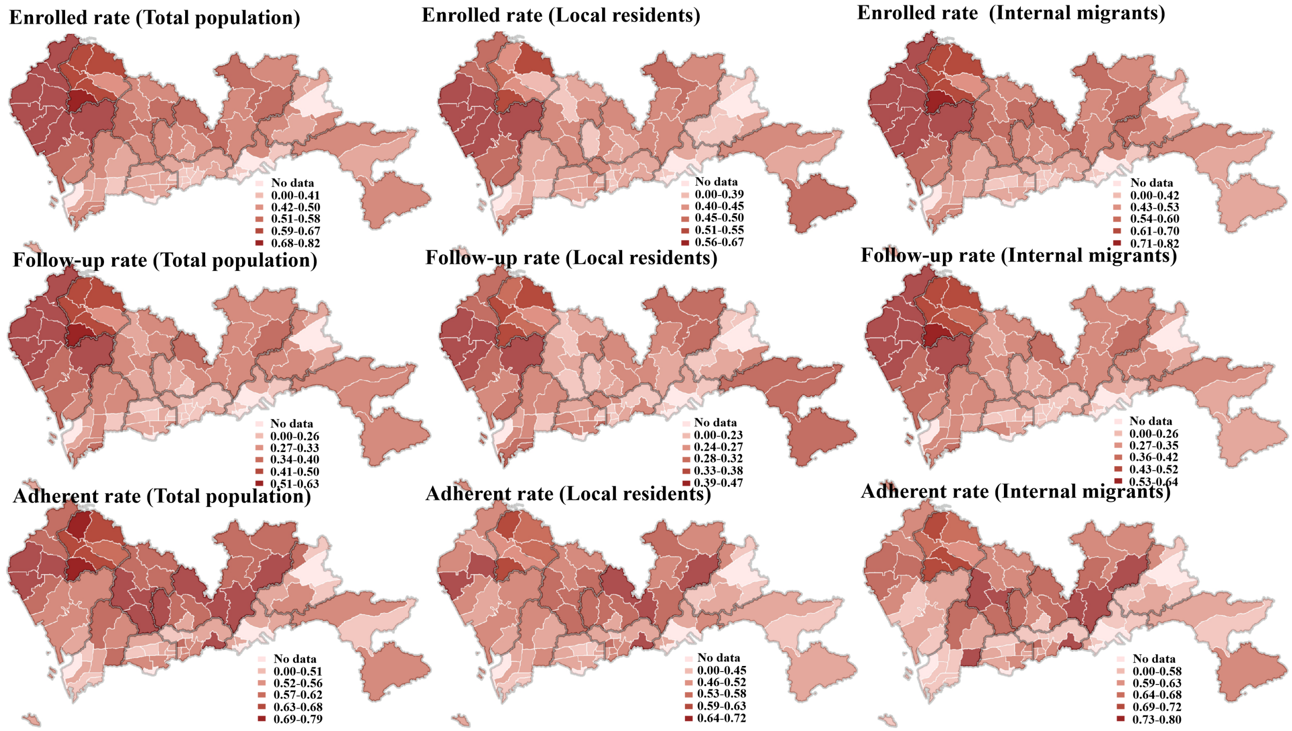
Figure 2. District-level hypertension management.



Community-Level Hypertension Management

There was considerable within-city and within-district variation in the community level of hypertension management. Generally speaking, the community-level enrolled rate, follow-up rate, and adherence rate were descending from east to west and from north to south. Although only the adherence rate was slightly different, some communities in the southern region showed a high rate, as shown in Figure 3.

Figure 3. Community-level hypertension management.



There also exists within-district variation such that a 24.30% gap between different communities was found in the enrollment rate of Guangming district. A 23.59% gap was detected in the follow-up rate of Baoan district. Similarly, a 23.59% gap was found in the adherence rate in Luohu district. In addition, rates of internal migrants were much higher than those of local residents in all districts.

Discussion

Principal Results

Regarding the crucial role of health service supply from primary health centers in achieving health equity in the domain of hypertension prevention, we summarized the enrolled rate, follow-up rate, and adherence rate of the Shenzhen population that participated in the HMP based on the data collected from CHCs. Among all diagnosed individuals, 55.36% (642,250/1,160,214) were enrolled in the HMP. Of those, 57.52% (369,439/642,250) followed up, and 62.59% (231,217/369,439) were adherents. Furthermore, we compared the health inequity in hypertension between internal migrants and local residents in Shenzhen.

Comparison With Prior Work

Overall, the data revealed that men exhibited higher rates of both follow-up and adherence compared to women, which is not consistent with the conclusion of some studies [19,20] related to hypertension treatment adherence. However, the medication adherence rates of men and women taking one medication were 61.04% and 55.86% in Beijing, respectively [21]. Despite the fact that the majority of individuals over the age of 65 years were enrolled in the program, this age group demonstrated the lowest rates of follow-up and adherence, although the rates were relatively high in Beijing and Shanghai [22]. Conversely, individuals aged between 18 and 44 years, who have the lowest enrollment rate, showed the highest rates of follow-up and adherence. It is not similar to common research results in which the oldest patients are the most adherent [19,20]. This phenomenon may be related to the fact that older individuals often have more complex health conditions, which CHCs cannot address well at present, leading them to generally prefer seeking treatment at higher-level medical institutions [23,24]. The observed low initial engagement of younger individuals with primary health care facilities may be attributed to a general mistrust or lower acceptance of these services among this demographic; the result is similar to that of a study conducted in Hong Kong [25]. Despite the low proportion of young patients who initially seek care at primary health centers following a hypertension diagnosis, those who do engage exhibit the highest follow-up and regular follow-up rates across all age groups. This indicates a high level of acceptance and satisfaction with primary health care management among younger patients once they have accessed these services.

Regarding marital status, the enrollment rate among widowed and divorced individuals was the highest, exceeding that of single individuals by more than 3 times. It is possible that this group, having experienced significant life events, became more concerned about their health. However, without a partner to provide encouragement and support [26], they may find it difficult to maintain consistent health behaviors and this group had higher mortality [27]. As a result, despite their high enrollment rates, their adherence tended to be relatively low. Interestingly, however, single individuals had the highest adherence rate. Perhaps this is because single individuals who enrolled in the HMP tend to pay more attention to their own health and are more likely to take action to keep healthy, leading to higher adherence; this is similar to a finding in the United States that never-married individuals reported better health habits and had higher rates of protective factors compared to married or divorced/separated individuals [28]. Previous studies showed that, in general, blue-collar employees appeared less likely to participate in health promotion programs compared to white-collar workers [29], and they may not have enough time to participate in the program [30,31]. Conversely, in our research, industrial workers stood out with the highest rates of enrollment, follow-up, and adherence across all categories; this may be related to the fact that in Shenzhen, many primary health care centers operate until 9 PM, providing this group with sufficient time to access the programs. Furthermore, the data indicated that, regardless of age, individuals with lower educational levels tended to have higher rates for all 3 metrics. Previous research has also found that individuals with higher education levels may be less likely to participate in cancer screening or achieve favorable screening outcomes [32], suggesting the impact of education on health behaviors is complex.

Importantly, internal migrants demonstrated a significantly higher participation in the HMP, especially in terms of project inclusion and regular follow-up. The participation rate of internal migrants was more than 10% higher than that of local residents. This result is inconsistent with previous global and national studies on the health service utilization of internal migrants. Previous research has indicated that internal migrants face long-standing disadvantages due to various institutional, social, cultural, and other exclusionary factors [33]. These disadvantages often extend to health and health care access, limiting their access to health resources [34,35]. A review focusing on migrants' and refugees' health status and health care in Europe highlighted the issue of underuse of primary health care services among these populations [36]. A study [37] based on the 2017 National Migrants Dynamic Monitoring Special Survey in China revealed that, compared to the registered population, the floating population is at a disadvantage in utilizing basic public health services. Most studies [38,39] focusing on hypertension management in China showed that the rate of hypertension management at primary health care centers among the floating population was less than 30% and as low as 2.3% in Hunan province [40], but our study showed that this situation does not exist in Shenzhen. In fact, among the hypertensive population in Shenzhen, on the contrary, internal migrants are better able to utilize primary health care services than locals. The Shenzhen municipal government has invested significantly in ensuring that all residents can access health care services within a 15-minute radius, making it highly convenient for the entire population. Additionally, the city provides free blood pressure screening at CHCs for all individuals aged ≥18 years, which promotes accessibility and equity in hypertension management services [41]. Certain districts in Shenzhen supply free medication to hypertensive patients enrolled in the program [42]. For most internal migrants, free health care services hold strong appeal due to their economic constraints. Some internal migrants came to Shenzhen to help their son/daughter look after children [43], so they or their son/daughter may care more about their health and they usually have enough time to go to a CHC. Insurance is considered an important factor for the utilization of health services for internal migrants and immediate reimbursement of medical insurance can significantly increase inpatient utilization, promoting health [44]. The statistical model showed that insurance type is an important factor influencing the 3 hypertension management rates studied. However, there were notable differences between local residents and internal migrants. For local residents, those with urban resident insurance were more likely to participate in the HMP compared to those with urban employee insurance. In contrast, for internal migrants, those with other insurance types were less likely to participate compared to those with urban employee insurance. It takes less money to get treatment and medicine in CHCs compared with upper health care facilities, especially for people who do not have native health insurance or cannot get immediate reimbursement.

Our analysis revealed several intriguing trends regarding HMP engagement among different demographic groups in Shenzhen. Previous studies conducted in Beijing and Shanghai have also found variations in hypertension management across different demographic factors, including age, gender, education level, workload, marital status, residential area, and annual household income [45,46]. Among individuals aged 18 - 44 years, the enrolled rate of internal migrants was nearly 18% higher than that of local residents. This discrepancy may be attributed to the comparatively better economic conditions of local residents, who do not face the housing issues often encountered by migrants [47,48]. They may just go to upper hospitals when they think necessary, while internal migrants are more likely to utilize CHCs as the first choice for their routine health problems [49,50]. Additionally, the presence of their parents may lead locals to prioritize their own health less than migrants, who often lack familial support [51] and thus might be more vigilant about their health. Interestingly, in the divorced or widowed demographic, the follow-up rate among local residents surpassed that of internal migrants. Moreover, among local residents, the follow-up rate for this group was higher than that of married individuals. This finding contrasts with conventional research, which typically suggests that marriage promotes greater health awareness and engagement with health care services while the effects of bereavement usually lead to less health care utilization [52]. For married migrant individuals, having a spouse provides them with companionship [53] and significant support from their partner, whereas local residents can receive support from other relatives. This may cause internal migrants, compared to local residents, to have a more limited support system, making the protective effects of marriage on their health more pronounced. In contrast, local residents who are divorced or widowed may have access to a broader range of social support. The adherence rate showed an increasing divergence between local residents and internal migrants as age groups progress. Internal migrants tended to exhibit greater concern for their health after relocating to a new environment, which may explain this trend. Despite the fact that, among those with an undergraduate education or higher, the enrolled rate and follow-up rate for local residents were higher than those for internal migrants, the adherence rate for local residents remained lower than that of internal migrants. This suggests that, while highly educated local residents may initially engage with health care services, they may not maintain the same level of adherence as their migrant counterparts.

Some research showed that there exists vast disparities in health insurance coverage and health service utilization between floating populations and local residents; most of the floating population has no timely access to primary or some other kinds of health care service [39], which cause the poor control of chronic diseases. In our study, it appears that health insurance played a less significant role in the disparity of participation in the HMP between local residents and internal migrants.

From a district and community level, hypertensive patients were experiencing health inequity in hypertension treatment and prevention. The potential reasons included population, economic reasons, and policy differences. The populations of Longgang, Longhua, and Nanshan in Shenzhen are large and concentrated; in addition, the health care services are mature and systematic. Hence, the HMP is better than in other regions such as Yantian and Pingshan districts. Similarly, the wealthy regions had higher enrolled, follow-up, and adherence rates compared to the relatively poor regions, namely Yantian and Pingshan districts. These finding were consistent with previous research. A large national survey during 2004 - 2018 found a provincial difference in awareness, treatment, and control is partly correlated with per capita gross domestic product [54]. Richer provinces are more successful in controlling hypertension. Globally, the proportions of hypertension awareness, treatment, and blood pressure control are low [55], particularly in low- and middle-income countries, and few comprehensive assessments of the economic impact of hypertension exist [56,57]. In terms of policy differences, the hypertensive patients who participate in the HMPs of CHCs could get free prescription medicines, blood pressure measurement services, and blood glucose measurement services once per quarter. The policy is different in different districts in Shenzhen, which may have caused the observed difference in enrolled, follow-up, and adherence rate. Furthermore, when it comes to health equity between internal migrants and native citizens in Shenzhen, the difference in enrolled, follow-up, and adherence rates within internal migrants was higher than that within local residents. This may be caused by different health care service supply situations toward internal migrants within different districts and communities [58]. The same problems were found in Ireland, Portugal, Spain, and the United States [59].

Limitations and Strengths

This study represents a pioneering effort to examine the participation of internal migrants and local residents in HMPs, thereby highlighting the equity of internal migrants' utilization of primary health care services under current Chinese hypertension management policies. By focusing on the specific involvement of diagnosed hypertension patients in HMPs, this research offers valuable insights into the accessibility and fairness of these services for internal migrants. The data utilized in this study were derived entirely from the primary health care system, ensuring its authenticity and reliability. However, this study has several limitations. It exclusively addressed the participation of hypertension patients in primary HMPs, which means that it did not account for patients who may receive regular treatment at other medical institutions without ever visiting CHCs. Consequently, the findings may not fully represent the hypertension management engagement of all patients citywide. Furthermore, the study was centered on participation rates within the HMP and did not capture the ultimate blood pressure control outcomes of the subjects. As a result, it is not possible to ascertain the differences in blood pressure control outcomes that may arise from varying levels of management participation.

Future research should consider integrating data from multiple health care providers and tracking long-term health outcomes to provide a more comprehensive understanding of hypertension management across diverse patient populations.

Conclusions

Based on our research, of the 1,160,214 hypertensive patients, 642,250 were enrolled in the HMP, 48,848 were included in the followed up population, and 34,186 were adherents. Notably, internal migrants demonstrated a significantly higher participation in the HMP. Furthermore, populations with different demographic and socioeconomic characteristics and in different regions showed different rates. Given these findings, there is the potential to enhance the outreach and engagement of native, younger, and single populations through targeted promotional strategies.

Acknowledgments

This project was supported by the Sanming Project of Medicine in Shenzhen (grant SZSM202111001). We would like to express our gratitude to the Shenzhen Government and the Shenzhen Health Development Research and Data Management Center for their invaluable support in providing data, funding, and an enabling environment for this project. We acknowledge the use of generative ChatGPT for grammatical checks.

Data Availability

The datasets generated during and/or analyzed during this study are not publicly available due to restrictions imposed by the local government. The data originate from the medical system and are managed by the local government. Access to the encrypted data stored on the cloud platform is limited to research institutions that have established formal collaborations with the local government.

Authors' Contributions

WL and NZ are the corresponding authors of this study, responsible for designing, guiding, organizing, and planning this study. JH was responsible for data analysis and manuscript writing. YL was responsible for data cleaning and provided helpful comments on several drafts of the manuscript. HZ, LW, HS, and ZZ provided advice and revised several drafts of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Two-proportion *z*-test for enrollment rates, follow-up rates, and adherence rates between local residents and internal migrants.
[XLSX File, 20 KB - publichealth_v11i1e65548_app1.xlsx]

Multimedia Appendix 2
Multivariable logistic regression analysis of the enrollment, follow up, and adherence of different groups in Shenzhen.
[XLSX File, 22 KB - publichealth_v11i1e65548_app2.xlsx]

References

1. Kyu HH, Abate D, Abate KH. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1859-1922. [doi: 10.1016/S0140-6736(18)32335-3] [Medline: 30415748]
2. Global report on hypertension: the race against a silent killer. : World Health Organization; 2023 URL: https://www.who.int/publications/i/item/9789240081062 [accessed 2024-12-16]
3. Feng XL, Pang M, Beard J. Health system strengthening and hypertension awareness, treatment and control: data from the China Health and Retirement Longitudinal Study. Bull World Health Organ 2014 Jan 1;92(1):29-41. [doi: 10.2471/BLT.13.124495] [Medline: 24391298]
4. Yin R, Yin L, Li L, et al. Hypertension in China: burdens, guidelines and policy responses: a state-of-the-art review. J Hum Hypertens 2022 Feb;36(2):126-134. [doi: 10.1038/s41371-021-00570-z] [Medline: 34215840]
5. Notice of the National Health and Family Planning Commission on the issuance of the "National Basic Public Health Service Standards (Third Edition)" [Website in Chinese]. National Health Commission of the PRC. 2017. URL: http://www.nhc.gov.cn/jws/s3578/201703/d20c37e23e1f4c7db7b8e25f34473e1b.shtml [accessed 2024-12-16]
6. Dover DC, Belon AP. The health equity measurement framework: a comprehensive model to measure social inequities in health. Int J Equity Health 2019 Feb 19;18(1):36. [doi: 10.1186/s12939-019-0935-0] [Medline: 30782161]
7. Zhang J, Lin S, Liang D, Qian Y, Zhang D, Hou Z. Public health services utilization and its determinants among internal migrants in China: evidence from a nationally representative survey. Int J Environ Res Public Health 2017 Sep 1;14(9):1002. [doi: 10.3390/ijerph14091002] [Medline: 28862682]
8. Lin Y, Zhang Q, Chen W, Ling L. The social income inequality, social integration and health status of internal migrants in China. Int J Equity Health 2017 Aug 4;16(1):139. [doi: 10.1186/s12939-017-0640-9] [Medline: 28778201]
9. Wang J, Tan F, Wang Z, et al. Understanding gaps in the hypertension and diabetes care cascade: systematic scoping review. JMIR Public Health Surveill 2024 Feb 16;10:e51802. [doi: 10.2196/51802] [Medline: 38149840]
10. Chaturvedi A, Zhu A, Gadela NV, Prabhakaran D, Jafar TH. Social determinants of health and disparities in hypertension and cardiovascular diseases. Hypertension 2024 Mar;81(3):387-399. [doi: 10.1161/HYPERTENSIONAHA.123.21354] [Medline: 38152897]
11. Doyle SK, Chang AM, Levy P, Rising KL. Achieving health equity in hypertension management through addressing the social determinants of health. Curr Hypertens Rep 2019 Jun 12;21(8):58. [doi: 10.1007/s11906-019-0962-7] [Medline: 31190099]
12. Jin Y, Jing M, Zhang L, Song S, Ma X. Internet access and hypertension management among the elderly population: a nationally representative cross-sectional survey in China. J Med Internet Res 2019 Jan 31;21(1):e11280. [doi: 10.2196/11280] [Medline: 30702439]
13. Zhang Z, Shi G, Xing Y, et al. Examining the potential impacts of intensive blood pressure treatment on the socioeconomic inequity in hypertension prevalence in China: a nationally representative cross-sectional study. Hypertens Res 2023 Dec;46(12):2746-2753. [doi: 10.1038/s41440-023-01441-5] [Medline: 37789112]
14. Liu Y, Liu N, Cheng M, et al. The changes in socioeconomic inequalities and inequities in health services utilization among patients with hypertension in Pearl River Delta of China, 2015 and 2019. BMC Public Health 2021 May 12;21(1):903. [doi: 10.1186/s12889-021-10879-6] [Medline: 33980187]
15. Cao D, Zhou Z, Si Y, et al. Prevalence and income-related equity in hypertension in rural China from 1991 to 2011: differences between self-reported and tested measures. BMC Health Serv Res 2019 Jul 1;19(1):437. [doi: 10.1186/s12913-019-4289-5] [Medline: 31262290]

16. Wang C. Local governance in China: realizing the citizens' rights of migrants. China Popul Dev Stud 2018 Mar;2(1):11-31. [doi: 10.1007/s42379-018-0002-3]
17. Who needs to get the flu vaccine? Do you need to get it again if you have had it before? [Website in Chinese]. Public Hygiene and Health Commission of Shenzhen Municipality. 2022. URL: https://wjw.sz.gov.cn/gzcy/ywzs/ggws/content/post_10268931.html [accessed 2024-12-16]
18. Free medical check-ups for Shenzhen hukou residents aged 60 or above [Website in Chinese]. Shenzhen News Network. 2020. URL: https://www.sznews.com/news/content/2020-06/06/content_23223870.htm [accessed 2024-12-16]
19. Khanam MA, Lindeboom W, Koehlmoos TLP, Alam DS, Niessen L, Milton AH. Hypertension: adherence to treatment in rural Bangladesh--findings from a population-based study. Glob Health Action 2014;7:25028. [doi: 10.3402/gha.v7.25028] [Medline: 25361723]
20. Gavrilova A, Bandere D, Rutkovska I, et al. Knowledge about disease, medication therapy, and related medication adherence levels among patients with hypertension. Medicina (Kaunas) 2019 Oct 28;55(11):715. [doi: 10.3390/medicina55110715] [Medline: 31661904]
21. Lan L, Hai P, Luo J, Li R, Wang Y. Medical behaviours and medication adherence of older hypertensive patients in different medical insurance programs in Beijing, China: a cross-sectional study. BMC Geriatr 2023 Dec 20;23(1):878. [doi: 10.1186/s12877-023-04476-y] [Medline: 38124122]
22. Yin M, Augustin B, Fu Z, Yan M, Fu A, Yin P. Geographic distributions in hypertension diagnosis, measurement, prevalence, awareness, treatment and control rates among middle-aged and older adults in China. Sci Rep 2016 Nov 14;6:37020. [doi: 10.1038/srep37020] [Medline: 27841326]
23. Yili Z, Xinyi H, Baoyu Q, Chuanrui S, Xiaoyang W, Xu W, et al. Multimorbidity in elderly population: challenges and coping strategies. Chin Gen Pract 2022;25(35):4363-4368. [doi: 10.12114/j.issn.1007-9572.2022.0514]
24. Na D, Jia-tang Z, Peng-fei W, Yuxia K, Xia T. Chronic diseases and comorbidity in the elderly in Beijing. Chin Gen Pract 2018;21(3):265-268. [doi: 10.3969/j.issn.1007-9572.2017.00.217]
25. Lim MK, Ha SCN, Luk KH, Yip WK, Tsang CSH, Wong MCS. Update on the Hong Kong Reference Framework for Hypertension Care for Adults in Primary Care Settings-review of evidence on the definition of high blood pressure and goal of therapy. Hong Kong Med J 2019 Feb;25(1):64-67. [doi: 10.12809/hkmj187701] [Medline: 30648830]
26. Scafato E, Galluzzo L, Gandin C, et al. Marital and cohabitation status as predictors of mortality: a 10-year follow-up of an Italian elderly cohort. Soc Sci Med 2008 Nov;67(9):1456-1464. [doi: 10.1016/j.socscimed.2008.06.026] [Medline: 18675500]
27. Dhindsa DS, Khambhati J, Schultz WM, Tahhan AS, Quyyumi AA. Marital status and outcomes in patients with cardiovascular disease. Trends Cardiovasc Med 2020 May;30(4):215-220. [doi: 10.1016/j.tcm.2019.05.012] [Medline: 31204239]
28. Kaplan RM, Kronick RG. Marital status and longevity in the United States population. J Epidemiol Community Health 2006 Sep;60(9):760-765. [doi: 10.1136/jech.2005.037606] [Medline: 16905719]
29. Glasgow RE, McCaul KD, Fisher KJ. Participation in worksite health promotion: a critique of the literature and recommendations for future practice. Health Educ Q 1993;20(3):391-408. [doi: 10.1177/109019819302000309] [Medline: 8307762]
30. Sorensen G, Linnan L, Hunt MK. Worksite-based research and initiatives to increase fruit and vegetable consumption. Prev Med 2004 Sep;39 Suppl 2:S94-100. [doi: 10.1016/j.ypmed.2003.12.020] [Medline: 15313078]
31. Wandel M, Roos G. Work, food and physical activity. A qualitative study of coping strategies among men in three occupations. Appetite 2005 Feb;44(1):93-102. [doi: 10.1016/j.appet.2004.08.002] [Medline: 15604036]
32. Korn AR, Walsh-Bailey C, Correa-Mendez M, et al. Social determinants of health and US cancer screening interventions: a systematic review. CA Cancer J Clin 2023;73(5):461-479. [doi: 10.3322/caac.21801] [Medline: 37329257]
33. Zhang J, Ye Y, Wang J. Social exclusion and psychopathologies of Chinese floating populations. J Affect Disord Rep 2021 Dec;6:100263. [doi: 10.1016/j.jadr.2021.100263]
34. Boardman J. Social exclusion and mental health – how people with mental health problems are disadvantaged: an overview. Ment Health Soc Incl 2011 Aug 15;15(3):112-121. [doi: 10.1108/20428301111165690]
35. Tang S, Long C, Wang R, Liu Q, Feng D, Feng Z. Improving the utilization of essential public health services by Chinese elderly migrants: strategies and policy implication. J Glob Health 2020 Jun;10(1):010807. [doi: 10.7189/jogh.10.010807] [Medline: 32257170]
36. Lebano A, Hamed S, Bradby H, et al. Migrants' and refugees' health status and healthcare in Europe: a scoping literature review. BMC Public Health 2020 Jun 30;20(1):1039. [doi: 10.1186/s12889-020-08749-8] [Medline: 32605605]
37. Song Y, Zhang G. Utilization of public health services and its influencing factors among migrant people with hypertension or diabetes in China. Chin J Public Health 2021;37(2):198-202. [doi: 10.11847/zgggws1123492]
38. Kui JZ. Analysis of hypertension morbidity among floating population of Sichuan in 2012. Chin J Prev Control Chronic Dis 2015;23(1):27-30. [doi: CNKI:SUN:ZMXB.0.2015-01-009]
39. Su L, Sun L, Xu L. Review on the prevalence, risk factors and disease management of hypertension among floating population in China during 1990-2016. Glob Health Res Policy 2018;3:24. [doi: 10.1186/s41256-018-0076-9] [Medline: 30123839]

40. Donghui J, Qiaohua X, Huilin L, et al. Management of hypertension and diabetes mellitus among floating population in Hunan. Pract Prev Med 2015;22(10):1211-1212. [doi: 10.3969/j.issn.1006-3110.2015.010.017]
41. Promote health management and prevent and control chronic diseases [Website in Chinese]. Public Hygiene and Health Commission of Shenzhen Municipality. 2024. URL: https://wjw.sz.gov.cn/ztzl/szsjmjkbps2022/jkszjs/content/post_11294466.html [accessed 2024-12-16]
42. Get free high blood pressure medicine! Come to these 12 community health centers in Shenzhen and get it before the end of the year [Website in Chinese]. Public Hygiene and Health Commission of Shenzhen Municipality. 2019. URL: https://wjw.sz.gov.cn/wzx/content/post_3196923.html [accessed 2024-12-16]
43. Yong T, Yipin W. The quality of life of the floating elderly and the establishment of social support network: a survey of the floating elderly in Shenzhen. Mgmt Admin 2014(4):132-134. [doi: 10.16517/j.cnki.cn12-1034/f.2014.04.011]
44. He W. Does the immediate reimbursement of medical insurance reduce the socioeconomic inequality in health among the floating population? Evidence from China. Int J Equity Health 2023 May 17;22(1):96. [doi: 10.1186/s12939-023-01913-7] [Medline: 37198632]
45. Jiang B, Liu H, Ru X, Zhang H, Wu S, Wang W. Hypertension detection, management, control and associated factors among residents accessing community health services in Beijing. Sci Rep 2014 May 2;4:4845. [doi: 10.1038/srep04845] [Medline: 24784167]
46. Zhang X, Zhang Y, Xiao X, Ma X, He J. The relation between health insurance and management of hypertension in Shanghai, China: a cross-sectional study. BMC Public Health 2016 Sep 10;16(1):959. [doi: 10.1186/s12889-016-3627-3] [Medline: 27613593]
47. Huang Y, Tao R. Housing migrants in Chinese cities: current status and policy design. Environ Plann C Gov Policy 2015 Jun;33(3):640-660. [doi: 10.1068/c12120]
48. Wang Y. Housing cost burden, homeownership, and self-rated health among migrant workers in Chinese cities: the confounding effect of residence duration. Cities 2023 Feb;133:104128. [doi: 10.1016/j.cities.2022.104128] [Medline: 37425220]
49. Li H, Chung RYN, Wei X, et al. Comparison of perceived quality amongst migrant and local patients using primary health care delivered by community health centres in Shenzhen, China. BMC Fam Pract 2014 Apr 29;15:76. [doi: 10.1186/1471-2296-15-76] [Medline: 24779564]
50. Wei X, Pearson S, Zhang Z, Qin J, Gerein N, Walley J. Comparing knowledge and use of health services of migrants from rural and urban areas in Kunming City, China. J Biosoc Sci 2010 Nov;42(6):743-756. [doi: 10.1017/S0021932010000313] [Medline: 20584352]
51. Bradby H, Liabo K, Ingold A, Roberts H. Visibility, resilience, vulnerability in young migrants. Health (Lond) 2019 Sep;23(5):533-550. [doi: 10.1177/1363459317739441] [Medline: 29090633]
52. Simeonova E. Marriage, bereavement and mortality: the role of health care utilization. J Health Econ 2013 Jan;32(1):33-50. [doi: 10.1016/j.jhealeco.2012.10.010] [Medline: 23202255]
53. Yang Y, Chen B, Huang P, Wang Y, Zhang L, Cai F. Prevalence and influencing factors of depressive symptoms among rural-to-urban migrant workers in China: a systematic review and meta-analysis. J Affect Disord 2022 Jun;307:11-19. [doi: 10.1016/j.jad.2022.03.061]
54. Wang W, Zhang M, Xu CD, et al. Hypertension prevalence, awareness, treatment, and control and their associated socioeconomic factors in China: a spatial analysis of a national representative survey. Biomed Environ Sci 2021 Dec 20;34(12):937-951. [doi: 10.3967/bes2021.130] [Medline: 34981717]
55. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. Nat Rev Nephrol 2020 Apr;16(4):223-237. [doi: 10.1038/s41581-019-0244-2] [Medline: 32024986]
56. Nguyen TN, Chow CK. Global and national high blood pressure burden and control. Lancet 2021 Sep 11;398(10304):932-933. [doi: 10.1016/S0140-6736(21)01688-3] [Medline: 34450082]
57. Kario K, Okura A, Hoshide S, Mogi M. The WHO Global report 2023 on hypertension warning the emerging hypertension burden in globe and its treatment strategy. Hypertens Res 2024 May;47(5):1099-1102. [doi: 10.1038/s41440-024-01622-w]
58. Zhou B, Carrillo-Larco RM, Danaei G. Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet 2021 Sep 11;398(10304):957-980. [doi: 10.1016/S0140-6736(21)01330-1] [Medline: 34450083]
59. Ledoux C, Pilot E, Diaz E, Krafft T. Migrants’ access to healthcare services within the European Union: a content analysis of policy documents in Ireland, Portugal and Spain. Glob Health 2018 Jun 15;14(1):57. [doi: 10.1186/s12992-018-0373-6] [Medline: 29903036]

Abbreviations

CHC: Community Health Center
HMP: hypertension management program

Edited by A Mavragani; submitted 19.08.24; peer-reviewed by L Zhang, O Olasunkanmi; revised version received 22.10.24; accepted 29.10.24; published 10.02.25.

Please cite as:
He J, Li Y, Zeng H, Sun H, Wu L, Zhu Z, Zhang N, Liang W
Health Equity of Hypertension Management Between Local Residents and Internal Migrants in Shenzhen, China: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e65548
URL: https://publichealth.jmir.org/2025/1/e65548
doi:10.2196/65548

Impact of Primary Health Care Data Quality on Infectious Disease Surveillance in Brazil: Case Study

Pilar Tavares Veras Florentino[1], PhD; Juracy Bertoldo Junior[1], MSc; George Caique Gouveia Barbosa[1], MSc; Thiago Cerqueira-Silva[2], PhD; Vinicius de Araújo Oliveira[1,3], MD; Marcio Henrique de Oliveira Garcia[4], MSc; Gerson Oliveira Penna[5,6], PhD; Viviane Boaventura[7], PhD; Pablo Ivan Pereira Ramos[1], PhD; Manoel Barral-Netto[1,7], PhD; Izabel Marcilio[1,8], PhD

[1]Centro de Integração de Dados e Conhecimento em Saúde (CIDACS), Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, R. Mundo, 121 - sala 315 - Trobogy, Salvador, Brazil

[2]Department of Medical Statistics, London School of Hygiene & Tropical Medicine, London, United Kingdom

[3]Secretaria de Atenção Primária, Ministério da Saúde, Brasília, Brazil

[4]Secretaria de Vigilância em Saúde e Ambiente, Ministério da Saúde, Brasília, Brazil

[5]Escola Fiocruz de Governo, Fundação Oswaldo Cruz (Fiocruz), Rio de Janeiro, Brazil

[6]Núcleo de Medicina Tropical, Universidade de Brasília, Brasília, Brazil

[7]Laboratório de Medicina e Saúde Pública de Precisão, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil

[8]Departamento de Epidemiologia, Escola Bahiana de Medicina e Saúde Pública, Salvador, Brazil

Corresponding Author:
Pilar Tavares Veras Florentino, PhD
Centro de Integração de Dados e Conhecimento em Saúde (CIDACS), Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, R. Mundo, 121 - sala 315 - Trobogy, Salvador, Brazil

Abstract

Background: The increase in emerging and re-emerging infectious disease outbreaks underscores the need for robust early warning systems (EWSs) to guide mitigation and response measures. Administrative health care databases provide valuable epidemiological insights without imposing additional burdens on health services. However, these datasets are primarily collected for operational use, making data quality assessment essential to ensure an accurate interpretation of epidemiological analysis. This study focuses on the development and implementation of a data quality index (DQI) for surveillance integrated into an EWS for influenza-like illness (ILI) outbreaks using Brazil's a nationwide Primary Health Care (PHC) dataset.

Objective: We aimed to evaluate the impact of data completeness and timeliness on the performance of an EWS for ILI outbreaks and establish optimal thresholds for a suitable DQI, thereby improving the accuracy of outbreak detection and supporting public health surveillance.

Methods: A composite DQI was established to measure the completeness and timeliness of PHC data from the Brazilian National Information System on Primary Health Care. Completeness was defined as the proportion of weeks within an 8-week rolling window with any register of encounters. Timeliness was calculated as the interval between the date of encounter and its corresponding registry in the information system. The backfilled PHC dataset served as the gold standard to evaluate the impact of varying data quality levels from the weekly updated real-time PHC dataset on the EWS for ILI outbreaks across 5570 Brazilian municipalities from October 10, 2023, to March 10, 2024.

Results: During the study period, the backfilled dataset recorded 198,335,762 ILI-related encounters, averaging 8,623,294 encounters per week. The EWS detected a median of 4 (IQR 2 - 5) ILI outbreak warnings per municipality using the backfilled dataset. Using the real-time dataset, 12,538 (65%) warnings were concordant with the backfilled dataset. Our analysis revealed that 100% completeness yielded 76.7% concordant warnings, while 80% timeliness resulted in at least 50% concordant warnings. These thresholds were considered optimal for a suitable DQI. Restricting the analysis to municipalities with a suitable DQI increased concordant warnings to 80.4%. A median of 71% (IQR 54%-71.9%) of municipalities met the suitable DQI threshold weekly. Municipalities with ≥60% of weeks achieving a suitable DQI demonstrated the highest concordance between backfilled and real-time datasets, with those achieving ≥80% of weeks showing 82.3% concordance.

Conclusions: Our findings highlight the critical role of data quality in improving the EWS' performance based on PHC data for detecting ILI outbreaks. The proposed framework for real-time DQI monitoring is a practical approach and can be adapted to other surveillance systems, providing insights for similar implementations. We demonstrate that optimal completeness and timeliness of data significantly impact the EWS' ability to detect ILI outbreaks. Continuous monitoring and improvement of data quality should remain a priority to strengthen the reliability and effectiveness of surveillance systems.

(JMIR Public Health Surveill 2025;11:e67050) doi:10.2196/67050

KEYWORDS

primary health care; data quality; infectious disease surveillance; Brazil; early warning system

Introduction

In recent decades, the world has witnessed an unprecedented surge of emerging and re-emerging infectious disease outbreaks, underscoring the need for stronger early warning systems (EWSs) [1,2]. The widespread and growing use of electronic health records (EHRs) has heightened the demand for automated processes in disease surveillance [3,4].

Systematic monitoring of administrative health care databases provides valuable epidemiological insights [5,6]. Importantly, using administrative data for health surveillance avoids the overburden of surveillance teams while ensuring timeliness, as no duplication of registry is required [3]. This cost-efficient approach enhances the ability to detect outbreaks, particularly in low-resource settings, thus contributing to global security [3].

However, an effective automated EWS based on administrative datasets requires that a real-time data quality assessment algorithm is set within the EWS pipeline. Since administrative data are primarily collected for operational purposes, assessing their quality is crucial to accurate interpretation of epidemiological analysis [7,8]. A systematic review on EHR data quality assessment studies found 14 articles describing dedicated data quality programs deployed in real-world settings, while only 4 produced results generally applicable in diverse settings. Ozonze et al [9] suggest there is an absence of comprehensive tools for facilitating reliable and consistent data quality assessments.

Moreover, despite existing methods for evaluating the quality of administrative health data, including EHR data quality assessment [8] and indicators for specific programs such as the Data Quality Audit and the Data Quality Self-Assessment for immunization data [10], there remains a gap in applying similar methods to data used for health surveillance. Although metrics for assessing the quality of surveillance systems are well established [3], to the best of our knowledge, these have not been applied to evaluate administrative data when used for epidemiological surveillance purposes.

This paper describes the development and implementation of a data quality index (DQI) to assess the quality of administrative data used in epidemiological surveillance systems. We focus on applying the DQI to nationwide Brazilian primary health care (PHC) administrative data integrated into an EWS for influenza-like illness (ILI) outbreaks. The study compares the EWS performance across different DQI levels, addressing a critical gap in current research by establishing metrics that ensure accurate and timely outbreak detection while leveraging the cost-efficiency of administrative health databases.

Methods

Study Design

We developed and implemented a data quality assessment algorithm within ÆSOP (Alert-Early System of Outbreaks with Pandemic Potential), a previously validated EWS [11]. This EWS applies aberration detection algorithms, such as the Early Aberration Reporting System (C2) [12], to a time series consisting of weekly counts of ILI-related PHC encounters per municipality, aiming at the early detection of outbreaks. To assess the data quality of the PHC data stream, we established the composite indicator DQI to measure the completeness and timeliness of the data. Using the backfilled PHC dataset as a gold standard, we evaluated the impact of data quality in the EWS' performance using different levels of data quality of the weekly updated real-time PHC dataset across all 5570 Brazilian municipalities from October 10, 2023, to March 10, 2024.

Data Source

Brazil is an upper middle-income country with approximately 212.6 million people living in 5570 municipalities [13], and we included all ILI-related PHC encounters occurring during the study period in our analysis. We analyzed data from the Brazilian Unified Health System (SUS), which stands as one of the largest public health systems globally, providing comprehensive and universal health care to the entire population. The effective management of SUS relies on diverse information systems, among which the Brazilian National Information System on Primary Health Care (SISAB [Sistema de Informação em Saúde para a Atenção Básica]) plays a crucial role. SISAB is a hierarchical, decentralized information system maintained and managed by the Ministry of Health (MoH), and harbors data on all publicly funded PHC encounters in the country. Data registration is mandatory for the allocation of financial resources from the federal to the municipal level. All encounters are coded by the *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) or the International Classification of Primary Care (ICPC-2).

According to the MoH's guidelines, municipalities are requested to update the system at least on a monthly basis, with a window of 4 months for amendments following each monthly submission. This operational guideline aligns with the SISAB's purpose of informing decision-making for the management of the PHC system in the country. However, the EWS uses weekly updates of the SISAB database to detect ILI outbreaks. Therefore, this real-time, weekly updated dataset may present incompleteness and a temporal lag between the dates of encounter and data registration into the system (Figure S1 in Multimedia Appendix 1).

The DQI

We defined the dimensions of completeness and timeliness to develop quantitative indicators for monitoring data quality in the EWS. Completeness is one of the most commonly used dimensions in data quality assessment and may be defined as the proportion of data filled with values for each attribute or entity in the database, while timeliness can be defined as the availability of data for decision-making, measured by the time interval between the occurrence of the measured event and its capture in an information system [14].

In our study, completeness refers to the proportion of weeks in each 8-week rolling window with any register of a PHC encounter. The indicator is measured as a fraction, with the numerator ranging from 0 to 8, and the denominator is 8 weeks, which is expressed as a percentage. Timeliness refers to the time interval, in number of weeks, between the date of the PHC encounter and its registry in the database. The indicator is represented by the proportion of registries occurring in 2 weeks or less from the PHC encounter in the same 8-week rolling window.

As it is recommended that the diverse quality dimensions should be collectively analyzed for a more comprehensive evaluation of data quality [15], we combined the 2 selected indicators in a composite measure, named DQI. The DQI is assessed weekly, for each municipality, once the PHC data are updated into the EWS pipeline.

Impact of DQI on the EWS' Performance

To decide on the minimum required threshold of completeness and timeliness to derive trustworthy results with the EWS, we applied the EWS algorithm to the retrospectively gathered, backfilled PHC dataset. We compared the results to those obtained when applying the EWS to the weekly updated, herein named real-time PHC dataset (Figure S1 in Multimedia Appendix 1). Using the backfilled dataset as a reference, we calculated the proportion of concordant warnings detected in the real-time dataset. Accordingly, the DQI is expressed as either "suitable" or "unsuitable" when the minimum threshold of both completeness and timeliness is reached, indicating that the data quality may not be adequate for reliable EWS outputs.

Analyses were performed using Python (version 3.9) and R (version 4.3.1) software. The database's description and the scripts are available on GitHub [16].

Ethical Considerations

The study protocol and procedures were reviewed and approved by the Ethical Review Board of Oswaldo Cruz Foundation – Fiocruz Bahia (protocol CAAE 61444122.0.0000.0040).

Data on publicly funded PHC encounters were collected and compiled by the MoH for funding reasons. No consent was needed for data collection at this administrative level. For this study, we accessed an aggregated database consisting of the number of encounters per epidemiological week, per municipality, and per diagnostic code. The accessed database has no information at the individual level, and given that this study involves secondary analysis of existing deidentified data and does not involve direct interaction with human participants, it is classified as exempt from the requirement for informed consent under applicable ethical guidelines.

Results

There were 198,335,762 recorded ILI-related encounters in the backfilled PHC dataset, which corresponds to an average of 8,623,294 encounters per week between October 10, 2023, and March 10, 2024. Using the backfilled dataset, the EWS detected a median of 4 (IQR 2 - 5) warnings of ILI outbreaks per municipality in the study period.

Figure 1 illustrates the impact of the DQI on the ability of the EWS to correctly identify potential ILI outbreaks. Using the real-time dataset, the EWS detected 12,538 (65%) warnings of ILI-outbreaks that were concordant with warnings detected in the backfilled dataset (Table S1 in Multimedia Appendix 1). The proportion of concordant warnings detected in the real-time dataset, based on different levels of completeness (Figure 2A and Table S1 in Multimedia Appendix 1) and timeliness (Figure 2B and Table S1 in Multimedia Appendix 1), indicated that 100% completeness and a minimum of 80% timeliness yielded the highest percentage of concordant warnings. Therefore, these values were established as the thresholds for grading the DQI as suitable or unsuitable for the EWS. Restricting the EWS analysis to municipalities with a suitable DQI, the proportion of warnings for ILI outbreaks concordant to the backfilled dataset increased to 80.4% (Table S2 in Multimedia Appendix 1). We found a median of 71% (IQR 54% - 71.9%) of Brazilian municipalities with a suitable DQI per week in the study period (Figure 3A and Table S2 in Multimedia Appendix 1).

Additionally, we analyzed concordant warnings by grouping municipalities based on the proportion of weeks in which they exhibited a suitable DQI (≤20%, 20% - 40%, 40% - 60%, 60% - 80%, and ≥80%). Our findings revealed that municipalities with over 60% of weeks featuring a suitable DQI had the highest proportion of concordant warnings between the backfilled and real-time datasets (Figure 3B, Table S2 in Multimedia Appendix 1).

Figure 1. Primary Health Care encounters due to influenza-like illness per week in (A) Sao Paulo (municipality with a suitable data quality index in less than 60% of the 23 weeks in study period) and (B) Vitoria da Conquista (municipality with a suitable data quality index for over 80% of the study period). Plots show backfilled (gray line) and real-time (green line) PHC datasets for influenza-like illness (ILI) encounters. Vertical dashed lines show all detected warnings with the Early Aberration Reporting System (EARS), red triangles show nonconcordant warnings between backfield and real-time datasets, and blue circles show concordant warnings between them. ILI: influenza-like illness; PHC: Primary Health Care.

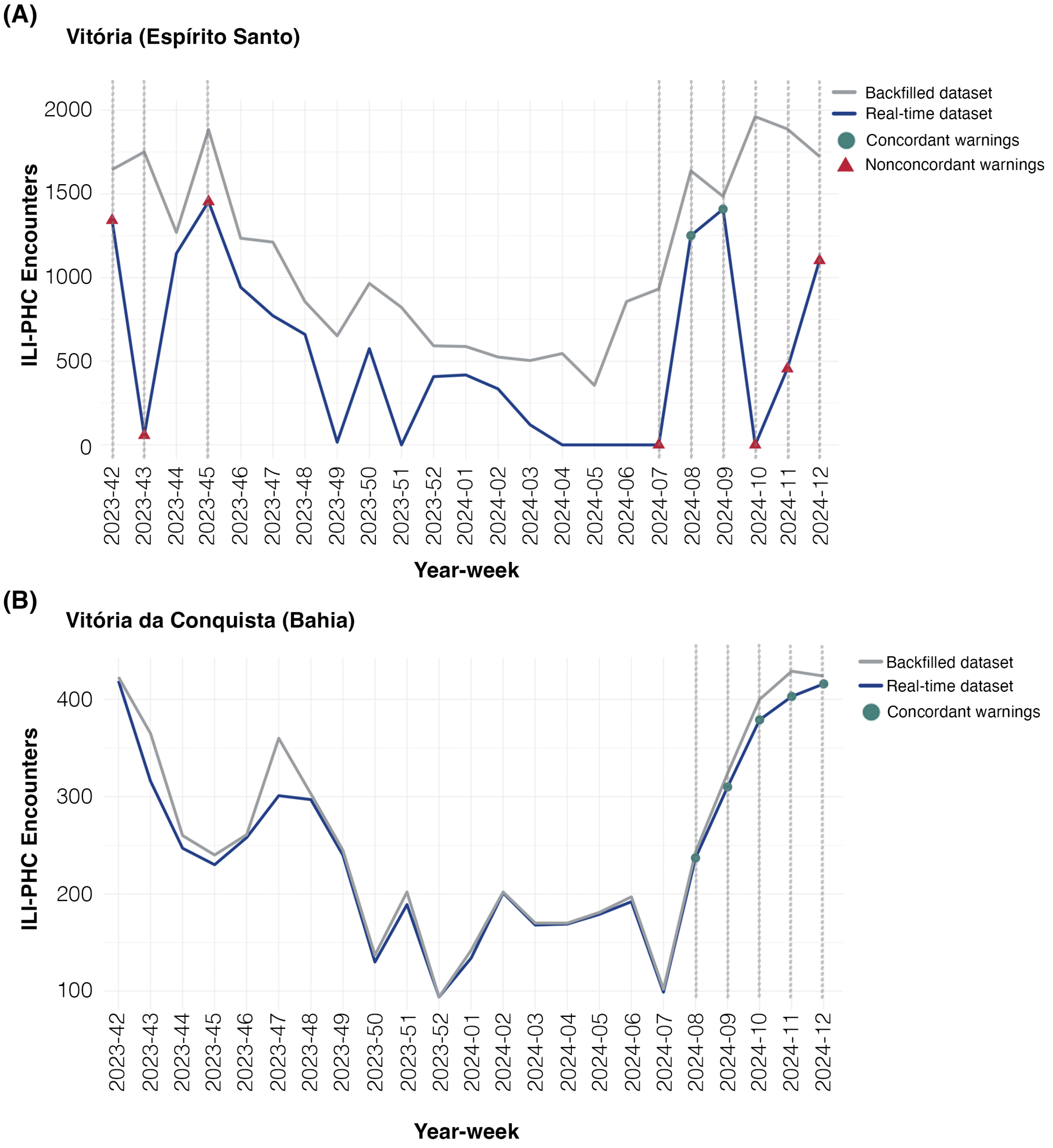


Figure 2. Proportion of concordant outbreak warnings detected in backfilled and real-time Brazilian Primary Health Care datasets. Outbreak warnings generated by the Early Aberration Reporting System (EARS) were identified in both backfilled and real-time datasets. Concordant warnings, detected in both datasets within the same week, are represented in dark blue, while nonconcordant warnings, identified only in the backfilled dataset, are shown in light blue. The analysis considers the proportion of concordant and nonconcordant warnings based on real-time dataset. (**A**) Completeness: the percentage of records from the real-time dataset in each 8-week rolling window (ranging from 0% to 100%) and (**B**) timeliness: the proportion of records registered within 2 weeks or less of the PHC encounter, measured within the same 8-week rolling window (ranging from 0% to 100%).

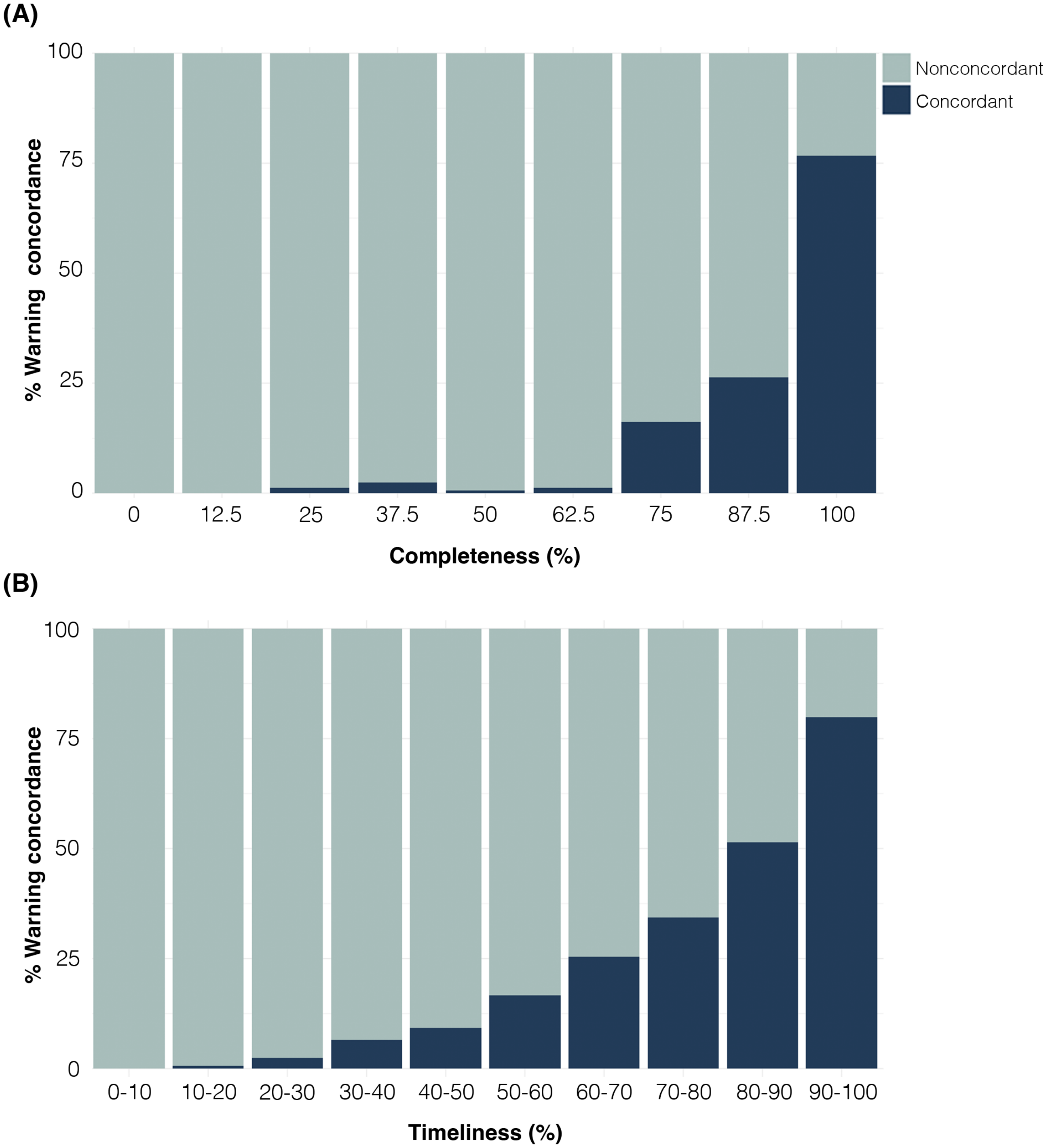
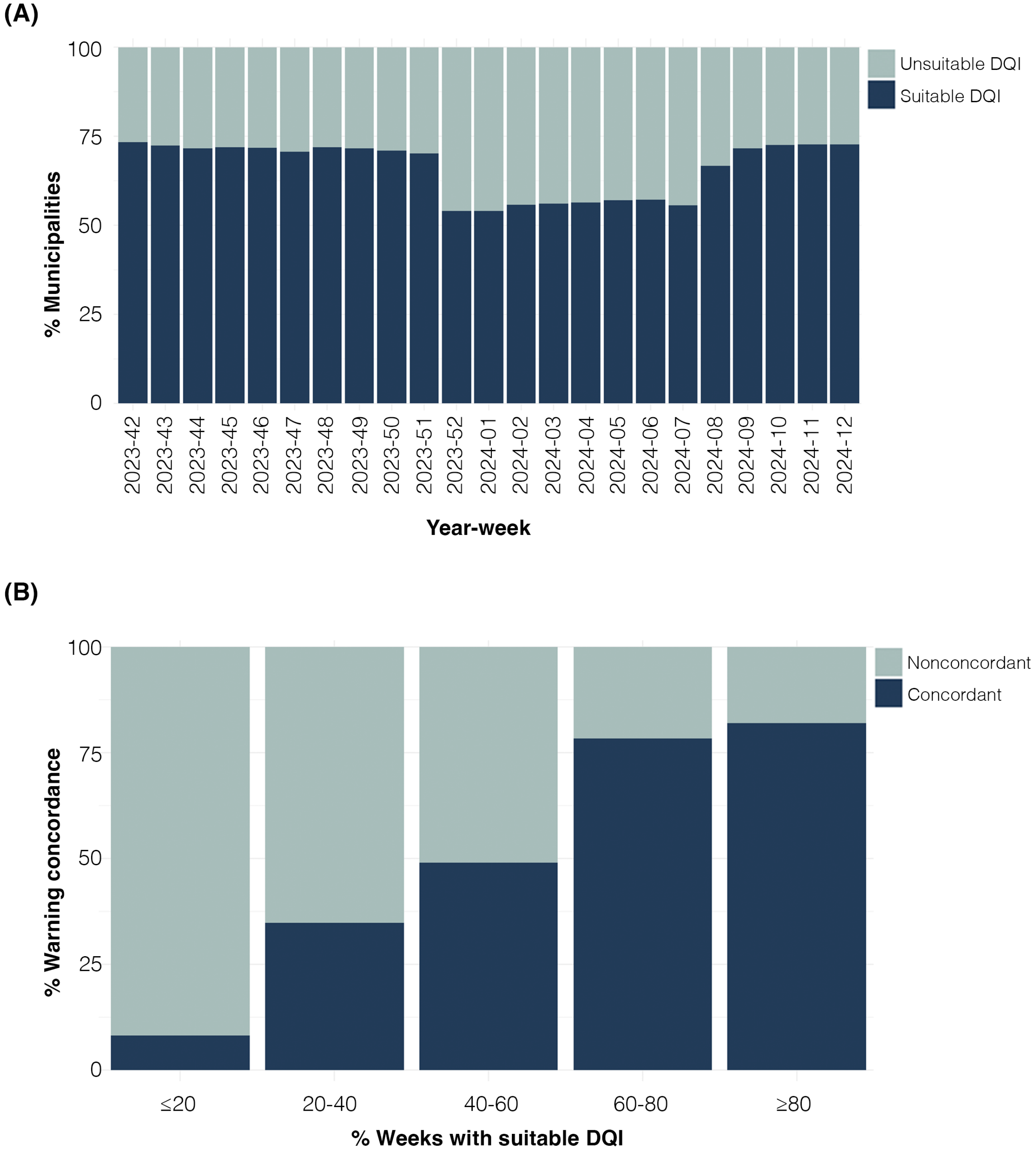


Figure 3. Proportion of Brazilian municipalities with suitable data quality index (DQI) and concordant warnings over time. (**A**) Weekly analysis of the proportion of municipalities with a suitable DQI from epidemiological week 42 of 2023 to week 12 of 2024. (**B**) Proportion of concordant warnings (dark blue), identified in both datasets within the same week, and nonconcordant warnings (light blue), detected only in the backfilled dataset. The analysis is based on the proportion of weeks with a suitable DQI in the real-time dataset from Brazilian municipalities.



Discussion

Principal Findings

Our study highlights the critical role of data quality in the performance of the EWS for infectious disease surveillance using PHC data. In addition, we provide a practical approach for monitoring data quality in real time, which can be adapted to other settings and data types. Our findings revealed that municipalities with over 60% of weeks featuring a suitable DQI had the highest proportion of concordant warnings between the backfilled and real-time datasets. Introducing the DQI as an algorithm integrated into the EWS can guide data management practices and inform decision-making processes.

Similar to our findings, a recent systematic review of the effectiveness of EWS found that the improvement of data is pivotal for emergency department–based surveillance [17]. However, efforts for automatization of data quality assessment are typically scattered [9], and the literature on the operationalization of data quality assessment remains scarce.

A study on data quality assessment for public health information systems found a lack of systematic procedures for quality assessment. While quality assessment of quantitative data generally used descriptive surveys, the authors argued about the importance of systematic scientific data quality assessment [18]. To the best of our knowledge, this is the first publication to assess the importance of integrating data quality monitoring into an EWS.

Fulcher et al [19] demonstrated how administrative health data were successfully used to implement a syndromic surveillance system during the COVID-19 pandemic. However, the process of cleaning data and handling missed data was carried out by a dedicated analyst once the updated database became available [19]. We anticipate that the framework for a data quality assessment integrated to the EWS pipeline presented here can be adapted to other surveillance systems and can provide insights for similar implementations.

Using a retrospectively gathered, backfilled PHC dataset, we evaluated the EWS based on optimal data quality conditions. However, administrative data usually exhibit incompleteness and delays, and the EWS should be capable of detecting outbreaks using the available dataset in real time. Our analysis revealed that high levels of completeness (100%) and timeliness (at least 80%) are necessary to achieve the highest proportion of concordant warnings between backfilled and real-time datasets. Additionally, our results indicate that even incremental data quality improvements substantially enhance the EWS' performance. Achieving such high standards may pose challenges, particularly in low-resource settings that potentially face limitations due to infrastructure such as unreliable internet connectivity and insufficient computer power. Despite these challenges, we found a weekly median of 71% of Brazilian municipalities achieving the threshold for a suitable DQI for the EWS. This result suggests that a significant proportion of municipalities met the minimum threshold for data quality even in constrained settings.

In this study, we used the SUS database, which covers approximately 75% of the Brazilian population, with great granularity, reaching underserved rural and remote regions [20]. This approach allowed us to assess the performance of the EWS across different regions and health service contexts. However, these findings may not be directly applicable to other countries. It is likely that the use of the EWS in different health system structures and data management practices will need adjustments and may require distinct data quality requirements [8].

Another limitation of this study is that we could not access other dimensions of data quality. Specifically, we could not access the accuracy of registers in the PHC dataset. Accuracy represents the extent to which the data are free of error and reliable [8,9]. We worked with aggregated, secondary data, and did not have access to the complete EHRs, which precluded us from verifying whether the diagnostic codes in the database accurately reflected patients' main clinical problems. It is our perspective that evaluating the accuracy of the *ICD-10* and ICPC-2 is of great importance. However, given the large numbers of PHC encounters registered weekly, misclassifications of the reason of encounter are likely to be nondifferential. Additionally, syndromic surveillance systems are designed to operate effectively even with some level of imprecision, as their primary purpose is to detect patterns and trends rather than to provide definitive diagnoses.

Conclusion

Our findings demonstrate that implementing a robust and integrated DQI analysis can significantly enhance the EWS' ability to detect ILI outbreaks, contributing to better public health outcomes and ultimately to global health security. Beyond contributing to the existing literature on EWS, this study highlights the importance of systematic data quality assessment. Continuous monitoring and improvement of data quality should be prioritized to ensure the reliability and effectiveness of surveillance systems. Additionally, our study suggests that similar frameworks can be adapted to different contexts. As health systems increasingly use digital health data for decision-making, our approach represents a model for integrating data quality monitoring into surveillance systems, ultimately enhancing the capacity to detect and respond to infectious disease outbreaks effectively.

Acknowledgments

This study was funded by the Rockefeller Foundation's Health Initiative (grant 2023-PPI-007 awarded to MBN). MBN, PIPR, and VB are research fellows from the National Council for Scientific and Technological Development (CNPq, Brazil). TCS acknowledges funding from the Royal Society (NIF\R1\231435). The funders did not interfere in the analysis, interpretation, or decision to submit the manuscript for publication. This study is part of the Alert-Early System of Outbreaks with Pandemic Potential (ÆSOP), an initiative under development by Brazil's Fundação Oswaldo Cruz (Fiocruz) and the Federal University of Rio de Janeiro, and financially supported by Rockefeller Foundation's Health Initiative. Figure S1 in Multimedia Appendix 1 was created with Biorender.com and was published with permission. Generative artificial intelligence was not used for ideation or any part of the study design. Additionally, it was not used for reference searches. Its use was limited to grammatical revisions in the manuscript, with the prompt, "Please check for readability and possible grammatical mistakes" (a transcript of the conversation with the chatbot is provided in Multimedia Appendix 2).

Data Availability

Our agreement with the Brazilian Ministry of Health (MoH) for accessing the referenced databases patently denies authorization of access to any third parties. All requests to access these databases must be addressed to the Brazilian MoH.

Authors' Contributions

Conceptualization: PTVF, IM, and MBN
Data acquisition: VdAO
Data curation and processing: VdAO, JBJ, and GCGB
Formal analysis: PTVF
Script verification: JBJ, GCGB, and TCS
Study design: PTVF, JBJ, GCGB, TCS, VdAO, MHdOG, GOP, VB, PIPR, MBN, and IM
Writing—original draft: PTVF and IM
Writing—review and editing: GOP, TCS, VB, MHdOG, VdAO, and PIPR

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional material.
[DOCX File, 141 KB - publichealth_v11i1e67050_app1.docx]

Multimedia Appendix 2
Conversation with the chatbot for grammatical revision.
[PDF File, 562 KB - publichealth_v11i1e67050_app2.pdf]

References

1. Baker RE, Mahmud AS, Miller IF, et al. Infectious disease in an era of global change. Nat Rev Microbiol 2022 Apr;20(4):193-205. [doi: 10.1038/s41579-021-00639-z] [Medline: 34646006]
2. Morgan OW, Aguilera X, Ammon A, et al. Disease surveillance for the COVID-19 era: time for bold changes. The Lancet 2021 Jun 19;397(10292):2317-2319. [doi: 10.1016/S0140-6736(21)01096-5] [Medline: 34000258]
3. Data quality monitoring and surveillance system evaluation - a handbook of methods and applications. European Centre for Disease Prevention and Control. 2014. URL: https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/Data-quality-monitoring-surveillance-system-evaluation-Sept-2014.pdf [accessed 2025-02-06]
4. Ramos PIP, Marcilio I, Bento AI, et al. Combining digital and molecular approaches using health and alternate data sources in a next-generation surveillance system for anticipating outbreaks of pandemic potential. JMIR Public Health Surveill 2024 Jan 9;10:e47673. [doi: 10.2196/47673] [Medline: 38194263]
5. Ulrich EH, So G, Zappitelli M, Chanchlani R. A review on the application and limitations of administrative health care data for the study of acute kidney injury epidemiology and outcomes in children. Front Pediatr 2021;9:742888. [doi: 10.3389/fped.2021.742888] [Medline: 34778133]
6. Tassi MF, le Meur N, Stéfic K, Grammatico-Guillon L. Performance of French medico-administrative databases in epidemiology of infectious diseases: a scoping review. Front Public Health 2023;11:1161550. [doi: 10.3389/fpubh.2023.1161550] [Medline: 37250067]
7. Shaw RJ, Harron KL, Pescarini JM, et al. Biases arising from linked administrative data for epidemiological research: a conceptual framework from registration to analyses. Eur J Epidemiol 2022 Dec;37(12):1215-1224. [doi: 10.1007/s10654-022-00934-w] [Medline: 36333542]
8. Lewis AE, Weiskopf N, Abrams ZB, et al. Electronic health record data quality assessment and tools: a systematic review. J Am Med Inform Assoc 2023 Sep 25;30(10):1730-1740. [doi: 10.1093/jamia/ocad120] [Medline: 37390812]
9. Ozonze O, Scott PJ, Hopgood AA. Automating electronic health record data quality assessment. J Med Syst 2023 Feb 13;47(1):23. [doi: 10.1007/s10916-022-01892-2] [Medline: 36781551]
10. Bloland P, MacNeil A. Defining & assessing the quality, usability, and utilization of immunization data. BMC Public Health 2019 Apr 4;19(1):380. [doi: 10.1186/s12889-019-6709-1] [Medline: 30947703]
11. Cerqueira-Silva T, Oliveira JF, de Araújo Oliveira V, et al. Early warning system using primary health care data in the post-COVID-19 pandemic era: Brazil nationwide case-study. Cad Saude Publica 2024;40(11):e00010024. [doi: 10.1590/0102-311XEN010024] [Medline: 39775767]
12. Zhu Y, Wang W, Atrubin D, Wu Y. Initial evaluation of the early aberration reporting system--Florida. MMWR Suppl 2005 Aug 26;54:123-130. [Medline: 16177703]
13. 2022 population census: main results. Brazilian Institute of Geography and Statistics. 2022. URL: https://www.ibge.gov.br/en/statistics/social/health/22836-2022-census-3.html [accessed 2024-12-18]
14. Fox C, Levitin A, Redman T. The notion of data and its quality dimensions. Inf Process Manag 1994 Jan;30(1):9-19. [doi: 10.1016/0306-4573(94)90020-5]
15. Hassenstein MJ, Vanella P. Data quality—concepts and problems. Encyclopedia 2022 Feb;2(1):498-510. [doi: 10.3390/encyclopedia2010032]

16. AESOP-data-documentation. GitHub. URL: https://github.com/cidacslab/AESOP-Data-Documentation/tree/main/DataPipeline [accessed 2025-02-06]
17. Meckawy R, Stuckler D, Mehta A, Al-Ahdal T, Doebbeling BN. Effectiveness of early warning systems in the detection of infectious diseases outbreaks: a systematic review. BMC Public Health 2022 Nov 29;22(1):2216. [doi: 10.1186/s12889-022-14625-4] [Medline: 36447171]
18. Chen H, Hailey D, Wang N, Yu P. A review of data quality assessment methods for public health information systems. Int J Environ Res Public Health 2014 May 14;11(5):5170-5207. [doi: 10.3390/ijerph110505170] [Medline: 24830450]
19. Fulcher IR, Boley EJ, Gopaluni A, et al. Syndromic surveillance using monthly aggregate health systems information data: methods with application to COVID-19 in Liberia. Int J Epidemiol 2021 Aug 30;50(4):1091-1102. [doi: 10.1093/ije/dyab094]
20. Macinko J, Harris MJ, Rocha MG. Brazil's National Program for Improving Primary Care Access and Quality (PMAQ): fulfilling the potential of the world's largest payment for performance system in primary care. J Ambul Care Manage 2017;40 Suppl 2 Supplement(2 Suppl):S4-S11. [doi: 10.1097/JAC.0000000000000189] [Medline: 28252498]

Abbreviations

DQI: data quality index
EHR: electronic health record
EWS: early warning system
***ICD-10*:** *International Statistical Classification of Diseases, Tenth Revision*
ICPC-2: International Classification of Primary Care
ILI: influenza-like illness
MoH: Ministry of Health
PHC: Primary Health Care
SISAB: Sistema de Informação em Saúde para a Atenção Básica
SUS: Brazilian Unified Health System
ÆSOP: Alert-Early System of Outbreaks with Pandemic Potential

---

Edited by A Mavragani; submitted 30.09.24; peer-reviewed by C Haridas, O Olasunkanmi; revised version received 19.12.24; accepted 20.12.24; published 21.02.25.

Please cite as:
Florentino PTV, Bertoldo Junior J, Barbosa GCG, Cerqueira-Silva T, Oliveira VDA, Garcia MHDO, Penna GO, Boaventura V, Ramos PIP, Barral-Netto M, Marcilio I
Impact of Primary Health Care Data Quality on Infectious Disease Surveillance in Brazil: Case Study
JMIR Public Health Surveill 2025;11:e67050
URL: https://publichealth.jmir.org/2025/1/e67050
doi:10.2196/67050

Effectiveness of Frequent Point-of-Care Molecular COVID-19 Surveillance in a Rural Workplace: Nonrandomized Controlled Clinical Trial Among Miners

Akshay Sood[1], MPH, MD; William “Cotton” Jarrell[2], CMSP; Xin W Shore[3], MSc; Nestor Sosa[3], MD; Alisha Parada[3], MD; Nicholas Edwardson[4], PhD; Alexandra V Yingling[3], PhD; Teah Amirkabirian[3], BSc; Qiuying Cheng[3], PhD; Ivy Hurwitz[3], PhD; Linda S Cook[5], PhD; Shuguang Leng[3], PhD; Orrin B Myers[3], PhD; Douglas J Perkins[6], PhD

[1]Department of Internal Medicine, School of Medicine, University of New Mexico Health Sciences Center, 1 University of New Mexico MSC 10 5550, Albuquerque, NM, United States

[2]Miners Colfax Medical Center, Raton, NM, United States

[3]University of New Mexico Health Sciences Center, Albuquerque, NM, United States

[4]School of Public Administration, University of New Mexico, Albuquerque, NM, United States

[5]Department of Epidemiology, University of Colorado School of Public Health, Aurora, CO, United States

[6]Center for Global Health, University of New Mexico Health Sciences Center, Albuquerque, NM, United States

Corresponding Author:
Akshay Sood, MPH, MD
Department of Internal Medicine, School of Medicine, University of New Mexico Health Sciences Center, 1 University of New Mexico MSC 10 5550, Albuquerque, NM, United States

Abstract

Background: Numerous studies have assessed the risk of SARS-CoV-2 exposure and infection among health care workers during the pandemic. However, far fewer studies have investigated the impact of SARS-CoV-2 on essential workers in other sectors. Moreover, guidance for maintaining a safely operating workplace in sectors outside of health care remains limited. Workplace surveillance has been recommended by the Centers for Disease Control and Prevention, but few studies have examined the feasibility or effectiveness of this approach.

Objective: The objective of this study was to investigate the feasibility and effectiveness of using frequent point-of-care molecular workplace surveillance as an intervention strategy to prevent the spread of SARS-CoV-2 at essential rural workplaces (mining sites) where physical distancing, remote work, and flexible schedules are not possible.

Methods: In this nonrandomized controlled clinical trial conducted from February 2021, to March 2022, 169 miners in New Mexico (intervention cohort) and 61 miners in Wyoming (control cohort) were enrolled. Investigators performed point-of-care rapid antigen testing on midnasal swabs (NSs) self-collected by intervention miners. Our first outcome was the intervention acceptance rate in the intervention cohort. Our second outcome was the rate of cumulative postbaseline seropositivity to SARS-CoV-2 nucleocapsid protein, which was analyzed in the intervention cohort and compared to the control cohort between baseline and 12 months. The diagnostic accuracy of detecting SARS-CoV-2 using rapid antigen testing on NSs was compared to laboratory-based reverse transcriptase polymerase chain reaction (RT-PCR) on nasopharyngeal swabs (NPSs) in a subset of 68 samples.

Results: Our intervention had a mean acceptance rate of 96.4% (11,413/11,842). The intervention miners exhibited a lower cumulative postbaseline incident seropositivity at 12 months compared to control miners (14/97, 14% vs 17/45, 38%; *P*=.002). Analysis of SARS-CoV-2 antigen detection in self-administered NSs revealed 100% sensitivity and specificity compared to laboratory-based RT-PCR testing on NPSs.

Conclusions: Our findings establish frequent point-of-care molecular workplace COVID-19 surveillance as a feasible option for keeping essential rural workplaces open and preventing SARS-CoV-2 spread. These findings extend beyond this study, providing valuable insights for designing interventions to maintain employees’ safety at other essential workplaces during an infectious disease outbreak.

Trial Registration: ClinicalTrials.gov NCT04977050; https://clinicaltrials.gov/study/NCT04977050

(JMIR Public Health Surveill 2025;11:e59845) doi:10.2196/59845

KEYWORDS

point-of-care; seroprevalence; SARS-CoV-2; coronavirus; COVID-19; surveillance; rural workplace; miners; infectious disease; pandemic; antigen testing; midnasal swabs; public health

Introduction

Thus far, an estimated 55 million US workers have provided essential services during the COVID-19 pandemic [1]. While the significant occupational risk of SARS-CoV-2 infection faced by essential health care workers has been well-documented [2-5], far fewer studies have investigated SARS-CoV-2 infection rates in other essential industries [6-10]. Moreover, guidance for maintaining a safely operating workplace in sectors outside of health care remains limited. Workplace surveillance has been recommended by the Centers for Disease Control and Prevention (CDC) [11], but few studies have examined the feasibility or effectiveness of this approach [12-14]. The lack of adequate data in this regard constitutes a critical gap in the literature that needs to be addressed to help devise pandemic intervention strategies for workplaces outside of health care.

Laboratory-based nucleic acid amplification tests, such as reverse transcriptase polymerase chain reaction (RT-PCR) assays performed on nasopharyngeal swabs (NPS) collected by trained professionals, are considered the gold standard for SARS-CoV-2 diagnostic testing based on accuracy, sensitivity, and specificity. However, essential workers outside the health care sector, particularly in rural and remote minority communities, may encounter issues with access, cost, and wait time for results associated with this gold standard testing. Consequently, inconsistent or absent gold standard testing has placed rural essential workers outside of health care at an increased risk for SARS-CoV-2 infection and associated complications [15].

One strategy to overcome challenges with gold standard testing is to use self-collected nasal swabs (NS) and point-of-care qualitative SARS-CoV-2 rapid antigen testing (QAT). QAT offers several advantages, such as low cost, enhanced testing accessibility, and rapid turnaround time, making it well-suited for resource-limited settings. Moreover, the use of self-collected NSs is comparable to clinician-collected NPSs for detecting SARS-CoV-2 in controlled settings [16,17]. In this study, we evaluate the diagnostic accuracy, reliability, feasibility, and effectiveness of frequent point-of-care molecular workplace COVID-19 surveillance in a diverse population of coal miners in a remote community in New Mexico, compared to a cohort of coal miners from a rural mining site in Campbell County, Wyoming. The novelty of the research question and the involvement of a diverse, understudied, and underserved population contribute to the innovation of the study. The long-term goal of the study was to inform strategies for future pandemic mitigation efforts in mining settings and other essential rural workplaces.

Methods

Study Design

To assess the effectiveness of implementing point-of-care QAT for SARS-CoV-2 on self-collected NS in rural workplaces, we conducted a nonrandomized controlled clinical trial involving two cohorts of coal miners from a mining site in McKinley County, New Mexico (intervention site) and from a mining site in Campbell County, Wyoming (control site). The primary recruitment period was between February 2021 and March 2022 with additional miners entering the study after this time. Inclusion criteria were as follows: male or female miners currently employed at the above-mentioned mines, at least 18 years of age, willing and able to consent to study participation, and willing and able to comply with study procedures. Exclusion criteria were as follows: unable or unwilling to consent and less than 18 years of age.

Both intervention and control sites were surface coal mines in rural and remote mountainous locations operated by the same company with similar working conditions and policies. They shared similar engineering and administrative controls and had similar protective measures on screening for SARS-CoV-2 symptoms and fever and for using face masks on-site. Workplace policy at both sites precluded miners with COVID-19 symptoms from reporting for duty, and symptom screening was conducted by mine safety personnel at the beginning of each work shift before miners could enter the mines.

Trained miner investigators collected information using phone- and computer-based web applications developed by Ingenuity Software Labs (Albuquerque, NM), paper-based questionnaires, and body temperature measurements using a no-touch forehead infrared thermometer. The information collected included a 24-hour history of SARS-CoV-2 symptoms, any provider diagnosis of infection, potential close contact with infected individuals, vaccination status, and use of cloth face covering in public settings outside the mine. Data were transferred manually by the miner investigators to the university-based investigators, who entered the information into a secure REDCap database. All information entered was double-checked for accuracy.

Miner Testing

Qualitative SARS-CoV-2 Rapid Antigen Testing

For rapid detection of the SARS-CoV-2 antigen, eligible intervention miners were provided with an NS at the mine entrance at the beginning of each alternate work shift, in a schedule which consisted of 4 days on and 4 days off, resulting in 2 NSs being collected in every 8-day cycle. NSs were self-collected by the miner subjects, with trained miner investigators supervising the process. Subsequently, the 10-minute Quidel Quickvue SARS-CoV-2 QAT (ie, intervention test; Quidel Corporation, San Diego, CA) was conducted on the collected swabs by trained miner investigators at a temporary laboratory created at the intervention mine entrance. Test results were available in about 10 minutes, roughly the time required to drive from the mine entrance to the parking lot. Based on the CDC guidelines, those who tested positive were requested to return home, isolate themselves, and contact their primary care providers [11].

Seropositivity Testing

Venipuncture blood samples of 8 mL were obtained from intervention and control miner subjects at the time of enrollment and at 3, 6, and 12 months. Whole blood was centrifuged at the mine site, and plasma was aspirated, aliquoted, and frozen by trained nurse technicians. Plasma samples were tested to determine the seroprevalence of COVID-19 using the immunoglobulin G antibody to the nucleocapsid protein of SARS-CoV-2, and tests were performed by a qualitative chemiluminescent immunoassay (ARUP Laboratory, Salt Lake City, UT), under the Emergency Use Authorization. Unlike the antibodies to spike protein, antibodies to nucleocapsid protein develop in response to natural infection and not due to COVID-19 vaccination.

Diagnostic Accuracy of NSs Versus NPSs

To determine whether self-collected NSs were comparable to NPSs collected by health care providers for detecting SARS-CoV-2, a subset of miners were requested to perform a repeat self-administered NS on both nostrils, while 2 NPSs were administered by a skilled health care provider from separate nostrils, all in one setting. All swabs were placed separately into appropriately labeled tubes with viral transport media for transportation to the laboratories. The first NPS sample was used for running the confirmatory qualitative SARS-CoV-2 RT-PCR test at TriCore Diagnostic Laboratory (Albuquerque, NM)—the gold standard—performed under Emergency Use Authorization from the Food and Drug Administration. Sensitivity, specificity, and predictive values were calculated for the intervention test compared to the gold standard test.

The NS and the second NPS were used for quantitative RT-PCR (RT-qPCR) tests. Aliquots of viral transport media were inactivated with DNA/RNA Shield (Zymo Research, CA) and stored at –80°C until bulk processing using the Quick-RNA Viral kit (Zymo Research). RT-qPCR assays were performed using the N1 and RP primer/probe sets from the CDC 2019-nCoV-01 diagnostic panel. No target controls were included in all amplification reactions. The N1 primer/probe set is designed to specifically detect and amplify a region within the nucleocapsid gene of SARS-CoV-2, while the RP primer/probe set targets a portion of the human RNAse P (RNP) gene. The RP primer/probe set is used in all clinical samples to assess specimen quality. Samples with a cycle threshold (Ct) values ≥40 were considered negative. For viral load quantification, standard curves were generated from the Ct values of N1 and RP using known concentrations of SARS-CoV-2 RNA and plasmids harboring the RNP gene fragment, respectively. Viral load per sample was determined as described by Perkins et al [18]. The Cohen κ coefficient was used to determine the intrarater agreement between the NS and NPS sample collection methodologies.

Statistical Analysis

Miner Characteristics

Characteristics of miners participating at the intervention and control mines were summarized as frequencies and percentages and means and SDs. Frequencies of categorical variables for the two study arms were compared using χ^2 tests and Fisher exact tests. Continuous age was compared using a Wilcoxon rank sum test.

Intervention Acceptance

Intervention test acceptance by miners was measured for each eligible work shift, and the outcome variable and cumulative acceptance rate were calculated over the study time frame (frequency of tests accepted/number of tests offered). Predictors influencing the outcome were examined in the intervention cohort, including self-reported educational status, racial and ethnic minority status, prior SARS-CoV-2 infection or vaccination, and evidence of previous infection (from a prior positive test). Nonparametric Wilcoxon and Kruskal-Wallis tests were used to assess the association of miner characteristics with cumulative acceptance rates, since acceptance rates had a skewed distribution. We also computed the total number of tests accepted per miner for each week of the study (total number of tests accepted in each week/number of consented miners). Weekly tests per miner were summarized graphically, and a test for trend was made assuming residual errors had a first-order autocorrelation.

Seropositivity Analysis

Seroprevalence patterns of miners enrolled and tested between February 23 and March 6, 2021 were compared between intervention and control miners over time intervals using χ^2 tests and time-to-first seropositivity by Kaplan-Meier analysis and log-rank tests. Kaplan-Meier analyses accounted for dropouts during follow-up after baseline. We used logistic regression to compare the following binary outcome variables: baseline (prevalence) seropositivity, postbaseline incident seropositivity, and combined cumulative seropositivity (as positive at any time during the study). Combined cumulative seropositivity analyses also included miners that were enrolled after March 6, 2021. Vaccination status and select host susceptibility and vulnerability factors were added to multivariable logistic regression models. Odds ratios (ORs) and 95% CIs were computed to describe association strength. SAS version 9.4 (SAS Institute) and SPSS version 28.0.0.0 (IBM Corp) were used for statistical analyses, and a *P* value of less than .05 was considered statistically significant.

Power Calculation

The study was planned to estimate nasal swab acceptance rates with precision of ±4.6% from a sample of up to 250 intervention group miners, which also would have at least 80% power for subgroup analyses. A comparison group with up to n=350 was estimated to have at least 80% power to detect group differences in baseline seropositivity and incident seropositivity. However, sample sizes recruited were lower than planned.

Ethical Considerations

The University of New Mexico Health Sciences Center’s Human Research Protections Office and Institutional Review Board (HRPO 20‐680) approved the study. All participants provided informed consent with an ability to opt out. The data were deidentified for analysis. Participants were compensated with a $10 electronic merchandise card for each specimen and survey collected.

Results

Study Design: Cohort Demographics

The study included 115 intervention miners and 60 miners recruited during February and March 2021 and another 54 intervention miners and 1 control miner recruited during the rest of the study (n=169 intervention and n=61 control miners). Baseline characteristics of subjects in the two cohorts are shown in Table 1. The two cohorts were similar in age distribution, with a mean age of 44.4 (SD 11.2) and 45.2 (SD 11.9) years in the intervention and control miner groups, respectively. When comparing the distribution of comorbidities, including asthma, chronic obstructive pulmonary disease, chronic lung diseases, hypertension, diabetes, and depression, 43.8% (74/169) of intervention miners reported at least 1 comorbidity, similar to 46% (28/61) of control miners.

Table . Comparison of baseline characteristics between New Mexico–based (intervention cohort) and Wyoming-based (control cohort) participants.

Characteristic	All (n=230)	New Mexico (n=169)	Wyoming (n=61)	*P* value
Age (years), mean (SD)	44.6 (11.3)	44.4 (11.2)	45.2 (12)	.57
Age group (years), n (%)				.34
<40	93 (40.4)	73 (43.2)	20 (33)	
40 - 50	57 (24.8)	41 (24.3)	16 (26)	
>50	80 (34.8)	55 (32.5)	25 (41)	
Education >12th grade, n (%)	121 (52.6)	69 (40.9)	52 (85)	<.001
Annual household income ≥$80,000, n (%)	111 (48.2)	70 (41.4)	41 (67)	.005
Race, n (%)				<.001
White	152 (66.1)	96 (56.8)	56 (92)	
American Indian or Alaska Native	29 (12.6)	27 (16)	2 (3)	
Asian	1 (0.4)	1 (0.6)	0 (0)	
Other race	34 (14.8)	31 18.3)	3 (5)	
Prefer not to answer or missing	14 (6.1)	14 (8.3)	0 (0)	
Ethnicity, n (%)				<.001
Not Hispanic, Latino, or Spanish origin	117 (50.9)	63 (37.3)	54 (88)	
Hispanic, Latino, or Spanish origin	101 (43.9)	97 (57.4)	4 (7)	
Prefer not to answer or missing	12 (5.2)	9 (5.3)	3 (5)	
Biological sex at birth, n (%)				.02
Male	202 (87.8)	154 (91.1)	48 (79)	
Female	25 (10.9)	14 (8.3)	11 (18)	
Intersex, prefer not to answer or missing	3 (1.2)	1 (0.6)	2 (3)	
COVID-19-related behavior and vaccination characteristics, n (%)				
Close contact with confirmed or suspected COVID-19 since the last shift	19 (8.3)	13 (7.7)	6 (10)	.76
Frequency of ride-sharing to and from mine site often or very often	81 (35.2)	75 (44.4)	6 (10)	<.001
Frequency of face covering outside the mine often or very often	162 (70.4)	142 (84.1)	20 (33)	<.001
COVID-19 vaccination (yes)	52 (22.6)	43 (25.4)	9 (15)	.07
Comorbidities, n (%)				
Previously tested positive for COVID-19	42 (18.3)	24 (14.2)	18 (30)	.08
Asthma, COPD[a], and other chronic lung diseases	30 (13.1)	23 (13.6)	7 (12)	.67
Hypertension	47 (20.4)	34 (20.1)	13 (21)	.87

Characteristic	All (n=230)	New Mexico (n=169)	Wyoming (n=61)	*P* value
Diabetes	21 (9.1)	17 (10.1)	4 (7)	.41
Depression	11 (4.8)	7 (4.1)	4 (7)	.46
Household wood smoke exposure (ever), n (%)	62 (26.9)	51 (30.3)	11 (18)	.08

[a] COPD: chronic obstructive pulmonary disease.

A significant difference was observed between the cohorts regarding racial and ethnic composition, with 72.8% (123/169) of intervention miners being from underrepresented racial and ethnic minorities (including American Indians and Hispanics), compared to 10% (6/61) of control miners (*P*<.001). The intervention cohort consisted of more male participants than the control cohort (154/169, 91% vs 48/61, 79%; *P*=.02). Furthermore, intervention miners had a lower annual income, with 58.6% (99/169) of intervention miners earning less than $80,000 annually versus 33% (20/61) of control miners (*P*=.005) and a lower education level, with 40.9% (69/169) of intervention miners having postsecondary education versus 85% (52/61) of control miners (*P*<.001). In terms of COVID-19–related behavior and vaccination characteristics, intervention miners were more likely to ride-share to and from the mining site (75/169, 44.4% vs 6/61, 10%; *P*<.001) but were also more likely to wear face coverings outside of the mine often or very often (142/169, 84.1% vs 30/61, 33%; *P*<.001). Moreover, intervention miners had a 25.4% (43/169) baseline vaccination rate against COVID-19 versus 15% (9/61) for control miners (*P*=.07). The intervention miners may have been more likely to have been exposed to household wood smoke than control miners (51/169, 30% vs 11/61, 18%; *P*=.08).

Intervention Acceptance

The mean intervention test acceptance rate was 96.4% (11,413/11,842). The high test acceptance was unaffected by demographic or medical history characteristics or baseline COVID-19 vaccine uptake (*P*>.05 for all, data not shown). The number of tests per miner per week (mean 1.64, SD 0.26) did not show a trend over the study (Figure 1; slope=0.004; SE=0.005; *P*=.46; temporal autocorrelation=–0.73). The total number of testing opportunities varied by the individual, with an average of 84 opportunities per miner over 267 study days. Participants voluntarily performed an average of 62.0 (SD 14.4) tests, ranging from 1 to 84 tests (Figure 2).

Figure 1. Number of SARS-CoV-2 antigen tests performed per miner per week. Intervention tests were performed at each alternate work shift (schedule 4 days on and 4 days off) during the 2021 calendar weeks 8-52 and 2022 calendar weeks 1-9 (ie, from the end of February 2021, to the end of February 2022; n=169 miners at the intervention site).

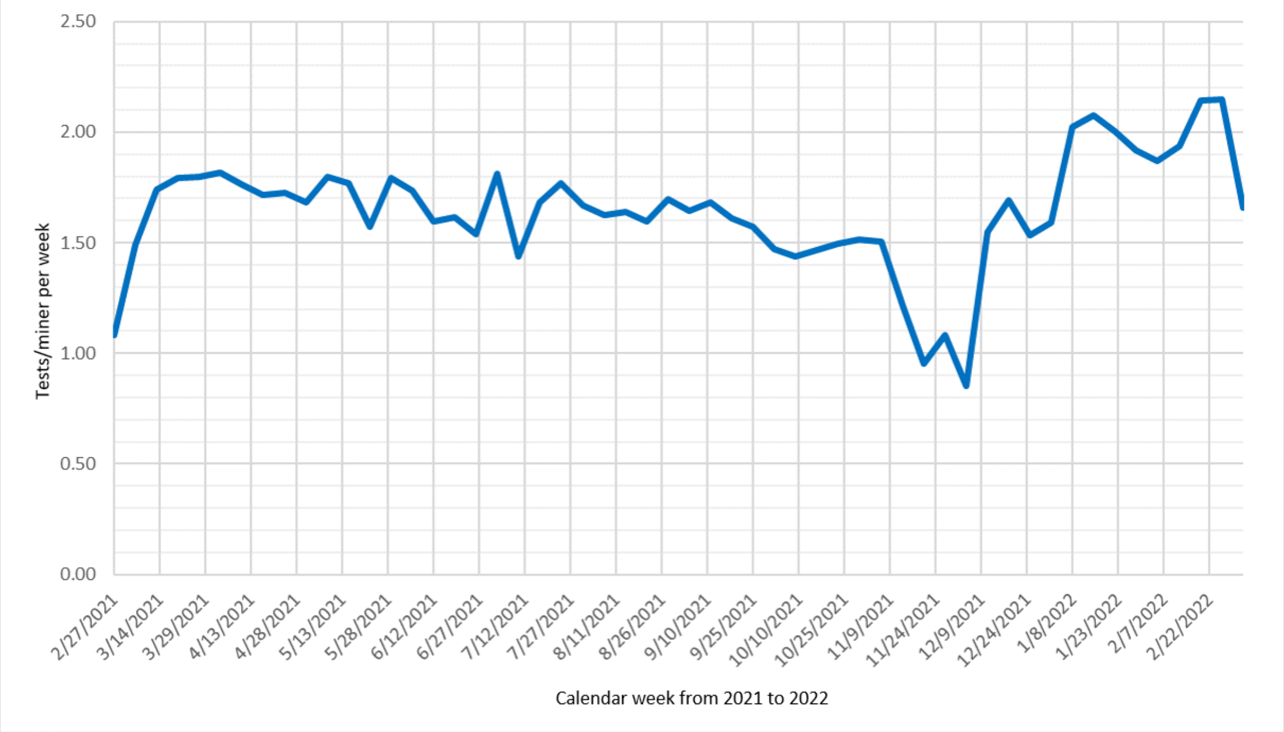
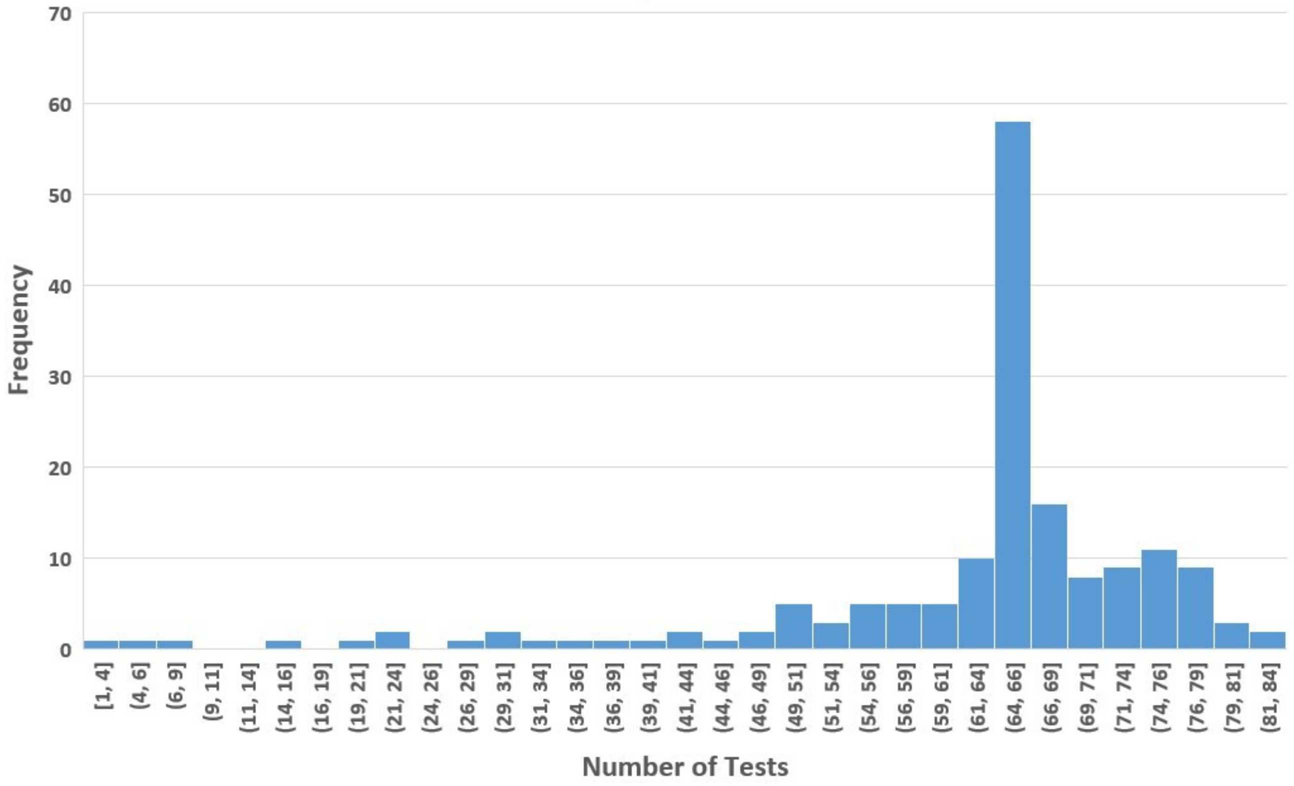


Figure 2. Distribution of the number of SARS-CoV-2 antigen tests per miner. There were approximately 84 opportunities over 267 study days. Participants performed an average of 62.0 (SD 14.4) tests and a median of 65 tests (IQR 68.25-61.75), ranging from 1 to 84 tests (n=169 miners at the intervention site).



Diagnostic Accuracy of the Intervention Test Compared to RT-PCR Gold Standard Testing

Sensitivity, specificity, and predictive values of the intervention (QAT on self-collected NSs at the mine site analyzed by trained miner investigators in the temporary field laboratory) were established through comparison with the gold standard test (ie, qualitative RT-PCR conducted on an NPS collected on the same day by trained health personnel and analyzed in a laboratory). Compared to gold standard testing on a sample of 68 pairs of tests, the intervention test in the field setting showed 100% sensitivity and 100% specificity (Tables 2 and 3).

Table . Comparison of test characteristics between the qualitative rapid antigen test on midnasal swabs in the field and the gold standard test in a laboratory, in a sample of 68 pairs of tests.

Test result	Gold standard test (qualitative RT-PCR[a] on NPSs[b])		
	Disease (positive)	Nondisease (negative)	Total
Antigen test on NS[c]			
Positive	38	0	38
Negative	0	30	30
Total	38	30	68

[a] RT-PCR:reverse transcriptase polymerase chain reaction.

[b] NPS: nasopharyngeal swab.

[c] NS: nasal swab.

Table . Sensitivity, specificity, and predictive values for the qualitative rapid antigen test on midnasal swabs in the field, compared to the gold standard test in a laboratory, in a sample of 68 pairs of tests.

Parameters	Equations	Calculated sensitivity (95% CI)
Sensitivity (%)	True positive/(true positive + false negative)	100 (91-100)
Specificity (%)	True negative/(true negative + false positive)	100 (88-100)
Positive predictive value (%)	True positive/(true positive + false positive)	100[a]
Negative predictive value (%)	true negative/(true negative + false negative)	100[a]

[a] 95% CI was not calculated for these values.

Diagnostic Accuracy of Self-Collected NS Versus NPS Collected by Health Personnel Using RT-qPCR

For this analysis, 46 paired samples (NS and NPS) were collected. Among these swabs, the human RNP gene was not detected in 4 samples: 3 from self-collected samples and 1 NPS collected from a health care professional following RT-qPCR conducted in the Center for Global Health Perkins laboratories. These samples were excluded from further analyses. The RT-qPCR results for the 42 remaining samples are shown in Table 4.

Table . Diagnostic performance of nasal swab versus nasopharyngeal swab using RT-qPCR[a] for SARS-CoV-2 (42 matched samples). Samples with a cycle threshold value of ≥40 were considered negative.

Test result	Nasopharyngeal swab		Concordance rate (%)	Kappa coefficient (95% CI)
	Negative	Positive		
Nasal swab			80.9	0.39 (0.04 - 0.73)
Negative	4	5		
Positive	3	30		

[a] RT-qPCR: quantitative reverse transcriptase polymerase chain reaction.

The concordance rate was 80.9% between the two sampling methods with a κ coefficient of 0.39 (95% CI 0.04 - 0.73; *P*=.01). Simple linear regression was performed to further investigate the relationship between viral loads and the two sample collection methodologies (Figure 3). A modest correlation (*R*=0.38; *P*=.04) for SARS-CoV-2 viral load was detected between 30 paired NP and NS samples. The Ct values and computed viral loads for the discordant samples are shown in Table 5. The NPS virus nucleoprotein gene N1 Ct values for the discordant samples were ≥35, which, after normalizing to the respective RNP sample [18], corresponded to viral loads between log 3 to log 6.5 copies of viral RNA per 1000 cells. This magnitude of viral load was not detectable in self-collected NS samples, suggesting that an NPS was better than an NS for detecting lower viral loads.

Figure 3. Modest correlation between SARS-CoV-2 viral load from nasopharyngeal swabs and nasal swabs. A linear regression model was performed from nasopharyngeal swab and nasal swab paired samples SARS-CoV-2 viral load (log10) to examine the relationship between viral loads and the two sample collection methodologies (n=30; y=0.2522x+7.2466; *R*=0.376). NPS: nasopharyngeal swab; NS: nasal swab.

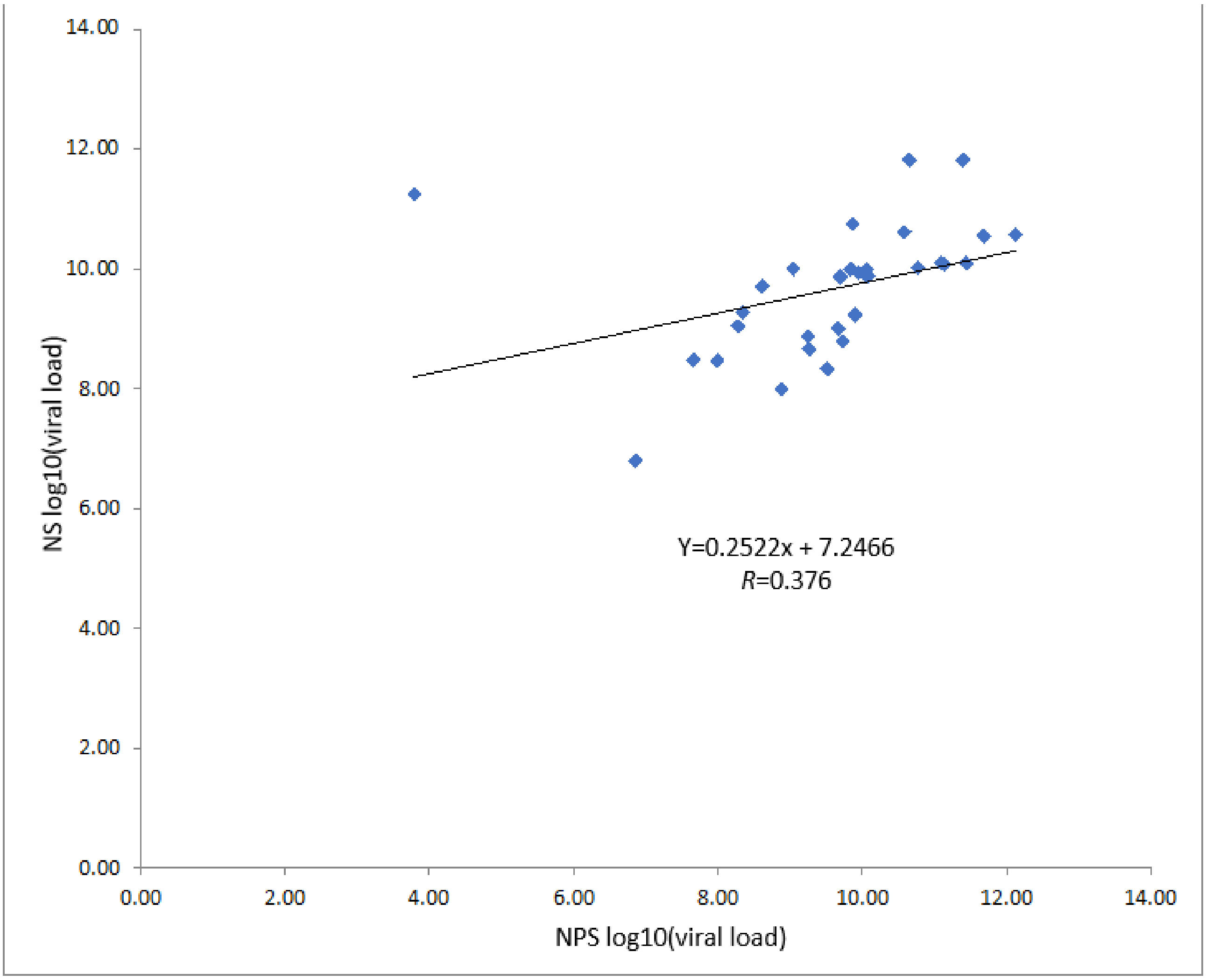


Table . Cycle threshold values and corresponding calculated viral loads of discordant self-collected nasal swab and nasopharyngeal samples collected by trained health care providers using RT-qPCR[a].

Sample ID	NS[b]		NPS[c]		Corresponding viral load (log$_{10}$)	
	N1 Ct[d]	RP[e] Ct	N1[f] Ct	RP Ct	NS	NP
221	>40	37.17	36.82	36.59	0[g]	7.31
263	>40	38.27	37.14	33.96	0	6.17
398	>40	30.85	38.43	25.07	0	3.09
420	>40	34.15	37.07	37.47	0	6.26
738	>40	35.64	37.14	33.01	0	5.78
749	38.52	31.72	>40	33.27	5.74	0
759	35.95	26.96	>40	24.58	4.42	0
967	28.59	31.39	>40	32.15	8.43	0

[a] RT-PCR: reverse transcriptase polymerase chain reaction.

[b] NS: nasal swab.

[c] NPS: nasopharyngeal swab.

[d] Ct: cycle threshold.

[e] RP: human RNase P gene assay control.

[f] N1: virus nucleoprotein gene N1 assay.

[g] Cycle threshold values ≥40 were considered negative.

Seropositivity Testing

Seropositivity patterns of miners enrolled and tested during the initial recruitment period (baseline, February to March 2021) were compared between intervention and control miners using χ^2 tests for prevalent seropositivity and cumulative postbaseline incidence during the study. At baseline, prevalent seropositivity was seen among 18 out of 115 (15.7%) intervention miners and 15 out of 60 (25%) control miners (OR 0.56, 95% CI 0.26-1.19; *P*=.13), and cumulative postbaseline incidence was found in 14 out of 97 (14%) intervention miners compared to 17 out of 45 (38%) control miners (OR 0.28, 95% CI 0.12-0.63; *P*=.002; Table 6). We also used Kaplan–Meier analysis and log-rank tests, which accounted for dropouts, to visualize and analyze whether seropositivity patterns were different for intervention and control miners. Figure 4 shows that seropositivity was greater in control miners than in intervention miners at all time periods (log-rank test *P*=.02). The 12-month cumulative incidence was 48.2% (95% CI 35.4-62.9) among intervention miners and 65.3% (95% CI 51.1-79.2) among control miners when estimated by Kaplan-Meier analysis. When only postbaseline incidence was analyzed, the log-rank test had a lower power and was not significant (*P*=.09) with a 12-month seropositivity of the intervention group of 38.6% (95% CI 24.9-56.4) and 53.8% (95% CI 37.2-72.1) for the control group.

Table . Prevalent and incident seropositivity patterns between the intervention and control miners enrolled and tested during the initial recruitment period (baseline, February to March 2021).

Study timepoint	Intervention (New Mexico–based) mine		Control (Wyoming-based) mine (n=35)	
	New seropositivity, n	Cumulative seropositivity, n/N (%)	New seropositivity, n	Cumulative seropositivity, n/N (%)
Prevalent seropositivity (baseline)	18	18/115 (15.7)	15	15/60 (25)
Postbaseline incident seropositivity[a]				
3 mo	0	0/97 (0)	0	0/45 (0)
6 mo	1	1/97 (1)	6	6/45 (13)
12 mo	13	14/97 (14)	11	17/45 (38)

[a] Postbaseline incident seropositivity was calculated using the total sample size minus the baseline sample size.

Figure 4. Cumulative incidence of SARS-CoV-2 infection in miners in intervention and control mine sites. Data shown are the cumulative incidences of seropositivity (1 – survival estimated by Kaplan-Meier analysis) for New Mexico–based miners from the intervention mine (solid blue line) and Wyoming-based miners from the control mine (dashed green line). Plus symbols indicate miners that were not followed in subsequent testing periods or were negative at the end of follow-up. Numbers at the bottom of the graph are the number of miners in the analysis at each time. NM: New Mexico; S: survival; WY: Wyoming.

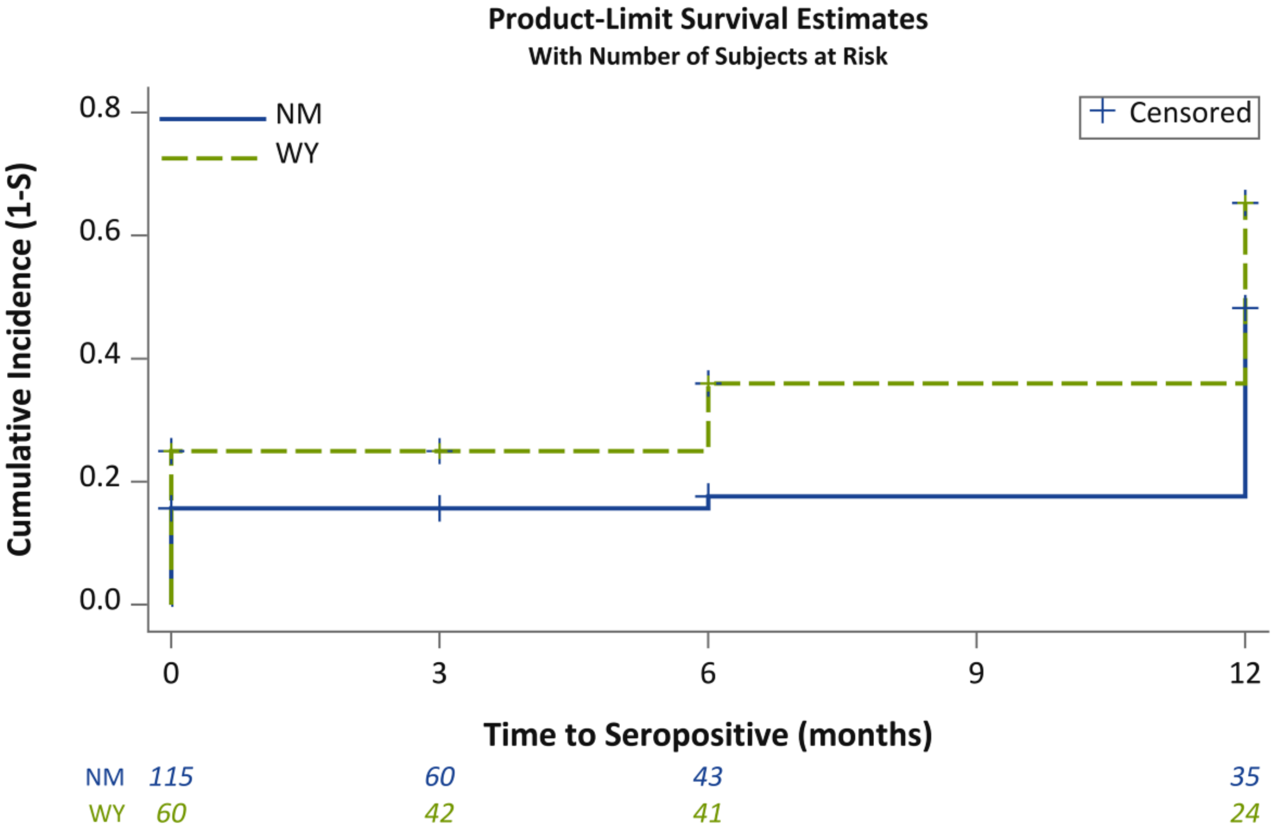


Table S1 in Multimedia Appendix 1 shows unadjusted logistic regression analyses for prevalent and postbaseline incidence, as well as the combined cumulative seropositivity. Combined cumulative seropositivity analyses included miners enrolled at any time during the study. When testing differences in the incidence of seropositivity after baseline, intervention miners (14/97, 14%) had a lower incidence compared to control miners (17/45, 38%; OR 0.28, 95% CI 0.12-0.64; *P*=.002). Incident seropositivity was lower in those who self-reported using masks very often (6/67, 9%) or often (11/39, 28%) compared to those using masks less often (13/35, 37%; *P*=.004; Table S1 in Multimedia Appendix 1). A previous positive COVID-19 test was associated with higher odds of baseline seropositivity (OR 12.72, 95% CI 5.00-32.31; *P*<.001) but not for incident seropositivity (OR 1.34, 95% CI 0.45-4.15; *P*=.59). Age, race and ethnicity, sex, ride-sharing, and COVID-19 vaccination status were not associated with baseline or incident seropositivity. Seropositivity at any time during the study (combined cumulative seropositivity) was lower among intervention miners (OR 0.23, 95% CI 0.12-0.43; *P*<.001), lower among miners who self-reported wearing masks very often (OR 0.19, 95% CI 0.09-0.38; *P*<.001), and lower among miners vaccinated at baseline (OR 0.41, 95% CI 0.19-0.89; *P*=.03). Combined cumulative seropositivity was lower among miners with any nasal testing, which was confounded by the intervention and baseline seropositivity, and odds of cumulative seropositivity did not show a dose-response relationship, with more tests having a lower OR than fewer tests (*P*>.05, data not shown). We also used multivariable logistic models with covariates for age, race and ethnicity, sex, masking frequency, ride-sharing, and COVID-19 vaccination status to assess factors associated with seropositivity (Table S2 in Multimedia Appendix 1). More frequent masking was significantly associated with lower odds of seropositivity for incident and combined cumulative seropositivity measures. Further, more frequent masking was more common in the intervention group, had a strong protective effect on seroprevalence, and confounded the protective effect of the intervention on incident seropositivity and combined cumulative seropositivity.

Discussion

Principal Results

Workers in essential sectors have experienced higher COVID-19 infection and mortality rates throughout the pandemic [19]. Much attention has been given to the significant occupational risk of infection among health care workers, but essential work extends beyond health care. Understanding the feasibility and effectiveness of implementing point-of-care testing in essential non–health care professions is crucial for preventing or slowing the transmission of infectious diseases.

Miners, deemed essential workers, face unique challenges that make them more vulnerable and susceptible to COVID-19 [20]. Mining environments often have decreased ventilation and limited availability for physical distancing, remote work, or flexible schedules. Miners are also exposed to higher levels of air pollutant particulates, which may increase their susceptibility to respiratory infections or worsen health outcomes from respiratory infections [21]. Moreover, mine sites are situated in rural areas with limited access to health care and testing facilities. Despite these disproportionate challenges, limited guidelines are available for maintaining mining and other non–health care operations effectively and safely during an infectious disease outbreak. Thus, the objective of this study was to investigate the feasibility and effectiveness of using frequent point-of-care molecular workplace surveillance as an intervention strategy to prevent the spread of SARS-CoV-2 at rural mining sites.

The intervention had a high mean acceptance rate of 96.4% (11,413/11,842), reflecting the involvement of mine safety personnel in the design and execution of the study. Additionally, frequent point-of-care workplace surveillance was associated with a lower cumulative SARS-CoV-2 seropositivity rate in intervention miners compared to control miners (P=.002). This study further examined the reliability of self-collected NS versus NPS samples collected by trained health care personnel to detect SARS-CoV-2 using RT-qPCR. The Cohen κ coefficient of 0.39 indicates fair agreement between the two collection methods. However, we observed that NPS samples that tested positive for SARS-CoV-2 by RT-qPCR but had lower viral loads were frequently associated with negative results when tested using NSs instead. While less reliable when compared to NPSs for detecting low levels of SARS-CoV-2, self-collected NS can provide diagnostic performance comparable to NPSs in participants with higher viral loads. As such, NS is a suitable and practical option for intervention test use, especially in resource-limited settings where point-of-care QAT is more accessible, cost-effective, and faster than the gold standard test.

Limitations

A limitation of this study is that protective behaviors influence COVID-19 health outcomes, including confounding seropositivity rates [22]. While the CDC has encouraged the public use of face coverings during high spikes in community infection rates and vaccination against COVID-19, the use of these protective measures can vary between mine sites based on individual and community factors [23]. We observed that intervention miners were more likely to wear face coverings in public and more willing to receive a vaccine against COVID-19. This difference may reflect partisan variation in official messaging related to COVID-19 protective behaviors to the public between Democratic-governed (New Mexico) and Republican-governed (Wyoming) states [24].

It is also important to note that our study did not involve randomization, so the unmatched distribution of covariates between the two cohorts may result in confounding bias which may have influenced our findings. However, several differences related to demographics and social determinants of health between the mine sites placed the intervention cohort at a higher risk for SARS-CoV-2 infection than the control miners. For example, intervention miners were more likely to be male, which is associated with an increased risk for SARS-CoV-2 infection [25]. Additionally, intervention miners were more likely to be racial and ethnic minorities. Given the disparities observed in minority populations during the pandemic [26,27], this selection bias may have placed the intervention miners at an increased inherent risk for infection and associated complications, which may have biased the protective effect of the intervention toward the null value. Intervention miners were more likely to ride-share and report previous exposure to woodsmoke, a risk factor for increased respiratory disease susceptibility [28]. Moreover, data previously published on our New Mexico cohort indicate that minority miners are at greater risk for developing lung disease than non-Hispanic White miners [29].

Intervention miners also had a lower mean income and educational status than control miners. A study analyzing the joint effects of socioeconomic position, race and ethnicity, and sex on mortality in the first year of the COVID-19 pandemic revealed that COVID-19 mortality was 5 times higher in individuals in low versus high socioeconomic positions [30]. The above-mentioned biases may impact the study's conclusions, making it difficult to determine whether the observed effects are genuinely due to the intervention or other factors. Although randomization is a key method to mitigate these biases and enhance the credibility of clinical research, community investigators did not consider this study design feasible during the pandemic. Challenges encountered by the study included the lack of easily available intervention tests and community research personnel in the field during the early stage of the pandemic, which limited study recruitment and expansion to other mine sites. The recruitment of a sample size that was lower than planned limited the power of the study. Additionally, the remote location of the study sites increased study costs, and the lack of cellphone and internet connectivity required adaptive strategies by the investigators. Although our intervention can be replicated in other geographic areas and workers in other public and private essential service settings, scaling the intervention would require careful consideration of these unique challenges and a study of its cost-effectiveness.

Comparison With Prior Work

Since the onset of the pandemic, several papers have been published supporting routine testing to limit the spread of SARS-CoV-2 in health care settings [31-33]. However, studies evaluating frequent point-of-care surveillance, particularly in sectors outside of health care, are still limited. One study, which included a variety of essential workers, including health care workers, found that regular testing of all key workers was associated with reduced transmission of approximately 67 individuals per 1000 tests, with high accuracy of testing (87.1% - 99.9%) [34]. In this study, researchers used RT-PCR rather than QAT, likely due to the limited availability of these tests at the study's onset. However, a later study evaluating the sensitivity and specificity of the Boson Rapid SARS-CoV-2 rapid antigen detection test revealed an overall sensitivity of 63.04% for anterior NSs and 73.33% for NPSs, with 100% test specificity [35]. A similar study, also showing 100% specificity, compared the Medomics SARS-CoV-2 antigen test device to gold standard testing and reported that rapid antigen detection tests using self-collected anterior NSs proved to be as sensitive as and more tolerable than professionally collected NPSs for Ct values up to 30, determined by RT-PCR [36]. Supporting these results and our own, additional studies have emerged revealing that NS self-sampling yields comparable results to NPS sampling using both RT-PCR [37,38] and QAT for analysis [39]. Studies have established the effectiveness of other field interventions, such as PCR-based wastewater surveillance of SARS-CoV-2 [40]. However, such samples are difficult to collect and analyze in rural areas and provide community-level and not individual-level surveillance data, limiting their impact.

Conclusion

Our data firmly establish SARS-CoV-2 QAT on self-collected NSs as a feasible alternative to laboratory-based RT-PCR on NPSs for preventing the spread of COVID-19 at essential workplaces within vulnerable communities at an increased risk of being adversely affected by the COVID-19 pandemic because of their occupation, demographics, and rural and remote locations. Additionally, our findings demonstrate the excellent performance of the intervention test in a real-world setting compared to the gold standard. The study findings support developing and implementing policy measures for workplace surveillance and other protective interventions against the spread of SARS-CoV-2 among rural non–health care essential workers. The implications of our findings extend beyond this study, providing valuable insights for designing interventions to maintain employees’ safety at other essential workplaces during an infectious disease outbreak of respiratory pathogens.

Acknowledgments

The research reported in this publication was supported in part by the National Institutes of Health (NIH) under Award Number 3U01GM132175. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Authors' Contributions

AS, WCJ, NS, AP, NE, IH, LSC, SL, OBM, and DJP made substantial contributions to the conception or design of the work. AS, WCJ, AVY, TA, QC, IH, and DJP contributed to data acquisition, and AS, XWS, OBM, QC, IH, and DJP contributed to data analysis. AS, WCJ, XWS, NS, AP, NE, IH, LSC, SL, OBM, and DJP contributed to the interpretation of the data. All authors contributed to drafting the work or reviewing it critically for important intellectual content and gave final approval of the version to be published. All authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. We would also like to thank Hesch Consulting LLC for their support during the preparation of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Univariate and multivariable analysis of the association between prevalence seropositivity, postbaseline incident seropositivity, combined cumulative seropositivity, and miner characteristics at baseline.

[DOCX File, 52 KB - publichealth_v11i1e59845_app1.docx]

References

1. Current population survey extracts. Economic Policy Institute. 2020. URL: https://microdata.epi.org [accessed 2023-08-06]
2. Dzinamarira T, Murewanhema G, Mhango M, et al. COVID-19 prevalence among healthcare workers. A systematic review and meta-analysis. Int J Environ Res Public Health 2021 Dec 23;19(1):146. [doi: 10.3390/ijerph19010146] [Medline: 35010412]
3. Gómez-Ochoa SA, Franco OH, Rojas LZ, et al. COVID-19 in health-care workers: a living systematic review and meta-analysis of prevalence, risk factors, clinical characteristics, and outcomes. Am J Epidemiol 2021 Jan 4;190(1):161-175. [doi: 10.1093/aje/kwaa191] [Medline: 32870978]
4. Manai MV, Shaholli D, La Torre G. Contact tracing as an essential prevention tool for the spreading of COVID-19 among healthcare workers. Clin Ter 2022;173(5):396-397. [doi: 10.7417/CT.2022.2452] [Medline: 36155722]
5. Al-Oraibi A, Naidu JS, Chaka A, et al. Prevalence of long COVID-19 among healthcare workers: a systematic review and meta-analysis protocol. BMJ Open 2022 Dec 6;12(12):e065234. [doi: 10.1136/bmjopen-2022-065234] [Medline: 36600349]
6. Song H, McKenna R, Chen AT, David G, Smith-McLallen A. The impact of the non-essential business closure policy on Covid-19 infection rates. Int J Health Econ Manag 2021 Dec;21(4):387-426. [doi: 10.1007/s10754-021-09302-9] [Medline: 33792808]
7. Naleway AL, Grant L, Caban-Martinez AJ, et al. Incidence of SARS-CoV-2 infection among COVID-19 vaccinated and unvaccinated healthcare personnel, first responders, and other essential and frontline workers: eight US locations, January-September 2021. Influenza Other Respir Viruses 2022 May;16(3):585-593. [doi: 10.1111/irv.12956] [Medline: 35023288]

8. Liu Y, Yang B, Liu L, Jilili M, Yang A. Occupational characteristics in the outbreak of the COVID-19 delta variant in Nanjing, China: rethinking the occupational health and safety vulnerability of essential workers. Int J Environ Res Public Health 2021 Oct 13;18(20):10734. [doi: 10.3390/ijerph182010734] [Medline: 34682477]
9. Lutrick K, Ellingson KD, Baccam Z, et al. COVID-19 infection, reinfection, and vaccine effectiveness in a prospective cohort of Arizona frontline/essential workers: the AZ HEROES research protocol. JMIR Res Protoc 2021 May 26;10(6):e28925. [doi: 10.2196/28925] [Medline: 34057904]
10. Alsharef A, Banerjee S, Uddin SMJ, Albert A, Jaselskis E. Early impacts of the COVID-19 pandemic on the United States construction industry. Int J Environ Res Public Health 2021 Feb 6;18(4):1559. [doi: 10.3390/ijerph18041559] [Medline: 33562127]
11. Guidance for SARS-Cov-2 rapid testing performed in point-of-care settings. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/covid/php/lab/point-of-care-testing.html [accessed 2024-01-17]
12. Larremore DB, Wilder B, Lester E, et al. Test sensitivity is secondary to frequency and turnaround time for COVID-19 surveillance. medRxiv 2020 Sep 8:2020.06.22.20136309. [doi: 10.1101/2020.06.22.20136309] [Medline: 32607516]
13. Paltiel AD, Zheng A, Walensky RP. Assessment of SARS-CoV-2 screening strategies to permit the safe reopening of college campuses in the United States. JAMA Netw Open 2020 Jul 1;3(7):e2016818. [doi: 10.1001/jamanetworkopen.2020.16818] [Medline: 32735339]
14. Connor BA, Rogova M, Garcia J, et al. Comparative effectiveness of single vs repeated rapid SARS-CoV-2 antigen testing among asymptomatic individuals in a workplace setting. JAMA Netw Open 2022 Mar 1;5(3):e223073. [doi: 10.1001/jamanetworkopen.2022.3073] [Medline: 35302635]
15. The Lancet. The plight of essential workers during the COVID-19 pandemic. Lancet 2020 May 23;395(10237):1587. [doi: 10.1016/S0140-6736(20)31200-9] [Medline: 32446399]
16. McCulloch DJ, Kim AE, Wilcox NC, et al. Comparison of unsupervised home self-collected midnasal swabs with clinician-collected nasopharyngeal swabs for detection of SARS-CoV-2 infection. JAMA Netw Open 2020 Jul 1;3(7):e2016382. [doi: 10.1001/jamanetworkopen.2020.16382] [Medline: 32697321]
17. Lim HJ, Baek YH, Park MY, et al. Performance analysis of self-collected nasal and oral swabs for detection of SARS-CoV-2. Diagnostics (Basel) 2022 Sep 21;12(10):2279. [doi: 10.3390/diagnostics12102279] [Medline: 36291968]
18. Perkins DJ, Yingling AV, Cheng Q, et al. Elevated SARS-CoV-2 in peripheral blood and increased COVID-19 severity in American Indians/Alaska Natives. Exp Biol Med (Maywood) 2022 Jul;247(14):1253-1263. [doi: 10.1177/15353702221091180] [Medline: 35491994]
19. Chen YH, Riley AR, Duchowny KA, et al. COVID-19 mortality and excess mortality among working-age residents in California, USA, by occupational sector: a longitudinal cohort analysis of mortality surveillance data. Lancet Public Health 2022 Sep;7(9):e744-e753. [doi: 10.1016/S2468-2667(22)00191-8] [Medline: 36057273]
20. Evans K, Lerch S, Boyce TW, et al. An innovative approach to enhancing access to medical screening for miners using a mobile clinic with telemedicine capability. J Health Care Poor Underserved 2016;27(4A):62-72. [doi: 10.1353/hpu.2016.0182] [Medline: 27818414]
21. Laney AS, Weissman DN. Respiratory diseases caused by coal mine dust. J Occup Environ Med 2014 Oct;56 Suppl 10(10):S18-S22. [doi: 10.1097/JOM.0000000000000260] [Medline: 25285970]
22. Green H, Fernandez R, MacPhail C. The social determinants of health and health outcomes among adults during the COVID-19 pandemic: a systematic review. Public Health Nurs 2021 Nov;38(6):942-952. [doi: 10.1111/phn.12959] [Medline: 34403525]
23. Fisher KA, Barile JP, Guerin RJ, et al. Factors associated with cloth face covering use among adults during the COVID-19 pandemic - United States, April and May 2020. MMWR Morb Mortal Wkly Rep 2020 Jul 17;69(28):933-937. [doi: 10.15585/mmwr.mm6928e3] [Medline: 32673303]
24. Golos AM, Hopkins DJ, Bhanot SP, Buttenheim AM. Partisanship, messaging, and the COVID-19 vaccine: evidence from survey experiments. Am J Health Promot 2022 May;36(4):602-611. [doi: 10.1177/08901171211049241] [Medline: 35232232]
25. Bwire GM. Coronavirus: why men are more vulnerable to Covid-19 than women? SN Compr Clin Med 2020;2(7):874-876. [doi: 10.1007/s42399-020-00341-w] [Medline: 32838138]
26. Feldman JM, Bassett MT. Variation in COVID-19 mortality in the US by race and ehnicity and educational attainment. JAMA Netw Open 2021 Nov 1;4(11):e2135967. [doi: 10.1001/jamanetworkopen.2021.35967] [Medline: 34812846]
27. Kakol M, Upson D, Sood A. Susceptibility of southwestern American Indian tribes to coronavirus disease 2019 (COVID-19). J Rural Health 2021 Jan;37(1):197-199. [doi: 10.1111/jrh.12451] [Medline: 32304251]
28. Assad NA, Kapoor V, Sood A. Biomass smoke exposure and chronic lung disease. Curr Opin Pulm Med 2016 Mar;22(2):150-157. [doi: 10.1097/MCP.0000000000000246] [Medline: 26814722]
29. Kocher E, Rendon KJ, Kesler D, et al. Uranium workers demonstrate lower lobe predominant irregular pneumoconiotic opacities on chest radiographs. J Health Care Poor Underserved 2016;27(4A):116-127. [doi: 10.1353/hpu.2016.0193] [Medline: 27818418]
30. Pathak EB, Menard JM, Garcia RB, Salemi JL. Joint effects of socioeconomic position, race/ethnicity, and gender on COVID-19 mortality among working-age adults in the United States. Int J Environ Res Public Health 2022 Apr 30;19(9):5479. [doi: 10.3390/ijerph19095479] [Medline: 35564872]

31. Maya S, Padda G, Close V, et al. Optimal strategies to screen health care workers for COVID-19 in the US: a cost-effectiveness analysis. Cost Eff Resour Alloc 2022 Jan 15;20(1):2. [doi: 10.1186/s12962-021-00336-x] [Medline: 35033100]
32. Sanchez-Taltavull D, Castelo-Szekely V, Murugan S, et al. Regular testing of asymptomatic healthcare workers identifies cost-efficient SARS-CoV-2 preventive measures. PLoS ONE 2021;16(11):e0258700. [doi: 10.1371/journal.pone.0258700] [Medline: 34739484]
33. Larese Filon F, Rui F, Ronchese F, De Michieli P, Negro C. Incidence of COVID-19 infection in hospital workers from March 1, 2020 to May 31, 2021 routinely tested, before and after vaccination with BNT162B2. Sci Rep 2022 Feb 15;12(1):2533. [doi: 10.1038/s41598-021-04665-y] [Medline: 35169127]
34. Sandmann FG, White PJ, Ramsay M, Jit M. Optimizing benefits of testing key workers for infection with SARS-CoV-2: a mathematical modeling analysis. Clin Infect Dis 2020 Dec 15;71(12):3196-3203. [doi: 10.1093/cid/ciaa901] [Medline: 32634823]
35. Leventopoulos M, Michou V, Papadimitropoulos M, et al. Evaluation of the Boson rapid Ag test vs RT-PCR for use as a self-testing platform. Diagn Microbiol Infect Dis 2022 Nov;104(3):115786. [doi: 10.1016/j.diagmicrobio.2022.115786] [Medline: 35998553]
36. Wölfl-Duchek M, Bergmann F, Jorda A, et al. Sensitivity and specificity of SARS-CoV-2 rapid antigen detection tests using oral, anterior nasal, and nasopharyngeal swabs: a diagnostic accuracy study. Microbiol Spectr 2022 Feb 23;10(1):e0202921. [doi: 10.1128/spectrum.02029-21] [Medline: 35107327]
37. Labhardt ND, González Fernández L, Katende B, et al. Head-to-head comparison of nasal and nasopharyngeal sampling using SARS-CoV-2 rapid antigen testing in Lesotho. PLoS ONE 2023;18(3):e0278653. [doi: 10.1371/journal.pone.0278653] [Medline: 36862684]
38. Nahar K, Begum MN, Tony SR, et al. Nasal swab as an alternative specimen for the detection of severe acute respiratory syndrome coronavirus 2. Health Sci Rep 2023 Apr;6(4):e1213. [doi: 10.1002/hsr2.1213] [Medline: 37077182]
39. Klein JAF, Krüger LJ, Tobian F, et al. Head-to-head performance comparison of self-collected nasal versus professional-collected nasopharyngeal swab for a WHO-listed SARS-CoV-2 antigen-detecting rapid diagnostic test. Med Microbiol Immunol 2021 Aug;210(4):181-186. [doi: 10.1007/s00430-021-00710-9] [Medline: 34028625]
40. Toledo DM, Robbins AA, Gallagher TL, et al. Wastewater-based SARS-CoV-2 surveillance in northern New England. Microbiol Spectr 2022 Apr 27;10(2):e0220721. [doi: 10.1128/spectrum.02207-21] [Medline: 35412387]

Abbreviations

CDC: Centers for Disease Control and Prevention
Ct: cycle threshold
NPS: nasopharyngeal swab
NS: nasal swab
OR: odds ratio
QAT: qualitative SARS-CoV-2 rapid antigen testing
RNP: RNAse P
RT-PCR: reverse transcriptase polymerase chain reaction
RT-qPCR: quantitative reverse transcriptase polymerase chain reaction

Edited by A Mavragani; submitted 23.04.24; peer-reviewed by D Laborde, X Sun; revised version received 19.11.24; accepted 20.11.24; published 27.01.25.

Please cite as:
Sood A, “Cotton” Jarrell W, Shore XW, Sosa N, Parada A, Edwardson N, Yingling AV, Amirkabirian T, Cheng Q, Hurwitz I, Cook LS, Leng S, Myers OB, Perkins DJ
Effectiveness of Frequent Point-of-Care Molecular COVID-19 Surveillance in a Rural Workplace: Nonrandomized Controlled Clinical Trial Among Miners
JMIR Public Health Surveill 2025;11:e59845
URL: https://publichealth.jmir.org/2025/1/e59845
doi:10.2196/59845

Original Paper

Updated Surveillance Metrics and History of the COVID-19 Pandemic (2020-2023) in East Asia and the Pacific Region: Longitudinal Trend Analysis

Alexander L Lundberg[1,2], PhD, MSc; Alan G Soetikno[3], BS; Scott A Wu[3], BA; Egon Ozer[4,5], MD, PhD; Sarah B Welch[1], MPH; Yingxuan Liu[1], MA; Claudia Hawkins[4,6], MD, MPH; Maryann Mason[1,2], PhD; Robert Murphy[4,7], MD; Robert J Havey[7,8], MD; Charles B Moss[9], PhD; Chad J Achenbach[4,7], MD, MPH; Lori Ann Post[1,2], PhD

[1]Buehler Center for Health Policy and Economics, Robert J Havey, MD Institute for Global Health, Northwestern University, Chicago, IL, United States

[2]Department of Emergency Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[3]Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[4]Department of Medicine, Division of Infectious Diseases, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[5]Center for Pathogen Genomics and Microbial Evolution, Robert J. Havey, MD Institute for Global Health, Northwestern University, Chicago, IL, United States

[6]Center for Global Communicable and Emerging Infectious Diseases, Robert J Havey, MD Institute for Global Health, Northwestern University, Chicago, IL, United States

[7]Robert J Havey, MD Institute for Global Health, Northwestern University, Chicago, IL, United States

[8]Department of Medicine, General Internal Medicine and Geriatrics, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[9]Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL, United States

Corresponding Author:
Lori Ann Post, PhD
Buehler Center for Health Policy and Economics
Robert J Havey, MD Institute for Global Health
Northwestern University
420 E Superior St
Chicago, IL, 60611
United States
Phone: 1 3125031706
Email: lori.post@northwestern.edu

Abstract

Background: This study updates the COVID-19 pandemic surveillance in East Asia and the Pacific region that we first conducted in 2020 with 2 additional years of data for the region.

Objective: First, we aimed to measure whether there was an expansion or contraction of the pandemic in East Asia and the Pacific region when the World Health Organization (WHO) declared the end of the COVID-19 *public health emergency of international concern* on May 5, 2023. Second, we used dynamic and genomic surveillance methods to describe the dynamic history of the pandemic in the region and situate the window of the WHO declaration within the broader history. Finally, we aimed to provide historical context for the course of the pandemic in East Asia and the Pacific region.

Methods: In addition to updates of traditional surveillance data and dynamic panel estimates from the original study, this study used data on sequenced SARS-CoV-2 variants from the Global Initiative on Sharing All Influenza Data to identify the appearance and duration of variants of concern. We used Nextclade nomenclature to collect clade designations from sequences and Pangolin nomenclature for lineage designations of SARS-CoV-2. Finally, we conducted a 1-sided *t* test to determine whether the regional weekly speed was greater than an outbreak threshold of 10. We ran the test iteratively with 6 months of data across the sample period.

Results: Several countries in East Asia and the Pacific region had COVID-19 transmission rates above an outbreak threshold at the point of the WHO declaration (Brunei, New Zealand, Australia, and South Korea). However, the regional transmission rate had remained below the outbreak threshold for 4 months. In the rolling 6-month window *t* test for regional outbreak status, the final *P* value ≤.10 implies a rejection of the null hypothesis (at the α=.10 level) that the region as a whole was not in an outbreak for the period from November 5, 2022, to May 5, 2023. From January 2022 onward, nearly every sequenced SARS-CoV-2 specimen in the region was identified as the Omicron variant.

Conclusions: While COVID-19 continued to circulate in East Asia and the Pacific region, transmission rates had fallen below outbreak status by the time of the WHO declaration. Compared to other global regions, East Asia and the Pacific region had the latest outbreaks driven by the Omicron variant. COVID-19 appears to be endemic in the region, no longer reaching the threshold for a pandemic definition. However, the late outbreaks raise uncertainty about whether the pandemic was truly over in the region at the time of the WHO declaration.

(JMIR Public Health Surveill 2025;11:e53214) doi:10.2196/53214

KEYWORDS

SARS-CoV-2; COVID-19; East Asia; Pacific; American Samoa; Australia; Brunei Darussalam; Cambodia; China; Fiji; French Polynesia; Guam; Hong Kong; Indonesia; Japan; Kiribati; People’s Democratic Republic of Korea; Republic of Korea; Lao People’s Democratic Republic; Macao; Malaysia; Marshall Islands; Federated States of Micronesia; Mongolia; Myanmar; Nauru, New Caledonia; New Zealand; Northern Mariana Islands; Palau; Papua New Guinea; Philippines; Samoa; Singapore; Solomon Islands; Thailand; Timor-Leste; Tonga; Tuvalu; Vanuatu; Vietnam; pandemic; surveillance; public health; speed; acceleration; deceleration; jerk; dynamic panel; generalized method of moments; Arellano-Bond; 7-day lag

Introduction

Background

COVID-19, the disease caused by SARS-CoV-2, was first detected in Wuhan, China, in the fall of 2019 [1-5]. Our research team conducted an analysis of the pandemic in East Asia and the Pacific region one year into the pandemic [6]. This study provides 2 additional years of updated surveillance and analysis for the region.

We adopted the World Bank’s definition of East Asia and the Pacific region, which is based on economic development and geographical proximity, encompassing American Samoa, Australia, Brunei Darussalam, Cambodia, China, Fiji, French Polynesia, Guam, Hong Kong, Indonesia, Japan, Kiribati, People’s Democratic Republic of Korea, Republic of Korea, Lao People’s Democratic Republic, Macao, Malaysia, Marshall Islands, Federated States of Micronesia, Mongolia, Myanmar, Nauru, New Caledonia, New Zealand, Northern Mariana Islands, Palau, Papua New Guinea, the Philippines, Samoa, Singapore, Solomon Islands, Thailand, Timor-Leste, Tonga, Tuvalu, Vanuatu, and Vietnam [7]. North Korea neither collects nor reports COVID-19 caseloads and deaths.

The World Health Organization (WHO) and Director-General Ghebreyesus declared the end of *COVID-19 as a public health emergency of international concern* on May 5, 2023 [8-10], based on the recommendation of the COVID-19 Emergency Committee [10]. We compared the progression of the pandemic before and after this declaration.

Empirical Definition of Pandemic Versus Epidemic Versus Outbreak Versus Endemic

Epidemiological terms, such as pandemic, epidemic, outbreak, and endemic, are used to describe the occurrence and spread of disease [11,12]. The distinctions between these terms lie in their scope, geographic extent, and severity. An epidemic refers to a sudden increase in the number of disease cases in a specific population or region. If the epidemic spreads across several countries or continents, it becomes a pandemic. An outbreak, on the other hand, describes a sudden increase in a concentrated setting, usually involving a more limited geographic area than an epidemic. Endemic refers to the constant presence of a disease in a particular geographic region or population, with no sudden increases in case volume [13,14]. Field epidemiology defines these terms based on transmission metrics and geographical distribution and uses transmission metrics and geographical distribution to differentiate the nuances of these terms. Although public health surveillance data may not capture all cases, they are the best proxy for monitoring and tracking disease spread within a population in near real time. Surveillance is crucial for the timely response to health threats [15].

Traditional Surveillance Versus Enhanced Surveillance

Public health surveillance is the “ongoing, systematic collection, analysis, and interpretation of health-related data essential to planning and evaluation of public health practice” [16]. Not only does surveillance explain the burden of death and disease transmission but also it generates research questions and guides investigators on topics that require further investigation [17-31]. Surveillance allows us to compare the burden of disease between geographical regions and to understand which regions are most impacted. The impact can be measured through per capita rates of how many people contract a disease, how many die, and affiliated costs.

However, traditional surveillance carries several limitations that this study has addressed. Traditional surveillance provides a snapshot of past events [17-31], meaning it is static and only reflects historical data. In the middle of a burgeoning pandemic, policy makers and public health practitioners also need to understand what is about to happen. Is an outbreak increasing? Will growth switch from linear to exponential? Are more people dying from that particular condition in one place than another? To inform health policy and practice, knowledge of what is about to happen is often more valuable than knowledge of what did happen. To that end, we have developed enhanced surveillance metrics that reflect the dynamics of a pandemic and inform imminent growth—most importantly, where along the epidemiological outbreak curve a particular region is situated. In addition, we included dynamic metrics about the speed of the pandemic at the national, regional, and global level. We measured how acceleration of speed one week compared to the prior week, as well as how novel infections in a prior week predicted new cases the following week. We can think of the latter measure as the echoing forward of cases. These metrics were tested and validated in prior research [6,32-42].

For the purpose of this study, standard surveillance metrics explain what has already happened in East Asia and the Pacific region, while enhanced surveillance metrics provide insights into future trends or the current position of a country along an epidemiological curve. We used both types of metrics to analyze the possible end to the pandemic in East Asia and the Pacific region.

Objectives

This study has 3 objectives. First, we aimed to measure whether there was an expansion or contraction in the pandemic in East Asia and the Pacific region at the time the WHO declared the end of the *COVID-19 pandemic as a public health emergency of international concern* on May 5, 2023 [43-51]. At both the region and country level, we used advanced surveillance and analytical techniques to describe the status of the pandemic in a 2-week window around the WHO declaration. From a public health perspective, we needed to know whether the rate of new COVID-19 cases was increasing, decreasing, or stable from week to week, and if any changes in the transmission rate indicated an acceleration or deceleration of the pandemic. Statistical insignificance is significant—it can signal the epidemiological *end* to the pandemic if the rate of new cases is 0 (or very low) and stable, meaning the number of new cases is neither accelerating nor decelerating.

Second, we used dynamic and genomic surveillance methods to describe the history of the pandemic in the region and situate the time window around the WHO declaration within the broader history. We included the ratio of COVID-19 deaths to the number of transmissions as a proxy for the mortality risk from infection at the population level. We also included a historical record of genomic surveillance from sequenced viral specimens to identify the appearance and spread of variants of concern in the region.

Third, we aimed to provide historical context for the course of the pandemic in East Asia and the Pacific region. We have addressed several questions. How did countries respond to the pandemic? How did the region fare in terms of disease burden? And what social, economic, and political factors shaped the course of COVID-19 in the region? This context can provide important lessons for disease prevention and mitigation in future pandemics.

Methods

Overview

We conducted trend analyses with longitudinal COVID-19 data from Our World in Data (OWID) [52]. This study provides updates of traditional surveillance data and dynamic panel estimates from the original work by Post et al [6,40,41,53-55]. For East Asia and the Pacific region, the data comprised an unbalanced panel of 36 countries and territories, running from August 14, 2020, to May 12, 2023. Because several countries around the world switched from daily to weekly reports at various points in 2023, we used a cubic spline to interpolate daily new cases and deaths if any country had 4 consecutive periods of nonzero new cases interspersed by 6 days of 0 new cases [56-58].

To identify the appearance and duration of variants of concern, we also used data on sequenced SARS-CoV-2 variants from the Global Initiative on Sharing All Influenza Data (GISAID), which is an effective and trusted web-based resource for sharing genetic, clinical, and epidemiological COVID-19 data [59-62]. We used Nextclade nomenclature [63] to collect clade designations from sequences and Pangolin nomenclature for lineage designations of SARS-CoV-2 [64,65]. Metadata for the study period were collected on June 22, 2023. To avoid low frequency or potentially erroneous samples, the dataset was further filtered to exclude months with fewer than 100 available samples, variant groups with fewer than 5 samples in a month, and variant groups representing <0.5% of the total samples in a month. The final dataset consisted of 184,386 total samples available on GISAID [59-62].

Traditional surveillance metrics include the *speed* of transmission or the rate of new COVID-19 cases per 100,000 population. Novel metrics include *acceleration*, *jerk*, and *1- and 7-day persistence* estimates. *Acceleration* is the change in speed from one time unit to the next. Acceleration can measure whether transmission rates are rising, falling, or stable. *Jerk* is the change in acceleration from one time unit to the next. The term *jerk* is adopted from physics nomenclature, as a large jerk can signal explosive growth in transmissions. The *1- and 7-day persistence* measures provide the impact of the 1- and 7-day lag of speed on current speed. For another interpretation, these estimates capture how COVID-19 cases echo forward to cases either 1 or 7 days later. They are coefficient estimates from an Arellano-Bond dynamic panel data model [66]:

in which the dependent variable is speed, the independent variables include weekend and recent week indicators, α_i is a country fixed effect, and u_{it} is the idiosyncratic error term. Advantages of the Arellano-Bond estimates include a correction for time-invariant omitted variables and the ability to assess both the predictive ability of the model and validity of its specification [41].

We analyzed the potential *statistical end* to the pandemic with a 1-sided t test for whether the mean of speed was equal to or greater than the outbreak threshold of 10. We ran the test on a rolling 6-month window over weekly speed for the region, and we plotted the P values from the test over time. The test can assess whether the region as a whole was experiencing an outbreak over a defined time window, during which the speed may sometimes exceed the outbreak threshold and at other times remain below it. All statistical analyses were conducted in R (version 4.2.1; R Foundation for Statistical Computing) with the *plm* package (version 2.6-2) [53,54].

Ethical Considerations

All data used in this study are publicly available and contain no identifiable, private information. Therefore, the study is deemed exempt research with human subjects as defined by the US Government Code of Federal Regulations 45CFR46 and so was not submitted the Northwestern University Institutional Review Board. However, the authors note that anonymized COVID-19 data surveillance systems can generate local and global ethical questions beyond the scope of this study [67].

Results

Tables 1-6 present traditional (ie, static) and novel (ie, dynamic) COVID-19 surveillance metrics with OWID transmission and death data for East Asia and the Pacific region. These results focus on the 2-week period around the time when the WHO declared COVID-19 was no longer a public health emergency of international concern. The week before the declaration was April 28, 2023, and the week after was May 5, 2023.

Table 1. Arellano-Bond dynamic panel data estimates of the number of daily COVID-19 infections reported by country in East Asia and the Pacific region from April 28 to May 12, 2023a,b.

Variables	Values	*P* values[c]
1-day lag coefficient	0.171	.46
7-day lag coefficient	0.877	<.001
Shift parameter, week of April 28	−0.488	<.001
Shift parameter, week of May 5	0.491	<.001
Weekend effect	−1.215	.28

[a] Wald test: χ^2_6=730,275, *P*≤.001.

[b] Sargan: χ^2_{540}=24, *P*=.99.

[c] Contains estimates from the model in equation (1).

Figure 1 plots the novel surveillance metrics over time, from the start of the pandemic up to the WHO declaration in May 2023. Figure 2 uses GISAID data on sequenced SARS-CoV-2 variants to present the proportion of variants of concern returned among specimens over a similar period. Figure 3 returns to regional transmission data to present the results of a rolling 6-month window *t* test for whether the region was in a state of outbreak over each period. Finally, Figure 4 presents a qualitative timeline of key events in the COVID-19 pandemic in East Asia and the Pacific region.

Table 1 presents the dynamic panel estimates for the most recent time window. The Wald test for the regression was significant (*P*<.001), and the Sargan test failed to reject the validity of the overidentification restrictions (*P*=.99). The 1-day lag coefficient was not statistically significant (*P*=.46), but the 7-day lag coefficient was (*P*<.001), suggesting a cluster effect in which cases on a given day impact cases 7 days later. The shift parameter for the week of April 28 was negative (−0.488), while the more recent shift parameter for the week of May 5 was positive (0.491). The roughly equal magnitudes suggest the echo-forward effect of cases remained roughly stable from the start to the end of the 2-week period.

Standard surveillance metrics for the weeks of April 28 and May 5, 2023, are provided in Tables 2 and 3. Although most countries in the region had a small number of new COVID-19 cases, several countries had high rates of transmission. For the week of April 28, Brunei had by far the highest rate of new COVID-19 cases at 1431 per 100,000 population. The next highest rate was 204 in New Zealand, followed by 115 in Australia and 36 in South Korea. These rates exceeded the outbreak threshold according to the Centers for Disease Control and Prevention [32-42,68]. Specifically, *low* transmission is defined as no more than 10 cases per 100,000 people per week, *moderate* transmission ranges from 10 to 50 cases per 100,000 people per week, and *substantial* transmission is between 50 to 100 cases per 100,000 people per week [68,69]. Thus, Brunei, New Zealand, and Australia were in greater-than-substantial outbreaks, while South Korea was in a moderate outbreak.

Outbreak status for countries in the region looked roughly the same for the week of May 5. Brunei had somewhat lower transmission rate than in the prior week, while Australia had a higher rate. The other exception is Guam, who entered an outbreak with a speed of 61 (Table 3). The speed in Guam was 17 for the week of April 21. The dip to zero in the intermediate week (Table 2) most likely represents a disruption in data reports, but it is also important to note that transmission speed can be more variable in small island territories. For the region as a whole, COVID-19 may have still met the pandemic definition at the time of the WHO declaration, with several countries experiencing sizeable outbreaks. However, transmission rates were low for the rest of the region, suggesting a transition from pandemic to endemic.

Table 2. Static COVID-19 surveillance metrics for East Asia and Pacific countries in the week of April 28, 2023.

Country	New COVID-19 cases, n	Cumulative COVID-19 cases, n	7-day moving average of new cases	Infection rate per 100,000 individuals	New weekly deaths, n	Cumulative deaths, n	7-day moving average of deaths	Death rate per 100,000 individuals	Conditional death rate
Australia	4301	11,272,355	4273.57	115.01	0	20,613	18	0.52	0
Brunei	918	292,644	721.29	1431.20	0	158	0.14	0.30	0
Cambodia	0	138,733	0.14	0	0	3056	0	0	0.02
China	462	99,250,200	511.14	0.03	7	120,984	5.86	0	0
Fiji	0	68,921	0	0	0	883	0	0	0.01
French Polynesia	6	78,518	9	0	0	649	0	0	0.01
Guam	0	51,240	8.71	0	0	413	0	0	0.01
Indonesia	2122	6,784,170	1871.14	0.77	20	161,404	20.86	0.01	0.02
Japan	7343	33,766,957	11,059.29	5.92	18	74,614	21	0.01	0
Laos	4	218,085	1.29	0.05	0	671	0	0	0
Malaysia	0	5,071,840	709	0	0	37,020	1.29	0	0.01
Mongolia	19	1,008,038	9.43	0	0	2136	0	0	0
Myanmar	208	635,660	140.57	0.38	0	19,492	0.14	0	0.03
New Caledonia	0	80,058	0	0	0	314	0	0	0
New Zealand	1509	2,261,126	1423.57	203.73	0	2762	3.71	0.51	0
Northern Mariana Islands	3	13,849	1.86	0	0	41	0	0	0
Palau	0	6000	0	5.53	0	9	0	0	0
Papua New Guinea	1	46,850	1.14	0.01	0	670	0	0	0.01
Philippines	1190	4,097,525	932.71	1.03	0	66,444	0	0	0.02
Singapore	3201	2,391,770	3316.71	0	0	1722	0	0	0
South Korea	18,752	31,251,203	15,477.57	36.19	6	34,518	6.71	0.01	0
Thailand	242	4,732,301	252.29	2.36	0	33,957	1.43	0.01	0.01
Timor	2	23,431	0.43	0.15	0	138	0	0	0.01
Vietnam	2233	11,567,728	1836.14	2.27	0	43,195	1	0	0

Table 3. Static COVID-19 surveillance metrics for East Asia and Pacific countries in the week of May 5, 2023.

Country	New COVID-19 cases, n	Cumulative COVID-19 cases, n	7-day moving average of new cases	Infection rate per 100,000 individuals	New weekly deaths, n	Cumulative deaths, n	7-day moving average of deaths	Death rate per 100,000 individuals	Conditional death rate
Australia	5076	11,303,671	4643.29	135.74	0	20,751	19.71	0.43	0
Brunei	816	299,505	927.86	1271.72	0	160	0.29	0.54	0
Cambodia	0	138,736	0.43	0	0	3056	0	0	0.02
China	572	99,254,488	612.57	0.04	12	121,048	9.14	0	0
Fiji	0	68,921	0	0	0	883	0	0	0.01
French Polynesia	3	78,545	3.57	0	0	649	0	0	0.01
Guam	105	51,345	15	61.12	0	413	0	0	0.01
Indonesia	1471	6,795,221	1578.71	0.53	22	161,574	24.29	0.01	0.02
Japan	833	33,803,572	5230.71	0.67	25	74,694	11.43	0.02	0
Laos	4	218,096	1.57	0.05	0	671	0	0	0
Malaysia	0	5,079,436	1085.14	0	0	37,028	1.14	0	0.01
Mongolia	49	1,008,265	36.43	0	0	2136	0	0	0
Myanmar	157	636,967	186.71	0.29	0	19,493	0.14	0	0.03
New Caledonia	0	80,058	0	0	0	314	0	0	0
New Zealand	1535	2,272,229	1564.43	207.28	0	2792	4.29	0.89	0
Northern Mariana Islands	4	13,872	3.57	0	0	41	0	0	0
Palau	1	6000	0.14	0	0	9	0	0	0
Papua New Guinea	0	46,864	2	0	0	670	0	0	0.01
Philippines	1940	4,108,914	1627	1.68	0	66,453	1.29	0	0.02
Singapore	3386	2,414,394	3277	0	0	1722	0	0	0
South Korea	19,989	31,371,347	17,163.43	38.58	8	34,591	10.43	0.02	0
Thailand	320	4,734,000	275.43	3.12	0	33,967	1.43	0.03	0.01
Timor	1	23,435	0.57	0.07	0	138	0	0	0.01
Vietnam	2823	11,585,390	2523.14	2.88	0	43,200	0.71	0	0

A comparison of Tables 2 and 3 demonstrates little to no change before and after the WHO declared an end to COVID-19 as a public health emergency. China, Japan, and South Korea had the most cumulative cases of COVID-19 transmissions, but these ranks are a function of population size. Thus, a better measure is the number of COVID-19 cases and deaths per 100,000 population. Moreover, death is often a better proxy for the state of an outbreak than transmissions because deaths are less likely to be undercounted [70]. Undercounting may be due to poor public health infrastructure, increased use of home antigen testing, or a dearth in polymerase chain reaction testing or other resources. China, Japan, and South Korea reported <.01 deaths per population. When we control for risk of death given the number of COVID-19 transmissions, we find that Myanmar had the highest conditional death rate, at 0.03 deaths per 100,000 population, while Cambodia, Indonesia, and the Philippines each reported a rate of 0.02.

Table 4. Novel COVID-19 surveillance metrics for East Asia and Pacific countries for the week of April 28, 2023.

Country	Speed	Acceleration	Jerk	7-day persistence effect on speed
Australia	114.28	0.31	−0.02	39.51
Brunei	1124.51	111.65	−4.97	134.57
Cambodia	0	0	0	0
China	0.04	0	0	0.01
Fiji	0	−0.05	−0.05	0.02
French Polynesia	3.26	0	0	0.51
Guam	5.07	−2.49	0.75	2.52
Indonesia	0.68	0	−0.03	0.15
Japan	8.92	−0.51	−0.90	2.88
Laos	0.02	0	−0.01	0.01
Malaysia	2.09	0	0	0.71
Mongolia	0.16	0	0	0.04
Myanmar	0.26	0.03	0	0.04
New Caledonia	0	0	0	0.02
New Zealand	192.18	2.01	1.31	72.76
Northern Mariana Islands	2.88	0	0	2.54
Palau	0.79	0.79	0.79	0
Papua New Guinea	0.01	0	0	0
Philippines	0.81	0.05	0.01	0.16
Singapore	59.34	0	0	24.83
South Korea	29.87	1.37	−0.17	8.68
Thailand	2.46	−0.01	−0.02	0.71
Timor	0.03	0.02	0.02	0.01
Vietnam	1.87	−0.11	0.12	0.85

Tables 4 and 5 present enhanced dynamic surveillance metrics for the weeks before and after May 5. For most countries, speed was low and stable, which implied values of acceleration and jerk close to 0. Four countries were in a state of outbreak: Australia, Brunei, New Zealand, and South Korea.

From the first to the second week, acceleration was positive for Australia but negative for Brunei, which means the outbreak was worsening in Australia but improving in Brunei. The near-zero value of jerk for Australia suggests an inflection point of the outbreak, which may have been on the verge of reduction. Because acceleration was negative for Brunei, the negative value of jerk suggests a continued, but slowed, downward trend in the transmission rate. However, the 7-day persistence effect remained high for both countries.

The outbreaks in New Zealand and South Korea remained stable, with near-zero acceleration and jerk, although the 7-day persistence effect showed a slight increase. Because 4 out of the 36 countries and territories were experiencing outbreaks, the distinction between pandemic and endemic for the region remains unclear. On the basis of trends in other global regions, these outbreaks are likely to be the last outbreaks of this magnitude for these 4 countries; however, the scale of these outbreaks means that transmissions could spread across borders, potentially leading to new outbreaks in neighboring countries.

Finally, we note that the figures in Tables 4 and 5 are not calculated as day-over-day averages across the week, as in Tables 2 and 3. Thus, the magnitudes of speed may not exactly match those presented in Tables 2 and 3.

Table 5. Novel COVID-19 surveillance metrics for East Asia and Pacific countries for the week of May 5, 2023.

Country	Speed	Acceleration	Jerk	7-day persistence effect on speed
Australia	124.16	2.96	0.61	44.39
Brunei	1446.57	−22.78	−25.05	436.81
Cambodia	0	0	0	0
China	0.04	0	0	0.01
Fiji	0	0	0	0
French Polynesia	1.26	0	0	1.27
Guam	8.73	8.73	8.73	1.97
Indonesia	0.57	−0.03	0.05	0.26
Japan	4.22	−0.75	1.17	3.47
Laos	0.02	0	0.01	0.01
Malaysia	3.20	0	0	0.81
Mongolia	0.95	0	0	0.06
Myanmar	0.34	−0.01	−0.01	0.10
New Caledonia	0	0	0	0
New Zealand	211.21	0.51	−1.12	74.65
Northern Mariana Islands	6.63	0	0	1.12
Palau	0	−0.79	−0.79	0.31
Papua New Guinea	0.02	0	0	0
Philippines	1.41	0.09	0.02	0.31
Singapore	57.34	0	0	23.05
South Korea	33.12	0.34	0.14	11.60
Thailand	2.69	0.11	0.03	0.96
Timor	0.04	−0.01	−0.01	0.01
Vietnam	2.57	0.09	−0.10	0.73

Table 6. East Asia and Pacific countries with the highest 7-day persistence estimate in the week of May 5, 2023.

Country	7-day persistence (May 5, 2023)
Brunei	436.81
New Zealand	74.65
Australia	44.39
Singapore	23.05
South Korea	11.60

Table 6 compares the 7-day persistence effect on speed for the top 5 countries for the week of May 5. Unsurprisingly, persistence was highest for Brunei, which had the largest outbreak at the time. Interestingly, Singapore appears on the list despite not having reached outbreak status. The persistence suggests that an outbreak could have been imminent for Singapore. Alternatively, the persistence could signify a strong, continued presence of COVID-19 as endemic in the area.

Figure 1 plots regional speed, acceleration, jerk, and 7-day persistence metrics from August 14, 2020, to May 12, 2023. The dashed gray line denotes the informal Centers for Disease Control and Prevention outbreak threshold of speed equal to 10. The region saw 4 outbreaks over the period, but the first and third were small, barely eclipsing a speed of 10. The first outbreak happened in August 2021. The second happened in February 2022, ultimately reaching a peak speed of 38. A third minor outbreak occurred in July 2022. The final outbreak began in December 2022, reaching a peak speed of 243 at the end of the year. This outbreak was one of the largest in the world over the course of the pandemic. However, since the end of January 2023, the region has not been in a state of outbreak.

Figure 1. Novel surveillance metrics (speed, acceleration, jerk, and 7-day persistence) for COVID-19 infections in East Asia and the Pacific region from August 2020 to May 2023.

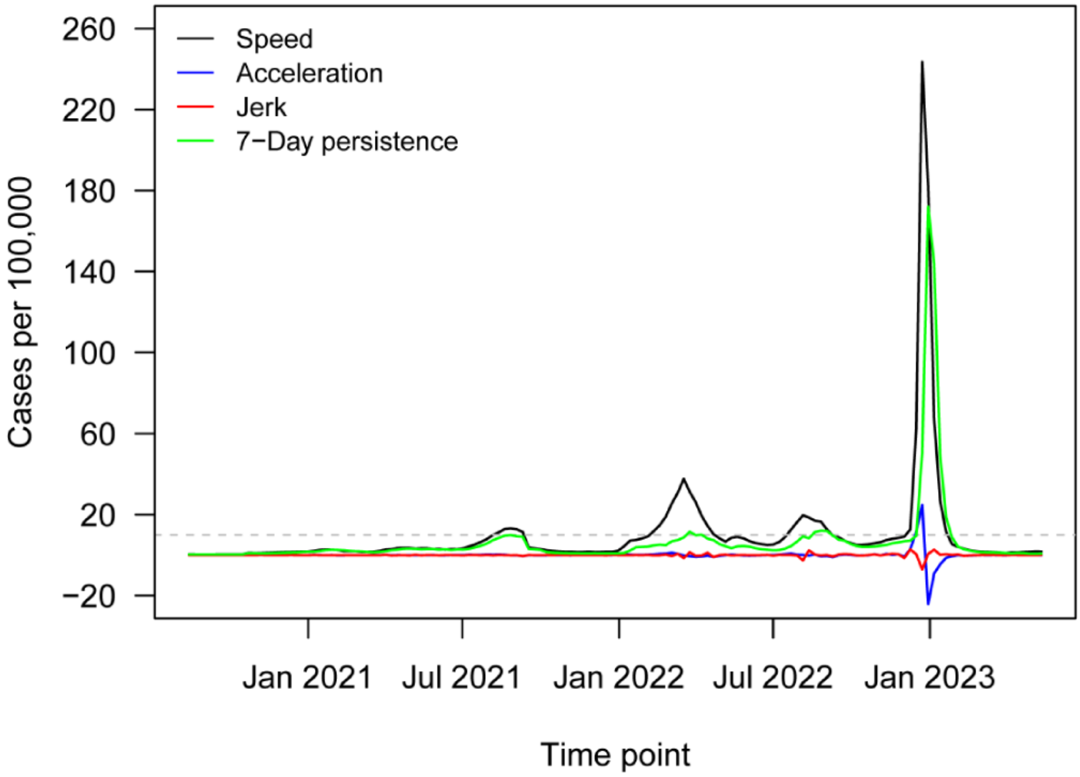


Figure 2 plots variant groups as a proportion of all viral specimens collected and sequenced in the region (and made available through GISAID) each month. The first outbreak occurred shortly after the appearance of the Delta variant. Each subsequent outbreak occurred after the dominant Omicron variant arrived. East Asia and the Pacific region, like much of the rest of the world, saw a surge in cases amid the heightened transmissibility of Omicron [71]. In contrast, the large outbreaks driven by Omicron did not occur until the end of 2022. This delay was achieved by the effective zero–COVID-19 policy in China, which was finally abandoned at that point. The outbreak hit in force shortly afterward.

Another potential indication of the end to the pandemic is the continued dominance of the Omicron variant. While the region saw a mixture of the ancestral, Alpha, Beta, and Delta variants before the arrival of Omicron in November of 2021, viral sequences have almost exclusively been identified as Omicron and its subvariants ever since.

Figure 3 plots *P* values from a series of 1-sided *t* tests of whether speed for the region was equal to or greater than the threshold outbreak of 10. These tests were conducted on a rolling 6-month window of weekly regional speed. The test can assess whether the region was in an outbreak over a period in which speed may sometimes eclipse an outbreak threshold and other times remain below the threshold. Statistical insignificance indicates no evidence against the null hypothesis of below-outbreak regional speed over a period. For example, at the α=.10 level, the final *P* value ≤.10 in the figure implies a rejection of the null hypothesis that the region as a whole was not in an outbreak for the period from November 5, 2022, to May 5, 2023.

In Figure 3, the dashed gray line denotes the least restrictive conventional significance level threshold of α=.10. The test first rejected the null in favor of the alternative for the 6-month period ending in early May of 2021. The test lost significance again in mid-September, before regaining significance at the very end of 2022. This second period of rejection was driven by the large Omicron outbreak around the time. More recently, the test was on the verge of insignificance at the time of WHO declaration. Although the test failed to reject the null at the α=.10 level, it did reject at the α=.05 level. Compared to other global regions, East Asia and the Pacific region are the latest to potentially transition from the pandemic to endemic COVID-19 status.

Within the historical context of enhanced surveillance metrics, the region may be at the end stage of the pandemic. Speed has not been this low for this long since the start of the pandemic. However, we do note that several countries in the region remained in substantial outbreaks at the time of the WHO declaration. This reality brings some uncertainty to the ostensible end to the pandemic in the region.

Figure 2. Variants of concern (VOCs) as a proportion of all sequenced SARS-CoV-2 specimens from March 2020 to May 2023 in East Asia and the Pacific region (n=1,234,493).

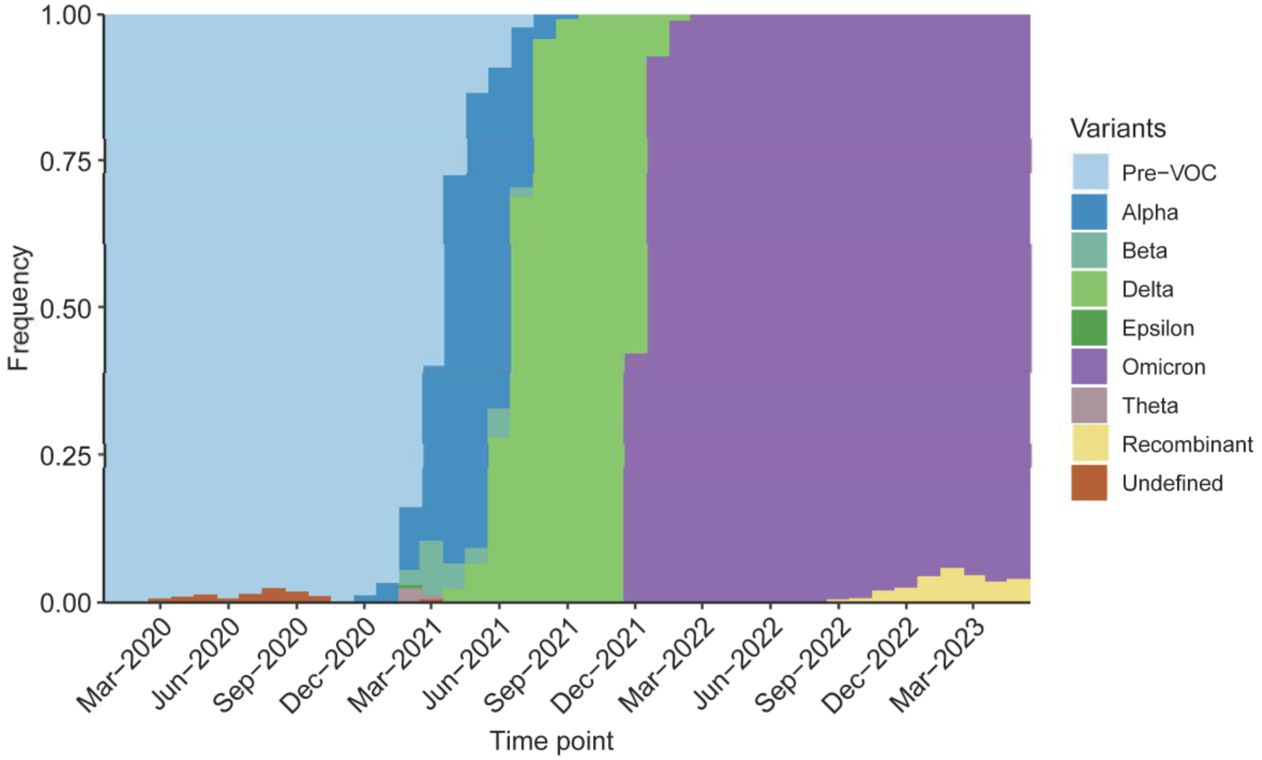


Figure 3. *P* values from t tests of weekly COVID-19 transmissions per 100,000 population equal to 10 over a rolling 6-month window in East Asia and the Pacific region.

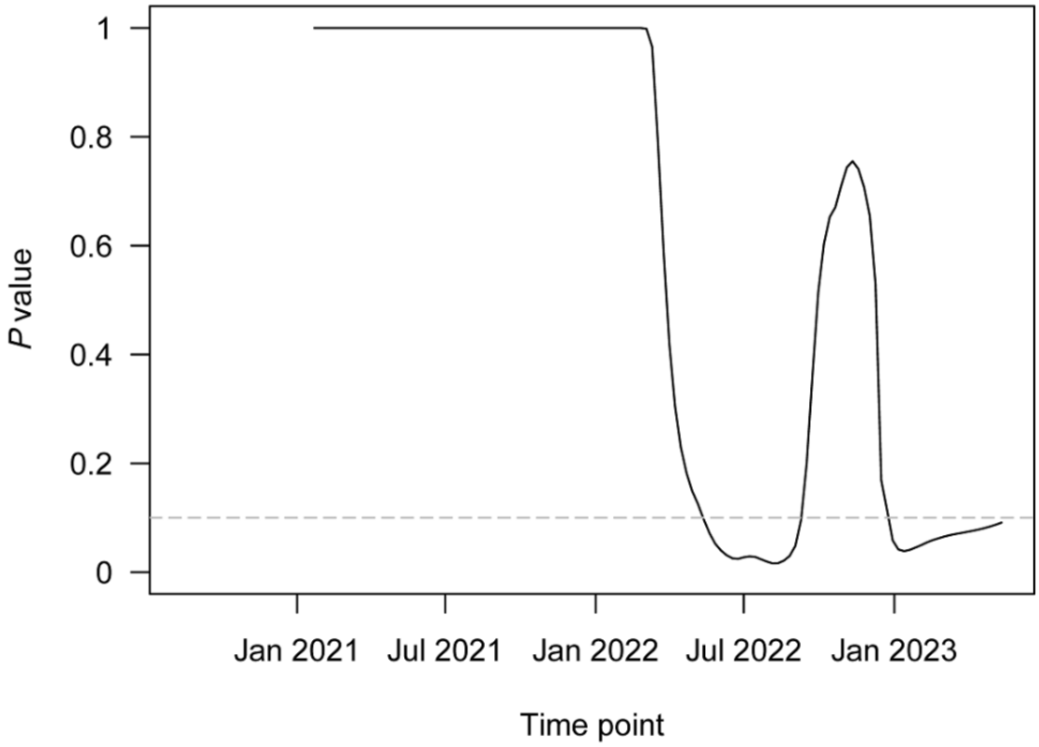
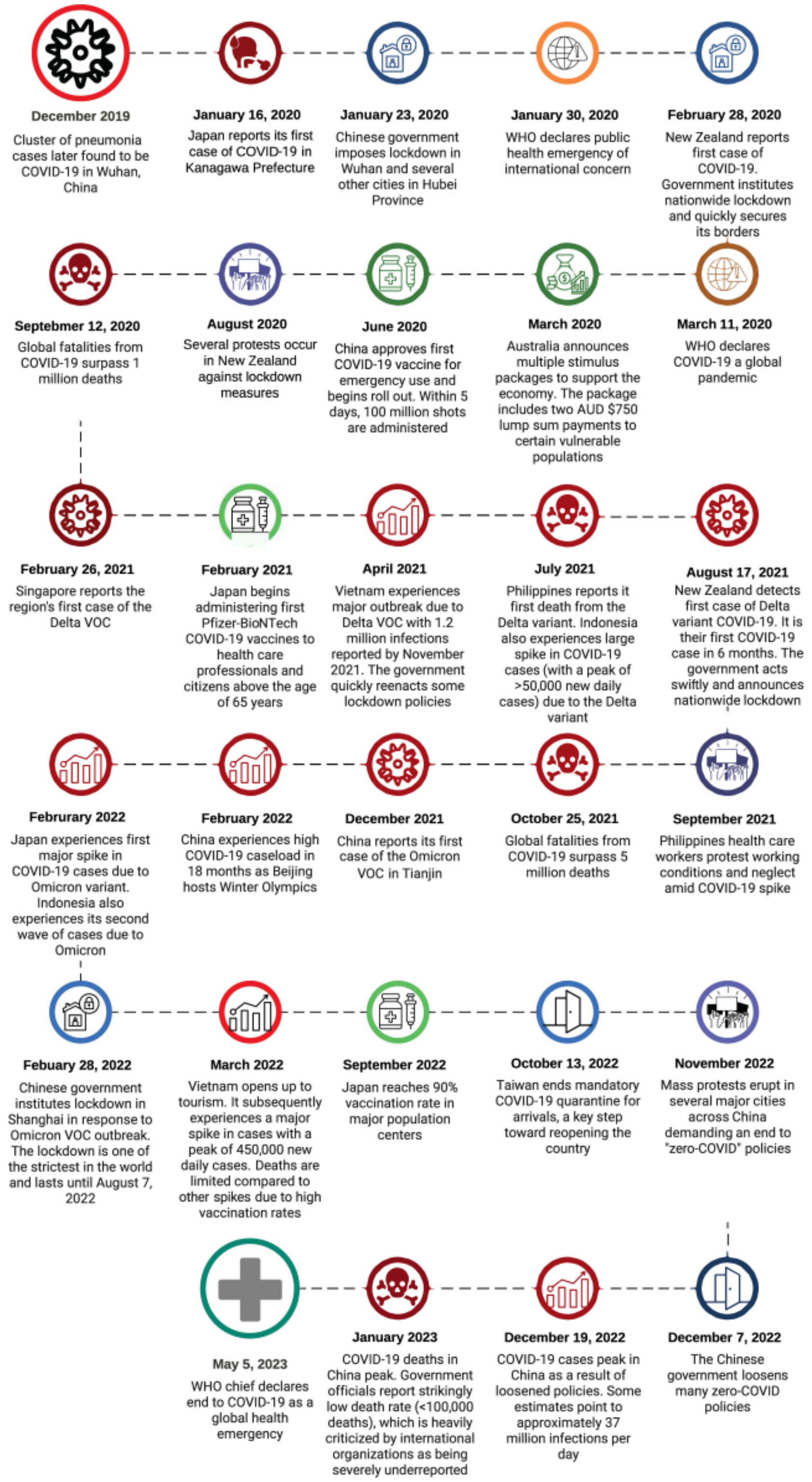


Figure 4 provides a timeline of the onset of COVID-19 in East Asia and the Pacific region as well as vaccination programs and major events that shaped the progression of the pandemic in the region. Most notably, the zero–COVID-19 policy in China effectively contained the disease in the country and therefore to an extent in the region. Upon the abandonment of the policy, a large wave of cases driven by Omicron finally hit the country, much later than in other parts of the world.

Figure 4. Timeline of the COVID-19 pandemic in East Asia and the Pacific region. VOC: variant of concern; WHO: World Health Organization.



Discussion

Principal Findings

COVID-19 has affected all countries in East Asia and the Pacific region. As of March 16, 2023, the WHO had reported a total of 213,883,614 COVID-19 cases and 622,018 deaths in East Asia and the Pacific region [72]. At the same point, the WHO reported 760,360,956 confirmed cases of COVID-19, including 6,873,477 deaths worldwide [72].

The results indicate a potential COVID-19 transition from pandemic to endemic at the time the WHO removed the designation of COVID-19 as a public health emergency of international concern on May 5, 2023. Regional speed had remained below the outbreak threshold of 10 for several months ahead of the WHO declaration. The rolling *t* test of speed, equal to the outbreak threshold, suggests that the region was in an outbreak over the six-month period starting at the end of 2022 when the Omicron variant drove a large increase in cases. However, the result was only significant at the α=.10 level, not the 0.05 level. However, several countries in the region had transmission rates consistent with a substantial outbreak in the 2 weeks around the WHO declaration. These outbreaks make an apparent end to the pandemic in the region unclear at the time.

As another indicator, nearly all sequenced SARS-CoV-2 specimens in the region were identified as the Omicron variant from early 2022 onward. This transition is consistent with a more clearly defined end to the pandemic in other global regions [73-75]. The relative delay for East Asia and the Pacific region may be attributable to a late Omicron outbreak in the region when China relaxed its stringent and effective zero–COVID-19 policy [76].

Policies Implemented to Control and Mitigate the Transmission of COVID-19

Many countries in the region experienced multiple waves of COVID-19 infections, with some waves more severe than others. During the initial outbreaks, countries in East Asia and the Pacific region were quick to respond with policies that were highly effective at curbing transmission rates and limiting mortality. Compared to other regions of the world, Asian countries experienced significantly lower mortality early in the pandemic (2.7 deaths per million in Asian countries vs 197 deaths per million in European countries) [77]. This disparity has been attributed to several factors.

First, countries in the region were able to quickly formulate effective strategies that were guided by their prior experiences with other respiratory pandemics, such as the 2003 outbreak of SARS-CoV-1 and the 2015 outbreak of Middle East respiratory syndrome [78-80].

Second, cultural differences between Eastern and Western societies may have allowed countries to more easily adopt face masking and social distancing policies [77]. Where Western countries such as the United States pride themselves on individual freedoms, many Asian countries have cultures that value collective societal well-being [81].

Third, the region as a whole tended to implement strict containment policies, which typically included mask mandates, social distancing, mass testing requirements, contact tracing, quarantine protocols, border lockdowns, and eventually mass vaccination programs to control the virus.

China initially implemented a strict zero–COVID-19 policy with the aim to keep cases as close to zero as possible. At times, entire cities, such as Shanghai, Chengdu, and Wuhan, entered citywide lockdown to prevent the spread of the virus [82-84]. Amid these lockdowns, economic activities were halted, public transportation was suspended, and residents were restricted to their homes. Meanwhile, individuals who tested positive for COVID-19 were taken to large isolation facilities to undergo quarantine [82,85].

New Zealand and Taiwan also initially implemented strict COVID-19 policies including mask mandates, tight border control, contact tracing, and tightly targeted quarantines for local outbreaks. These strategies were highly effective in containing the virus during major outbreaks for both countries [86,87]. However, both eventually shifted away from these strategies with the emergence of more contagious variants such as Delta and Omicron. Instead, they focused more on achieving herd immunity through widespread vaccination campaigns and protecting populations [86,88].

Vietnam, having learned lessons from the SARS outbreak in 2003, began preparing for COVID-19 weeks before the first cases were reported in their country. Knowing that their health care resources were limited, Vietnam's strategic goal was to keep COVID-19 cases as low as possible. They began stockpiling personal protective equipment and created local medical response teams to address outbreaks. When Vietnam detected its first cases in the community, it acted decisively and swiftly by locking down areas and performing meticulous contact tracing [89,90]. Vietnam never relied on widespread testing to curb infections [89].

The Philippines and Indonesia were COVID-19 hotspots throughout the pandemic due to limited public health responses by their governments. Both countries were late to institute COVID-19 containment measures, such as contact tracing, testing, and isolation, leading to high rates of infections and deaths relative to other countries in the region [91-94]. In the Philippines, the country's limited health care infrastructure left it ill-prepared for the pandemic, leading to an exponential growth in cases during the first wave [94,95].

Many Pacific Island countries were quick to implement strategies to prevent COVID-19 infections. Given their geographical isolation, these countries went almost completely free of COVID-19 cases during the initial wave of infections through the use of strict border closures and restrictions on international travel [96-98]. For example, the Solomon Islands reported its first COVID-19 case in October 2020 in a student who was returning from the Philippines [99]. However, it was not until 2022 that the Solomon Islands faced its first widespread community outbreak [100]. These strict border closures, while effective, came with a natural economic cost as many island nations rely heavily on tourism, which virtually evaporated during the pandemic [97].

Subsequent waves of infection were primarily driven by local easing of COVID-19 precautions and the emergence of new and more contagious variants such as Delta and Omicron. For example, Japan has experienced several waves of COVID-19 infections, with the largest wave occurring in mid-2022 due to the Omicron variant [101]. China, on the other hand, has reported fewer waves of COVID-19 infections due to its strict containment measures, although it is not clear how accurate their data are given their lack of consistent reporting. However, when China eased their zero–COVID-19 policy in December 2022, it subsequently experienced an intense surge in cases and related deaths [76]. Other countries in Asia, such as Vietnam and Taiwan, have been relatively successful in controlling the virus and have experienced fewer waves of infection.

While countries enjoyed relative success early on in controlling disease, they had to grapple with the economic and social consequences of their restrictive containment policies. These outbreaks and their associated governmental responses led to substantial sociopolitical turmoil throughout the region. Protests were widespread and focused on the government's mishandling of economic recovery and overly restrictive COVID-19 policies. Thailand and Malaysia saw protests break out in 2021 due to the government's perceived lackluster pandemic response [102,103]. The Philippines saw protesters calling for an end to quarantine rules, which led to President Duterte to issue a "shoot to kill" order against the protesters [104,105]. In China, antigovernment protests broke out throughout the country in late 2022 calling for an end to China's zero–COVID-19 policy, which used strict and draconian city-wide lockdowns to curb the spread of the virus. In response, China eased their policy in December 2022 and subsequently experienced an intense surge in cases and deaths. The WHO criticized China for underrepresenting the true number of deaths, with one senior official stating that China was using too narrow of a definition to count deaths [106,107]. New Zealand saw protesters converge on Parliament to oppose COVID-19 vaccines mandates [108].

In terms of the economy, East Asia and the Pacific region experienced a 0.2% decline in gross domestic product (GDP) in 2020 driven by containment measures that limited economic activity. However, as vaccines became more prevalent and countries began easing restrictions, economies rebounded with a GDP growth of 5.8% in 2021 [109].

To combat the economic ramifications of their containment strategies, governments in the region focused on reopening their economies and using large fiscal packages to stimulate economic recovery. These fiscal packages typically included direct cash support for affected workers, tax relief for individuals and corporations, support for their strained health care systems, and economic investment in specific economic sectors and public works. In 2020, the Chinese government launched several fiscal measures estimated to be worth approximately 4.9 trillion RMB (4.7% of GDP; US $672.3 billion). Key aspects of these measures included tax relief, investment in epidemic prevention and control, and medical equipment [110]. In Vietnam, the government introduced a fiscal support package worth 291.7 trillion VND (3.6% of GDP; US $11.6 billion) with measures that included deferment of value-added, personal income, and corporate income taxes [110]. The Australian government approved a substantial fiscal stimulus package worth AUS $312 billion (15.75% of GDP; US $196 billion) to be used through the financial year 2025, which included an AUS $20 billion (US $12.5 billion) health response package to secure vaccines and strengthen the existing health care system [110]. Singapore announced an SGD $92 billion (US $67.9 billion) fiscal package that included a cash payout to all Singaporeans [110]. In 2020, Japan approved multiple stimulus packages, including the Emergency Economic Package against COVID-19 (worth ¥117.1 trillion; US $750.5 billion) and the Comprehensive Economic Measures to Secure People's Lives and Livelihoods toward Relief Hope package (worth ¥73.6 trillion; US $471.7 billion). These measures were intended to protect businesses and support wages while also focusing on building a more resilient future economy [110].

Vaccination Efforts

Efforts to create a vaccine started immediately. SinoVac, a Chinese pharmaceutical company, began development of an inactivated COVID-19 vaccine called CoronaVac as early as January 2020. By fall 2020, SinoVac launched phase 3 trials in several countries, including Brazil, Indonesia, the Philippines, and Turkey [111]. The large phase 3 trial in Brazil, which used a 2-dose series administered 14 days apart, demonstrated an efficacy of 51% against symptomatic COVID-19 infections and 100% against severe infections [112]. By early 2021, many countries in East Asia and the Pacific region, as well as in South America and Eastern Europe, began mass vaccination programs using CoronaVac [113-118]. By April 2021, SinoVac had announced that it had production capacity to produce 2 billion doses per year [111]. On June 1, 2021, the WHO validated the vaccine for emergency use. Other countries, such as Taiwan, Japan, and Korea, purchased millions of vaccines from western countries, including the AstraZeneca, Pfizer-BioNTech, and Moderna vaccines in early 2021 [119-122].

With many countries now having reached herd immunity through a combination of several available vaccines, there has been widespread easing of border restrictions, as seen in Japan, China, Taiwan, Singapore, Australia, and New Zealand [123-127].

Conclusions

Concerns about potential resurgences of the virus remain valid. As long as COVID-19 continues to spread and mutate, the possibility of new variants emerging remains. These variants could potentially be more transmissible, be resistant to vaccines, or cause more severe illness. This underlines the importance of continued vigilance, vaccination efforts, and global cooperation to control the spread of the virus [39].

Limitations

COVID-19 data became less frequently reported worldwide by the time the WHO declared an end to the pandemic as a public health emergency of international concern [128]. In addition, more people began using at-home tests as the pandemic evolved [129], and experts believe the Chinese government has underreported COVID-19 deaths [130]. Because the enhanced surveillance metrics of speed, acceleration, jerk, and 7-day persistence are based on rates and not total counts, statistical bias caused by countries dropping in or out of the sample is mitigated. However, if a nonincluded country is unrepresentative of the region in terms of disease burden, its omission can still influence historical data comparisons. Viral specimen tests for variants of concern in GISAID depend on testing and sequencing capacity, which varied by country across the region.

Acknowledgments

Research reported in this publication was generously supported by the Robert J Havey, MD, Institute for Global Health's Innovation Challenge Award at Northwestern University Feinberg School of Medicine.

The authors gratefully acknowledge all data contributors, including the authors and their originating laboratories responsible for obtaining the specimens and their submitting laboratories for generating the genetic sequence and metadata and sharing them via the Global Initiative on Sharing All Influenza Data, on which this research is based.

Conflicts of Interest

None declared.

References

1. Muralidar S, Ambi SV, Sekaran S, Krishnan UM. The emergence of COVID-19 as a global pandemic: understanding the epidemiology, immune response and potential therapeutic targets of SARS-CoV-2. Biochimie 2020 Dec;179:85-100 [FREE Full text] [doi: 10.1016/j.biochi.2020.09.018] [Medline: 32971147]
2. Sharma A, Ahmad Farouk I, Lal SK. COVID-19: a review on the novel coronavirus disease evolution, transmission, detection, control and prevention. Viruses 2021 Jan 29;13(2):202 [FREE Full text] [doi: 10.3390/v13020202] [Medline: 33572857]
3. Chilamakuri R, Agarwal S. COVID-19: characteristics and therapeutics. Cells 2021 Jan 21;10(2):206 [FREE Full text] [doi: 10.3390/cells10020206] [Medline: 33494237]
4. Hu B, Guo H, Zhou P, Shi Z. Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol 2021 Mar 06;19(3):141-154 [FREE Full text] [doi: 10.1038/s41579-020-00459-7] [Medline: 33024307]
5. Seyed Hosseini E, Riahi Kashani N, Nikzad H, Azadbakht J, Hassani Bafrani H, Haddad Kashani H. The novel coronavirus disease-2019 (COVID-19): mechanism of action, detection and recent therapeutic strategies. Virology 2020 Dec;551:1-9 [FREE Full text] [doi: 10.1016/j.virol.2020.08.011] [Medline: 33010669]
6. Post LA, Lin JS, Moss CB, Murphy RL, Ison MG, Achenbach CJ, et al. SARS-CoV-2 wave two surveillance in east Asia and the Pacific: longitudinal trend analysis. J Med Internet Res 2021 Feb 01;23(2):e25454 [FREE Full text] [doi: 10.2196/25454] [Medline: 33464207]
7. East Asia and Pacific. The World Bank Group. URL: https://data.worldbank.org/region/east-asia-and-pacific [accessed 2024-04-29]
8. Smith-Schoenwalder C. When Does the COVID-19 pandemic end?: neither WHO nor any other organization decides when a pandemic begins or ends. U.S. News & World Report. 2023. URL: https://www.usnews.com/news/health-news/articles/2023-05-12/when-does-the-covid-19-pandemic-end [accessed 2024-04-29]
9. Burki T. WHO ends the COVID-19 public health emergency. Lancet Respir Med 2023 Jul;11(7):588 [FREE Full text] [doi: 10.1016/S2213-2600(23)00217-5] [Medline: 37247628]
10. WHO chief declares end to COVID-19 as a global health emergency. United Nations. URL: https://news.un.org/en/story/2023/05/1136367 [accessed 2024-04-29]
11. Lesson 1: introduction to epidemiology. Centers for Disease Control and Prevention. URL: https://archive.cdc.gov/#/details?url=https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section11.html [accessed 2024-04-29]
12. Epidemic, endemic, pandemic: what are the differences? Columbia University. 2021 Feb 19. URL: https://www.publichealth.columbia.edu/public-health-now/news/epidemic-endemic-pandemic-what-are-differences [accessed 2024-04-29]
13. The Lancet Infectious Diseases. Transitioning to endemicity with COVID-19 research. Lancet Infect Dis 2022 Mar;22(3):297 [FREE Full text] [doi: 10.1016/S1473-3099(22)00070-6] [Medline: 35151377]
14. Britto D, Delorme J, Miller B, Sleat D, Summers H, Wain R. Pandemic to endemic: the race against time. Tony Blair Institute for Global Change. 2021 Sep 19. URL: https://institute.global/policy/pandemic-endemic-race-against-time [accessed 2024-04-29]
15. Gilbert R, Cliffe SJ. Public health surveillance. In: Regmi K, Gee I, editors. Public Health Intelligence. Cham, Switzerland: Springer; Apr 27, 2016:91-110.
16. Hong R, Walker R, Hovan G, Henry L, Pescatore R. The power of public health surveillance. Dela J Public Health 2020 Jul;6(2):60-63 [FREE Full text] [doi: 10.32481/djph.2020.07.016] [Medline: 34467112]
17. Teutsch SM, Churchill RE. Principles and Practice of Public Health Surveillance 2nd Edition. Oxford, UK: Oxford University Press; 2000.
18. Teutsch SM. Considerations in planning a surveillance system. In: Lee LM, Teutsch SM, Thacker SB, St. Louis ME, editors. Principles & Practice of Public Health Surveillance. 3rd edition. Oxford, UK: Oxford Academic Press; 2010:18-28.
19. Teutsch SM, Thacker SB. Planning a public health surveillance system. Epidemiol Bull 1995 Mar;16(1):1-6. [Medline: 7794696]
20. Thacker SB, Qualters JR, Lee LM, Centers for Disease ControlPrevention. Public health surveillance in the United States: evolution and challenges. MMWR Suppl 2012 Jul 27;61(3):3-9. [Medline: 22832990]

21. Centers for Disease Control (CDC). Guidelines for evaluating surveillance systems. MMWR Suppl 1988 May 06;37(5):1-18. [Medline: 3131659]
22. Lee LM, Thacker SB. Public health surveillance and knowing about health in the context of growing sources of health data. Am J Prev Med 2011 Dec;41(6):636-640. [doi: 10.1016/j.amepre.2011.08.015] [Medline: 22099242]
23. Davis AM, Dunet DO, Keaton R, Snider DE, Sosin DM, Stroup DF, et al. CDC guidelines: improving the quality. Centers for Disease Control and Prevention. 1996. URL: https://stacks.cdc.gov/view/cdc/24407#moretextPAmods.subject_name [accessed 2024-04-29]
24. Thacker SB, Stroup DF. Future directions for comprehensive public health surveillance and health information systems in the United States. Am J Epidemiol 1994 Sep 01;140(5):383-397. [doi: 10.1093/oxfordjournals.aje.a117261] [Medline: 8067331]
25. Nsubuga P, White M, Thacker S, Anderson M, Blount S, Broome C, et al. Public health surveillance: a tool for targeting and monitoring interventions. In: Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, et al, editors. Disease Control Priorities in Developing Countries. 2nd edition. New York, NY: Oxford University Press; 2006.
26. Lee LM, Thacker SB, St. Louis ME, Teutsch SM. Principles and Practice of Public Health Surveillance 3rd Edition. New York, NY: Oxford University Press; 2010.
27. Thacker SB, Stroup DF, Rothenberg RB. Public health surveillance for chronic conditions: a scientific basis for decisions. Stat Med 1995 Feb 28;14(5-7):629-641. [doi: 10.1002/sim.4780140520] [Medline: 7792453]
28. Perry HN, McDonnell SM, Alemu W, Nsubuga P, Chungong S, Otten MW, et al. Planning an integrated disease surveillance and response system: a matrix of skills and activities. BMC Med 2007 Aug 15;5(1):24 [FREE Full text] [doi: 10.1186/1741-7015-5-24] [Medline: 17697387]
29. Koo D, Thacker SB. In snow's footsteps: commentary on shoe-leather and applied epidemiology. Am J Epidemiol 2010 Sep 15;172(6):737-739. [doi: 10.1093/aje/kwq252] [Medline: 20720100]
30. Romaguera RA, German RR, Klaucke DN. Evaluating public health surveillance. In: Teutsch SM, Churchill RE, editors. Principles and Practice of Public Health Surveillance. Oxford, UK: Oxford University Press; 2000:176-193.
31. Pappaioanou M, Malison M, Wilkins K, Otto B, Goodman RA, Churchill R, et al. Strengthening capacity in developing countries for evidence-based public health: the data for decision-making project. Soc Sci Med 2003 Nov;57(10):1925-1937. [doi: 10.1016/s0277-9536(03)00058-3] [Medline: 14499516]
32. Post L, Boctor MJ, Issa TZ, Moss CB, Murphy RL, Achenbach CJ, et al. SARS-CoV-2 surveillance system in Canada: longitudinal trend analysis. JMIR Public Health Surveill 2021 May 10;7(5):e25753 [FREE Full text] [doi: 10.2196/25753] [Medline: 33852410]
33. Post L, Culler K, Moss CB, Murphy RL, Achenbach CJ, Ison MG, et al. Surveillance of the second wave of COVID-19 in Europe: longitudinal trend analyses. JMIR Public Health Surveill 2021 Apr 28;7(4):e25695 [FREE Full text] [doi: 10.2196/25695] [Medline: 33818391]
34. Post L, Marogi E, Moss CB, Murphy RL, Ison MG, Achenbach CJ, et al. SARS-CoV-2 surveillance in the Middle East and North Africa: longitudinal trend analysis. J Med Internet Res 2021 Jan 15;23(1):e25830 [FREE Full text] [doi: 10.2196/25830] [Medline: 33302252]
35. Post L, Ohiomoba RO, Maras A, Watts SJ, Moss CB, Murphy RL, et al. Latin America and the Caribbean SARS-CoV-2 surveillance: longitudinal trend analysis. JMIR Public Health Surveill 2021 Apr 27;7(4):e25728 [FREE Full text] [doi: 10.2196/25728] [Medline: 33852413]
36. Post LA, Argaw ST, Jones C, Moss CB, Resnick D, Singh LN, et al. A SARS-CoV-2 surveillance system in Sub-Saharan Africa: modeling study for persistence and transmission to inform policy. J Med Internet Res 2020 Nov 19;22(11):e24248 [FREE Full text] [doi: 10.2196/24248] [Medline: 33211026]
37. Post LA, Benishay ET, Moss CB, Murphy RL, Achenbach CJ, Ison MG, et al. Surveillance metrics of SARS-CoV-2 transmission in central Asia: longitudinal trend analysis. J Med Internet Res 2021 Feb 03;23(2):e25799 [FREE Full text] [doi: 10.2196/25799] [Medline: 33475513]
38. Post LA, Issa TZ, Boctor MJ, Moss CB, Murphy RL, Ison MG, et al. Dynamic public health surveillance to track and mitigate the US COVID-19 epidemic: longitudinal trend analysis study. J Med Internet Res 2020 Dec 03;22(12):e24286 [FREE Full text] [doi: 10.2196/24286] [Medline: 33216726]
39. Post LA, Lorenzo-Redondo R. Omicron: fewer adverse outcomes come with new dangers. Lancet 2022 Apr 02;399(10332):1280-1281 [FREE Full text] [doi: 10.1016/S0140-6736(22)00514-1] [Medline: 35305298]
40. Oehmke JF, Moss CB, Singh LN, Oehmke TB, Post LA. Dynamic panel surveillance of COVID-19 transmission in the United States to inform health policy: observational statistical study. J Med Internet Res 2020 Oct 05;22(10):e21955 [FREE Full text] [doi: 10.2196/21955] [Medline: 32924962]
41. Oehmke JF, Oehmke TB, Singh LN, Post LA. Dynamic panel estimate-based health surveillance of SARS-CoV-2 infection rates to inform public health policy: model development and validation. J Med Internet Res 2020 Sep 22;22(9):e20924 [FREE Full text] [doi: 10.2196/20924] [Medline: 32915762]
42. Oehmke TB, Post LA, Moss CB, Issa TZ, Boctor MJ, Welch SB, et al. Dynamic panel data modeling and surveillance of COVID-19 in metropolitan areas in the United States: longitudinal trend analysis. J Med Internet Res 2021 Feb 09;23(2):e26081 [FREE Full text] [doi: 10.2196/26081] [Medline: 33481757]

43. Zarocostas J. With the COVID-19 PHEIC over, what next? Lancet 2023 May 20;401(10389):1642-1643 [FREE Full text] [doi: 10.1016/S0140-6736(23)01003-6] [Medline: 37211013]
44. Sarker R, Roknuzzaman AS, Hossain MJ, Bhuiyan MA, Islam MR. The WHO declares COVID-19 is no longer a public health emergency of international concern: benefits, challenges, and necessary precautions to come back to normal life. Int J Surg 2023 Sep 01;109(9):2851-2852 [FREE Full text] [doi: 10.1097/JS9.0000000000000513] [Medline: 37222700]
45. Fan VY, Cash R, Bertozzi S, Pate M. The when is less important than the what: an epidemic scale as an alternative to the WHO's Public Health Emergency of International Concern. Lancet Glob Health 2023 Oct;11(10):e1499-e1500 [FREE Full text] [doi: 10.1016/S2214-109X(23)00314-5] [Medline: 37598696]
46. Torner N. The end of COVID-19 public health emergency of international concern (PHEIC): and now what? Vacunas 2023 Jul;24(3):164-165. [doi: 10.1016/j.vacune.2023.05.001]
47. Wise J. COVID-19: WHO declares end of global health emergency. BMJ 2023 May 09;381:1041. [doi: 10.1136/bmj.p1041] [Medline: 37160309]
48. Harris E. WHO declares end of COVID-19 global health emergency. JAMA 2023 Jun 06;329(21):1817. [doi: 10.1001/jama.2023.8656] [Medline: 37195734]
49. Okesanya OJ, Manirambona E, Buban J, Olabode ON, Lucero-Prisno DE. Coronavirus disease 2019 emergency is over but the pandemic is not: implications for a new global order. Int J Surg Glob Health 2023;6(4):e0207. [doi: 10.1097/gh9.0000000000000207]
50. Lenharo M. WHO declares end to COVID-19's emergency phase. Nature 2023 May 05:882. [doi: 10.1038/d41586-023-01559-z] [Medline: 37147368]
51. Sarker R, Roknuzzaman AS, Nazmunnahar, Hossain MJ, Islam MR. Benefits and probable ill effects of WHO's declaration of end of COVID-19 pandemic: a way back to pandemic-free normal life. Ann Med Surg (Lond) 2023 Jun;85(6):3199-3201 [FREE Full text] [doi: 10.1097/MS9.0000000000000848] [Medline: 37363491]
52. Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Gavrilov D, Giattino C, et al. Coronavirus pandemic (COVID-19). Our World in Data. URL: https://ourworldindata.org/coronavirus [accessed 2024-04-29]
53. Croissant Y, Millo G. Panel data econometrics in R: the plm package. J Stat Soft 2008;27(2):1-43. [doi: 10.18637/jss.v027.i02]
54. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2023. URL: http://www.R-project.org/ [accessed 2024-04-14]
55. Hansen LP. Large sample properties of generalized method of moments estimators. Econometrica 1982 Jul;50(4):1029. [doi: 10.2307/1912775]
56. Eales O, Ainslie KE, Walters CE, Wang H, Atchison C, Ashby D, et al. Appropriately smoothing prevalence data to inform estimates of growth rate and reproduction number. Epidemics 2022 Sep;40:100604 [FREE Full text] [doi: 10.1016/j.epidem.2022.100604] [Medline: 35780515]
57. Elliott P, Eales O, Steyn N, Tang D, Bodinier B, Wang H, et al. Twin peaks: the Omicron SARS-CoV-2 BA.1 and BA.2 epidemics in England. Science 2022 Jun 24;376(6600):eabq4411 [FREE Full text] [doi: 10.1126/science.abq4411] [Medline: 35608440]
58. Shaddick G, Zidek JV, Schmidt AM. Spatio–Temporal Methods in Environmental Epidemiology with R. New York, NY: CRC Press; 2023.
59. Global influenza surveillance and response system (GISRS). Global Initiative on Sharing All Influenza Data (GISAID). URL: https://www.gisaid.org/ [accessed 2024-04-29]
60. Khare S, Gurry C, Freitas L, Schultz MB, Bach G, Diallo A, et al. GISAID's role in pandemic response. China CDC Wkly 2021 Dec 03;3(49):1049-1051 [FREE Full text] [doi: 10.46234/ccdcw2021.255] [Medline: 34934514]
61. Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data - from vision to reality. Euro Surveill 2017 Mar 30;22(13):30494 [FREE Full text] [doi: 10.2807/1560-7917.ES.2017.22.13.30494] [Medline: 28382917]
62. Nasereddin A, Golan Berman H, Wolf DG, Oiknine-Djian E, Adar S. Identification of SARS-CoV-2 variants of concern using amplicon next-generation sequencing. Microbiol Spectr 2022 Aug 31;10(4):e0073622 [FREE Full text] [doi: 10.1128/spectrum.00736-22] [Medline: 35758686]
63. Huddleston J, Hadfield J, Sibley T, Lee J, Fay K, Ilcisin M, et al. Augur: a bioinformatics toolkit for phylogenetic analyses of human pathogens. J Open Source Softw 2021 Jan;6(57):2906 [FREE Full text] [doi: 10.21105/joss.02906] [Medline: 34189396]
64. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol 2020 Nov 15;5(11):1403-1407 [FREE Full text] [doi: 10.1038/s41564-020-0770-5] [Medline: 32669681]
65. O'Toole Á, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. Virus Evol 2021;7(2):veab064 [FREE Full text] [doi: 10.1093/ve/veab064] [Medline: 34527285]
66. Arellano M, Bond S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev Econ Stud 1991 Apr;58(2):277-297. [doi: 10.2307/2297968]
67. Hu M. Pandemic Surveillance: Privacy, Security, and Data Ethics Elgar Law, Technology and Society Series. New York, NY: Monograph Book; 2022.

68. Levenson E, Firger J. What the CDC’s ‘substantial’ and ‘high’ levels of COVID-19 transmission actually mean. CNN Health. 2021. URL: https://edition.cnn.com/2021/07/28/health/substantial-or-high-covid-19-transmission-wellness/index.html [accessed 2024-04-29]
69. Christie A, Brooks JT, Hicks LA, Sauber-Schatz EK, Yoder JS, Honein MA, et al. Guidance for implementing COVID-19 prevention strategies in the context of varying community transmission levels and vaccination coverage. MMWR Morb Mortal Wkly Rep 2021 Jul 27;70(30):1044-1047 [FREE Full text] [doi: 10.15585/mmwr.mm7030e2] [Medline: 34324480]
70. Stoto MA. Public health assessment in the 1990s. Annu Rev Public Health 1992;13:59-78. [doi: 10.1146/annurev.pu.13.050192.000423] [Medline: 1599602]
71. Lundberg AL, Lorenzo-Redondo R, Ozer EA, Hawkins CA, Hultquist JF, Welch SB, et al. Has Omicron changed the evolution of the pandemic? JMIR Public Health Surveill 2022 Jan 31;8(1):e35763 [FREE Full text] [doi: 10.2196/35763] [Medline: 35072638]
72. Coronavirus disease (COVID-19) dashboard. World Health Organization. URL: https://covid19.who.int/table [accessed 2024-04-29]
73. Post L, Wu SA, Soetikno AG, Ozer EA, Liu Y, Welch SB, et al. Updated surveillance metrics and history of the COVID-19 pandemic (2020-2023) in Latin America and the Caribbean: longitudinal trend analysis. JMIR Public Health Surveill 2024 May 17;10:e44398 [FREE Full text] [doi: 10.2196/44398] [Medline: 38568194]
74. Lundberg AL, Wu SA, Soetikno AG, Hawkins C, Murphy RL, Havey RJ, et al. Updated surveillance metrics and history of the COVID-19 pandemic (2020-2023) in Europe: longitudinal trend analysis. JMIR Public Health Surveill 2024 Jun 21;10:e53551 [FREE Full text] [doi: 10.2196/53551] [Medline: 38568186]
75. Soetikno AG, Lundberg AL, Ozer EA, Wu SA, Welch SB, Mason M, et al. Updated surveillance metrics and history of the COVID-19 pandemic (2020-2023) in the Middle East and North Africa: longitudinal trend analysis. JMIR Public Health Surveill 2024 Jun 12;10:e53219 [FREE Full text] [doi: 10.2196/53219] [Medline: 38568184]
76. Woodruff J, Warsi Z. COVID rapidly spreads in China as government eases strict quarantine rules. Public Broadcasting Service. URL: https://tinyurl.com/yc3xb293 [accessed 2024-04-29]
77. Landoni G, Maimeri N, Fedrizzi M, Fresilli S, Kuzovlev A, Likhvantsev V, et al. Why are Asian countries outperforming the western world in controlling COVID-19 pandemic? Pathog Glob Health 2021 Feb 26;115(1):70-72 [FREE Full text] [doi: 10.1080/20477724.2020.1850982] [Medline: 33241776]
78. Acter T, Uddin N, Das J, Akhter A, Choudhury TR, Kim S. Evolution of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as coronavirus disease 2019 (COVID-19) pandemic: a global health emergency. Sci Total Environ 2020 Aug 15;730:138996 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.138996] [Medline: 32371230]
79. To KK, Sridhar S, Chiu KH, Hung DL, Li X, Hung IF, et al. Lessons learned 1 year after SARS-CoV-2 emergence leading to COVID-19 pandemic. Emerg Microbes Infect 2021 Dec;10(1):507-535 [FREE Full text] [doi: 10.1080/22221751.2021.1898291] [Medline: 33666147]
80. Khan M, Adil SF, Alkhathlan HZ, Tahir MN, Saif S, Khan M, et al. COVID-19: a global challenge with old history, epidemiology and progress so far. Molecules 2020 Dec 23;26(1):39 [FREE Full text] [doi: 10.3390/molecules26010039] [Medline: 33374759]
81. Jennings R. How cultural differences help Asian countries beat COVID-19, while US struggles. Voice of America (VOA). URL: https://www.insidevoa.com/p/5831.html [accessed 2024-04-29]
82. Brant R. Shanghai Covid: China announces largest city-wide lockdown. BBC News. 2022 Mar 27. URL: https://www.bbc.com/news/world-asia-china-60893070 [accessed 2024-04-29]
83. Woo R. China's Chengdu exits full citywide Covid lockdown on Thursday. Reuters. 2022 Sep 14. URL: https://www.reuters.com/world/china/chinas-chengdu-lift-citywide-covid-lockdown-thursday-2022-09-14/ [accessed 2022-09-14]
84. China ends 76-day lockdown of Wuhan city, where the virus outbreak started. CNBC. 2020 Apr 07. URL: https://www.cnbc.com/2020/04/08/coronavirus-china-ends-lockdown-of-wuhan-city-where-outbreak-started.html [accessed 2024-04-29]
85. What is China's 'zero covid' policy and why did it trigger protests? The Washington Post. 2022 Nov 09. URL: https://www.washingtonpost.com/world/2022/11/29/china-zero-covid-policy-protests-explained/ [accessed 2024-04-29]
86. Chen SC. Taiwan's experience in fighting COVID-19. Nat Immunol 2021 Apr 26;22(4):393-394. [doi: 10.1038/s41590-021-00908-2] [Medline: 33772224]
87. Qin A, Chien AC. Covid? What Covid? Taiwan thrives as a bubble of normality. The Washington Post. 2021 Aug 23. URL: https://www.nytimes.com/2021/03/13/world/asia/taiwan-covid.html [accessed 2024-04-29]
88. Frost N. Battling delta, New Zealand abandons its zero-covid ambitions. The NewYork Times. 2021 Oct 04. URL: https://www.nytimes.com/2021/10/04/world/australia/new-zealand-covid-zero.html [accessed 2024-04-29]
89. Gan N. How Vietnam managed to keep its coronavirus death toll at zero. CNN World. 2020 May 30. URL: https://edition.cnn.com/2020/05/29/asia/coronavirus-vietnam-intl-hnk/index.html [accessed 2024-04-29]
90. Willoughby E. An ideal public health model? Vietnam’s state-led, preventative, low-cost response to COVID-19. Brookings. 2021 Jun 29. URL: https://www.brookings.edu/blog/order-from-chaos/2021/06/29/an-ideal-public-health-model-vietnams-state-led-preventative-low-cost-response-to-covid-19/ [accessed 2024-04-29]
91. Indonesia: weak public health response to COVID-19. Human Rights Watch. 2021 Jan 13. URL: https://www.hrw.org/news/2021/01/13/indonesia-weak-public-health-response-covid-19 [accessed 2024-04-29]

92. Sullivan M. The Philippines becomes coronavirus hot spot in Southeast Asia. National Public Radio. 2020 Oct 08. URL: https://www.npr.org/2020/08/10/900279722/the-philippines-becomes-coronavirus-hot-spot-in-southeast-asia [accessed 2024-04-29]
93. Philippines now has worst COVID-19 outbreak in Southeast Asia. The Japan Times. 2020 Aug 07. URL: https://www.japantimes.co.jp/news/2020/08/07/asia-pacific/philippines-coronavirus-southeast-asia/ [accessed 2024-04-29]
94. Amit AM, Pepito VC, Dayrit MM. Early response to COVID-19 in the Philippines. Western Pac Surveill Response J 2021 Mar 31;12(1):56-60 [FREE Full text] [doi: 10.5365/wpsar.2020.11.1.014] [Medline: 34094626]
95. -. COVID-19: an ongoing public health crisis in the Philippines. Lancet Reg Health West Pac 2021 Apr;9:100160 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100160] [Medline: 33937888]
96. Ryall J. Many Pacific Island nations have kept Covid at bay. Deutsche Welle. 2020 May 19. URL: https://www.dw.com/en/zero-cases-how-pacific-islands-kept-coronavirus-at-bay/a-53495263 [accessed 2024-04-29]
97. Addressing COVID-19's impacts and structural crisis in the Pacific. The Food and Agriculture Organization. 2020 May 06. URL: https://www.fao.org/newsroom/detail/Addressing-COVID-19-s-impacts-and-structural-crisis-in-the-Pacific/en [accessed 2024-04-29]
98. McClure H. Surviving the pandemic: COVID-19 in the Pacific and health security. Development Bulletin. URL: https://pacificsecurity.net/wp-content/uploads/2021/10/DB82_Part17.pdf [accessed 2024-04-29]
99. The Solomon Islands reports its first coronavirus case, a student returning from the Philippines. SBS News. 2020 Oct 04. URL: https://www.sbs.com.au/news/article/the-solomon-islands-reports-its-first-coronavirus-case-a-student-returning-from-the-philippines/x20eytmn1 [accessed 2024-04-30]
100. Rising D. Solomon Islands' 1st virus outbreak causes growing concern. The Associated Press. 2022 Feb 17. URL: https://apnews.com/article/coronavirus-pandemic-health-solomon-islands-honiara-30d3df8cb0f809c9e3e5a0cc5c6f886f [accessed 2024-04-29]
101. McCurry J. Japan urges ‘highest level of vigilance’ as Omicron subvariant drives record Covid surge. The Gurdian. 2022 Jul 22. URL: https://www.theguardian.com/world/2022/jul/22/japan-urges-highest-level-of-vigilance-as-omicron-subvariant-drives-record-covid-surge [accessed 2024-04-29]
102. Suhartono M, Beech H. Thai protesters are back, and angrier, as government fumbles on Covid. The NewYork Times. 2021 Sep 05. URL: https://www.nytimes.com/2021/09/05/world/asia/thailand-protesters-covid.html [accessed 2024-04-29]
103. Malaysia: hundreds take to the streets in anti-government protest. Al Jazeera Media Network. 2021 Jul 21. URL: https://www.aljazeera.com/news/2021/7/31/hundreds-take-to-streets-in-anti-government-protest-in-malaysia [accessed 2024-04-29]
104. Philippines: president Duterte gives “shoot to kill” order amid pandemic response. Amnesty International. 2020 Apr 02. URL: https://www.amnesty.org/en/latest/news/2020/04/philippines-president-duterte-shoot-to-kill-order-pandemic/ [accessed 2024-04-29]
105. Lo B, Ott H. The Philippines has the worst coronavirus outbreak in Southeast Asia, and critics warn the government could use it to crack down on dissent. CBS News. 2020 Aug 07. URL: https://www.cbsnews.com/news/covid-19-philippines-worst-coronavirus-outbreak-in-southeast-asia-crackdown-on-dissent/ [accessed 2024-04-29]
106. Branswell H. Senior WHO official faults China for undercounting Covid deaths. STAT News. 2023 Jun 04. URL: https://www.statnews.com/2023/01/04/who-official-faults-china-undercounting-covid-deaths/ [accessed 2024-04-29]
107. China might be underreporting its COVID-related death toll. Euronews. 2022 Dec 21. URL: https://www.euronews.com/2022/12/21/china-might-be-underreporting-its-covid-related-death-toll-experts [accessed 2024-04-29]
108. Anti-mandate protesters converge on New Zealand parliament. VOA News. 2022 Aug 23. URL: https://www.voanews.com/a/anti-mandate-protesters-converge-on-new-zealand-parliament-/6712836.html [accessed 2024-04-29]
109. GDP growth (annual %) - East Asia and Pacific. World Bank Group. URL: https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=Z4 [accessed 2024-04-29]
110. Policy responses to COVID-19. International Monetary Fund. URL: https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19#C [accessed 2024-04-29]
111. Corum J, Zimmer C. How the Sinovac vaccine works. The NewYork Times. 2021 May 07. URL: https://www.nytimes.com/interactive/2020/health/sinovac-covid-19-vaccine.html [accessed 2024-04-29]
112. The Sinovac-CoronaVac COVID-19 vaccine: what you need to know. World Health Organization. 2022 Jun 10. URL: https://www.who.int/news-room/feature-stories/detail/the-sinovac-covid-19-vaccine-what-you-need-to-know#:~:text=The%20vaccine%20is%20safe%20and,COVID%2D19%20in%20the%20past [accessed 2024-04-29]
113. Lyudmila K. Vaccination with CoronaVac launched in Ukraine on April 13 – Health minister. UNIAN. 2021 Apr 13. URL: https://www.unian.info/society/covid-19-vaccination-with-coronavac-launched-in-ukraine-on-april-13-health-minister-11386060.html [accessed 2024-04-29]
114. Miranda NA. Chile receives 2 million-dose first delivery of Sinovac COVID-19 vaccine. Reuters. 2021 Jan 28. URL: https://www.reuters.com/business/healthcare-pharmaceuticals/chile-receives-2-million-dose-first-delivery-sinovac-covid-19-vaccine-2021-01-28/ [accessed 2024-04-29]

115. Mano A, Simões E. Chinese vaccine draws demand across Latin America, say Brazilian officials. Reuters. 2020 Dec 10. URL: https://www.reuters.com/article/business/healthcare-pharmaceuticals/chinese-vaccine-draws-demand-across-latin-america-say-brazilian-officials-idUSKBN28K1IJ/ [accessed 2024-04-29]
116. Nguyen A. Thailand finally kicks off COVID vaccinations: 5 things to know. NIKKEI Asia. 2021 Mar 01. URL: https://asia.nikkei.com/Spotlight/Coronavirus/COVID-vaccines/Thailand-finally-kicks-off-COVID-vaccinations-5-things-to-know [accessed 2024-04-29]
117. Indonesia starts mass COVID vaccinations over vast territory. The Associated Press. 2021 Jan 13. URL: https://apnews.com/article/asia-pacific-indonesia-coronavirus-pandemic-joko-widodo-3ff30c85dc6943e2a06db470a9d04931 [accessed 2024-04-29]
118. Brazil health ministry says plans to order 30 million more Coronavac doses. Reuters. 2021 Feb 19. URL: https://www.reuters.com/article/business/healthcare-pharmaceuticals/brazil-health-ministry-says-plans-to-order-30-million-more-coronavac-doses-idUSKBN2AJ1D3/ [accessed 2024-04-29]
119. Takenaka K. Japan begins COVID-19 vaccination in 'first major step' to halt pandemic. Reuters. 2021 Feb 16. URL: https://www.reuters.com/article/world/japan-begins-covid-19-vaccination-in-first-major-step-to-halt-pandemic-idUSKBN2AH00O/ [accessed 2024-04-29]
120. Cha S. S.Korea to secure more mRNA vaccines for COVID-19 booster shot. Reuters. 2021 Jun 29. URL: https://www.reuters.com/world/asia-pacific/skorea-secure-more-mrna-vaccines-covid-19-booster-shot-2021-06-29/ [accessed 2024-04-29]
121. Lee IC. Taiwan begins COVID-19 vaccinations. Taipei Times. 2021 Mar 23. URL: https://www.taipeitimes.com/News/front/archives/2021/03/23/2003754315 [accessed 2024-04-24]
122. Taiwan clears AstraZeneca vaccine, shots might start on Monday. Al Jazeera. 2021 Mar 18. URL: https://www.aljazeera.com/news/2021/3/18/taiwan-clears-astrazeneca-vaccine-shots-might-start-on-monday [accessed 2024-04-29]
123. Travelers rush to enter China as it eases pandemic border restrictions. PBS News. 2023 Jan 08. URL: https://www.pbs.org/newshour/world/travelers-rush-to-enter-china-as-it-eases-pandemic-border-restrictions [accessed 2024-04-29]
124. Coronavirus (COVID-19) measures by the government. Japan National Tourism Organization. 2023 May 09. URL: https://www.japan.travel/en/practical-coronavirus-information/government-measures/ [accessed 2024-04-29]
125. COVID-19 and the border. Government of Australia. 2023 Mar 11. URL: https://www.homeaffairs.gov.au/portfolio-functions/travel-and-crossing-the-border/covid-19-and-the-border [accessed 2024-04-29]
126. You want to travel to New Zealand. New Zealand Customs. URL: https://covid19.govt.nz/international-travel/#:~:text=Travellers%20and%20crew%20do%20not,still%20require%20proof%20of%20vaccination [accessed 2024-04-29]
127. COVID-19 travel information. Ministry of Foreign Affairs Singapore. URL: https://www.mfa.gov.sg/Consular-Services/Singapore-Citizens/COVID-19-Information [accessed 2024-04-29]
128. Stein R. As the pandemic ebbs, an influential COVID tracker shuts down. KERA News. 2023 Feb 10. URL: https://www.keranews.org/2023-02-10/as-the-pandemic-ebbs-an-influential-covid-tracker-shuts-down [accessed 2024-04-29]
129. Ritchey MD, Rosenblum HG, Del Guercio K, Humbard M, Santos S, Hall J, et al. COVID-19 self-test data: challenges and opportunities - United States, October 31, 2021-June 11, 2022. MMWR Morb Mortal Wkly Rep 2022 Aug 12;71(32):1005-1010 [FREE Full text] [doi: 10.15585/mmwr.mm7132a1] [Medline: 35951486]
130. Brown RB. Biases in COVID-19 case and death definitions: potential causes and consequences. Disaster Med Public Health Prep 2022 Dec 12;17:e313 [FREE Full text] [doi: 10.1017/dmp.2022.281] [Medline: 36503702]

Abbreviations

GDP: gross domestic product
GISAID: Global Initiative on Sharing All Influenza Data
OWID: Our World in Data
VOC: variant of concern
WHO: World Health Organization

Edited by A Mavragani; submitted 03.10.23; peer-reviewed by S Thawillarp, A Coristine; comments to author 09.10.24; revised version received 06.01.25; accepted 12.01.25; published 21.02.25.

Please cite as:
Lundberg AL, Soetikno AG, Wu SA, Ozer E, Welch SB, Liu Y, Hawkins C, Mason M, Murphy R, Havey RJ, Moss CB, Achenbach CJ, Post LA
Updated Surveillance Metrics and History of the COVID-19 Pandemic (2020-2023) in East Asia and the Pacific Region: Longitudinal Trend Analysis
JMIR Public Health Surveill 2025;11:e53214
URL: https://publichealth.jmir.org/2025/1/e53214
doi:10.2196/53214
PMID:39804185

Availability and Key Characteristics of National Early Warning Systems for Emerging Profiles of Antimicrobial Resistance in High-Income Countries: Systematic Review

Jessica Iera[1,2], MD; Claudia Isonne[1,3], MD; Chiara Seghieri[2], PhD; Lara Tavoschi[4], PhD; Mariateresa Ceparano[1], BSc; Antonio Sciurti[1], MD; Alessia D'Alisera[5], MD; Monica Sane Schepisi[5], MSc, MD; Giuseppe Migliara[6], MD, PhD; Carolina Marzuillo[1], BSc; Paolo Villari[1], MPH, MD; Fortunato D'Ancona[3], MD; Valentina Baccolini[1], MPH, MD

[1]Department of Public Health and Infectious Diseases, Sapienza University of Rome, Rome, Italy
[2]Department EMbeDS, Sant'Anna School of Advanced Studies, Management and Health Laboratory, Institute of Management, Pisa, Italy
[3]Department of Infectious Diseases, Istituto Superiore di Sanità, Rome, Italy
[4]Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy
[5]Italian Ministry of Health, General Directorate for Health Prevention, Rome, Italy
[6]Department of Life Sciences, Health and Health Professions, Link Campus University, Rome, Italy

Corresponding Author:
Jessica Iera, MD
Department of Public Health and Infectious Diseases, Sapienza University of Rome, Rome, Italy

Abstract

Background: The World Health Organization (WHO) recently advocated an urgent need for implementing national surveillance systems for the timely detection and reporting of emerging antimicrobial resistance (AMR). However, public information on the existing national early warning systems (EWSs) is often incomplete, and a comprehensive overview on this topic is currently lacking.

Objective: This review aimed to map the availability of EWSs for emerging AMR in high-income countries and describe their main characteristics.

Methods: A systematic review was performed on bibliographic databases, and a targeted search was conducted on national websites. Any article, report, or web page describing national EWSs in high-income countries was eligible for inclusion. EWSs were identified considering the emerging AMR-reporting WHO framework.

Results: We identified 7 national EWSs from 72 high-income countries: 2 in the East Asia and Pacific Region (Australia and Japan), 3 in Europe and Central Asia (France, Sweden, and the United Kingdom), and 2 in North America (the United States and Canada). The systems were established quite recently; in most cases, they covered both community and hospital settings, but their main characteristics varied widely across countries in terms of the organization and microorganisms under surveillance, with also different definitions of emerging AMR and alert functioning. A formal system assessment was available only in Australia.

Conclusions: A broader implementation and investment of national surveillance systems for the early detection of emerging AMR are still needed to establish EWSs in countries and regions lacking such capabilities. More standardized data collection and reporting are also advisable to improve cooperation on a global scale. Further research is required to provide an in-depth analysis of EWSs, as this study is limited to publicly available data in high-income countries.

(JMIR Public Health Surveill 2025;11:e57457) doi:10.2196/57457

KEYWORDS

early warning system; surveillance; emerging AMR; high-income countries; antimicrobial resistance

Introduction

Antimicrobial resistance (AMR) is a global public health emergency [1,2] that increases the number of deaths, length of hospital stays, and health care costs expenditure [3-8]. Worldwide, a total of 4.95 million deaths per year are associated with bacterial AMR, 1.27 million of which are directly attributable to it [9]. Gram-negative bacteria resistant to last-resort antibiotics such as carbapenem-resistant *Acinetobacter baumannii* and carbapenem-resistant and third-generation cephalosporin-resistant *Enterobacterales* have been listed as of critical priority for public health measures, due to their significant global burden and the ability to transfer resistance genes [10]. As the threat of AMR continues to grow, more efforts are needed to strengthen infection prevention and control strategies as well as to enhance surveillance systems [11].

AMR surveillance plays a pivotal role in strengthening health system resilience and preparedness, thereby supporting effective antibiotic stewardship and leading to optimized patient health outcomes [12]. Enhanced investments are considered essential to improve the comparability, quantity, and quality of AMR data [13]. Current trends, such as the rising prevalence of carbapenem-resistant *Acinetobacter spp.* isolates, underline the need to intensify efforts for the early detection of drug resistance [13]. Several international initiatives aimed at improving epidemiological and microbiological discussion for coordinated actions are ongoing [14,15]. Within this context, the World Health Organization (WHO) launched the Global Antimicrobial Resistance and Use Surveillance System (GLASS) in 2015 [16,17], included the development of national surveillance systems capable of timely detecting and reporting emerging resistance among the goals outlined in the Global Plan of Action on AMR published in 2016 [18], and more recently, developed a focused surveillance with a specific reporting component for novel (emerging) AMR (ie, GLASS-EAR), in order to support early data sharing and coordinated actions among member states [19].

However, while well-established routine AMR surveillance systems are available in high-income countries (HIC), implementing AMR surveillance has been particularly challenging in low- and middle-income countries, where the number of surveillance sites contributing to national surveillance is often not representative [20,21], and a limited number of referral hospitals report AMR data to GLASS [22]. Therefore, studies have mainly focused on countries where the laboratory capabilities required to strengthen national surveillance systems are more likely to be in place (ie, HIC) and publicly available information on the implementation of tools for the early detection and reporting of emerging AMR profiles is often fragmented [23]. Even when findings are available, understanding these systems is challenging due to different approaches used for data collection, reporting, and interpretation of definitions [24,25]. Given the lack of a clear and comprehensive overview on the topic, the aim of our study was to map existing national early warning systems (EWSs) for emerging AMR in HIC as well as describe and compare their main characteristics. Findings from our study could provide relevant information to stakeholders engaged in AMR surveillance, informing the development and/or optimization of EWSs in several countries, including low- and middle-income countries.

Methods

Ethical Considerations

This study was performed according to the Cochrane Handbook for Systematic Reviews and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement (Checklist 1) [26,27]. The protocol was registered on the Open Science Framework via the Center for Open Science (210 Ridge McIntire Road, Suite 500, Charlottesville, VA 22903 - 5083; identifier: pr6a8). As the study did not involve the collection of primary data, it did not require informed consent or the submission for institutional review board approval.

Search Strategy and Study Selection

Three bibliographic databases (ie, PubMed, Web of Science, and Scopus) were searched. Search strings were adapted to fit the search criteria of each database (Table S1 in Multimedia Appendix 1). Taking into account the topic under investigation and the expected challenges in retrieving information, we intentionally developed broad search strings, prioritizing sensitivity over specificity. The search was conducted among records published from database inception to July 3, 2024, without restrictions such as language or date. After the removal of duplicate articles, language restriction was applied in the context of the title and abstract of all the retrieved records during screening, according to the eligibility criteria described below. The full texts of potentially relevant articles were examined by three researchers, who resolved disagreements through discussion and recorded reasons for exclusion. In addition, relevant national websites (ie, the Ministries of Health and/or National Health Institutes websites) were also explored, focusing on thematic areas involving AMR surveillance (Table S2 in Multimedia Appendix 1).

Eligibility Criteria

Considering the Emerging Antimicrobial Resistance Reporting framework within the GLASS developed by the WHO [19], EWSs were defined, for the purposes of this study, as any system able to timely detect, provide verification, and report emerging AMR events. This broad definition was adopted as we expected to encounter systems with diverse characteristics, including variations in design, data flow requirements, implementation, and functionality. Emerging AMR was defined as unusual AMR findings in bacteria and fungi causing infections in humans with the potential impact on public health, such as new types of phenotypic resistance (not previously reported or very rare) and new genetic determinants of AMR that may have high potential for spread. As a specific resistant phenotype/genotype may be emerging in one area but may already be endemic in another one, a list of AMR events was not developed by considering both newly detected AMR and known resistance profiles detected for the first time in new geographical areas.

We included articles, web pages, and/or technical reports with the following characteristics: (1) reported in English, French, Spanish, or Italian, aiming at broadening our investigation by taking into account the language capabilities of the co-authors; (2) describing national EWSs for emerging AMR currently implemented or under development; and (3) referring to HIC, as identified by the World Bank classification of countries by income [28]. We excluded articles, web pages, and/or technical reports that (1) described only standard national AMR surveillance systems, for example routine surveillance systems that retrospectively analyzed data on an annual basis and/or that lacked active alert components; (2) focused on EWSs limited to subnational levels; or (3) reported data in languages other than English, French, Spanish, or Italian.

Data Collection and Synthesis

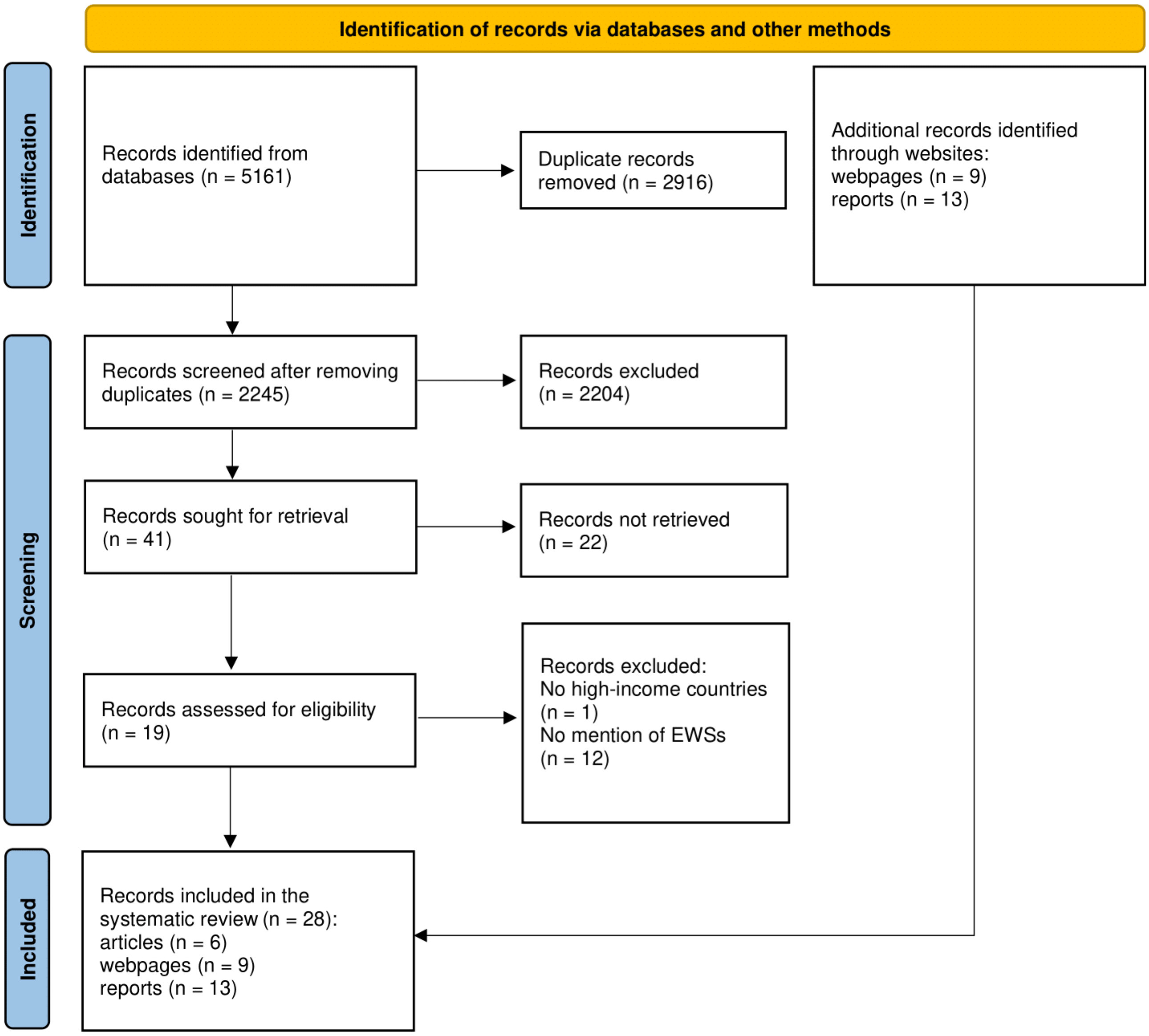
For each record, two reviewers independently retrieved the following information using a standardized data abstraction form: identification of the report/study, if applicable (title, first author, year of publication, and DOI); country; key characteristics of the EWS (eg, coordinating institution, year of establishment, sector involved—human/animal/food/environment, setting where specimens are collected, specimen types, and availability of genomic data); key characteristics of the microorganisms under surveillance (eg, resistance profiles); alert timing (eg, real-time, daily, or weekly); and conduction of performance evaluation of the EWS. Countries were grouped by region, according to the World Bank classification. Data were descriptively synthesized through tables and graphical representations using R software (version 4.3.1; R Foundation for Statistical Computing).

Results

Records Identification and Screening

Overall, 5161 records were identified via database searching (Figure 1). After duplicate removal and screening by the title and abstract, 41 articles were eligible for full-text analysis. Of these, 35 were excluded with reasons, providing a total of 6 articles ultimately included in the systematic review [24,29-33]. A targeted search on relevant national websites allowed for the identification of thematic areas involving AMR from 72 out of 86 HIC (15 countries were excluded due to language restrictions or the inability to retrieve websites), with 22 records finally included in the systematic review (ie, 9 web pages [34-42] and 13 reports [25,43-54]).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the review process. EWSs: early warning systems.

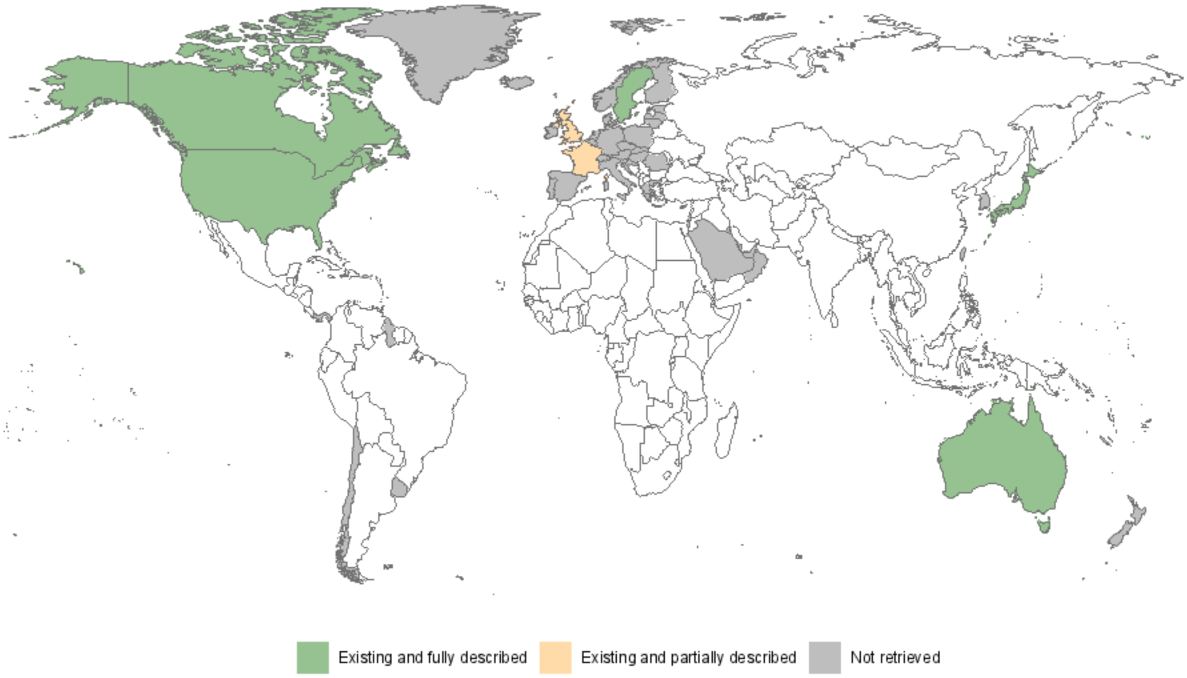


Availability of National EWSs for Emerging AMR in HIC

Overview

Nine web pages and 19 publications described 7 national EWSs in 72 (9.7%) HIC (Figure 2). Of these, 2 were in the East Asia and Pacific Region (ie, Australia [29,34,44,45,53,54] and Japan [24,30,31,33]), 3 were in Europe and Central Asia (ie, France [36-38,48,49], Sweden [24,39,51,52], and the United Kingdom [32,40,46]), and 2 were in North America (ie, United States [25,41,42,50] and Canada [35,43,47]). A partial description of the systems was available in 2 cases (ie, the United Kingdom and France).

Figure 2. Availability of early warning systems for emerging antimicrobial resistance at the national level in high-income countries.



East Asia and Pacific Region

Australia: National Alert System for Critical Antimicrobial Resistances

The National Alert System for Critical Antimicrobial Resistances (CAR-Alert) was established by the Australian Commission on Safety and Quality in Health Care in March 2016 as part of the Antimicrobial Use and Resistance in Australia (ie, AURA) Surveillance System and reviewed in 2018 and 2022 (Table 1) [54]. By collecting data on nationally agreed priority microorganisms with critical resistance to last-line antimicrobial agents, it is a coordinated national system that allows both the collection and the communication of information on confirmed critical antimicrobial resistance (CAR) and potential outbreaks, as close as possible to the time of confirmation [29,34]. At the national level, 28 confirming laboratories provided data to the system in 2021 and 2022, with at least 1 confirming laboratory in each state and territory [53]. In 2023, 26 confirming laboratories participated in the CAR-Alert [54].

Table . National EWSs[a] for emerging AMR[b] identified in high-income countries.

Country	EWS	Institution	Year of establishment	Brief description
East Asia and Pacific Region				
Australia	CAR-Alert National Alert System for Critical Antimicrobial Resistances	Australian Commission on Safety and Quality in Health Care	2016 Revised in 2018 and 2022	CAR-Alert allows for the timely surveillance of critical nationally agreed antimicrobial resistances, integrating the AURA[c] surveillance system. The system involves laboratories conducting confirmatory susceptibility testing on human isolates.
Japan	JANIS-CL Japan nosocomial infection surveillance-Clinical Laboratory division	National Institute of Infectious Diseases	2000 Revised in 2007	The system involves several laboratories across the country. Alert emails are automatically sent to contact personnel of facilities that report unusual combinations of bacterial species and antimicrobial susceptibility, defined as “unusual AMR.”
North America				
Canada	AMRnet Canadian Antimicrobial Resistance Network	Public Health Agency of Canada	NA[d]	Laboratory-based surveillance system capturing information on AMR susceptibility testing from laboratory information systems in both public and private clinical and veterinary laboratories, including reference laboratories. The system aims at collecting and disseminating real time data to allow stakeholders response to emerging AMR issues.
United States	AR Lab Network Antimicrobial Resistance Laboratory Networks	Center for Disease Control and Prevention	2016	Provides nationwide lab capacity to rapidly detect antimicrobial resistance. Laboratories cooperate to rapidly identify AMR profiles causing hard-to-treat or potentially untreatable infections. Protocols to immediately notify health department, health care provider, and infection control staff of unusual resistance are provided.
Europe and Central Asia				

Country	EWS	Institution	Year of establishment	Brief description
France	e-SIN Signalement Externe des Infections Nosocomiales	Ministry of Health	2011 Revised in 2017	Application that allows the monitoring of infectious events in health care facilities, providing data on isolated cases/outbreaks of bacteria with particular AMR profiles. A specific reporting form for emerging highly resistant bacteria is available since 2017. The National Reference Center for Antibiotic Resistance is integrated into this alert system, in particular for events involving emerging mechanisms of resistance. Reports may result in national or international health alerts.
Sweden	Svebar	Public Health Agency of Sweden	2015	Svebar consists of an IT system for early alerts and continuous resistance monitoring. All culture findings from the country's laboratories are automatically transferred on a daily basis, allowing an early warning on findings of serious antibiotics resistance.
United Kingdom	Antimicrobial resistance alerts (ARAs)	Public Health England Antimicrobial resistance and healthcare associated infections reference unit (AMRHAI)	NA	AMRHAI detects new and emerging AMR via interpretive analysis of MIC[e] profiles/molecular investigation. Unusual isolates are sent for testing from diagnostic laboratories to AMRHAI, where an appropriate reference unit confirms relevant isolates resistance. If deemed of public health impact, ARAs inform microbiologists of emerging AMR pathogen strains that could spread in the United Kingdom health service.

[a] EWS: early warning system.

[b] AMR: antimicrobial resistance.

[c] AURA: Antimicrobial Use and Resistance in Australia.

[d] NA: not available.

[e] MIC: minimum inhibitory concentration.

In order to identify and confirm potential CAR, pathology laboratories perform a routine three-step process based on (1) isolate collection from the patient and routine testing; (2) confirmation by a laboratory with the capacity to identify CAR, if a critical AMR profile is suspected by the originating laboratory; and (3) data submission to the CAR-Alert web portal by the confirming laboratory and feedback to the originating laboratory, allowing the implementation of proper infection prevention and control by the health service caring for the patient [44,45].

Japan: Japan Nosocomial Infections Surveillance - Clinical Laboratory Division

The Ministry of Health, Labour, and Welfare of Japan provided two national surveillance systems to monitor AMR in bacteria, including the Japan Nosocomial Infections Surveillance (ie, JANIS) (Table 1) [31]. JANIS was launched in 2000 as a voluntary surveillance system focusing on infections in health care settings and includes several divisions [30,33]. Among them, the JANIS clinical laboratory division (JANIS-CL) specifically focuses on AMR bacteria, and it collects comprehensive specimen-based data from participating hospitals on a monthly basis through a member-restricted website.

Submitted files from participating hospitals are automatically processed to check data structure, interpret antimicrobial susceptibility, remove duplicates, and calculate the prevalence of AMR [30]. The system has the capacity to detect unusual combinations of bacterial species and antimicrobial susceptibility, defined as unusual AMR, and generates alert emails that are automatically sent to the contact personnel of the facilities [31].

North America

Canada: Antimicrobial Resistance Network

The Antimicrobial Resistance Network (AMRNet) is a collaboration between the Public Health Agency of Canada, provincial and territorial public health organizations, and clinical and veterinary laboratories across the country (Table 1). The system has been designed to detect the emergence and spread of AMR and to timely disseminate data to stakeholders, in order to address emerging AMR issues in Canada [35,47].

Originally started as a series of pilot projects, the AMRNet transitioned to collecting routine data (antimicrobial susceptibility testing results from bacterial and fungal pathogens, together with a defined set of patient or animal characteristics) from a select group of provinces in 2022. Currently, data collection involves approximately 1.5 million bacterial and fungal isolates per year from Ontario, Saskatchewan, and Prince Edward Island [43].

United States: The Antimicrobial Resistance Laboratory Network

The Antimicrobial Resistance Laboratory Network (AR Lab Network) was established and is coordinated by the Centers for Disease Control and Prevention (CDC) to rapidly detect existing and emerging resistance nationwide (Table 1). The network includes 7 regional laboratories, the National Tuberculosis Molecular Surveillance Center, and laboratories across 50 states, 5 cities, and Puerto Rico [25,50]. Laboratories cooperate, from the local to the national level, to provide the rapid identification and response to antimicrobial-resistant genes and microorganisms that cause hard-to-treat or potentially untreatable infections. When unusual resistance is detected, the Healthcare and Clinical Laboratories notify the Health Department. Public health laboratories then confirm the bacterial species identity, perform additional testing to characterize the isolates received, provide notification of unusual resistance to CDC and, when appropriate, send the isolates to regional laboratories for additional testing [42]. Every regional laboratory performs core testing, including the detection of new and emerging threats. If needed, regional laboratories send unusual isolates to the CDC to confirm testing and conduct additional tests. When threats are reported, the CDC provides infection prevention and control recommendations and supports outbreak responses [41].

Europe and Central Asia

France: E-SIN (Signalement Externe Des Infections Nosocomiales - External Reporting on Healthcare-Associated Infections)

Since 2011, the reporting of health care–associated infections, initially introduced in 2001, has been based on a digital tool called e-SIN (Signalement Externe Des Infections Nosocomiales - External Reporting on Healthcare-Associated Infections; Table 1) [48]. The e-SIN application monitors infectious cases reported by health care facilities and allows for the identification of national or even international health alerts, thereby providing data on isolated cases or outbreaks of bacteria with exceptional AMR profiles such as emerging highly resistant bacteria (defined as microorganisms with susceptibility to only one or two classes of antibiotics) [37,38,49]. The reporting process, moreover, enables a simultaneous reporting to both the Regional Coordination Centers for Healthcare-Associated Infections and the Regional Health Agency and includes an internal alert system that allows the notification of the members of the Epidemiology and Hospital Hygiene Team [36].

Sweden: Svebar

The Public Health Agency of Sweden is responsible for national AMR monitoring and analysis (Table 1) [39]. The IT system Svebar was developed in 2015 to expand and strengthen existing surveillance systems both at the national and local levels [24,52]. This system relies on the voluntary participation of clinical microbiology laboratories. All data on clinical isolates from humans are transferred to Svebar by connected laboratories on a daily basis. Every night, the participating laboratories automatically send to the system a file, saved in a short-term storage, with the culture findings from the past 14 days. Data are reported according to standard definitions and saved in a short-term storage for 13 days before being transferred to a long-term storage, which retains the data from the previous night’s report, thereby receiving more processed data [51]. The system allows the generation of early alerts and the continuous monitoring of serious antibiotic resistance or other undesired changes in resistance conditions, through pre-set alert functions [51]. Currently, 22 laboratories provide data to Svebar [52].

United Kingdom: Antimicrobial Resistance Alerts

The UK antimicrobial resistance alerts (ARAs) provide microbiologists with information of emerging antimicrobial-resistant pathogen strains that could spread in the UK health service (Table 1) [32]. When a suspicious resistant strain is identified, isolates are sent to Public Health England’s AMRHAI reference unit for further testing. The AMRHAI reference unit includes the national reference laboratory for AMR investigation and detection of new and emerging resistances: the unit investigate isolates found to have unusual resistances by diagnostic laboratories, aiming to identify emerging resistance of public health importance, underlying resistance mechanisms and clonal spread of resistant strains [46].

Proposals for ARAs can be submitted by any interested party and should be directed to the co-chairs of a standing committee, which review the proposals and determine its public health implications, in order to assess the necessity of issuing a national alert [40].

Characteristics of National EWSs for Emerging AMR in HIC

Health Sectors Involved, Specimen Type, and Setting

The human sector was involved in all the identified EWSs, and, in the case of Canada, the system also included the animal sector [47]. All the systems reported data from clinical isolates, while 2 also included screening samples (ie, Australia and the United States; Table 2) [45,50].

Table . Main characteristics of national early warning systems for emerging AMR[a] in high-income countries.

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
East Asia and Pacific Region							

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
Australia (CAR-Alert)	Human	Hospital-based Community-based	Clinical isolate screening		Yes	Weekly summary alert emails	Yes

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
				• Carbapenemase-producing *Acinetobacter baumannii* complex • *Candida auris* • Carbapenemase-producing *Enterobacterales* and/or ribosomal methyltransferase-producing *Enterobacterales*, with transmissible colistin-resistance • Linezolid-resistant *Enterococcus* species • Multidrug-resistant *Mycobacterium tuberculosis* – resistant to at least rifampicin and isoniazid • Ceftriaxone- and/or azithromycin-nonsusceptible *Neisseria gonorrhoeae*; gentamicin-resistant *Neisseria gonorrhoeae* • Ciprofloxacin-nonsusceptible *Neisseria meningitidis* • Carbapenemase-producing *Pseudomonas aeruginosa* • Ceftriaxone-nonsusceptible *Salmonella* species			

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
				- Multidrug-resistant *Shigella* species - Vancomycin- or linezolid-nonsusceptible *Staphylococcus aureus* complex (*S argenteus* and *S schweitzeri*) - Penicillin-reduced susceptibility *Streptococcus pyogenes*			
Japan (Janis-CL)	Human	Hospital-based Community-based	Clinical isolates	- AMR never officially reported (eg, vancomycin-resistant *Staphylococcus aureus*) - AMR reported but rare (eg, multidrug-resistant *Acinetobacter spp*. and vancomycin-resistant enterococci)	Yes	Alert emails within 1 - 2 h	NA[b]
North America							

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
Canada (AMRnet)	Human Animal	Hospital-based Community-based	Clinical isolates	- All bacterial and fungal pathogens tested for AMR, including priority pathogens[c], as classified by the World Health Organization	NA	NA	NA
United States (AR Lab Network)	Human	Hospital-based Community-based	Clinical isolate screening	- Carbapenemase-producing organism (CPO), including carbapenem-resistant *Enterobacterales* (CRE) - *Candida* species, to identify emerging resistance - Emerging and concerning threats, such as mcr-1 and carbapenem-resistant *Acinetobacter baumannii*, and changes in known threats, including MRSA - Pandrug-resistant (PDR) bacteria to new antibiotics	Yes	NA	NA
Europe and Central Asia							

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
France (e-SIN)	Human	Hospital-based	Clinical isolates	• Carbapenemase-producing *Enterobacterales* • Glycopeptide-resistant *Enterococcus faecium*	NA	NA	NA
Sweden (Svebar)	Human	Hospital-based Community-based	Clinical isolates	• Bacterial species resistant to one or more antibiotics (ie, *E coli* resistant to carbapenems) • A trend (ie, resistance to ampicillin in more than 30 percent of cases of *E coli*).	Yes	Daily	NA
United Kingdom (ARAs)	Human	Hospital-based	Clinical isolates		Yes	NA	NA

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
				The AMRHAI[d] advises the referral of isolates with exceptional resistance phenotypes. A list of unusual combinations of resistance and organisms is available for diagnostic laboratories[e], including the following microorganisms: • *Acinetobacter spp*. • Coagulase-negative *staphylococci* • *Enterobacterales* • *Enterococci* • *Haemophilus influenzae* • *Moraxella catarrhalis* • Organisms or antibiotics for which there are no EUCAST[f] clinical breakpoints (invasive sites) • *Pseudomonas aeruginosa* • *Staphylococcus aureus* • *Streptococci* (groups A, B, C, and G, β-haemolytic) • *Streptococcus pneumoniae*			

---

[a] AMR: antimicrobial resistance.

[b] NA: not available.

[c] *Acinetobacter* spp; *Escherichia coli*; *Klebsiella pneumoniae*; *Neisseria gonorrhoeae*; *Salmonella* spp; *Shigella* spp; *Staphylococcus aureus*; *Streptococcus pneumoniae*.

[d] AMRHAI: antimicrobial resistance and healthcare-associated infections reference unit.

[e] United Kingdom Health Security Agency Reference Laboratories Colindale. Bacteriology Reference Department user manual. Appendix 1 (Antimicrobial resistance and mechanisms service). Version 15, October 10, 2022.

[f] EUCAST: European Committee on Antimicrobial Susceptibility Testing.

The setting where specimens are collected was hospital based and community based in all the identified systems except for e-SIN (France) and ARAs (United Kingdom), where the collection was mainly hospital based [36,46]. More specifically, CARAlert (Australia) was the most comprehensive alert system, including public and private hospitals, general practices, aged care homes, community health services, and hospital nonadmitted care services [54]; JANIS-CL (Japan) included outpatients and inpatients data from large (usually tertiary care) and small hospitals (usually privately owned hospitals with less than 200 beds, providing both acute and long-term care) [31]; and AMRnet (Canada) included AMR data from clinical laboratories in both public and private facilities [43]. Detailed published information was not available referring to the remaining EWSs.

AMR Events Reported and Availability of Genomic Data

Reported AMR events varied widely among the identified EWSs (Table 2). In the majority of instances, emerging AMR was classified in terms of broader categories of unusual events (eg, AMR never previously officially reported, AMR reported but rarely, or a trend). Notably, among the 7 systems examined, 3 systems (ie, Australia, France, and the UK) provided a list of unusual combinations of microorganisms and resistance, including in most cases, carbapenemase-producing *Acinetobacter baumannii* and carbapenemase-producing *Enterobacterales* [38,44,46]. Two countries (ie, Australia and the United States) included emerging fungal pathogens such as *Candida auris* in their systems [25,44].

Genomic data were provided by 5 out of 7 EWSs, while this information was not retrievable for 2 countries, specifically Canada and France [25,33,44,46,52].

Alert Functions

Alert timing and functioning were clearly defined in 3 out of 7 EWSs (ie, Australia, Japan, and Sweden), with several differences among these systems. Concerning CAR-Alert (Australia), data are submitted within 7 days of the isolate being confirmed as a CAR (Table 2). The system generates a weekly summary email alert to report information on confirmed CARs to state and territory health authorities, the Australian Government Department of Health, and confirming laboratories, supporting timely responses to CARs by hospitals and state and territory health departments [45].

In the case of JANIS-CL (Japan), a data format was developed as a unified standard for collecting electronic data from microbiology laboratories. Microbiological alerts can be checked on the JANIS member-restricted website within a few hours after data submission. Furthermore, alert emails reporting unusual profiles of AMR are automatically sent within 1 or 2 hours to facilities contact personnel while monthly feedback emails are sent within 48 hours [30].

Regarding Svebar (Sweden), preset alerting algorithms search through the short-term storage and generate an alert in specific conditions, for example when the system discovers a trend. National alerts are sent via e-mail to the contact person at the local laboratory and to administrators at the Public Health Agency of Sweden, and if needed, they can discuss the alert, while local alerts only reach the contact person in the local laboratory. By systematically tracking the progression of resistance, the detection of early-stage changes might justify modifications to the alert system’s settings [51].

Performance Evaluation

Based on publicly available information, the formal assessment of national EWSs has been conducted in only 1 country, specifically Australia, where the Department of Health and Aged Care maintains a routine process of evaluating national surveillance systems to ensure they align with their intended purpose and objectives (Table 2). The evaluation of CAR-Alert took place in 2022 - 2023, with the aim of assessing the system’s operational efficiency in fulfilling its objectives, evaluating their appropriateness, and identifying potential enhancements to improve the system’s ability to achieve its goals [54].

To assess the system’s overall effectiveness and performance, the evaluation process was based on the “Updated guidelines for evaluating public health surveillance systems” from the CDC [55].

Discussion

Principal Findings

Our systematic review allowed the identification of national EWSs for emerging profiles of AMR in a limited number of HIC, approximately 10%. In addition, only half of the 6 World Bank regions that we investigated had functional EWSs in place, revealing a disparity in the availability of these systems at the global level. Given that several regions, including Latin America and the Caribbean, are currently facing alarming developments with national authorities raising concerns about the emergence of carbapenemase-producing *Enterobacterales* not previously described, or about the increasing number of isolates that coexpress two or more of these enzymes [56], the lack of available information on national EWSs in these countries is concerning. Even in the European region, where countries are mostly high-income and have advanced surveillance and health care systems, national EWSs were not retrievable in those areas where AMR rates have been found to be extremely high by the European Center for Disease Prevention and Control (ECDC) [13]. Such an absence of EWSs in high AMR burden countries reveals a critical gap in the region’s preparedness and response to the growing challenge of AMR [54], making it difficult for these nations to not only track the evolution of resistance patterns over time, but also early detect and mitigate the spread of unusual resistant bacteria, implement targeted interventions in high-risk settings, and contribute to shared data on a global scale [19]. In 2021, the ECDC launched an online portal (EpiPulse) for European public health authorities, aiming at enhancing the early detection and assessment of threats due to infectious diseases. However, its effectiveness relies on information from member states concerning microbiological alerts, and without national EWSs in place, the management of threats at the international level becomes challenging [15].

Comparison With Previous Studies

Notably, while identified EWSs were established quite recently and shared the overarching goal of monitoring and responding to emerging AMR threats, our study showed that their main characteristics varied significantly across countries. This heterogeneity might be attributable to several factors, including the unique health care infrastructure, different public health policies and funding, and the specific challenges and priorities faced by each country [57-59]. The heterogeneity in identified EWSs also reflects the specific AMR trends and epidemiological situation within each state, considering that AMR profiles can vary significantly, being unusual in one country and endemic in another [60]. Nevertheless, some common ground in monitoring, understanding, and responding to AMR emerged in a few aspects, especially if we consider the inclusion of the community setting and the availability of genomic data in most of the identified EWSs. Several reasons can explain this approach: community-based surveillance is crucial because resistant microorganisms can spread within health care and community settings, with early detection allowing the proper prevention and control of outbreaks [61], while genomic data play a pivotal role in understanding AMR trends and developing targeted prevention and control interventions [62]. As for the availability of performance evaluation of EWSs, which is crucial to assess their effectiveness, we found that Australia was the only country known to conduct a system evaluation in accordance with the CDC guidelines, even if related data are not yet available [54]. Considering that regular evaluations can lead to improvements in the system design, data collection, and response mechanisms, and ultimately enhance the ability to respond effectively to emerging AMR profiles [55], more efforts should be made to institutionalize regular assessments as an integral part of any strategy aimed to manage AMR.

Lastly, although the existence of EWSs for emerging AMR was not clearly reported in most cases, this review found that many countries indicated, in their National Action Plans against AMR, the goal of strengthening national surveillance through the implementation of these systems in the next years. This is comforting, as this underlines their commitment in addressing this issue, in line with the WHO recommendations [18,63-65]. Furthermore, as the landscape of infectious disease surveillance is rapidly evolving owing to the advancements in new technologies, there might be a significant potential for progress in bolstering the efforts against emerging AMR [66]. Artificial intelligence, for example, could enhance AMR surveillance systems given its ability to integrate and analyze data from various sources, such as clinical data and microbiology reports, recognize AMR patterns, develop predictive models that signal the early emergence of resistant strains or the spread of resistance in specific regions or health care facilities, or provide real-time insights into AMR trends [67]. The incorporation of machine learning algorithms to predict trends in resistance development based on historical data could also improve AMR surveillance systems capacity in early warning [68]. Therefore, as artificial intelligence technologies continue to improve, they are definitely an option to be considered as a supporting tool in the early detection of emerging AMR in the near future, enabling a more timely and effective response also by facilitating the sharing and analysis of data on a global scale [69].

Limitations

To the best of our knowledge, this is the first synthesis review of publicly available information on national EWSs for emerging AMR in HIC. The limitations of our study mostly rely on the inclusion criteria we adopted. First, by focusing on human surveillance, our study did not include other relevant sectors in the context of the “One Health” approach [70,71]. Second, by limiting evidence collection to HIC, we did not address challenges faced by countries, such as low- and middle-income countries, where AMR has a disproportionate impact due to higher burden of infections, reduced laboratory capabilities, and limited regulations involving antimicrobial use [72] Third, the restriction to the national level did not account for regional or local-level EWSs, which could be available in some countries [23]. However, the decision to focus on national rather than subnational EWSs relies on the aim of mapping systems capable of centralizing AMR-related data through a comprehensive and coordinated approach, facilitating global data sharing and response strategies, in accordance with the WHO recommendations [18]. Moreover, we are aware that, as this study relies on publicly available information, and not all EWSs might be documented or publicly disclosed [73], our review may not fully represent the extent of EWSs across different countries and across various levels of implementation, such as national and subnational levels. Similarly, the partial availability of information on a few systems limited our ability to fully describe the different strategies adopted in the investigated areas. Nevertheless, our interest was focused on mapping the efforts made by countries at the national level, and we simultaneously searched multiple data sources as an attempt to retrieve all available information. Findings from our study could support countries currently lacking EWSs in the process of strengthening national surveillance systems and contribute to global effort, through improved detection, reporting, and data sharing, in tackling AMR. Additional information is certainly required to provide a complete overview on available EWSs. To this end, our review could be integrated with a country consultation and/or international survey aimed at retrieving additional findings from relevant stakeholders, such as the Ministries of Health and National Health Institutes representatives. Further investigations could also provide an in-depth analysis on the availability of EWSs at subnational (local or regional) levels and in low- and middle-income countries.

Conclusions

These findings highlight the urgent need for a broader implementation of surveillance systems that allow for the early detection of emerging AMR, with increased investments and collaborative efforts to establish EWSs in countries and regions lacking such capabilities to date. Platforms such as EpiPulse within the European region could enhance the collection and analysis of AMR surveillance data, taking into account that each country should implement tools capable of detecting and managing alert signals. In addition, given the heterogeneity in national health care systems and their different epidemiological contexts, tailored approaches that enable the collection of standardized and comparable AMR data are strongly encouraged to help promote the global preparedness for AMR.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategies used in the systematic review and high-income countries' institutional websites explored.

[PDF File, 197 KB - publichealth_v11i1e57457_app1.pdf]

Checklist 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[PDF File, 81 KB - publichealth_v11i1e57457_app2.pdf]

References

1. Antimicrobial resistance. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance [accessed 2024-09-30]
2. About antimicrobial resistance. Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/antimicrobial-resistance/about/index.html [accessed 2024-09-30]
3. O'Neill J. Tackling drug-resistant infections globally: final report and recommendations. : Review on Antimicrobial Resistance; 2016 URL: https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf [accessed 2024-12-24]
4. O'Neill J. Antimicrobial resistance: tackling a crisis for the health and wealth of nations the review on antimicrobial resistance. : Review on Antimicrobial Resistance; 2014 URL: https://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf [accessed 2024-12-24]
5. Antibiotic resistance threats in the United States, 2019. : Centers for Disease Control and Prevention (CDC); 2019 URL: https://stacks.cdc.gov/view/cdc/82532 [accessed 2024-09-30]
6. Naylor NR, Atun R, Zhu N, et al. Estimating the burden of antimicrobial resistance: a systematic literature review. Antimicrob Resist Infect Control 2018;7:58. [doi: 10.1186/s13756-018-0336-y] [Medline: 29713465]
7. Cassini A, Högberg LD, Plachouras D, et al. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. Lancet Infect Dis 2019 Jan;19(1):56-66. [doi: 10.1016/S1473-3099(18)30605-4] [Medline: 30409683]
8. Migliara G, Di Paolo C, Barbato D, et al. Multimodal surveillance of healthcare associated infections in an intensive care unit of a large teaching hospital. Ann Ig 2019;31(5):399-413. [doi: 10.7416/ai.2019.2302] [Medline: 31304521]
9. Murray CJL, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet 2022 Feb;399(10325):629-655. [doi: 10.1016/S0140-6736(21)02724-0]
10. WHO Bacterial Priority Pathogens List, 2024: bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance. : World Health Organization; 2024.
11. Hay SI, Rao PC, Dolecek C, et al. Measuring and mapping the global burden of antimicrobial resistance. BMC Med 2018 Jun 4;16(1):78. [doi: 10.1186/s12916-018-1073-z] [Medline: 29860943]
12. Tacconelli E, Sifakis F, Harbarth S, et al. Surveillance for control of antimicrobial resistance. Lancet Infect Dis 2018 Mar;18(3):e99-e106. [doi: 10.1016/S1473-3099(17)30485-1] [Medline: 29102325]
13. European Centre for Disease Prevention and Control. Antimicrobial resistance in the EU/EEA (EARS-net) - annual epidemiological report 2022. Stockholm. : ECDC; 2023.
14. Emerging antimicrobial resistance reporting: guide for emerging AMR event sharing. : World Health Organization; 2018 URL: https://www.who.int/publications/i/item/9789241514583 [accessed 2024-12-24]
15. EpiPulse - the European surveillance portal for infectious diseases. European Centre for Disease Prevention and Control. 2021. URL: https://www.ecdc.europa.eu/en/publications-data/epipulse-european-surveillance-portal-infectious-diseases [accessed 2024-09-30]
16. Global antimicrobial resistance and use surveillance system (GLASS) report 2022. : World Health Organization; 2022.
17. GLASS manual for antimicrobial resistance surveillance in common bacteria causing human infection. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/9789240076600 [accessed 2024-09-30]
18. Global action plan on antimicrobial resistance. World Health Organization. URL: https://www.who.int/publications/i/item/9789241509763 [accessed 2024-09-30]
19. GLASS emerging antimicrobial resistance reporting framework (GLASS-EAR). World Health Organization. 2018. URL: https://www.who.int/publications/i/item/9789241514590 [accessed 2024-09-30]

20. Sulis G, Sayood S, Gandra S. Antimicrobial resistance in low- and middle-income countries: current status and future directions. Expert Rev Anti Infect Ther 2022 Feb;20(2):147-160. [doi: 10.1080/14787210.2021.1951705] [Medline: 34225545]
21. Gandra S, Alvarez-Uria G, Turner P, Joshi J, Limmathurotsakul D, van Doorn HR. Antimicrobial resistance surveillance in low- and middle-income countries: progress and challenges in eight South Asian and Southeast Asian countries. Clin Microbiol Rev 2020 Sep 16;33(3):e00048-19. [doi: 10.1128/CMR.00048-19] [Medline: 32522747]
22. Implementing the global action plan on antimicrobial resistance: first quadripartite biennial report. : World Health Organization, Food and Agriculture Organization of the United Nations, United Nations Environment Programme and World Organisation for Animal Health; 2023 URL: https://www.who.int/publications/i/item/9789240074668 [accessed 2024-12-23]
23. Iera J, Seghieri C, Tavoschi L, et al. Early warning systems for emerging profiles of antimicrobial resistance in Italy: a national survey. Int J Environ Res Public Health 2023 Apr 24;20(9):5623. [doi: 10.3390/ijerph20095623] [Medline: 37174143]
24. Vong S, Anciaux A, Hulth A, et al. Using information technology to improve surveillance of antimicrobial resistance in South East Asia. BMJ 2017 Sep 5;358:j3781. [doi: 10.1136/bmj.j3781] [Medline: 28874344]
25. About CDC’s AR lab network. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/drugresistance/solutions-initiative/ar-lab-networks.html [accessed 2024-09-30]
26. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions version 6.2 (updated February 2021). : Cochrane; 2021.
27. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
28. Fantom NJ, Serajuddin U. The World Bank’s classification of countries by income. Policy research working paper, no. WPS 7528 Washington, D.C. World Bank Group. URL: http://documents.worldbank.org/curated/en/408581467988942234/The-World-Banks-classification-of-countries-by-income [accessed 2024-09-30]
29. Turnidge JD, Meleady KT. Antimicrobial Use and Resistance in Australia (AURA) surveillance system: coordinating national data on antimicrobial use and resistance for Australia. Aust Health Rev 2018;42(3):272. [doi: 10.1071/AH16238]
30. Tsutsui A, Suzuki S. Japan nosocomial infections surveillance (JANIS): A model of sustainable national antimicrobial resistance surveillance based on hospital diagnostic microbiology laboratories. BMC Health Serv Res 2018 Oct 20;18(1):799. [doi: 10.1186/s12913-018-3604-x] [Medline: 30342499]
31. Suzuki S. A view on 20 years of antimicrobial resistance in Japan by two national surveillance systems: the national epidemiological surveillance of infectious diseases and Japan nosocomial infections surveillance. Antibiotics (Basel) 2021 Sep 30;10(10):1189. [doi: 10.3390/antibiotics10101189] [Medline: 34680770]
32. Johnson AP. Surveillance of antibiotic resistance. Philos Trans R Soc Lond B Biol Sci 2015 Jun 5;370(1670):20140080. [doi: 10.1098/rstb.2014.0080] [Medline: 25918439]
33. Kajihara T, Yahara K, Hirabayashi A, Shibayama K, Sugai M. Japan Nosocomial Infections Surveillance (JANIS): current status, international collaboration, and future directions for a comprehensive antimicrobial resistance surveillance system. Jpn J Infect Dis 2021 Mar 24;74(2):87-96. [doi: 10.7883/yoken.JJID.2020.499] [Medline: 32863357]
34. National alert system for critical antimicrobial resistances (caralert). Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/our-work/antimicrobial-resistance/antimicrobial-use-and-resistance-australia-surveillance-system/national-alert-system-critical-antimicrobial-resistances-caralert [accessed 2024-09-30]
35. One-health surveillance of antimicrobial resistance in Canada: spotlight on AMRNet. Government of Canada. URL: https://health-infobase.canada.ca/amrnet/?utm_source=stakeholder-weekly-newsletter&utm_medium=link&utm_content=en-02-07-23&utm_campaign=srt-stakeholder-weekly-newsletter&wbdisable=true [accessed 2024-09-30]
36. E-sin: signalement externe des infections nosocomiales [Article in French]. Santé publique France. URL: https://www.santepubliquefrance.fr/maladies-et-traumatismes/infections-associees-aux-soins-et-resistance-aux-antibiotiques/infections-associees-aux-soins/articles/e-sin-signalement-externe-des-infections-nosocomiales [accessed 2024-09-30]
37. Un système d’alerte: le signalement externe des infections associées aux soins [Article in French]. Santé publique France. URL: https://www.santepubliquefrance.fr/maladies-et-traumatismes/infections-associees-aux-soins-et-resistance-aux-antibiotiques/resistance-aux-antibiotiques/articles/un-systeme-d-alerte-le-signalement-externe-des-infections-associees-aux-soins [accessed 2024-09-30]
38. Les bhre, bactéries hautement Résistantes émergentes [Article in French]. Centre d’appui pour la prévention des infections associées aux soins. URL: https://www.cpias-ile-de-france.fr/appui/bhre.php [accessed 2024-09-30]
39. Multisector surveillance of antibiotic sales and resistance. The Public Health Agency of Sweden. URL: https://www.folkhalsomyndigheten.se/the-public-health-agency-of-sweden/communicable-disease-control/antibiotics-and-antimicrobial-resistance/overview-of-swedens-one-health-response-to-antibiotic-resistance/swedish-work-against-antibiotic-resistance--a-one-health-approach/multisector-surveillance-of-antibiotic-sales-and-resistance/ [accessed 2024-09-30]

40. UK public health antimicrobial resistance alerts. Public Health England. URL: https://www.gov.uk/government/publications/uk-public-health-antimicrobial-resistance-alerts [accessed 2024-09-30]
41. Antibiotic resistance laboratory network. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/drugresistance/ar-lab-networks/domestic.html [accessed 2024-09-30]
42. How labs work together. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/drugresistance/ar-lab-networks/domestic/how-labs-work-together.html [accessed 2024-09-30]
43. Rudnick W, Mukhi SN, Reid-Smith RJ, et al. Overview of Canada's Antimicrobial Resistance Network (AMRNet): A data-driven One Health approach to antimicrobial resistance surveillance. Can Commun Dis Rep 2022 Nov 3;48(11-12):522-528. [doi: 10.14745/ccdr.v48i1112a05] [Medline: 38173468]
44. National alert system for critical antimicrobial resistances (CARAlert) laboratory handbook. Australian Commission on Safety and Quality in Health Care. 2022. URL: https://www.safetyandquality.gov.au/publications-and-resources/resource-library/caralert-laboratory-handbook-0 [accessed 2024-09-30]
45. CARAlert standard operating procedures. : Australian Commission on Safety and Quality in Health Care; 2022 URL: https://www.safetyandquality.gov.au/publications-and-resources/resource-library/caralert-standard-operating-procedures [accessed 2024-12-24]
46. Doshi N. Bacteriology reference department user manual. GOV-13422. : UK Health Security Agency; 2022.
47. Canadian antimicrobial resistance surveillance system (CARSS) report 2022. Government of Canada. URL: https://www.canada.ca/en/public-health/services/publications/drugs-health-products/canadian-antimicrobial-resistance-surveillance-system-report-2022.html [accessed 2024-09-30]
48. Berger-Carbonne A, et al. La lettre du signalement. Infections associées aux soins. : Santé publique France; 2024.
49. Actualisation des recommandations relatives à la maîtrise de la diffusion des bactéries hautement résistantes aux antibiotiques émergentes (BHRe) [Article in French]. Conseil de la santé publique. URL: https://www.hcsp.fr/explore.cgi/avisrapportsdomaine?clefr=758 [accessed 2024-09-30]
50. National action plan for combating antibiotic-resistant bacteria. Year 5 report. : Task Force for Combating Antibiotic-Resistant Bacteria; 2022.
51. Swedish work on containment of antibiotic resistance. Tools, methods and experiences. : Public Health Agency of Sweden; 2014.
52. Aspevall O, Dobric S, Jagdmann J, Nilsson O, Pringle M. A report on Swedish antibiotic sales and resistance in human medicine (Swedres) and Swedish veterinary antibiotic resistance monitoring (Svarm) editors. Swedish Veterinary Agency. URL: https://www.sva.se/en/what-we-do/antibiotics/svarm-resistance-monitoring/swedres-svarm-reports/ [accessed 2024-09-30]
53. AURA 2023: fifth Australian report on antimicrobial use and resistance in human health. Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/publications-and-resources/resource-library/aura-2023-fifth-australian-report-antimicrobial-use-and-resistance-human-health-report [accessed 2024-09-30]
54. CARAlert annual report: 2023. Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/publications-and-resources/resource-library/caralert-annual-report-2023 [accessed 2024-09-30]
55. Updated guidelines for evaluating public health surveillance systems recommendations from the guidelines working group. Centers for Disease Control and Prevention. 2001. URL: https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm [accessed 2024-09-30]
56. Emergence and increase of new combinations of carbapenemases in enterobacterales in Latin America and the Caribbean. Pan American Health Organization. 2021. URL: https://www.paho.org/sites/default/files/2021-10/2021-October_PHE-Epi%20Alert_carbapenemases_EN.pdf [accessed 2024-12-24]
57. Lessa FC, Sievert DM. Antibiotic resistance: a global problem and the need to do more. Clin Infect Dis 2023 Jul 5;77(Suppl 1):S1-S3. [doi: 10.1093/cid/ciad226] [Medline: 37406051]
58. Saltman RB, Bankauskaite V, Vrangbaek K, editors. Decentralization in Health Care: Strategies and Outcomes: European Observatory on Health Systems and Policies; 2007:1-293 URL: https://eurohealthobservatory.who.int/publications/m/decentralization-in-health-care-strategies-and-outcomes [accessed 2024-12-24]
59. James C, Beazley I, Penn C, Philips L, Dougherty S. Decentralisation in the health sector and responsibilities across levels of government. OECD J Budget 2019 Dec 13;19(3) [FREE Full text] [doi: 10.1787/c2c2058c-en]
60. Bonomo RA, Burd EM, Conly J, et al. Carbapenemase-producing organisms: a global scourge. Clin Infect Dis 2018 Apr 3;66(8):1290-1297. [doi: 10.1093/cid/cix893] [Medline: 29165604]
61. Embracing a one health framework to fight antimicrobial resistance. Organisation for Economic Co-operation and Development (OECD). 2023. URL: https://www.oecd-ilibrary.org/social-issues-migration-health/embracing-a-one-health-framework-to-fight-antimicrobial-resistance_ce44c755-en [accessed 2024-09-30]
62. ECDC public health microbiology strategy 2018 – 2022. European Centre for Disease Prevention and Control (ECDC). URL: https://www.ecdc.europa.eu/sites/default/files/documents/ECDC-public-health-microbiology-strategy-2018-2022.pdf [accessed 2024-09-30]
63. Piano nazionale di contrasto all'Antimicrobico-resistenza (PNCAR) 2022-2025 [Article in Italian]. : Ministero della Salute; 2022 URL: https://www.salute.gov.it/imgs/C_17_pubblicazioni_3294_allegato.pdf [accessed 2024-09-30]

64. New Zealand antimicrobial resistance action plan. Ministry of Health. URL: https://www.health.govt.nz/strategies-initiatives/programmes-and-initiatives/antimicrobial-resistance/nz-antimicrobial-resistance-action-plan [accessed 2024-09-30]
65. Ireland’s national action plan on antimicrobial resistance 2017 - 2020. : Department of Health; 2017 URL: https://assets.gov.ie/9519/afcba9bce7c54bf9bcbe9a74f49fdaf2.pdf [accessed 2024-09-30]
66. Digital technologies for infectious disease surveillance, prevention and control - a scoping review of the research literature 2015−2019. : European Centre for Disease Prevention and Control URL: https://www.ecdc.europa.eu/en/publications-data/digital-technologies-surveillance-prevention-and-control-infectious-diseases [accessed 2024-09-30]
67. Brownstein JS, Rader B, Astley CM, Tian H. Advances in artificial intelligence for infectious-disease surveillance. N Engl J Med 2023 Apr 27;388(17):1597-1607. [doi: 10.1056/NEJMra2119215] [Medline: 37099342]
68. Sakagianni A, Koufopoulou C, Feretzakis G, et al. Using machine learning to predict antimicrobial resistance-a literature review. Antibiotics (Basel) 2023 Feb 24;12(3):452. [doi: 10.3390/antibiotics12030452] [Medline: 36978319]
69. MacIntyre CR, Chen X, Kunasekaran M, et al. Artificial intelligence in public health: the potential of epidemic early warning systems. J Int Med Res 2023 Mar;51(3):3000605231159335. [doi: 10.1177/03000605231159335] [Medline: 36967669]
70. A One Health Priority Research Agenda for Antimicrobial Resistance: World Health Organization, Food and Agriculture Organization of the United Nations, United Nations Environment Programme, World Organisation for Animal Health; 2022. URL: https://www.who.int/publications/i/item/9789240075924 [accessed 2024-09-30]
71. One Health High-Level Expert Panel (OHHLEP), Adisasmito WB, Almuhairi S, et al. One Health: a new definition for a sustainable and healthy future. PLoS Pathog 2022 Jun;18(6):e1010537. [doi: 10.1371/journal.ppat.1010537] [Medline: 35737670]
72. Loftus M, Stewardson A, Naidu R, et al. Antimicrobial resistance in the Pacific Island countries and territories. BMJ Glob Health 2020 Apr;5(4):e002418. [doi: 10.1136/bmjgh-2020-002418] [Medline: 32349993]
73. Núñez-Núñez M, Navarro MD, Palomo V, et al. The methodology of surveillance for antimicrobial resistance and healthcare-associated infections in Europe (SUSPIRE): a systematic review of publicly available information. Clin Microbiol Infect 2018 Feb;24(2):105-109. [doi: 10.1016/j.cmi.2017.07.014] [Medline: 28750921]

Abbreviations

AMR: antimicrobial resistance
AMRNet: Antimicrobial Resistance Network
AR Lab Network: Antimicrobial Resistance Laboratory Network
ARAs: antimicrobial resistance alerts
CAR: critical antimicrobial resistances
CAR-Alert: National Alert System for Critical Antimicrobial Resistances
CDC: Centers for Disease Control and Prevention
ECDC: European Center for Disease Prevention and Control
EWSs: early warning systems
GLASS: Global Antimicrobial Resistance and Use Surveillance System
HIC: high-income countries
JANIS-CL: Japan Nosocomial Infections Surveillance - Clinical Laboratory Division
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
WHO: World Health Organization

Edited by AS Bhagavathula; submitted 17.02.24; peer-reviewed by A Rosso, B Tornimbene, FC Dezza; revised version received 30.09.24; accepted 04.10.24; published 15.01.25.

Please cite as:
Iera J, Isonne C, Seghieri C, Tavoschi L, Ceparano M, Sciurti A, D'Alisera A, Sane Schepisi M, Migliara G, Marzuillo C, Villari P, D'Ancona F, Baccolini V
Availability and Key Characteristics of National Early Warning Systems for Emerging Profiles of Antimicrobial Resistance in High-Income Countries: Systematic Review
JMIR Public Health Surveill 2025;11:e57457
URL: https://publichealth.jmir.org/2025/1/e57457
doi:10.2196/57457

Monitoring Public Health Through a Comprehensive Primary Care Database in the Netherlands: Overview of the Nivel Syndromic Surveillance System

Christos Baliatsas[1], PhD; Jojanneke van Summeren[1], PhD; Sander van Beusekom[1]; Amy Matser[1,2,3,4], PhD; Mariette Hooiveld[1], PhD

[1]Nivel – Netherlands Institute for Health Services Research, Otterstraat 118, Utrecht, The Netherlands
[2]Infectious Diseases, Amsterdam UMC, location University of Amsterdam, Amsterdam, The Netherlands
[3]Infectious Diseases, Amsterdam Institute for Infection and Immunity, Amsterdam, The Netherlands
[4]Amsterdam Public Health, Amsterdam, The Netherlands

Corresponding Author:
Christos Baliatsas, PhD
Nivel – Netherlands Institute for Health Services Research, Otterstraat 118, Utrecht, The Netherlands

Abstract

Background: Syndromic surveillance systems are crucial for the monitoring of population health and the early detection of emerging health problems. Internationally, there are numerous established systems reporting on different types of data. In the Netherlands, the Nivel syndromic surveillance system provides real-time monitoring on all diseases and symptoms presented in general practice.

Objective: The present article introduces the national syndromic surveillance system in primary care, emphasizing its role in providing real-time information on infectious diseases and various health problems at the population level, in the Netherlands. In addition, we report on the central role of the participating general practices in data provision, and discuss the applicability of the syndromic surveillance data in different contexts of public health research.

Methods: The Nivel syndromic surveillance system is part of the Nivel Primary Care Database (Nivel-PCD) that collects routinely recorded data from electronic health records of about 10% of the Dutch population, on the basis of approximately 500 practices. This translates to approximately 1.9 million citizens. Since 2010, the surveillance system relies on representative, pseudonymized data collected on a weekly basis from a subset of about 400 practices in the Nivel-PCD, for the entire practice population. Health problems are registered according to the International Classification of Primary Care, applied in all general practices in the Netherlands. Prevalence rates are recalculated and reported every week in the form of figures, also stratified by age, sex, and region. Weekly rates are defined as the number of people that consulted the general practitioner in a certain week for a specific health problem, divided by the total number of registered individuals in the practice.

Results: While utilizing data from general practitioners' electronic health records, the system allows for the timely monitoring and identification of symptom and disease patterns and trends, not only among individuals who seek primary health care, but the entire registered population. Besides their use in disease monitoring, syndromic surveillance data are useful in various public health research contexts, such as environmental health and disaster research.

Conclusions: The Nivel syndromic surveillance system serves as a valuable tool for health monitoring and research, offering valuable insights into public health.

(JMIR Public Health Surveill 2025;11:e58767) doi:10.2196/58767

KEYWORDS

surveillance; monitoring; general practice; public health

Introduction

Syndromic surveillance comprises the systematic, automated collection, analysis, and evaluation of real-time health data for the early monitoring and detection of changes in health indicators [1-3]. Data can be routinely collected from different sources, such as primary (general practitioners; GPs) or secondary care (emergency department visits) [4-6], depending on the health system structure. Other information sources can also be considered, such as helpline calls, medication sales, and internet searches, although they are less commonly used [4,6,7]. The broad scope of syndromic surveillance and the rapid data collection procedures can provide great flexibility for timely assessment of different health outcomes on a large population scale [6-8]. Many countries, also in Europe, have successfully incorporated syndromic surveillance into national surveillance systems [5,9-12], the role of which was crucial during the COVID-19 pandemic, despite inherent limitations [7,8,13,14].

There are well-established national networks that make use of primary care data under strict privacy regulations, for health (services) research and monitoring, for instance in the United Kingdom and Belgium [5,15,16], offering important insights into population health. In the United States, the integration of routine anonymized primary care data is more limited, primarily due to regulatory restrictions, with greater reliance on secondary health care data [17]. In the Netherlands, the Nivel (Netherlands Institute for Health Services Research) Primary Care Database (Nivel-PCD; formerly known as the Netherlands Information Network of General Practice) [17,18] was developed in the early 1990s.

Within this infrastructure, the Nivel syndromic surveillance system provides real-time information on all diseases and symptoms presented in general practice, utilizing data from a large network of GPs. Despite its important role nationwide, there is currently relatively limited international visibility of this infrastructure as well as its applicability in different research contexts. The present paper has the following main objectives: (1) to provide an overview of the Nivel syndromic surveillance system, detailing its methodology; (2) to discuss the important role of the participating general practices in providing data; and (3) to outline the diverse applications of syndromic surveillance data in public health research, highlighting its utility in identifying emerging health threats and supporting research on the health impacts of various exposures.

Methods

Standardized Registration in Dutch General Practices

In the Netherlands, GPs play a vital role as the first point of contact for health care services, acting as gatekeepers for specialized secondary care. Almost every individual is registered at a general practice located in the broader vicinity of their residence. GPs keep detailed electronic health records (EHRs) of consultations, diagnoses, prescribed medications, and referrals. Dutch GPs adhere to the NHG (Dutch College of General Practitioners) guidelines, specifically the ADEPD (adequate record keeping in electronic patient records) guideline, which ensures consistency, reproducibility, and standardization in medical registries, on the basis of a well-structured classification system [19,20]. More specifically, health problems and symptoms documented in general practice adhere to the International Classification of Primary Care (ICPC) [21,22], while medication prescriptions are classified according to the anatomical therapeutic chemical (ATC) classification system [23]. This EHR registration system ensures that information on both consulting and nonconsulting populations is available, which is of high importance for epidemiological research and monitoring. The data recorded by general practices, therefore, provide a comprehensive view of the health characteristics of not only primary care patients but also the broader population.

Nivel-PCD: A Large, Nationally Representative Primary Care Database on Routinely Recorded Data

The Nivel is an independent, nonprofit, research institute that is part of the Dutch Ministry of Health’s knowledge infrastructure. The Nivel-PCD [18] is a large research infrastructure based on real world data extracted from routinely kept EHRs of GPs; primary care out-of-hours services; and paramedical care physicians such as physiotherapists, exercise therapists, and dieticians. There are approximately 500 general practices across the Netherlands that participate yearly in the Nivel-PCD [18], while the total number of general practices in the country is nearly 5000, comprising about 12,000 GPs [24].

Data Collection and Processing

The Nivel gathers data from approximately 10% of the Dutch population (~1.9 million citizens) registered in general practices and about two-thirds of the population considering the use of out-of-hours services. Additionally, data are collected from sentinel general practices (Nivel Peilstations), where a subset of patients with influenza-like illness or other acute respiratory infections provide samples for laboratory testing. These samples are analyzed by the National Institute for Public Health and the Environment (RIVM) to identify the presence of influenza and other respiratory viruses.

For the Nivel syndromic surveillance system, EHR data are collected, extracted, and delivered by the EHR software providers used by the general practices. The system relies on pseudonymized data, available weekly since 2010, on patients consulting for certain health problems and conditions recorded in approximately 400 general practices from all over the country that provide data on a weekly basis. Pseudonymization refers to assigning nonidentifiable identification numbers to the registered practice population. The focus lies on observing and documenting trends in symptoms and diseases, presenting a timely and realistic view of health problems presented in primary care, including influenza-like illness, COVID-19, and other infectious and noninfectious diseases in the Dutch population.

The Nivel-PCD adheres to Dutch data protection regulations and laws for health data use in epidemiological research (Dutch Civil Law, Article 7:458). Medical and personal information are separated with Trusted Third Party support, preventing access to identifiable patient details. Participating health care providers can withdraw from the Nivel-PCD anytime. An opt-out system is available for patients who do not want their data to be used. The Nivel-PCD is ISO27001 certified.

Ethical Considerations

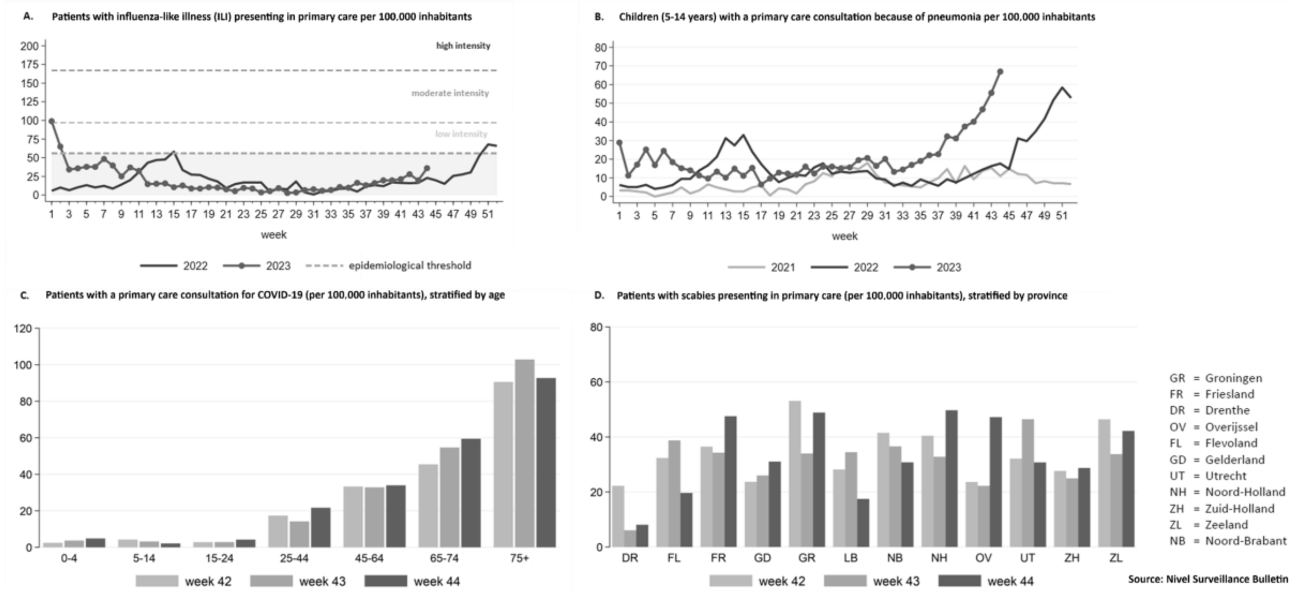
The use of EHRs for research purposes is allowed under certain conditions. When these conditions are fulfilled, neither obtaining informed consent from patients nor approval by a medical ethics committee is obligatory for this type of observational studies containing no directly identifiable data (art. 24 GDPR Implementation Act jo art. 9.2 sub j GDPR).

Results

Analysis and Reporting

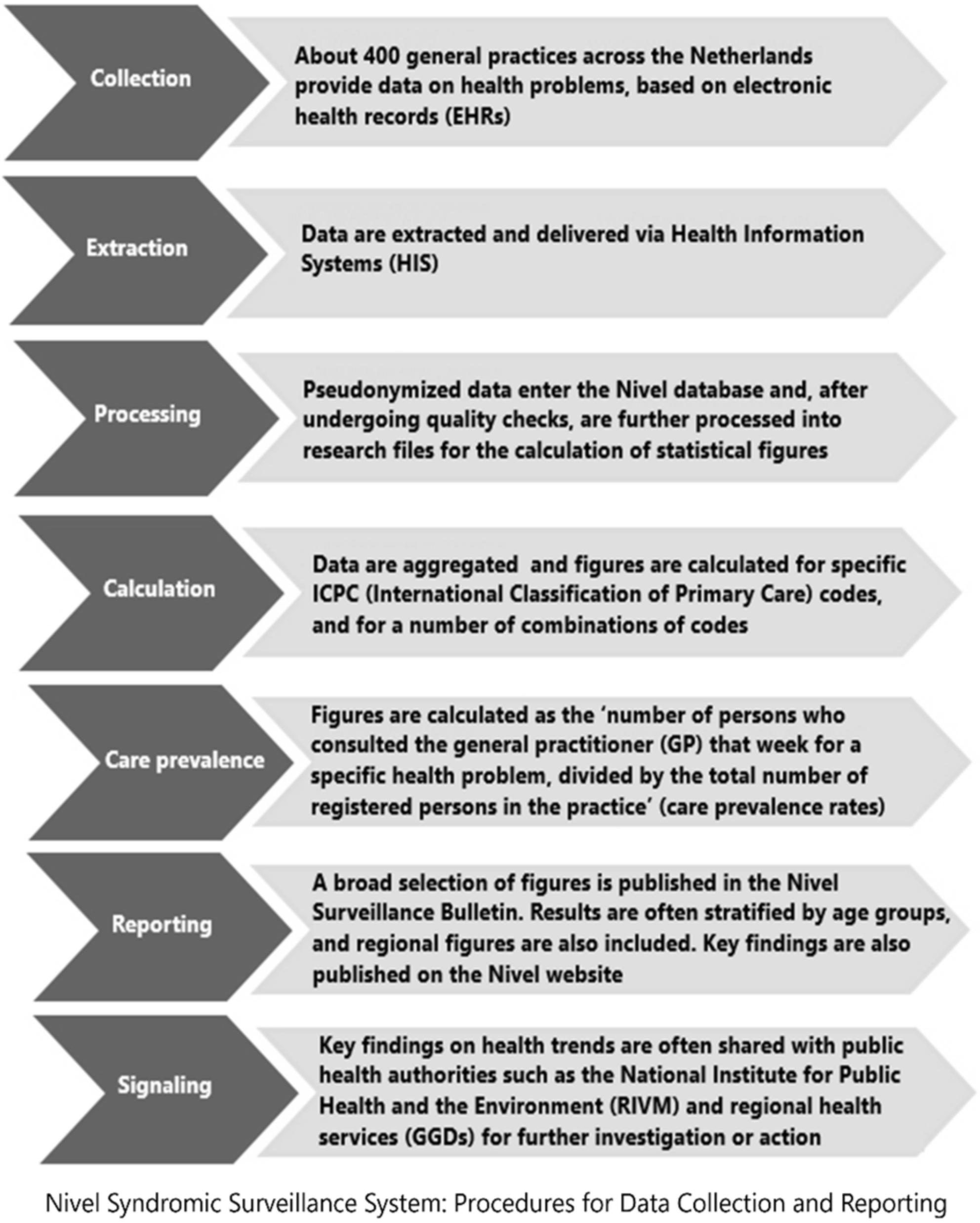
Rates are generated for specific diagnostic codes and for a number of ICPC code combinations (eg, R74, R75, R77, R78, R80, and R83.03 for the cluster “acute respiratory infections excluding pneumonia”); the rates are defined as the number of people who have consulted the GP in that week for a certain health problem, divided by the total number of registered people in the practice. The weekly care prevalence rates are also stratified by age and sex, and by province or public health regions where the local Public Health Services operate, to monitor regional differences. The figures obtained in the previous week are recalculated weekly with any additional data provided, in order to be as up-to-date and complete as possible. The key findings are summarized weekly and published in the Nivel Surveillance Bulletin, accessible on the Nivel website [25] (see example in Figure 1). The report includes selected figures, often categorized by age groups and regions, with additional figures available upon request. An example is the notable increase in the number of children with pneumonia, reported by the Nivel syndromic surveillance system (Figure 1). This early warning initiated further investigation by the Nivel and National Institute for Public Health and the Environment (RIVM) to determine the cause of this increase. Since the start of the COVID-19 pandemic in early 2020, the system has been monitoring cases in primary care. First, the system used the free text describing the disease episode and manually categorized patients clinically diagnosed with COVID-19 or COVID-19-like illness. After the introduction of the ICPC subcode R83.03 in November 2020, the monitoring continued using the registration of this diagnosis code. Figure 2 summarizes the main steps in data collection, analysis and reporting of the Nivel syndromic surveillance system.

Figure 1. Example of four data representations from the Nivel Surveillance Bulletin as of week 44, 2023: **A.** Number of patients with influenza-like illness presenting in primary care per week; **B.** Weekly number of patients in the age group 5 ‐ 14 years with a primary care consultation because of pneumonia; **C.** Weekly number of patients with a primary care consultation because of COVID-19, by age group; **D.** Number of patients with scabies presenting in primary care per week, by province. All numbers are expressed per 100,000 inhabitants.



Source: Nivel Surveillance Bulletin

Figure 2. Overview of the main steps and tasks involved in data collection and report of the Nivel syndromic surveillance system.



Diverse Applications of Syndromic Surveillance in Different Population Research Contexts

Syndromic surveillance also extends beyond monitoring health problems reported in the Surveillance bulletin. Its infrastructure has been used for studies on different topics and contexts, for instance on the physical and psychosocial impact of the COVID-19 pandemic [26,27] as well as to monitor the presentation of symptoms in primary care in relation to environmental exposures and heat waves [28,29]. The comprehensive patient registry in the Nivel-PCD enables continuous tracking of patients over time, linked with additional sources such as socioeconomic data from Statistics Netherlands (CBS), geographic information, and environmental exposure data. Integration occurs across aggregated levels using pseudonyms to ensure patient privacy.

Discussion

Study Findings and Comparisons With Previous Works

The Nivel syndromic surveillance system provides large scale monitoring of diseases and health symptoms, based on representative population health data routinely registered in general practice in the Netherlands, on the basis of unified registration guidelines and strict privacy regulations. The system's infrastructure can also be used in different contexts of public health research, including studies on the physical and psychosocial impact of pandemics as well as environmental exposures. It provides information to the stakeholders through the openly accessible Surveillance Bulletin and signals early warnings to the national and local public health authorities.

Nivel syndromic surveillance allows for the observation of patterns and trends not only among individuals who seek care, but the entire registered population. This is an important strength compared to syndromic surveillance systems that solely rely on emergency department visits [10], which provide a rapid picture of health trends but are limited to the severe cases. The monitoring of broad categories of health problems such as respiratory infections enables syndromic surveillance to serve as an early warning mechanism for novel threats, providing information earlier than laboratory confirmation [6]. The data of the Nivel syndromic surveillance system can also be used in different public health research contexts, providing insight into the extent of public health impacts, identify vulnerable populations, and guide appropriate mitigation strategies.

Limitations

It is important to acknowledge the limitations and inherent complexities associated with syndromic surveillance. One of them is the reliance of syndromic surveillance on reporting clinical diagnoses or symptoms rather than diagnoses confirmed by a laboratory. This could introduce some degree of uncertainty, given that symptoms may not always be associated with a specific disease. In addition, it could increase the possibility of a delayed diagnosis registration compared to surveillance infrastructures that use laboratory-confirmed cases or emergency visits. Identifying rare or emerging diseases without pre-established diagnostic codes or historical data poses another challenge; a recent example is the miscoding and misclassification of COVID-19 cases in primary care, during the pandemic [30]. Establishing a reliable baseline for certain health problems can also be difficult due to different factors such as seasonality and changes in health care–seeking behavior. Technical aspects such as updates in guidelines and coding practices in primary care may also disrupt trends. Moreover, there is currently a lack of quantitative data from other surveillance systems to support a structured comparative evaluation [31] of the system's effectiveness regarding indicators such as early detection.

Future Directions

As the field of syndromic surveillance evolves, there are several challenges to be addressed that could substantially strengthen the utility of the existing system in public health monitoring. Despite the satisfactory national coverage rate of general practices, participation of practices in the system is voluntary and can vary per region [32]. The expansion of the system's geographical coverage with the inclusion of more general practices across the country would strengthen the sample power as well as representativeness, allowing for a more accurate detection of disease trends and variation at the regional and population group levels. In light of climate change and emerging crises, linking syndromic surveillance with publicly available environmental exposure data could allow for the monitoring of the physical and mental impact on the population's health more accurately, especially during seasonal events such as heatwaves or acute incidents of hazardous emissions. Another aspect would be the integration of additional data sources in the syndromic surveillance system such as those on emergency visits, similar to other established surveillance systems [9], in order to provide a more comprehensive monitoring of the population's health. Strengthening and expanding international collaborations and networks with other syndromic surveillance systems could also help identify existing caveats and develop best practices to address them and would improve early warning of potential health threats across countries [33]. Finally, innovation should be an indispensable part of syndromic surveillance.

In particular, machine learning approaches as well as artificial intelligence are promising areas that could contribute to more accurate and efficient health monitoring, by facilitating analysis processes, improving detection algorithms, and even the development of predictive models [34].

Conclusions

The Nivel syndromic surveillance system is an important infrastructure in providing monitoring of real world data on the occurrence of all symptoms and diseases presented in general practice, at the population level. Further implementation and refinement of syndromic surveillance methodologies towards data completeness, quality, and representativeness will enhance the system's effectiveness in contributing to public health research and response strategies.

Acknowledgments

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Data Availability

Data are available upon request (zorgregistraties@nivel.nl).

Conflicts of Interest

None declared.

References

1. Triple S Project. Assessment of syndromic surveillance in Europe. Lancet 2011 Nov 26;378(9806):1833-1834. [doi: 10.1016/S0140-6736(11)60834-9] [Medline: 22118433]

2. Ziemann A, Fouillet A, Brand H, Krafft T. Success factors of European syndromic surveillance systems: a worked example of applying qualitative comparative analysis. PLoS ONE 2016;11(5):e0155535. [doi: 10.1371/journal.pone.0155535] [Medline: 27182731]
3. Elliot AJ, Hughes HE, Harcourt SE, Morbey RA, Smith S, Smith GE. Real-time syndromic surveillance. In: WHO Guidance on Research Methods for Health Emergency and Disaster Risk Management: World Health Organization; 2021. URL: https://extranet.who.int/kobe_centre/sites/default/files/chapters/Chapter%204.9%20Slideshow.pdf
4. Duijster JW, Doreleijers SDA, Pilot E, et al. Utility of emergency call centre, dispatch and ambulance data for syndromic surveillance of infectious diseases: a scoping review. Eur J Public Health 2020 Aug 1;30(4):639-647. [doi: 10.1093/eurpub/ckz177] [Medline: 31605491]
5. Bednarska NG, Smith S, Bardsley M, et al. Trends in general practitioner consultations for hand foot and mouth disease in England between 2017 and 2022. Epidemiol Infect 2025 Jan 13;153:e22. [doi: 10.1017/S095026882400181X] [Medline: 39801026]
6. Adedire O, Love NK, Hughes HE, Buchan I, Vivancos R, Elliot AJ. Early detection and monitoring of gastrointestinal infections using syndromic surveillance: a systematic review. Int J Environ Res Public Health 2024 Apr 17;21(4):489. [doi: 10.3390/ijerph21040489] [Medline: 38673400]
7. Corrao G, Bonaugurio AS, Bagarella G, et al. Does syndromic surveillance assist public health practice in early detecting respiratory epidemics? Evidence from a wide Italian retrospective experience. J Infect Public Health 2025 Feb;18(2):102621. [doi: 10.1016/j.jiph.2024.102621] [Medline: 39729672]
8. Vos B, Debouverie L, Doggen K, et al. Monitoring COVID-19 in Belgian general practice: a tool for syndromic surveillance based on electronic health records. Eur J Gen Pract 2024 Dec;30(1):2293699. [doi: 10.1080/13814788.2023.2293699] [Medline: 38186340]
9. Smith GE, Elliot AJ, Lake I, et al. Syndromic surveillance: two decades experience of sustainable systems - its people not just data!. Epidemiol Infect 2019 Jan;147:e101. [doi: 10.1017/S0950268819000074] [Medline: 30869042]
10. Hughes HE, Edeghere O, O’Brien SJ, Vivancos R, Elliot AJ. Emergency department syndromic surveillance systems: a systematic review. BMC Public Health 2020 Dec 9;20(1):1891. [doi: 10.1186/s12889-020-09949-y] [Medline: 33298000]
11. Boender TS, Cai W, Schranz M, et al. Using routine emergency department data for syndromic surveillance of acute respiratory illness, Germany, week 10 2017 until week 10 2021. Euro Surveill 2022 Jul;27(27):2100865. [doi: 10.2807/1560-7917.ES.2022.27.27.2100865] [Medline: 35801521]
12. de Fougerolles TR, Damm O, Ansaldi F, et al. National influenza surveillance systems in five European countries: a qualitative comparative framework based on WHO guidance. BMC Public Health 2022 Jun 9;22(1):1151. [doi: 10.1186/s12889-022-13433-0] [Medline: 35681199]
13. Thiam MM, Pontais I, Forgeot C, et al. Syndromic surveillance: a key component of population health monitoring during the first wave of the COVID-19 outbreak in France, February-June 2020. PLoS ONE 2022;17(2):e0260150. [doi: 10.1371/journal.pone.0260150] [Medline: 35143501]
14. Brainard J, Lake IR, Morbey RA, Jones NR, Elliot AJ, Hunter PR. Comparison of surveillance systems for monitoring COVID-19 in England: a retrospective observational study. Lancet Public Health 2023 Nov;8(11):e850-e858. [doi: 10.1016/S2468-2667(23)00219-0] [Medline: 37832574]
15. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: clinical practice research datalink (CPRD). Int J Epidemiol 2015 Jun;44(3):827-836. [doi: 10.1093/ije/dyv098] [Medline: 26050254]
16. Bartholomeeusen S, Kim CY, Mertens R, Faes C, Buntinx F. The denominator in general practice, a new approach from the Intego database. Fam Pract 2005 Aug;22(4):442-447. [doi: 10.1093/fampra/cmi054] [Medline: 15964863]
17. Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible sources of bias in primary care electronic health record data use and reuse. J Med Internet Res 2018 May 29;20(5):e185. [doi: 10.2196/jmir.9134] [Medline: 29844010]
18. Vanhommerig JW, Verheij RA, Hek K, Ramerman L, Hooiveld M, Veldhuijzen NJ, et al. Data Resource Profile: Nivel Primary Care Database (Nivel-PCD). Int J Epidemiol 2024. [doi: 10.1093/ije/dyaf017]
19. Duineveld B, Kole HM, Werven H, NHG-richtlijn. Adequate dossiervorming met het elektronisch patiëntendossier (ADEPD). Volledig gereviseerde versie 2019. [Adequate record keeping in electronic patient records (In Dutch)]. : Nederlands Huisartsen Genootschap; 2019 URL: https://www.nhg.org/wp-content/uploads/2022/11/NHG-Richtlijn-ADEPD-2019.pdf
20. Smeets HM, Kortekaas MF, Rutten FH, et al. Routine primary care data for scientific research, quality of care programs and educational purposes: the Julius General Practitioners’ Network (JGPN). BMC Health Serv Res 2018 Sep 25;18(1):735. [doi: 10.1186/s12913-018-3528-5] [Medline: 30253760]
21. Sean E. The anatomical therapeutic chemical classification system. J Pharmacol Ther 2021;1(5):1 [FREE Full text]
22. Soler JK, Okkes I, Wood M, Lamberts H. The coming of age of ICPC: celebrating the 21st birthday of the International Classification of Primary Care. Fam Pract 2008 Aug;25(4):312-317 [FREE Full text] [doi: 10.1093/fampra/cmn028] [Medline: 18562335]
23. World Health Organization. WHO collaborating centre for drug statistics methodology. guidelines for ATC classification and DDD assignment. : Norwegian Institute of Public Health; 2024. URL: https://atcddd.fhi.no/filearchive/publications/1_2024_guidelines__final_web.pdf [accessed 2025-02-20]

24. Batenburg RS, Flinterman LE, Vis EB, et al. Cijfers uit de nivel- registratie van huisartsen en huisartsenpraktijken: een actualisering voor de periode 2020-2022. [figures from the nivel registration of general practitioners and general practices: an update for the period 2020-2022 (in dutch)]. : Nivel; 2022 URL: https://www.nivel.nl/sites/default/files/bestanden/1004340.pdf [accessed 2025-02-20]
25. Hooiveld M, Baliatsas C, Summeren J, van J, Stelma FF. Nivel-cijfers Ziekten per week in Nederland - actueel: wat heerst er? [Nivel Figures Weekly Diseases in the Netherlands (In Dutch)]. URL: https://www.nivel.nl/nl/zorg-en-ziekte-in-cijfers/actuele-cijfers-ziekten-per-week [accessed 2024-03-10]
26. van Duinkerken A, Bosmans M, Baliatsas C, et al. The integrated health monitor COVID-19: a protocol for a comprehensive assessment of the short- and long-term health impact of the pandemic in the Netherlands. Methods Protoc 2023 Dec 2;6(6):117. [doi: 10.3390/mps6060117] [Medline: 38133137]
27. Rahmon I, Bosmans M, Baliatsas C, Hooiveld M, Marra E, Dückers M. COVID-19 health impact: a use case for syndromic surveillance system monitoring based on primary care patient registries in the Netherlands. JMIR Public Health Surveill 2024 Sep 26;10:e53368. [doi: 10.2196/53368] [Medline: 39588660]
28. Baliatsas C, Hooiveld M, Kenens R, et al. Gezondheidsproblemen in gebieden met hoge plaagdruk van de eikenprocessierups: de grote overlast in 2019 [Health issues in areas heavily infested with the oak processionary: the significant nuisance in 2019] (In Dutch).: Nivel; 2022. URL: https://www.nivel.nl/nl/publicatie/gezondheidsproblemen-gebieden-met-hoge-plaagdruk-van-de-eikenprocessierups-de-grote [accessed 2025-02-20]
29. Hondema L, Hoek G, Kraaij-Dirkzwager M, Fischer P, Hooiveld M. Heat wave effects on morbidity in the Netherlands using general practice data. Environmental Epidemiology 2019 Oct;3(Supplement 1):161-162. [doi: 10.1097/01.EE9.0000607552.35945.d9]
30. de Lusignan S, Williams J. To monitor the COVID-19 pandemic we need better quality primary care data. BJGP Open 2020;4(2):bjgpopen20X101070. [doi: 10.3399/bjgpopen20X101070] [Medline: 32295793]
31. European Centre for Disease Prevention and Control (ECDC). Data quality monitoring and surveillance system evaluation: A handbook of methods and applications. : Stockholm: ECDC; 2014. URL: https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/Data-quality-monitoring-surveillance-system-evaluation-Sept-2014.pdf [accessed 2025-02-20] [doi: 10.2900/35329]
32. Smith S, Morbey R, de Lusignan S, Pebody RG, Smith GE, Elliot AJ. Investigating regional variation of respiratory infections in a general practice syndromic surveillance system. J Public Health (Oxf) 2021 Jun 7;43(2):e153-e160. [doi: 10.1093/pubmed/fdaa014] [Medline: 32009178]
33. Zucs P, Beauté J, Palm D, Spiteri G. Focus, vigilance, resilience: towards stronger infectious disease surveillance, threat detection and response in the EU/EEA. Euro Surveill 2024 Aug;29(34):2400066. [doi: 10.2807/1560-7917.ES.2024.29.34.2400066] [Medline: 39176987]
34. Chowdhury AT, Newaz M, Saha P, Pedersen S, Khan MS, Chowdhury MEH. Use of artificial intelligence in the surveillance of seasonal respiratory infections. In: Chowdhury MEH, Kiranyaz S, editors. Surveillance, Prevention, and Control of Infectious Diseases: Springer; 2024. [doi: 10.1007/978-3-031-59967-5_10]

Abbreviations

ADEPD: adequate record keeping in electronic patient records
ATC: anatomical therapeutic chemical
CBS: Statistics Netherlands
EHR: electronic health record
GP: general practitioner
ICPC: International Classification of Primary Care
NHG: Dutch College of General Practitioners
Nivel: Netherlands Institute for Health Services Research
Nivel-PCD: Netherlands Institute for Health Services Research-primary care database
RIVM: National Institute for Public Health and the Environment

Edited by A Mavragani; submitted 24.03.24; peer-reviewed by S Mukhida, S Jonathan; revised version received 29.01.25; accepted 29.01.25; published 12.03.25.

Please cite as:
Baliatsas C, van Summeren J, van Beusekom S, Matser A, Hooiveld M
Monitoring Public Health Through a Comprehensive Primary Care Database in the Netherlands: Overview of the Nivel Syndromic Surveillance System
JMIR Public Health Surveill 2025;11:e58767
URL: https://publichealth.jmir.org/2025/1/e58767
doi:10.2196/58767

Wastewater Monitoring During the COVID-19 Pandemic in the Veneto Region, Italy: Longitudinal Observational Study

Honoria Ocagli[1*], PhD; Marco Zambito[1*], MSc; Filippo Da Re[2], MSc; Vanessa Groppi[2], BSc; Marco Zampini[3], MSc; Alessia Terrini[3], BSc; Franco Rigoli[3], MSc; Irene Amoruso[4], PhD; Tatjana Baldovin[4], PhD; Vincenzo Baldo[4], MD; Francesca Russo[2*], MD; Dario Gregori[1*], PhD

[1]Unit of Biostatistics, Epidemiology and Public Health, Department of Cardio-Thoraco-Vascular Sciences and Public Health, University of Padova, Via Loredan 18, Padova, Italy

[2]Directorate of Prevention, Food Safety and Veterinary Public Health-Veneto Region, Venice, Italy

[3]ARPAV - Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto, U.O. Biologia, Padova, Italy

[4]Unit of Hygiene and Public Health, Department of Cardiac, Thoracic, Vascular Sciences, and Public Health, University of Padova, Padova, Italy

[*]these authors contributed equally

Corresponding Author:
Dario Gregori, PhD
Unit of Biostatistics, Epidemiology and Public Health, Department of Cardio-Thoraco-Vascular Sciences and Public Health, University of Padova, Via Loredan 18, Padova, Italy

Abstract

Background: As the COVID-19 pandemic has affected populations around the world, there has been substantial interest in wastewater-based epidemiology (WBE) as a tool to monitor the spread of SARS-CoV-2. This study investigates the use of WBE to anticipate COVID-19 trends by analyzing the correlation between viral RNA concentrations in wastewater and reported COVID-19 cases in the Veneto region of Italy.

Objective: We aimed to evaluate the effectiveness of the cumulative sum (CUSUM) control chart method in detecting changes in SARS-CoV-2 concentrations in wastewater and its potential as an early warning system for COVID-19 outbreaks. Additionally, we aimed to validate these findings over different time periods to ensure robustness.

Methods: This study analyzed the temporal correlation between SARS-CoV-2 RNA concentrations in wastewater and COVID-19 clinical outcomes, including confirmed cases, hospitalizations, and intensive care unit (ICU) admissions, from October 2021 to August 2022 in the Veneto region, Italy. Wastewater samples were collected weekly from 10 wastewater treatment plants and analyzed using a reverse transcription–quantitative polymerase chain reaction. The CUSUM method was used to detect significant shifts in the data, with an initial analysis conducted from October 2021 to February 2022, followed by validation in a second period from February 2022 to August 2022.

Results: The study found that peaks in SARS-CoV-2 RNA concentrations in wastewater consistently preceded peaks in reported COVID-19 cases by 5.2 days. Hospitalizations followed with a delay of 4.25 days, while ICU admissions exhibited a lead time of approximately 6 days. Notably, certain health care districts exhibited stronger correlations, with notable values in wastewater anticipating ICU admissions by an average of 13.5 and 9.5 days in 2 specific districts. The CUSUM charts effectively identified early changes in viral load, indicating potential outbreaks before clinical cases increased. Validation during the second period confirmed the consistency of these findings, reinforcing the robustness of the CUSUM method in this context.

Conclusions: WBE, combined with the CUSUM method, offers valuable insight into the level of COVID-19 outbreaks in a community, including asymptomatic cases, thus acting as a precious early warning tool for infectious disease outbreaks with pandemic potential.

(JMIR Public Health Surveill 2025;11:e58862) doi:10.2196/58862

KEYWORDS

wastewater-based epidemiology; SARS-CoV-2; COVID-19; CUSUM; WBE; cumulative sum chart

Introduction

The COVID-19 pandemic has had a profound impact on public health, with more than 200 million confirmed cases and more than 4 million deaths worldwide as of August 2021. As the virus spread, concerns about possible transmission through wastewater became a significant public health issue [1]. Wastewater-based epidemiology (WBE) plays a pivotal role in various health-related studies. Research highlights several applications. The estimation of global nicotine consumption through WBE was reviewed in the work of Asadi et al [2], highlighting its diverse applications, shedding light on trends in nicotine consumption worldwide and the associated risks, and stressing the urgency of global action. Antibiotic-resistant bacteria are another topic of interest in WBE, as shown in the work of Tiwari et al [3], whose systematic review underscores the need for active monitoring of antibiotic-resistant bacteria, offering insights into their clinical utility, sensitivity, and uniformity. In addition, antimicrobial concentration was investigated. The systematic review by Chau et al [4] examines the concordance between estimates of the prevalence of antimicrobial resistance in wastewater and humans, while Holton et al [5] explore community-wide antimicrobial use via WBE, evaluating 16 antimicrobials and their metabolites, providing insight for improving precision in assessing drug intake. Wastewater-based antimicrobial resistance surveillance, as demonstrated in the most recent literature, appears to be an effective and promising approach to confronting such phenomena in support of public health actions. In addition, wastewater has long been useful for the surveillance of different pathogens.

Recent studies have shown that SARS-CoV-2, a COVID-19 viral strain, can be detected in wastewater samples from infected individuals, even when they are asymptomatic or have mild symptoms [1,6]. Wastewater surveillance is based on the assumption that viral RNA can be detected in the fecal matter of asymptomatic people, as the shedding of viral particles can occur before the onset of symptoms [7] and since diarrhea is a symptom reported by a significant proportion of patients with SARS-CoV-2 [8]. Furthermore, wastewater can also contain fragments of viral RNA secreted by oral or nasal routes that can enter wastewater when hands are washed. The virus can persist in wastewater for several days, although higher temperatures and lower pH levels can reduce its survival [9].

In the literature, various studies offer information on multiple aspects of COVID-19 and wastewater, including the detection of SARS-CoV-2 in wastewater [10], the persistence of the virus in wastewater [11], and the potential use of WBE as a tool to monitor the prevalence of the virus in communities [12]. The study by Ciannella et al [13] focuses on analytical procedures and epidemiological modeling in WBE for COVID-19. They highlight the lack of standardization in wastewater analytical methods, with reverse transcription–quantitative polymerase chain reactions (PCRs) being the most widely used technique for the detection and quantification of viral RNA. They suggest exploring the solid portion of wastewater due to its higher viral load and advocate the development of cost-effective portable detection devices [13]. The study by Peccia et al [14] investigates the potential of wastewater surveillance as a tool for the early detection of COVID-19 outbreaks in a community. The results show that WBE could provide early warnings of outbreaks before clinical cases are detected.

In Italy, the study by La Rosa et al [10] suggests that SARS-CoV-2 was present in northern Italy even in December 2019. The researchers analyzed wastewater samples from October 2019 to February 2020 and found positive results for the virus in multiple cities. This challenged the prior understanding that the geo-warning system primarily served to detect the virus' presence and to monitor outbreaks before they were reported to the health care system.

In WBE, a variety of statistical methods have traditionally been employed to analyze the relationship between viral RNA concentrations in wastewater and epidemic incidence, along with other public health outcomes. Common techniques include correlation analysis, regression models [15], and principal component analysis [16], which are often used to establish associations between wastewater parameters and COVID-19 case numbers, hospitalizations, or deaths. For example, studies have frequently utilized methods such as ANOVA [17], Gaussian distribution models [18], and autoregressive integrated moving average (ARIMA) time series models to understand the temporal dynamics and predict trends in viral loads [19]. Advanced techniques include the Efficient and Practical Virus Identification System with Enhanced Sensitivity for Membrane (EPISENS-M) method, which combines highly sensitive RNA detection with mathematical modeling to predict cases of COVID-19 [20]. Additionally, the framework utilized by Dai et al [21] uses functional data analysis to detect true trends in viral concentration, incorporating known covariates such as sample storage temperature and influent volume, and employs Markov chain Monte Carlo methods to forecast future viral concentrations.

However, while these approaches are effective for retrospective analysis and prediction, they often focus on broad trend identification rather than real-time detection. Our study introduces the cumulative sum (CUSUM) chart method, a statistical process control tool commonly used in quality control [22], which offers unique advantages in WBE by detecting small shifts in viral load trends. Unlike traditional methods, CUSUM charts are particularly useful for identifying early deviations from established baselines, thus allowing earlier detection of potential outbreaks.

This study aims to investigate whether the time series of viral RNA concentrations in wastewater, expressed as genome copies (gc) per microliter, can serve as an early indicator of COVID-19–positive cases, including total positives, hospitalized cases, and those admitted to intensive care units (ICUs), in the Veneto region using CUSUM charts.

Methods

Overview

This study was conducted in the Veneto region of Italy, focusing on a specific timeframe during the COVID-19 pandemic. While the number of data points is limited, the study aims to provide a detailed analysis of the correlation between the viral RNA load of wastewater and the COVID-19 clinical outcomes, offering valuable information for regional public health strategies.

Our research evaluated this hypothesis by comparing the temporal trends of these case categories to SARS-CoV-2 concentrations in wastewater samples.

Ethical Considerations

This study was conducted using publicly available, anonymized, and deidentified data, ensuring that no personal information of individuals was accessed or processed. Wastewater data were derived from environmental samples (ie, untreated wastewater) and COVID-19 epidemiological data were aggregated at the population level, preventing any direct identification of individuals. Additionally, the analysis adhered to ethical guidelines for research using secondary data, ensuring that all data were used responsibly and in accordance with the applicable privacy laws. No personal or sensitive information was collected and the study design did not involve any interventions or direct interaction with human subjects, thus exempting it from requiring formal ethical approval.

Data Sources

This study used data from multiple sources to analyze the correlation between SARS-CoV-2 RNA concentrations in urban wastewater and COVID-19 clinical outcomes. The primary data sources included analytical data from raw wastewater samples and clinical data, including COVID-19 case data, hospitalizations, and ICU admissions, from the Presidency of the Council of Ministers' Civil Protection Department.

These datasets were analyzed to assess temporal correlations and to identify potential early warning signals for COVID-19 surges based on wastewater monitoring.

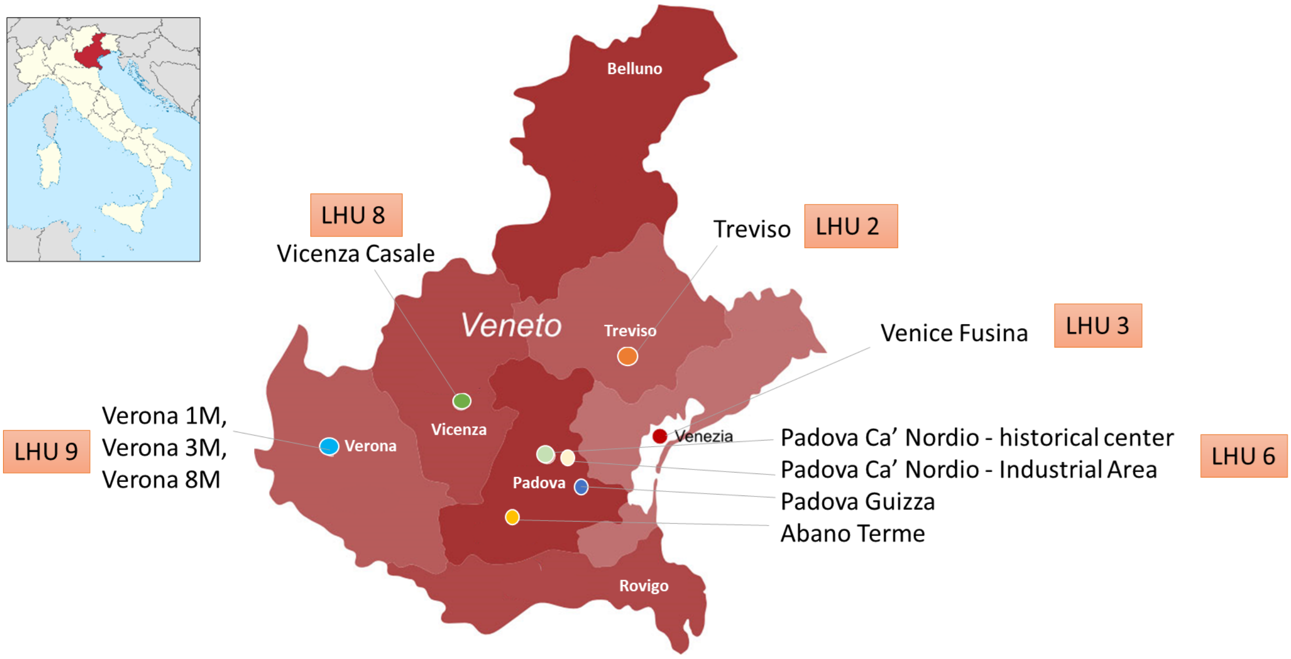
Wastewater Data and Study Design

The study focused on the Veneto region, selecting 10 wastewater treatment plants (WTPs) in 5 provinces, collectively serving more than 4 million inhabitants. Weekly samples were collected from October 5, 2021, to August 2, 2022.

Sampling Equipment and Sites

The samples were collected using an automated sampler (compact portable sampler 6712C, ISCO) to obtain 100 mL of a 24-hour composite raw wastewater sample [10], ensuring representative capture of daily variations in wastewater composition. Each sample consisted of 20 mL subsamples of raw wastewater collected every 30 minutes for the 24-hour period. After collection, the samples were immediately refrigerated at 4 °C to preserve viral RNA integrity until analysis. Sample concentration, extraction, and viral RNA detection protocols are described in the study by Baldovin et al [23]. The selected WTPs included: Abano Terme, Padova Ca' Nordio - historical center, Padova Ca' Nordio - Industrial Area, Padova Guizza, Treviso, Venice Fusina, Verona 1M, Verona 3M, Verona 8M, and Vicenza Casale (Figure 1).

Figure 1. Geographic distribution of wastewater treatment plants and correspondence with LHUs in the Veneto region, Italy. LHU: local health unit.



RNA Detection and Laboratory Methods

Following collection, wastewater samples were concentrated using the Epidemiological Surveillance for SARS-CoV-2 in Urban Sewage in Italy, third revision (SARI_rev3) method [24], which includes a sample concentration step involving polyethylene glycol precipitation. This process yielded a concentrated pellet from a 45 mL sample volume, which was then subjected to RNA extraction using 2 different commercial kits, following the manufacturer's instructions. For the University of Padova samples, RNA extraction was performed using the QIAamp viral RNA mini kit (Qiagen). The samples were collected at the following locations: Abano Terme, Padova Ca' Nordio - historical center, Padova Ca' Nordio - Industrial Area, and Padova Guizza. For the Regional Agency for Environmental Protection and Prevention of Veneto samples, RNA extraction was conducted using the eGENE-UP system (Biomerieux), a well-established method for SARS-CoV-2 RNA recovery. These samples were collected from Treviso, Venice Fusina, Verona 1M, Verona 3M, Verona 8M, and Vicenza Casale. For all samples, RNA was treated with the OneStep PCR Inhibitor Removal Kit (Zymo Research) after extraction to remove PCR inhibitors. The final volume of the RNA elution was standardized to 100 µL.

Reverse transcription–quantitative PCR assays targeting the SARS-CoV-2 open reading frames 1ab (ORF1ab) (nonstructural protein 14) gene were used for quantification. A limit of detection was established at 2 gc/µL. For measurements below this threshold, the values were recorded as half of the limit of detection to account for the presence of the low level. PCR inhibition was routinely assessed using an internal Mengovirus control and recovery efficiency was confirmed for each sample using this control as well.

Replicates and Quality Control

Each wastewater sample was processed in duplicate to ensure the reproducibility and accuracy of viral load quantification. Additionally, a spiking experiment was performed periodically using known concentrations of synthetic SARS-CoV-2 RNA to verify the precision of the assay on different matrices. This approach enabled the detection of any potential assay failures or matrix effects that could impact quantification. To account for the variability in influent flow rates and population size, viral load data were normalized per liter of wastewater and per capita. This normalization facilitated comparison across different WTPs and over time. Detailed records of each WTP's daily influent volume, as well as specific operational characteristics, were maintained and integrated into the data analysis to ensure that any observed variations in viral concentrations could be accurately interpreted.

Positive COVID-19 Case Data

The Presidency of the Council of Ministers' Civil Protection Department diligently managed, processed, and provided data on active and cumulative confirmed COVID-19 cases and daily COVID-19-related deaths for public use under the Creative Commons Attribution 4.0. The population of confirmed COVID-19 cases in the area served by each WTP was identified based on local health units (LHUs) as follows: LHU 2 (Treviso collector); LHU 3 (Venice Fusina collector); LHU 6 (Abano Terme, Padova Ca' Nordio - historical center, Padova Ca' Nordio - Industrial Area, and Padova Guizza collectors); LHU 8 (Vicenza Casale collector); and LHU 9 (Verona 1M, 3M, and 8M collectors) (Figure 1).

Statistical Analysis

Continuous variables were reported as median and first and third quartiles, and categorical variables as percentages and absolute numbers. The trend of wastewater was reported as an absolute value for each measurement, and the trend in the number of positive cases was shown as positives per province population multiplied by 100,000.

CUSUM Charts

To compare COVID-19 cases and trends in wastewater, we used CUSUM charts. Originally developed by Page in 1954 [25], CUSUM charts have since been widely adopted across various fields, including health care, for quality control purposes. The key advantage of a CUSUM chart lies in its cumulative nature, where each plotted point reflects not only the current value but also the cumulative effect of all previous data points [22]. This makes CUSUM charts especially sensitive to small deviations that could indicate a gradual trend change, positive or negative, within a dataset [26].

This graphical tool detects slight deviations from a target value (in this case, the mean) by considering current and previous values in the series. Specifically, the CUSUM chart is based on a statistic derived from the standard deviations of the process, assuming that the empirical distribution follows a normal distribution with a mean of zero. Statistic values that exceed the upper and lower limits (typically set at 4σ, where σ represents the standard deviation of the process) are marked as "out of control" points. These points can be interpreted as "structural breakpoints" or "turning points," indicating significant shifts in the series' trend, either positive or negative. The "differential" signifies the difference between the time ("t") and the value at the values of time ("t-1"). The CUSUM charts for positive case differentials employ upper and lower limits at 4σ, while those for gc/µL concentrations use a threshold at ±2.782 [27] due to lower data variability (varpositive=16,048,128; varRNA=3070.525 gc/µL).

The analyses were performed using R [28], with the *qcc* package used for the CUSUM chart [29]; *ggplot* [30] and *ggplot2* [31] were used to generate the graphs.

Results

Overview

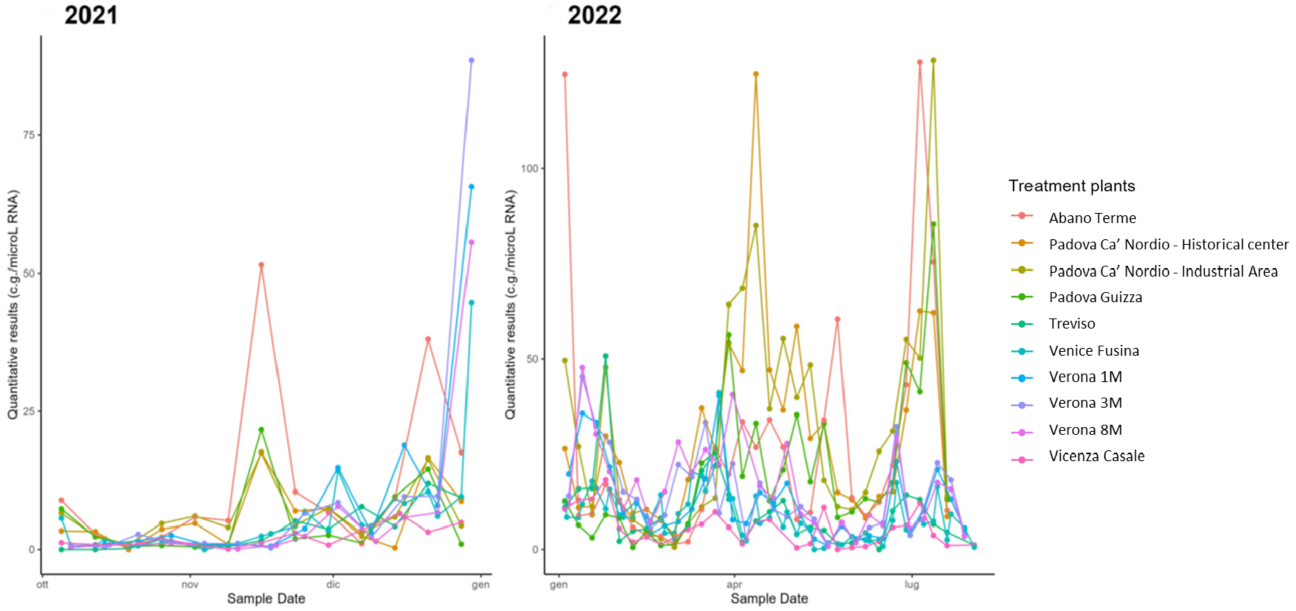
Table 1 provides an overview of the WTPs considered during the sampling period. It offers information on their key characteristics, including flow rates (m^3/24 h) and total suspended solids concentrations (mg/L). These parameters exhibit variations at different locations, indicating the diverse compositions of the wastewater. The table also highlights the volume of collected samples (mL) and the corresponding distribution across different sample sizes. The concentrations of total suspended solids showed diversity, ranging from 94.25 mg/L in Treviso to 207.00 mg/L in Verona 8M.

Figure 2 shows the trend of wastewater for 2021 and 2022, divided by the different collectors. It should be noted that there are different peaks in wastewater load during specific periods, including November 2021, late December 2021, and January 2022, followed by recurring peaks in April and July 2022.

Table . Descriptive characteristics of wastewater treatment plants in the Veneto region, Italy, considered in the study. Continuous variables are reported as I, II, and III quartiles, whereas categorical variables are reported as percentages and absolute numbers.

	Flow rates (m^3/24 h), I/II/III quartile	Total suspended solids (mg/L), I/II/III quartile	Sample volume (125 mL), n (%)	Sample volume (250 mL), n (%)	Sample volume (500 mL), n (%)
Abano Terme (n=44)	4927.00/5428.50/6059.50	170.00/232.00/276.00	0 (0)	0 (0)	44 (100)
Padova Ca’ Nordio - historical center (n=44)	22,074.25/24,677.00/27,760.75	112.00/172.00/234.00	0 (0)	0 (0)	44 (100)
Padova Ca’ Nordio - Industrial Area (n=44)	21,626.75/24,677.00/27,760.75	112.00/170.00/235.00	0 (0)	0 (0)	44 (100)
Padova Guizza (n=44)	3087.25/3275.50/3406.25	16.00/47.00/74.50	0 (0)	0 (0)	44 (100)
Treviso (n=44)	13,477.50/14,365.00/14,830.00	94.25/123.50/169.75	0 (0)	0 (0)	44 (100)
Venice Fusina (n=87)	72,906.50/87,562.00/100,155.25	68.50/116.00/193.00	34 (39)	51 (59)	2 (2)
Verona 1M (n=43)	12,657.50/12,874.00/13,089.50	117.00/181.00/230.00	0 (0)	0 (0)	43 (100)
Verona 3M (n=43)	11,232.00/14,688.00/17,237.00	132.50/183.00/203.50	0 (0)	0 (0)	43 (100)
Verona 8M (n=43)	33,739.50/34,474.00/39,571.00	207.00/231.00/265.00	0 (0)	0 (0)	43 (100)
Vicenza Casale (n=44)	27,052.75/30,018.50/32,466.75	85.75/101.00/119.25	0 (0)	0 (0)	44 (100)

Figure 2. Temporal trends of SARS-CoV-2 RNA concentrations in wastewater throughout the Veneto region, Italy, during 2021 and 2022.



CUSUM Chart Results in the First Period

Daily total COVID-19–positive cases, hospitalized cases, and admissions to the ICU were compared with daily viral concentrations. The analyses considered recovery efficiency (percentage recovery value), which was used as a weighting factor. The results obtained, particularly with respect to the peak concentrations, remained consistent even after adjustment for recovery values.

Figure 3 illustrates the CUSUM charts that show the interaction between positive cases and wastewater trends, plotting progressive observations on the x-axis, while the y-axis shows the cumulative deviation from the target, highlighting shifts and trends over time. Notable spikes indicate significant increases, suggesting potential outbreaks or changes in viral concentration.

In Figure 3, we observe a significant positive differential value beyond the upper limit at the 106th data point in the CUSUM Positive Cases Trend chart, which corresponds to the observation made on January 19, 2022. This indicates a substantial increase in positive cases. When comparing this with the CUSUM Wastewater Trend chart, several positive and negative peaks appear. In particular, there is a peak in wastewater effluent between December 30, 2021, and January 25, 2022, and each of these effluent peaks aligns with a subsequent peak in positive case differentials between January 3 and 31, 2022.

Table 2 presents data on the correspondence between significant peaks in wastewater SARS-CoV-2 RNA concentrations and subsequent peaks in COVID-19 cases, distinguishing between total cases, hospitalized cases, and admissions to the ICU.

Figure 3. CUSUM analysis of COVID-19–positive cases (top) and SARS-CoV-2 RNA detected in wastewater (bottom) in the Veneto region, covering the period from October 5, 2021, to February 14, 2022. CUSUM: cumulative sum.

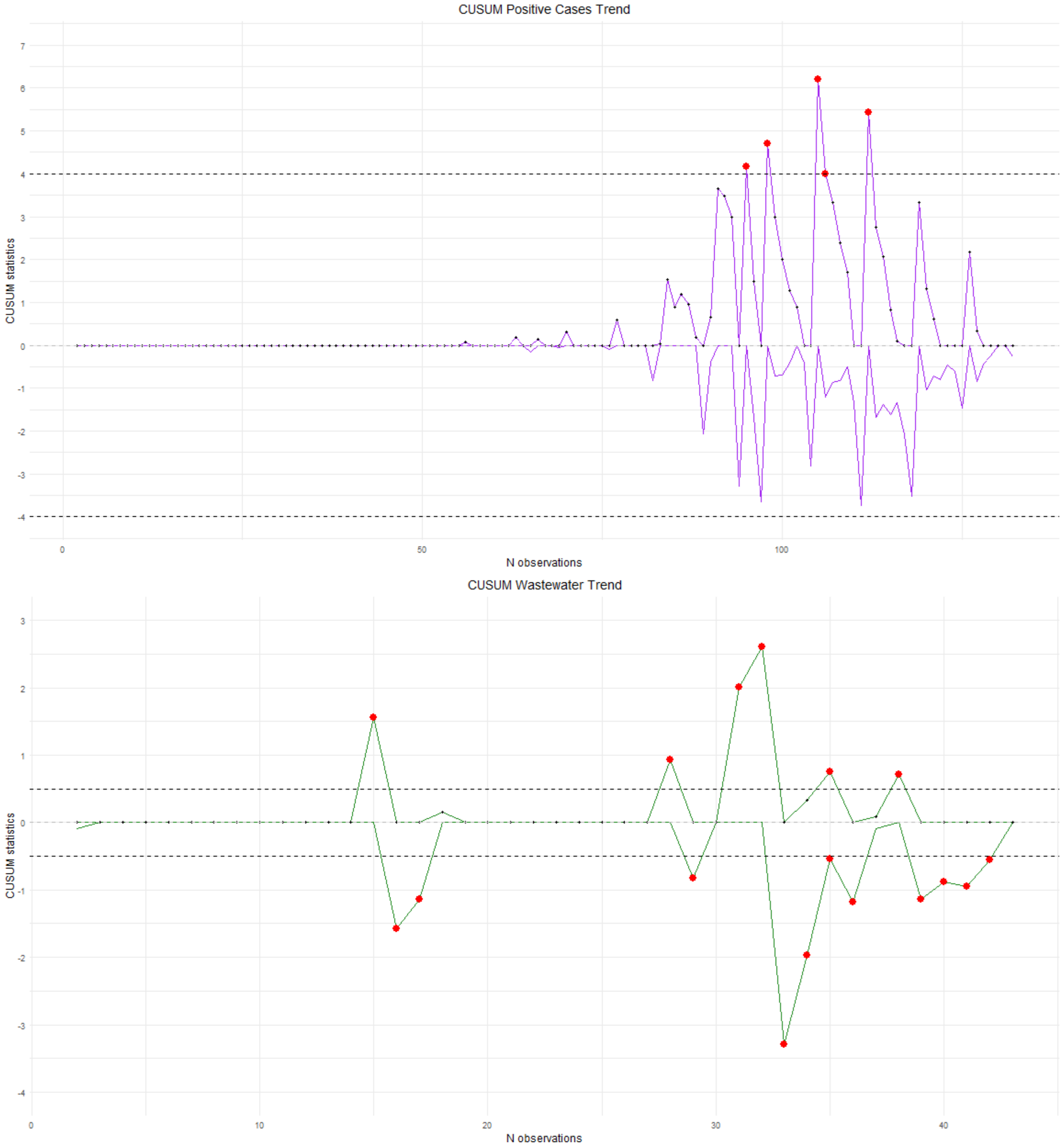


Table . Temporal correspondence between peaks in SARS-CoV-2 RNA concentrations in wastewater and COVID-19 cases in the Veneto region (October 5, 2021, to February 14, 2022).

Dates of wastewater peaks	SARS-CoV-2 RNA concentration, gc/μL	Dates of COVID-19 peak cases	COVID-19 cases, n	Difference in days, n
Total positive COVID-19 cases				
12/21/2021	100.75	12/27/2021	8883	6
12/30/2021	188.73	1/03/2022	16,848	4
1/04/2022	247.79	1/10/2022	22,024	6
1/13/2022	140.62	1/17/2022	25,283	4
1/25/2022	172.71	1/31/2022	15,372	6
Hospitalized COVID-19 cases				
12/21/2021	100.75	12/27/2021	284	6
12/30/2021	188.73	1/03/2022	265	4
1/04/2022	247.79	1/07/2022	253	3
1/13/2022	140.62	1/17/2022	254	4
Hospitalized in intensive care units COVID-19 cases				
12/21/2021	100.75	12/27/2021	18	6

For example, for total positive cases, it should be noted that effluent peaks on December 30, 2021, and January 4, 2022, were followed by significant increases in positive cases, with a delay of 4 to 6 days. This suggests a correlation between wastewater changes and surges in total cases of COVID-19.

Analysis of the CUSUM charts reveals that effluent concentration peaks consistently preceded positive case peaks by an average of 5.2 days for total positive cases. In the case of hospitalized patients, the average time lag between the detection of a significant effluent value and the subsequent detection of a significant value in hospitalizations was 4.25 days.

Similarly, for positive cases that required hospitalization, effluent peaks on December 30, 2021, and January 4, 2022, were followed by case peaks, although with a slightly shorter delay of 3 to 4 days. This close temporal relationship between effluent peaks and hospitalizations is a significant finding.

Furthermore, data indicate that effluent peaks on December 21 and 27, 2021, coincided with the increase in ICU admissions 6 days later, underlining a consistent trend in which wastewater changes anticipated changes in the need for critical care.

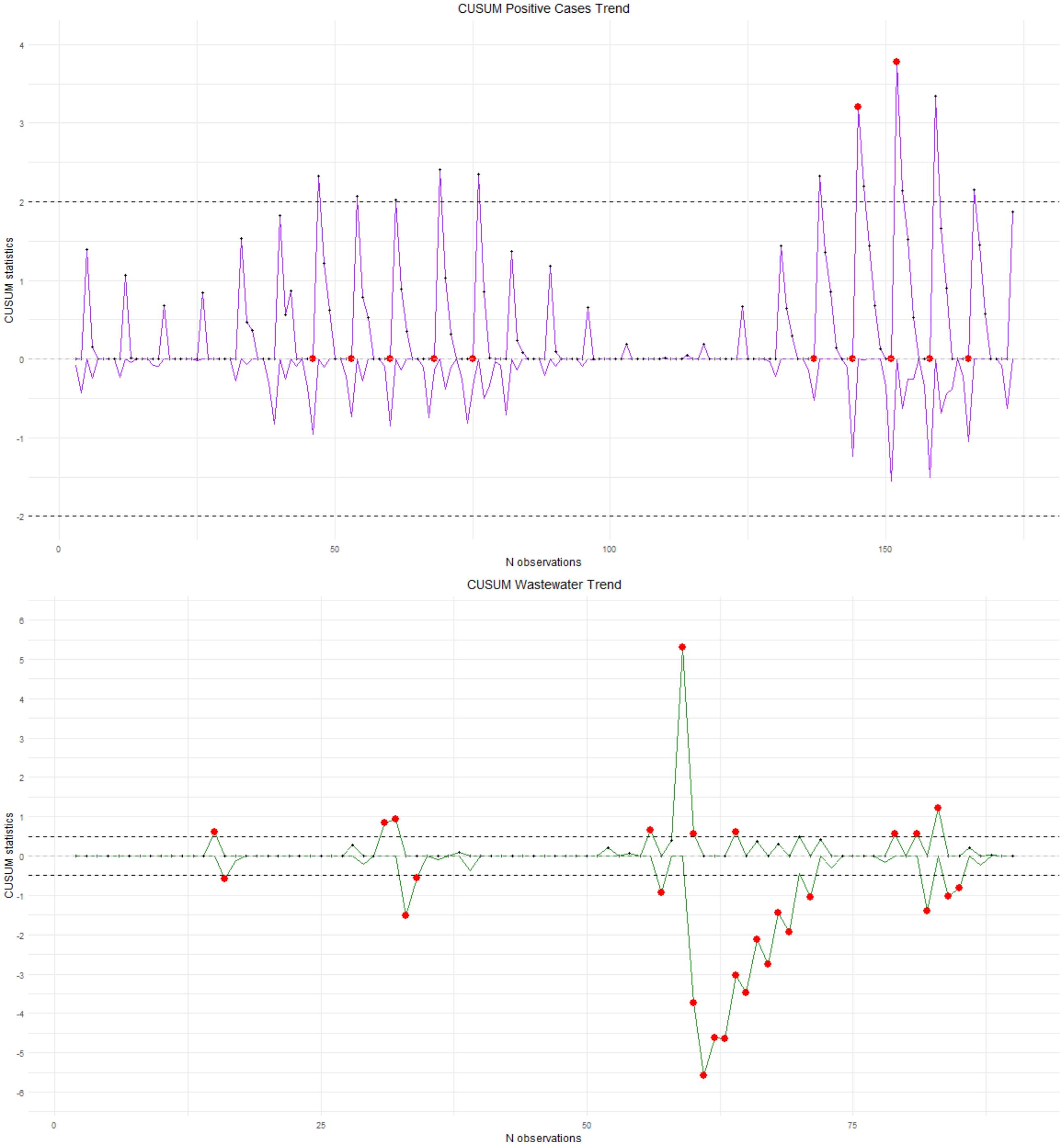
CUSUM Chart Results in the Second Period

To validate the hypotheses and results, the same analysis was performed on the series of positive cases within the time frame from January 14 to August 2, 2022, (Figure 4). Furthermore, Figure 4 shows the complete series of the differential of viral load in wastewater effluent.

As shown in Table 3, seven turning points in the effluent series corresponded to turning points in the positive case series. The average number of days between a wastewater peak and a clinically confirmed case peak was 6. In particular, a peak in the effluent dated April 7, 2022, was associated with an elevated positive case count, although it did not differ significantly from other outbreak values.

The results of the comparison of the CUSUM charts for total positive cases and effluents until August 2, 2022, align with the findings observed in the analysis up to February 13, 2022. The average number of days between an effluent peak and a positive case peak was 6.

Figure 4. CUSUM analysis of positive cases for COVID-19 (top) and SARS-CoV-2 RNA trends in wastewater (bottom) in the Veneto region, Italy, February 14, 2022, to August 2, 2022. CUSUM: cumulative sum.


CUSUM Positive Cases Trend

CUSUM Wastewater Trend

Table . Temporal correspondence between peaks in SARS-CoV-2 RNA concentrations in wastewater and COVID-19 cases in the Veneto region (February 14, 2022, to August 2, 2022).

Dates of wastewater peaks	SARS-CoV-2 RNA concentration, gc/μL	Dates of COVID-19 peak cases	Total positive COVID-19 cases, n	Difference in days, n
3/29/2022	227.96	4/05/2022	8205	7
4/07/2022	715.86	4/12/2022	7854	5
4/12/2022	290.86	4/20/2022	8922	8
4/26/2022	158.38	4/27/2022	8814	1
6/21/2022	109.60	6/28/2022	8753	7
6/28/2022	209.81	7/05/2022	10,884	7
7/05/2022	189.48	7/12/2022	12,571	7

Discussion

Principal findings

This study provides valuable information on the use of WBE as an early warning system for COVID-19 outbreaks in the Veneto region of Italy. Using CUSUM charts, we were able to detect significant changes in viral RNA concentrations in wastewater and assess their correlation with COVID-19 clinical outcomes, such as positive cases, hospitalizations, and admissions to the ICU.

Our findings confirm the potential of WBE to anticipate COVID-19 trends, with viral peaks in wastewater consistently preceding peaks in reported cases by an average of 5.2 days. Similarly, wastewater data provided early indications of increases in hospitalizations and admissions to the ICU, with a lag time of 4.25 and 6 days, respectively. These lead times are crucial for public health preparedness, allowing authorities to mobilize resources and implement containment measures before the number of clinical cases surges.

Our results align with previous WBE studies for the monitoring of SARS-CoV-2. Henriques et al [32], in their 2023 study, underscore the importance of WBE in assessing the presence and prevalence of SARS-CoV-2, particularly in regions with limited clinical data. WBE is a crucial complementary tool for the early detection and surveillance of COVID-19 outbreaks. It helps to estimate clinical cases and evaluate the effectiveness of vaccination programs [32].

The systematic review and meta-analysis by Li et al [33] emphasize the varying correlation between SARS-CoV-2 RNA concentrations in wastewater and clinically confirmed cases of COVID-19. Environmental factors, WBE sampling designs, and epidemiological conditions influence the strength of this correlation. Their study underscores the need to understand the dynamics of viral shedding, in-sewer decay, and sampling strategies to accurately estimate COVID-19 cases through WBE [33]. Furthermore, a study by Sodhi and Singh in 2022 [34] explores SARS-CoV-2 in wastewater, highlighting the presence of the virus even before symptoms. They advocate revised guidelines and advanced viral remediation techniques.

The temporal correlations between wastewater trends and clinical outcomes underscore the potential usefulness of WBE as a supplementary tool for public health surveillance. However, our analysis revealed a lead time of only 5 days when considering COVID-19–positive cases. Although a 5-day warning can help prepare health care resources, it comes with certain limitations. Relatively short lead times can lead to a faster but less reliable response, especially given the complexities of reorganizing health care resources and managing intensive care. Therefore, while we acknowledge the value of short-term alerts, it is crucial to strike a balance between timeliness and reliability when implementing preventive and resource management measures. To better focus on the increase in health care resources, it is essential to obtain more data on hospitalized cases and admissions to the ICU for COVID-19. However, during the considered period, these data are not robust enough to suggest a mean number of days to prepare for an increase in the number of beds for admitted patients.

The analysis in Table 2 shows a 13-day gap between the peak in the SARS-CoV-2 RNA concentration in wastewater on January 4, 2022, and the peak in reported cases of COVID-19 on January 17, 2022. This longer-than-expected delay highlights the variability inherent in WBE. Several factors could explain this discrepancy, including complex disease dynamics, inconsistent lag times due to varying public health responses, and potential delays in testing and reporting during periods of high case numbers. These factors underscore the need for careful interpretation of the correlations between wastewater data and clinical case trends, as the relationship may not always be linear or immediate.

The comparison between the fall of 2021 and the spring of 2022 reveals a potential gap in clinical surveillance, identified through the analysis of divergent trends between wastewater concentrations and reported cases of COVID-19. Specifically, while wastewater data indicated fluctuations in viral RNA concentrations, the number of reported COVID-19 clinical cases began to decline from April 2022, suggesting either a decrease in testing or reporting accuracy or a genuine decrease in cases not fully captured by the wastewater analysis. This divergence highlights the importance of integrating multiple data sources, including wastewater and clinical data, to provide a comprehensive overview of disease trends.

Despite current challenges, future research aims to improve sampling and analysis protocols, understand viral dynamics in sewer systems, and exploit artificial intelligence and big data to better monitor wastewater. The study by Nayak et al [18] highlights the potential of machine learning, particularly in diverse sewer sheds, using time series models such as long short-term memory and various data parameters for accurate trend forecasting. In parallel, the work by Jiang et al [35] underscores the crucial role of WBE in tracking COVID-19 transmission.

In the field of WBE, several methodologies have been developed to predict future COVID-19 cases, including our method, the EPISENS-M technique of Ando et al, and the statistical framework proposed by Dai et al. Although all 3 methods share the objective of using wastewater data to monitor and forecast the evolution of the pandemic, they differ significantly in their approach, robustness, and applicability.

The EPISENS-M method, developed by Ando et al [20], represents a significant advancement in the sensitive detection of SARS-CoV-2 RNA in wastewater. This method utilizes membrane adsorption and direct RNA extraction to enhance detection sensitivity, allowing for accurate prediction of clinical cases even in low-prevalence settings. However, a key limitation of this approach is its reliance on recent clinical data for calibration, which may reduce its effectiveness in situations where such data are not readily available.

On the other hand, Dai et al [21] developed a sophisticated statistical framework based on Bayesian models and functional principal component analysis to address the challenges posed by noisy and sparsely sampled data. Although their approach is extremely useful for interpreting viral concentration data in wastewater and mitigating fluctuations due to technical errors, it is not designed to directly predict clinical cases of COVID-19. This limits its applicability as a direct epidemiological forecasting tool.

The use of the CUSUM method in this study marks a novel approach to identifying subtle shifts in SARS-CoV-2 trends, offering a more sensitive detection of deviations from established baselines. Unlike EPISENS-M, our method is less dependent on recent clinical data, making it potentially more robust in contexts with limited clinical monitoring. Furthermore, compared to the method of Dai et al, our approach focuses specifically on epidemiological prediction, providing a direct advantage for public health planning.

Limitations

Despite its potential, our study has limitations. The reliance on a single region and a specific time frame restricts the generalizability of our findings.

Furthermore, the significant heterogeneity observed in flow rates, suspended solid concentrations, and volume distribution provides information on the differences in the sites considered. The sites reflect the data on specific cities; so, while we have the number of positive cases per province, we need to exercise caution when generalizing these findings to the whole region. Additionally, the study's reliance on SARS-CoV-2 RNA concentrations as an early indicator may overlook other potential contributors to wastewater dynamics and require further validation through complementary approaches.

This study focused on the correlation between SARS-CoV-2 RNA concentrations in wastewater and clinical indicators of COVID-19, without accounting for the dynamics of population mobility or additional biomarkers in wastewater. While these factors can influence wastewater analysis and provide deeper insights, they were beyond the scope of the current study. Future research could integrate mobility data and additional biomarkers to improve the predictive capacity of WBE.

Validation analyses were performed only for positive cases due to a low number of peaks for hospitalized and ICU cases. Therefore, the study acknowledges the need for further validation and standardization of analytical methods in wastewater analysis. Additionally, while the study highlights correlations, the cause cannot be definitively established. Future research could explore the generalizability of these findings in different regions and populations.

Conclusions

This study confirms that WBE, even when applied at a regional level with limited data points, can provide early warnings for COVID-19 outbreaks. The use of CUSUM charts enabled the sensitive and timely detection of shifts in viral load trends, offering crucial lead times for public health interventions. Our findings highlight the value of integrating WBE into public health surveillance systems, especially in areas where clinical data may not be immediately available.

The results confirm the reliability of wastewater analysis in anticipating COVID-19 trends, particularly at the LHU level, which is vital for informed public health planning and proactive responses. The potential of WBE to monitor the spread of SARS-CoV-2, including emerging variants, underscores the importance of continued efforts to improve wastewater monitoring and treatment.

More research is essential to refine the detection and quantification methods for pathogens and viruses in wastewater, as well as to evaluate their feasibility in different contexts. Ultimately, this approach offers a robust and valuable tool for epidemiological surveillance, contributing significantly to the early detection and management of infectious disease outbreaks with pandemic potential.

Acknowledgments

This work was supported by the scope of the Centro Nazionale per la Prevenzione e il Controllo delle Malattie of the Italian Ministry of Health 2020 Project “Modelli organizzativi per l'efficientamento delle attività di contact tracing” (Organizational models to improve the efficiency of contact tracing activities) funded by the Italian Ministry of Health. We appreciate the assistance of the staff of the Consorzio per la Ricerca Sanitaria, who helped us with the administrative aspects of the project.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

HO and DG contributed to the conception and design of the study. IA, TB, FR, MZ, and VG collected the data. IA, TB, FR, and MZ conducted the laboratory analysis. HO and MZ conducted the statistical analysis. HO, MZ, and all other authors interpreted the results. HO prepared the draft manuscript. FDR and VB provided resources. All authors reviewed the results and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Ahmad J, Ahmad M, Usman ARA, Al-Wabel MI. Prevalence of human pathogenic viruses in wastewater: a potential transmission risk as well as an effective tool for early outbreak detection for COVID-19. J Environ Manage 2021 Nov 15;298:113486. [doi: 10.1016/j.jenvman.2021.113486] [Medline: 34391102]
2. Asadi A, Fakhri Y, Salimi Y, Daglioglu N, Tahmasebifard M, Aghajarinezhad M. Nicotine consumption rate through wastewater-based epidemiology: a systematic review, meta-analysis and probabilistic risk assessment. Environ Sci Pollut Res Int 2023 May;30(23):63416-63426. [doi: 10.1007/s11356-023-27017-x] [Medline: 37084052]
3. Tiwari A, Kurittu P, Al-Mustapha AI, et al. Wastewater surveillance of antibiotic-resistant bacterial pathogens: a systematic review. Front Microbiol 2022 Dec 15;13:977106. [doi: 10.3389/fmicb.2022.977106] [Medline: 36590429]
4. Chau KK, Barker L, Budgell EP, et al. Systematic review of wastewater surveillance of antimicrobial resistance in human populations. Environ Int 2022 Apr;162:107171. [doi: 10.1016/j.envint.2022.107171] [Medline: 35290866]
5. Holton E, Sims N, Jagadeesan K, Standerwick R, Kasprzyk-Hordern B. Quantifying community-wide antimicrobials usage via wastewater-based epidemiology. J Hazard Mater 2022 Aug 15;436:129001. [doi: 10.1016/j.jhazmat.2022.129001] [Medline: 35594673]
6. La Rosa G, Iaconelli M, Mancini P, et al. First detection of SARS-CoV-2 in untreated wastewaters in Italy. Sci Total Environ 2020 Sep 20;736:139652. [doi: 10.1016/j.scitotenv.2020.139652] [Medline: 32464333]
7. Giacobbo A, Rodrigues MAS, Zoppas Ferreira J, Bernardes AM, de Pinho MN. A critical review on SARS-CoV-2 infectivity in water and wastewater. What do we know? Sci Total Environ 2021 Jun 20;774:145721. [doi: 10.1016/j.scitotenv.2021.145721] [Medline: 33610994]
8. Han C, Duan C, Zhang S, et al. Digestive symptoms in COVID-19 patients with mild disease severity: clinical presentation, stool viral RNA testing, and outcomes. Am J Gastroenterol 2020 Jun;115(6):916-923. [doi: 10.14309/ajg.0000000000000664] [Medline: 32301761]
9. Rimoldi SG, Stefani F, Gigantiello A, et al. Presence and infectivity of SARS-CoV-2 virus in wastewaters and rivers. Sci Total Environ 2020 Nov 20;744:140911. [doi: 10.1016/j.scitotenv.2020.140911] [Medline: 32693284]
10. La Rosa G, Mancini P, Bonanno Ferraro G, et al. SARS-CoV-2 has been circulating in northern Italy since December 2019: evidence from environmental monitoring. Sci Total Environ 2021 Jan 1;750:141711. [doi: 10.1016/j.scitotenv.2020.141711] [Medline: 32835962]
11. Randazzo W, Truchado P, Cuevas-Ferrando E, Simón P, Allende A, Sánchez G. SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. Water Res 2020 Aug 15;181:115942. [doi: 10.1016/j.watres.2020.115942] [Medline: 32425251]
12. Wu F, Xiao A, Zhang J, et al. Wastewater surveillance of SARS-CoV-2 across 40 U.S. states from February to June 2020. Water Res 2021 Sep 1;202:117400. [doi: 10.1016/j.watres.2021.117400] [Medline: 34274898]
13. Ciannella S, González-Fernández C, Gomez-Pastora J. Recent progress on wastewater-based epidemiology for COVID-19 surveillance: a systematic review of analytical procedures and epidemiological modeling. Sci Total Environ 2023 Jun 20;878:162953. [doi: 10.1016/j.scitotenv.2023.162953] [Medline: 36948304]
14. Peccia J, Zulli A, Brackney DE, et al. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. Nat Biotechnol 2020 Oct;38(10):1164-1167. [doi: 10.1038/s41587-020-0684-z] [Medline: 32948856]
15. Vallejo JA, Trigo-Tasende N, Rumbo-Feal S, et al. Modeling the number of people infected with SARS-COV-2 from wastewater viral load in Northwest Spain. Sci Total Environ 2022 Mar 10;811:152334. [doi: 10.1016/j.scitotenv.2021.152334] [Medline: 34921882]
16. Chekkala A, Atasoy M, Williams C, Cetecioglu Z. Statistical analysis of SARS-CoV-2 using wastewater-based data of Stockholm, Sweden. Int J Environ Res Public Health 2023 Feb 26;20(5):4181. [doi: 10.3390/ijerph20054181] [Medline: 36901194]
17. Howard I, Espigares E, Lardelli P, Martín JL, Espigares M. Evaluation of microbiological and physicochemical indicators for wastewater treatment. Environ Toxicol 2004 Jun;19(3):241-249. [doi: 10.1002/tox.20016] [Medline: 15101039]

18. Nayak SR, Arora V, Sinha U, Poonia RC. A statistical analysis of COVID-19 using Gaussian and probabilistic model. J Interdiscip Math 2020 Nov 25;24(1):19-32. [doi: 10.1080/09720502.2020.1833442]
19. Karthikeyan S, Ronquillo N, Belda-Ferre P, et al. High-throughput wastewater SARS-CoV-2 detection enables forecasting of community infection dynamics in San Diego County. mSystems 2021 Mar 2;6(2):e00045-21. [doi: 10.1128/mSystems.00045-21] [Medline: 33653938]
20. Ando H, Murakami M, Ahmed W, Iwamoto R, Okabe S, Kitajima M. Wastewater-based prediction of COVID-19 cases using a highly sensitive SARS-CoV-2 RNA detection method combined with mathematical modeling. Environ Int 2023 Mar;173:107743. [doi: 10.1016/j.envint.2023.107743] [Medline: 36867995]
21. Dai X, Champredon D, Fazil A, et al. Statistical framework to support the epidemiological interpretation of SARS-CoV-2 concentration in municipal wastewater. Sci Rep 2022 Aug 5;12(1):13490. [doi: 10.1038/s41598-022-17543-y] [Medline: 35931713]
22. Noyez L. Control charts, Cusum techniques and funnel plots. A review of methods for monitoring performance in healthcare. Interact Cardiovasc Thorac Surg 2009 Sep;9(3):494-499. [doi: 10.1510/icvts.2009.204768] [Medline: 19509097]
23. Baldovin T, Amoruso I, Fonzo M, et al. Trends in SARS-CoV-2 clinically confirmed cases and viral load in wastewater: a critical alignment for Padua city (NE Italy). Heliyon 2023 Oct;9(10):e20571. [doi: 10.1016/j.heliyon.2023.e20571] [Medline: 37822618]
24. La Rosa G, Bonadonna L, Suffredini E. Protocollo della sorveglianza di SARS-cov-2 in reflui urbani (SARI) - rev. 3 [Article in Italian]. Zenodo. 2021. URL: https://zenodo.org/records/5758725 [accessed 2024-12-09]
25. Page ES. Continuous Inspection Schemes. Biometrika 1954 Jun;41(1-2):100-115. [doi: 10.1093/biomet/41.1-2.100]
26. Novoa NM, Varela G. Monitoring surgical quality: the cumulative sum (CUSUM) approach. Mediastinum 2020 Mar 25;4:4. [doi: 10.21037/med.2019.10.01] [Medline: 35118272]
27. Govindaraju K, Lai CD. Run length variability and three sigma control limits. Stoch Qual Control 2004;19(2):175-184. [doi: 10.1515/EQC.2004.175]
28. The R Project for Statistical Computing. URL: https://www.R-project.org/ [accessed 2024-11-19]
29. Scrucca L. Qcc: an R package for quality control charting and statistical process control. R News 2004 Jun;4(1):11-17 [FREE Full text]
30. Wickham H. Data analysis. In: ggplot2: Elegant Graphics for Data Analysis, 2nd edition: Springer; 2016:189-201. [doi: 10.1007/978-3-319-24277-4_9]
31. Wickham H. ggplot2: Elegant Graphics for Data Analysis, 2nd edition: Springer; 2016. [doi: 10.1007/978-3-319-24277-4]
32. Henriques TB, Cassini ST, de Pinho Keller R. Contribution of wastewater-based epidemiology to SARS-CoV-2 screening in Brazil and the United States. J Water Health 2023 Mar;21(3):343-353. [doi: 10.2166/wh.2023.260] [Medline: 37338314]
33. Li X, Zhang S, Sherchan S, et al. Correlation between SARS-CoV-2 RNA concentration in wastewater and COVID-19 cases in community: a systematic review and meta-analysis. J Hazard Mater 2023 Jan 5;441:129848. [doi: 10.1016/j.jhazmat.2022.129848] [Medline: 36067562]
34. Sodhi KK, Singh CK. A systematic review on the occurrence, fate, and remediation of SARS-CoV-2 in wastewater. Int J Environ Sci Technol (Tehran) 2023;20(7):8073-8086. [doi: 10.1007/s13762-022-04326-1] [Medline: 35755183]
35. Jiang G, Liu Y, Tang S, et al. Moving forward with COVID-19: future research prospects of wastewater-based epidemiology methodologies and applications. Curr Opin Environ Sci Health 2023 Jun;33:100458. [doi: 10.1016/j.coesh.2023.100458] [Medline: 37034453]

Abbreviations

ARIMA: autoregressive integrated moving average
CUSUM: cumulative sum
EPISENS-M: Efficient and Practical Virus Identification System with Enhanced Sensitivity for Membrane
gc: genome copies
ICU: intensive care unit
LHU: local health unit
ORF1ab: open reading frames 1ab
PCR: polymerase chain reaction
WBE: wastewater-based epidemiology
WTP: wastewater treatment plant

Edited by A Mavragani; submitted 27.06.24; peer-reviewed by A Berzins, RK Mishra, RDF Bueno; revised version received 31.08.24; accepted 02.09.24; published 14.01.25.

Please cite as:
Ocagli H, Zambito M, Da Re F, Groppi V, Zampini M, Terrini A, Rigoli F, Amoruso I, Baldovin T, Baldo V, Russo F, Gregori D
Wastewater Monitoring During the COVID-19 Pandemic in the Veneto Region, Italy: Longitudinal Observational Study
JMIR Public Health Surveill 2025;11:e58862
URL: https://publichealth.jmir.org/2025/1/e58862
doi:10.2196/58862

Characterizing US Spatial Connectivity and Implications for Geographical Disease Dynamics and Metapopulation Modeling: Longitudinal Observational Study

Giulia Pullano[1], PhD; Lucila Gisele Alvarez-Zuzek[2], PhD; Vittoria Colizza[3], PhD; Shweta Bansal[1], PhD

[1]Department of Biology, Georgetown University, 37th and O Streets NW, Washington, DC, United States

[2]Fondazione Bruno Kessler, Trento, Italy

[3]Sorbonne Université, INSERM, Institut Pierre Louis d’Epidémiologie et de Santé Publique (IPLESP), Paris, France

Corresponding Author:
Shweta Bansal, PhD
Department of Biology, Georgetown University, 37th and O Streets NW, Washington, DC, United States

Abstract

Background: Human mobility is expected to be a critical factor in the geographic diffusion of infectious diseases, and this assumption led to the implementation of social distancing policies during the early fight against the COVID-19 emergency in the United States. Yet, because of substantial data gaps in the past, what still eludes our understanding are the following questions: (1) How does mobility contribute to the spread of infection within the United States at local, regional, and national scales? (2) How do seasonality and shifts in behavior affect mobility over time? (3) At what geographic level is mobility homogeneous across the United States?

Objective: This study aimed to address the questions that are critical for developing accurate transmission models, predicting the spatial propagation of disease across scales, and understanding the optimal geographical and temporal scale for the implementation of control policies.

Methods: We analyzed high-resolution mobility data from mobile app usage from SafeGraph Inc, mapping daily connectivity between the US counties to grasp spatial clustering and temporal stability. Integrating this into a spatially explicit transmission model, we replicated SARS-CoV-2’s first wave invasion, assessing mobility’s spatiotemporal impact on disease predictions.

Results: Analysis from 2019 to 2021 showed that mobility patterns remained stable, except for a decline in April 2020 due to lockdowns, which reduced daily movements from 45 million to approximately 25 million nationwide. Despite this reduction, intercounty connectivity remained seasonally stable, largely unaffected during the early COVID-19 phase, with a median Spearman coefficient of 0.62 (SD 0.01) between daily connectivity and gravity networks. We identified 104 geographic clusters of US counties with strong internal mobility connectivity and weaker links to counties outside these clusters. These clusters were stable over time, largely overlapping state boundaries (normalized mutual information=0.82) and demonstrating high temporal stability (normalized mutual information=0.95). Our findings suggest that intercounty connectivity is relatively static and homogeneous at the substate level. Furthermore, while county-level, daily mobility data best captures disease invasion, static mobility data aggregated to the cluster level also effectively models spatial diffusion.

Conclusions: Our work demonstrates that intercounty mobility was negligibly affected outside the lockdown period in April 2020, explaining the broad spatial distribution of COVID-19 outbreaks in the United States during the early phase of the pandemic. Such geographically dispersed outbreaks place a significant strain on national public health resources and necessitate complex metapopulation modeling approaches for predicting disease dynamics and control design. We thus inform the design of such metapopulation models to balance high disease predictability with low data requirements.

(JMIR Public Health Surveill 2025;11:e64914) doi:10.2196/64914

KEYWORDS

geographical disease dynamics; spatial connectivity; mobility data; metapopulation modeling; COVID-19; human mobility; infectious diseases; social distancing; epidemic; mobile apps; SafeGraph; SARS-CoV-2; coronavirus; pandemic; spatio-temporal; US; public health; mobile health; mHealth; digital health; health informatics

Introduction

Human mobility plays a crucial role in the spread of respiratory diseases [1]. The combination of regional travel and local commuting represents the spatial connectivity between locations, serving as the main driver in the geographic diffusion of infectious diseases. Characterizing the spatial dynamics of pathogen transmission is, therefore, intricately tied to unraveling human mobility patterns. Such a task has proven to be challenging due to the inherent complexity and privacy-related limitations on collecting mobility data [2]. Over the past few decades, researchers have extensively relied on mobility data obtained from census records, surveys, transportation statistics, commuting data, and international air traffic data. Such datasets have widely contributed to a better understanding of human mobility patterns and their impact on the epidemic spread [3-6], but can be limited in their resolution or scale. More recently, this gap has been filled by the use of mobile phone data [7,8], primarily based on phone records, but no such data have been available in the United States.

The global health crisis triggered by COVID-19 has underscored the critical need for swift access to mobility to help mitigate the spread of the virus. The urgency of the situation prompted an unprecedented sharing of data by private companies worldwide, through legally and ethically compliant agreements. These data were based on mobile location–based app usage and thus provided incomparable access to high-resolution, large-scale, and near-real-time mobility data that have expanded human mobility science [9] and computational epidemiology [10,11]. The availability of these data has especially represented a shift in the US public health and it has been used to inform epidemic models and reveal the impact of mitigation strategies on behavior [12-16]. Although the association between mobility patterns and COVID-19 transmission in the United States has been extensively explored, for example, in studies by Pei et al [14], Badr et al [16], Xiong et al [17], Tokey [18], Zheng et al [19], and Gao et al [20], no studies have been devoted to assessing when the underlying mobility network needs to be embedded into models to characterize epidemic spread.

Furthermore, the effects of control measures such as social distancing and travel restrictions as well as the most suitable geographical and temporal granularity for implementing these measures still lack clarity. This gap in understanding the characteristic spatio-temporal scale of mobility not only limits target control policies but also our ability to model transmission dynamics effectively. To date, mobility data have been integrated into epidemic models without due consideration for the optimal geographical (eg, municipalities, regions, and states) and temporal resolution (eg, day, week, and month) required to accurately capture epidemic spread. The level of granularity used in these models has consistently been dictated by a priori assessments from data providers [21,22].

To address these gaps, this study made 3 key contributions. First, for characterizing spatiotemporal scales, we systematically analyzed human mobility patterns across the United States using high-resolution mobile app–based location data to characterize intercounty connectivity at different spatial and temporal scales before and after the COVID-19 pandemic. Second, for quantifying mobility's impact on the geographic spread of disease, we integrated these mobility data into spatially explicit transmission models to assess its role in the geographic diffusion of SARS-CoV-2 during the first wave. Finally, for evaluating the predictive power across scales, we analyzed how the predictive ability of epidemic models changes with varying resolutions of mobility data, identifying the spatial and temporal scales of intercounty connectivity that significantly influence disease dynamics. This analysis allowed us also to explore the trade-offs between using fine-grained mobility data and aggregated data at coarser scales.

While our approach yields valuable insights into human behavior and disease dynamics and makes a theoretical contribution to the field, we acknowledge its limitations—most notably the simplifications in the disease model, such as assuming homogeneous mixing within counties. However, our focus was on understanding heterogeneity at scales larger than counties, as most public health data are typically available at the county level. This approach represents a significant step toward understanding the role of mobility in geographical disease diffusion and optimizing the integration of mobility data into epidemic models to better inform public health policies.

Methods

Study Design and Population

This study investigated the role of human mobility at various temporal and spatial scales in the spread of COVID-19 across the US counties, using mobility and disease incidence data. First, we examined human displacements and the resulting intercounty connectivity patterns in the United States based on the daily number of visits between census blocks. We also analyzed disease incidence during the early phases of the COVID-19 pandemic, from January 2019 to July 2020, and used an inference framework to estimate underreporting based on the number of new deaths. In the second step, we developed a metapopulation model to simulate disease transmission, incorporating intercounty mobility patterns and epidemiological data, adjusted for underreporting. The study focused on US counties with populations greater than 10,000 to minimize biases in mobility and incidence data.

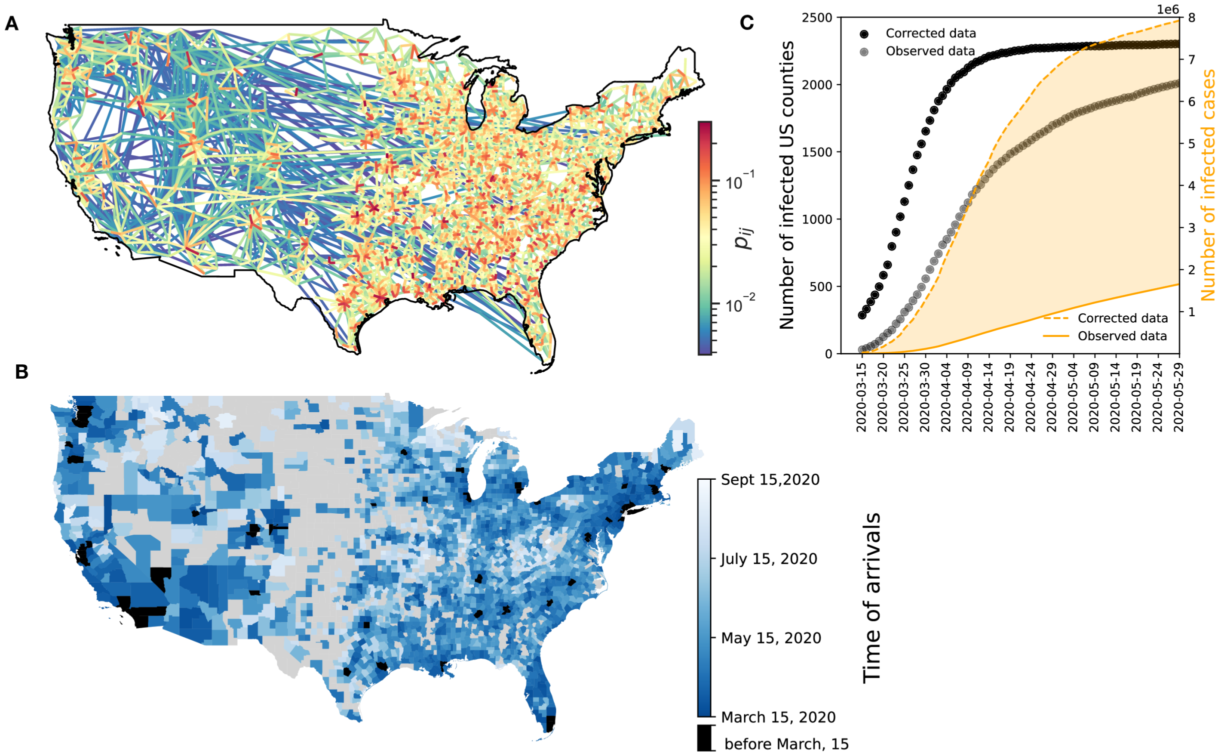
Data Collection

Characterizing Intercounty Connectivity With Mobility Data

We used data from SafeGraph [23] (now Advan Patterns [Advan Inc]), a platform gathering mobility data from location-based mobile app usage. Specifically, we used the daily Social Distancing dataset provided by SafeGraph (refer to Multimedia Appendix 1 for dataset details). This dataset includes information on the number of mobile devices with a home from a census block group visiting another census block group or staying in the originating one for at least a minute. The data covered the period from January 2019 to April 2021 on a daily basis. These data were aggregated to the US county level to ensure consistency in geographic scale for disease surveillance and public health decision-making. A correction factor addresses spatial and temporal variations (Multimedia Appendix 1).

We quantified monthly intercounty connectivity by normalizing visits from an origin county to destination counties and calculated average daily visits for all county pairs from January 2019 to March 2021. We then normalized the results so that the sum of probabilities for each county equals 1. The resulting time-evolving network reflected daily coupling probabilities. Figure 1A shows the connectivity network for March 2020. Figure S2 in Multimedia Appendix 1 compares monthly and daily datasets.

Figure 1. Data sources and epidemic context. (A) Mobility data. The figure shows the spatial connectivity network between US counties in March 2020. The map shows only links with the top 1% of coupling probability (p_{ij}>.0038). (B-C) Public health data. (B) The map shows the time of arrival by county, defined as the time when the total number of cases (observed+unreported) in counties that reached at least 10 cases. Black-colored counties are counties that have been infected before March 15, 2020. (C) Black dots show the corrected number of infected counties over time nationally. Gray dots show the observed number of infected counties over time nationally. The orange solid line shows the daily number of new confirmed cases nationally, while the orange dotted line shows the real number of cases nationally accounting for underreporting.



Characterizing Early Phase of COVID-19 in the United States With Public Health Data

The first confirmed case of COVID-19 in the United States emerged in Washington state on January 21, 2020, quickly leading to local transmission. Guidelines promoting social distancing and discouraging gatherings were issued on March 16, 2020. While many European countries enacted nationwide restrictions, the United States implemented localized measures at varying times. Lockdowns peaked in the United States in April 2020, with over 40 states enforcing stay-at-home or shelter in-place orders [24]. Despite these efforts, COVID-19 continued to spread, with most US counties reporting cases by June 2020. As cases surged again in October 2020, recommendations for social distancing were made to keep epidemic activity in check.

The COVID-19 disease incidence data were derived from the Centers for Disease Control and Prevention (CDC) data [24]. We used daily reports of new cases and the timing of their arrival in each US county, defined as the day when at least 10 cases were recorded in that area. Daily new-reported cases and time of arrivals were adjusted for potential underreporting at the county level, estimated using global data on COVID-19 cases and fatalities, as outlined in a study by Russell et al [25] (Figures 1B and 1C).

Statistical Analysis

Describing Temporal and Spatial Variability in the Mobility Network

We examined the monthly network structure to evaluate the temporal mobility patterns, quantifying each county's degree (number of connections to other counties). We also defined link persistence as the probability links with nonzero mobility in a month of 2019 that remained in the same month of 2020 and 2021. We also fit a gravity model to the intercounty connectivity network for each month. Community detection analysis using the stochastic InfoMap algorithm developed by Rosvall et al identifies regions with more frequent internal movements [26]. We used a bootstrap resampling method to account for stochasticity (refer to Multimedia Appendix 1 for details). Urban and rural classifications are based on the National Center for Health Statistics Urban-Rural Classification Scheme. All network analyses were performed using Python's NetworkX library [27], which is widely used for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks. Gravity model fitting was conducted using the scikit-mobility library, a Python package designed for mobility data analysis [28].

Incorporating Human Mobility Into Infectious Disease Models

We used a stochastic non-Markovian transmission model with a metapopulation structure at the US county level [21]. The model accounts for disease transmission proportional to (1) infected residents not moving, (2) infected visitors coming from other counties, and (3) returning residents previously infected in other counties. The resulting force of infection in the county i is defined as follows:

λi=λii+∑i≠jλjiv+∑i≠jλijr

λii=βpii2IiNi^

λjiv=βpiipjiIjN^i;λijr=βpijI^jN^j

where p_{ij} is the coupling probability extracted from the intercounty connectivity network. The effective population, and effective number of infections are, respectively, defined as follows:

Ni^=piiNi+∑i≠jpjiNj

Ii^=piiIi+∑i≠jpjiIj

We considered the susceptible-exposed-infectious-recovered epidemic dynamics specific to COVID-19. The epidemics parameters were described by Pullano et al [29].

The detailed mathematical framework, model calibration, and implementation details can be found in Multimedia Appendix 1. The model and inference framework were developed in C++ to optimize performance and efficiency, particularly when working with large-scale data.

Inference Framework and Goodness of Fit

To calibrate the epidemic pathway, we used the cumulative number of infected US counties (Figure 1C). Calibration covered March 14 to July 15, 2020, when all counties reported infections. Parameter estimates β_{pre-LD} (March 15 ‑ 31) and $\beta_{post-LD}$ (March 31-May 15) were derived using maximum likelihood, assuming a Poisson distribution for reported cases.

We assessed model performance by comparing the modeled invasion probability $p_{i,inv}(t)$ with the observed early phase COVID-19 spatial invasion. $p_{i,inv}(t)$ denotes the probability for a county i reporting at least 10 infected cases on day t in the simulation. The goodness of fit is defined as follows:

G(t)=∑i(Iilog pi,inv(t)+(1−Ii)log (1−pi,inv(t)))

I_i=1 if the county i reported at least 10 infected cases at the day t; I_i=0 otherwise.

Comparing Models Across Geographical Scales

In designing the metapopulation structure at different spatial scales, we aimed to better understand the role of mobility heterogeneity that matters in disease diffusion. Beyond this primary goal, this approach also highlighted the data requirements necessary for accurate modeling at different spatial scales. Specifically, it explored the implications of using mobility data at a coarser regional scale (eg, cluster or state), assuming higher-resolution flow data are unavailable.

For a given spatial scale R (eg, cluster and state), we randomized the mobility links among counties within a region R, preserving the number of links. Coupling probabilities for connected counties within each region R are equal to the average coupling probability of the links within that region, as follows:

pijR=∑l,k∈RplkN

with i,j, ∈, and R, while coupling probabilities across regions are not changed.

Ethical Considerations

This study used publicly available, deidentified, and aggregated data from SafeGraph (now Advan Patterns) and the CDC. SafeGraph ensures user privacy by implementing differential privacy techniques, which involve adding Laplacian noise to anonymize data at the Census Block level. In addition, SafeGraph excluded connections associated with census block groups containing data from fewer than 2 devices. Ethical review for this study was sought from the Institutional Review Board at Georgetown University and the study was approved on October 14, 2020 (STUDY00003041).

Results

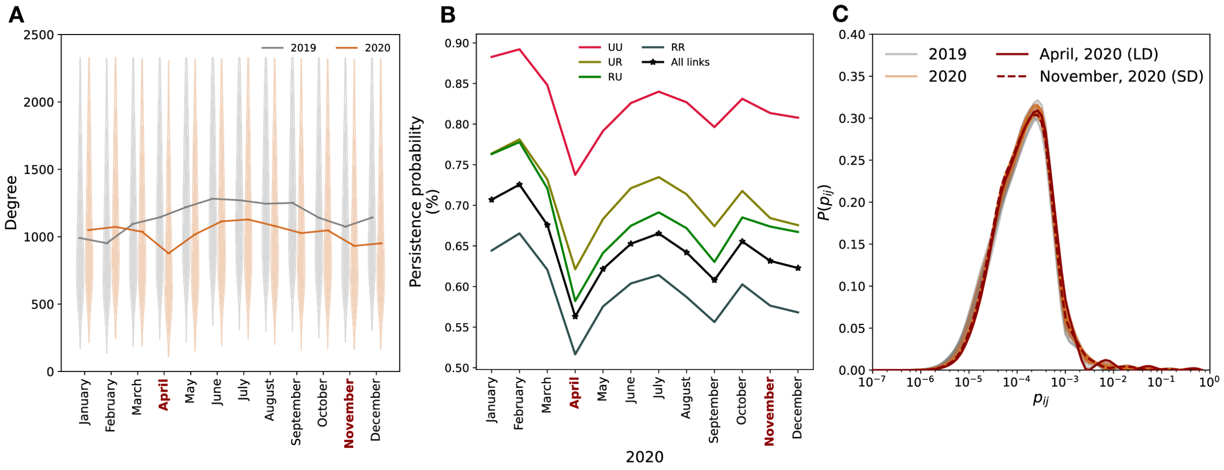
Analyzing the US county connectivity via mobility data revealed temporal and geographical stability, identifying the key scale driving COVID-19's initial spread. Our findings addressed public health needs and determined the optimal scale for metapopulation model design.

Temporal Stability of the Intercounty Connectivity Network

From January 2019 to March 2021, mobility remained relatively stable, except for a significant drop in April 2020 coinciding with lockdown measures, reducing daily visits from 45 million to around 25 million nationwide (Figure 2A). The mobility shock extended throughout the month, encompassing a transitional period (Figures 2A and 2B). Analysis of the intercounty connectivity network's temporal evolution revealed a consistent seasonal pattern in degree distribution and link persistence. Only April 2020 showed local variations, with a 23% reduction in degree (from 1144 to 877) compared with 2019 and a 20% reduction in link persistence (from 0.70 to 0.56) compared to previous months in 2020. Surprisingly, no variation occurred in November 2020, despite strong social distancing recommendations preceding the winter surge of SARS-CoV-2. The reduction in rural-urban connections was particularly pronounced, with a 22% decrease (from 0.75 to 0.58). The decrease stabilized in May and beyond. Notably, urban-urban connections exhibited greater resilience over time when compared with connections involving rural areas. We hypothesized that rural-rural connections were systematically less persistent due to the inherently less stable nature of these links. Factors such as lower population density, reduced economic activity, and less frequent interconnectivity in rural areas contributed to this instability. Furthermore, while coupling probabilities stayed consistent over the study period (Figure 1C), the probability of staying in the home location exhibited larger variability (Multimedia Appendix 1).

Despite occasional local fluctuations, the intercounty connectivity network demonstrated temporal stability and high predictability through a gravity fit model (Figures S7-S9 in Multimedia Appendix 1). The Spearman coefficient between the original and modeled intercounty connectivity network remained constant over time, with a median of 0.62 (SD 0.02).

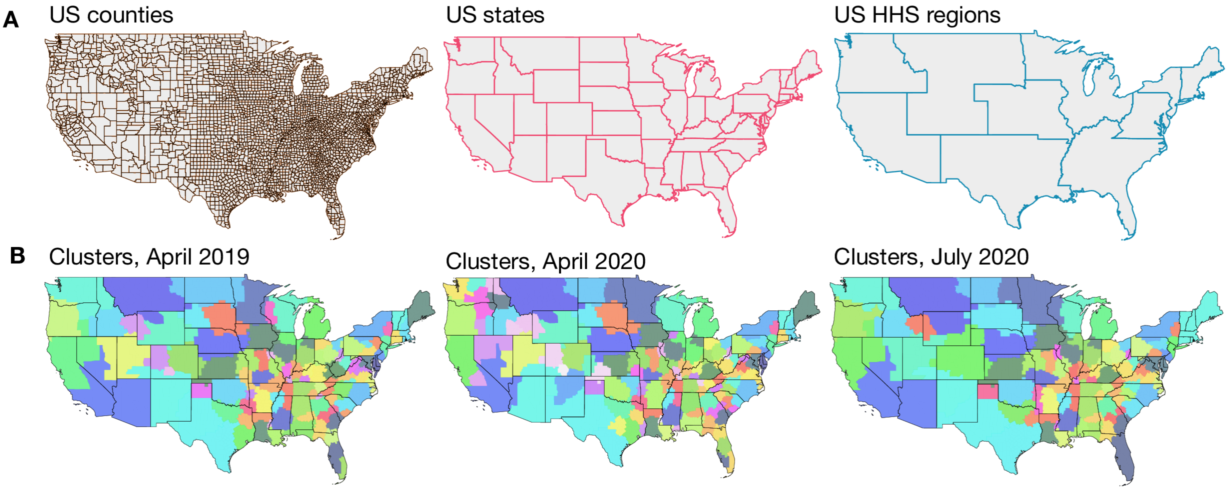
Figure 2. Temporal stability in intercounty mobility. (A) The monthly degree distribution of the intercounty network is shown as a violin plot with 95% CIs. It demonstrates significant variability while remaining consistent across months. (B) The persistence probability of links is illustrated, denoting the likelihood that a connection existing in 2019 remains present in 2020 and 2021. The plot provides a breakdown for different link types: urban-urban (UU), urban-rural (UR), rural-urban (RU), and rural-rural (RR) links. (C) Distribution of coupling probabilities in the connectivity network by month. We highlighted in dark red: April, 2020 (LD) that represents the peak time of number of US states in lockdown, and November, 2020 (SD) that represents the period when social distancing recommendations were in place. LD: lockdown; RR: rural-rural; RU: rural-urban; SD: social distancing; UR: urban-rural; UU: urban-urban.



Spatial Stability of the Intercounty Connectivity Network

To identify the geographic scale at which mobility is highly connected, we detected clusters of counties that were more connected via mobility within the cluster than outside the clusters, we used a network community detection algorithm. Our hypothesis was that this partitioning of the United States would be at a geographic scale larger than 3143 US counties but smaller than 50 US states or 10 Health and Human Services (HHS) regions. Indeed, we found that based on human mobility, the United States can be partitioned into around 100 regions that split most US states into multiple clusters (Figure 3). We also found that these clusters were highly and spatially contiguous and respected state boundaries (with a similarity measured by normalized mutual information as 0.82). Furthermore, these regions demonstrated stability over time (normalized mutual information=0.95) despite the perturbations of the early phases of the COVID-19 pandemic (Figure 3B and in Figures S9-S10 in Multimedia Appendix 1). Thus, we identified a persistent geographic partitioning of the United States in which clusters were more connected within than between, and hypothesized that the relevance of mobility to the spatial diffusion of infectious diseases occured at a mesoscale.

Figure 3. Analysis of spatial stability. (A) Geographical subdivisions at the county and state levels within the United States, as well as the division into Health and Human Services regions, used for health administration purposes. (B) We partition the intercounty mobility network so that each cluster of counties is more strongly connected to each other via mobility than to counties in other clusters. We find highly consistent portioning based on mobility networks from April 2019, April 2020 (during the mandated lockdown period of the early COVID-19 pandemic), and July 2020 (after the lockdown period of the early COVID-19 pandemic). Clusters are colored to delineate cluster boundaries and do not represent any other information. Counties colored in gray have populations of fewer than 11,000 inhabitants and are excluded from the analysis. HHS: Health and Human Services.

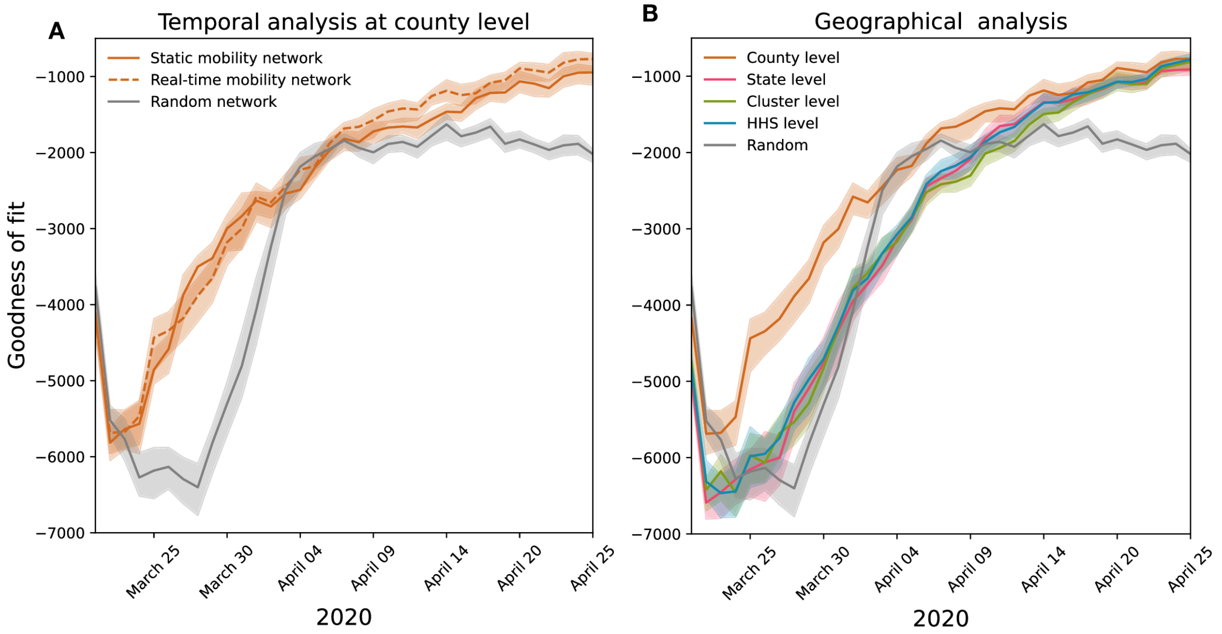


Implications for Metapopulation Disease Models

After examining the stability of mobility patterns over time and space, we evaluated how spatiotemporal scales of mobility affect disease modeling in a metapopulation framework by integrating connectivity networks into a model simulating the initial SARS-CoV-2 spread in the United States. We investigated the influence of geographic scale by homogenizing networks at different spatial levels, maintaining the county resolution. In addition, to gauge the impact of temporal scale, we provided the model with either a time-evolving or static connectivity network representing mobility patterns since March 2020, considering the network's temporal stability. Goodness of fit was assessed by comparing predicted and observed disease arrival times in counties.

Figure 4A shows that a county-level metapopulation model using empirical connectivity networks accurately predicts early COVID-19 diffusion, outperforming a Erdős–Rényi random network. Indeed, the empirical mobility network had a stronger goodness of fit throughout the early phase of the pandemic. This emphasized the crucial role of human mobility in the spatial spread of the initial SARS-CoV-2 wave and underscored the importance of mobility data for modeling SARS-CoV-2 spread. In addition, predictions based on static and time-varying mobility networks were comparable, suggesting that static data were sufficient to accurately capture epidemic spatial heterogeneity. Following the initial invasion phase in March, transmission became more widespread, reducing the importance of spatial connections. At this stage, random networks and county-level networks yielded similar performance, indicating that local epidemic growth outweighed the influence of importations and exportations. Figure 4B shows that county-level mobility data predicted spatial diffusion better than data at larger scales, such as US HHS regions, states, and clusters (as defined in Figure 3). All larger-scale data performed similarly to a random network, indicating they lack the granularity needed to capture diverse mobility patterns effectively. Sensitivity analyses on the definition of the time of arrivals and its impact on the goodness of fit are reported in Multimedia Appendix 1, more precisely in Figure S14B.

Figure 4. Implications of temporal and spatial scale of mobility data for the prediction with metapopulation disease models. (A) The goodness of fit (median and 95% CI) for the time of arrivals for metapopulation models at a county level, informed by a time-evolving intercounty connectivity network, a static intercounty connectivity network, and a random intercounty network. (B) Solid lines show the goodness of the fit (median and 95% CI) for metapopulation models informed with state-, cluster-, and Health and Human Services–level static mobility network. HHS: Health and Human Services.



Discussion

Principal Results

Our findings revealed significant insights into the dynamics of human mobility and their implications for infectious disease modeling. In contrast to findings from other countries (eg, France [29], India [30], Germany [31], and Spain [32]), we observed that despite the implementation of local social distancing measures and lockdowns, intercounty connectivity remained largely unperturbed, leading to rapid geographic diffusion of SARS-CoV-2. Mobility patterns experienced only marginal changes before and after the early-stage COVID-19 pandemic. The most notable disruption occurred during the first lockdown period in April 2020, when mobility sharply declined. Although a temporary reduction in mobility was observed, it proved to be short-lived as mobility patterns quickly returned to prerestriction levels. Importantly, this decline in mobility did not alter the underlying intercounty connectivity structure, potentially diminishing the overall effectiveness of the implemented travel restrictions. Even when mobility is limited—resulting in low-strength edges between locations—these connections can still facilitate the spread of disease, introducing pathogens into new areas [33]. This underscores the challenge of relying solely on travel restrictions as a mitigation strategy, as they may fail to disrupt the pathways that enable disease transmission, particularly when the connectivity network remains intact. Such insights emphasize the need for more nuanced approaches that address the persistence of network structures during public health interventions. Notably, even during periods of social distancing recommendations, the mobility network remains relatively stable. Assuming the lockdown represents the most extreme form of mobility disruption, the temporal stability findings suggest that global human mobility demonstrates resilience against short-term changes.

We also assessed the spatial stability of the intercounty connectivity network by detecting spatial communities based on mobility patterns. Our results indicate that mobility-driven clusters align closely with state boundaries, reflecting the influence of administrative and geographical factors on human movement, accordingly with Steinegger et al [31]. These clusters exhibited remarkable stability over time, reinforcing the idea that spatial mobility patterns are deeply ingrained and relatively resistant to abrupt changes. The fact that mobility patterns are highly correlated with state boundaries suggests that state-level structures could be effective for designing target public health interventions based on travel reductions. Our findings underscore the importance of considering mobility patterns when designing interventions, resource allocation, and disease control strategies.

As shown in the context of COVID-19 pandemic in France [34], we also demonstrated that incorporating high-resolution human mobility data are crucial for accurately capturing the spatial spread of infectious diseases. Our findings indicate that county-level, daily mobility data offer the most accurate representation of the spatial spread of disease in the United States. Notably, static county-level mobility data achieve similar model performance to real-time data, suggesting that an undisturbed representation of reality is adequate for reproducing spatial spread. More interestingly, our exploration of various spatial scales for metapopulation models underscores the significance of aligning the model's structure with the inherent spatial scale of human mobility. While county-level mobility data yield the most accurate depiction, mobility databased clusters, US states, and HHS regions do not capture the heterogeneity of the COVID-19 geographical diffusion.

Limitations

While our study provides valuable insights, it is not without limitations. Our work focused on the early phase of the pandemic, during which response measures (eg, social distancing, closures, and lack of masking) were largely homogeneous in the United States, and pharmaceutical measures (eg, vaccination and antivirals) were not available; thus, these findings are not generalizable to later stages. Furthermore, we assume homogeneity within the US counties. In addition, Safegraph mobility data, like all mobile app–based location data, exhibit sampling biases. Ongoing efforts to comprehend these biases are crucial for developing better correction methods [35]. An independent analysis by Safegraph revealed the underrepresentation of older and non-White individuals in point of interest-specific analyses, though the panel is representative of race, education, and income [35,36].

Comparison With Previous Work

Since the onset of the COVID-19 pandemic, mobile phone data have played a crucial role in addressing the public health crisis [10,11,13-15,22,29,30,34,37]. During this period, numerous network operators and private enterprises have made considerable efforts to swiftly share their data within the confines of legal regulations. Consequently, researchers worldwide have embarked on working with this data, monitoring human behavior caused by containment measures and adaptive responses to the epidemic, and using it to enhance epidemic models in order to increase their reliability [14,15,22,29,34].

While static mobility data have predominantly been analyzed and integrated into models over the past decades [3,4], the current accessibility to real-time human behavior data prompts an essential investigation into the optimal scenarios for using this dynamic information versus relying solely on static representations of reality [38]. Equally important is the exploration of the characteristic mobility scale to comprehensively capture the intricate coupling between different locations, a consideration with potential implications for target control policies to reduce epidemic activity, and for improving epidemic model forecasting. Furthermore, numerous researchers have emphasized the pressing necessity to implement standardized strategies that facilitate rapid data access while upholding stringent data privacy measures [2]. To address this gap in the literature, in this study, we investigated the spatial connectivity of the US counties during the early phase of the COVID-19 pandemic using high-resolution real-time human mobility data obtained from mobile phone usage.

Conclusions

While characterizing the key role of mobility in the spatial invasion of the COVID-19 pandemic in the United States, our study sheds light on the global stability of human mobility patterns, and the relevant information needed to design a reliable predictive model. This result may be specific to countries, such as the United States, in which mobility restrictions were not stringent, specified for intercounty mobility, nor enforced. Metapopulation models that incorporate accurate mobility data can provide valuable insights into disease dynamics and enhance our ability to predict and control the spread of future infectious disease outbreaks. Furthermore, standardized data extraction and sharing that we introduced might help facilitate the timelines associated with legal agreements for data sharing, which do not always align with the rapid spread of epidemics, thus diminishing the feasibility of timely responses to such outbreaks.

Acknowledgments

The authors thank the teams at Safegraph (now Advan Patterns) for sharing mobility data. The research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under awards R01GM123007 and R35GM153478.

Authors' Contributions

GP and SB conceived of and designed the study. GP and LGA-Z analyzed the data and did the analysis. GP, VC, and SB interpreted the results. GP drafted the manuscript. All authors contributed to the writing of the final version of the manuscript.

Conflicts of Interest

SB has received funding from the Merck Pharmaceuticals Investigator Studies Program.

Multimedia Appendix 1
Supplementary material.
[DOCX File, 8862 KB - publichealth_v11i1e64914_app1.docx]

References

1. Findlater A, Bogoch II. Human mobility and the global spread of infectious diseases: a focus on air travel. Trends Parasitol 2018 Sep;34(9):772-783. [doi: 10.1016/j.pt.2018.07.004] [Medline: 30049602]
2. de Montjoye YA, Gambs S, Blondel V, et al. On the privacy-conscientious use of mobile phone data. Sci Data 2018 Dec 11;5(1):180286. [doi: 10.1038/sdata.2018.286] [Medline: 30532052]

3. Colizza V, Barrat A, Barthélemy M, Vespignani A. Predictability and epidemic pathways in global outbreaks of infectious diseases: the SARS case study. BMC Med 2007 Nov 21;5(1):34-7015. [doi: 10.1186/1741-7015-5-34] [Medline: 18031574]
4. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. Proc Natl Acad Sci U S A 2009 Dec 22;106(51):21484-21489. [doi: 10.1073/pnas.0906910106] [Medline: 20018697]
5. Barbosa H, Barthelemy M, Ghoshal G, et al. Human mobility: models and applications. Phys Rep 2018 Mar;734:1-74. [doi: 10.1016/j.physrep.2018.01.001]
6. Viboud C, Bjørnstad ON, Smith DL, Simonsen L, Miller MA, Grenfell BT. Synchrony, waves, and spatial hierarchies in the spread of influenza. Science 2006 Apr 21;312(5772):447-451. [doi: 10.1126/science.1125237] [Medline: 16574822]
7. Blondel VD, Decuyper A, Krings G. A survey of results on mobile phone datasets analysis. EPJ Data Sci 2015 Dec;4(1):10-1127. [doi: 10.1140/epjds/s13688-015-0046-0]
8. Maxmen A. Can tracking people through phone-call data improve lives? Nature New Biol 2019 May 30;569(7758):614-617. [doi: 10.1038/d41586-019-01679-5]
9. Pappalardo L, Manley E, Sekara V, Alessandretti L. Future directions in human mobility science. Nat Comput Sci 2023 Jul;3(7):588-600. [doi: 10.1038/s43588-023-00469-4] [Medline: 38177737]
10. Oliver N, Lepri B, Sterly H, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. Sci Adv 2020 Jun;6(23):eabc0764. [doi: 10.1126/sciadv.abc0764] [Medline: 32548274]
11. Grantz KH, Meredith HR, Cummings DAT, et al. The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. Nat Commun 2020;11(1):4961-1723. [doi: 10.1038/s41467-020-18190-5]
12. Garnier R, Benetka JR, Kraemer J, Bansal S. Socioeconomic disparities in social distancing during the COVID-19 pandemic in the United States: observational study. J Med Internet Res 2021 Jan 22;23(1):e24591. [doi: 10.2196/24591] [Medline: 33351774]
13. Susswein Z, Rest EC, Bansal S. Disentangling the rhythms of human activity in the built environment for airborne transmission risk: An analysis of large-scale mobility data. elife 2023 Apr 4;12:e80466. [doi: 10.7554/eLife.80466] [Medline: 37014055]
14. Pei S, Kandula S, Shaman J. Differential effects of intervention timing on COVID-19 spread in the United States. Sci Adv 2020 Dec 4;6(49):2375-2548. [doi: 10.1126/sciadv.abd6370]
15. Chang S, Pierson E, Koh PW, et al. Mobility network models of COVID-19 explain inequities and inform reopening. Nature New Biol 2021 Jan;589(7840):82-87. [doi: 10.1038/s41586-020-2923-3] [Medline: 33171481]
16. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. Lancet Infect Dis 2020 Nov;20(11):1247-1254. [doi: 10.1016/S1473-3099(20)30553-3] [Medline: 32621869]
17. Xiong C, Hu S, Yang M, Luo W, Zhang L. Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. Proc Natl Acad Sci USA 2020 Nov 3;117(44):27087-27089. [doi: 10.1073/pnas.2010836117]
18. Tokey AI. Spatial association of mobility and COVID-19 infection rate in the USA: a county-level study using mobile phone location data. J Transp Health 2021 Sep;22:101135. [doi: 10.1016/j.jth.2021.101135] [Medline: 34277349]
19. Zheng Z, Xie Z, Qin Y, Wang K, Yu Y, Fu P. Exploring the influence of human mobility factors and spread prediction on early COVID-19 in the USA. BMC Public Health 2021 Dec;21(1):1-13. [doi: 10.1186/s12889-021-10682-3]
20. Gao S, Rao J, Kang Y, et al. Association of mobile phone location data indications of travel and stay-at-home mandates with COVID-19 infection rates in the US. JAMA Netw Open 2020 Sep 1;3(9):e2020485. [doi: 10.1001/jamanetworkopen.2020.20485] [Medline: 32897373]
21. Pullano G. Human mobility and epidemics. PhD thesis.: Sorbonne Université; 2021 URL: https://theses.hal.science/tel-04099742
22. Zhang M, Wang S, Hu T, et al. Human mobility and COVID-19 transmission: a systematic review and future directions. Ann GIS 2022 Oct 2;28(4):501-514. [doi: 10.1080/19475683.2022.2041725]
23. Safegraph. URL: https://www.safegraph.com [accessed 2025-02-10]
24. CDC. COVID Data Tracker. 2020. URL: https://covid.cdc.gov/covid-data-tracker [accessed 2025-02-10]
25. Russell TW, Golding N, Hellewell J, et al. Reconstructing the early global dynamics of under-ascertained COVID-19 cases and infections. BMC Med 2020 Oct 22;18(1):332. [doi: 10.1186/s12916-020-01790-9] [Medline: 33087179]
26. Rosvall M, Axelsson D, Bergstrom CT. The map equation. Eur Phys J Spec Top 2009 Nov;178(1):13-23. [doi: 10.1140/epjst/e2010-01179-1]
27. NetworkX Python Library. URL: https://networkx.org/ [accessed 2025-02-10]
28. Scikit-Mobility Python Library. URL: https://scikit-mobility.github.io/scikit-mobility/ [accessed 2025-02-10]
29. Pullano G, Valdano E, Scarpa N, Rubrichi S, Colizza V. Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: a population-based study. Lancet Digit Health 2020 Dec;2(12):e638-e649. [doi: 10.1016/S2589-7500(20)30243-0] [Medline: 33163951]

30. Saha J, Barman B, Chouhan P. Lockdown for COVID-19 and its impact on community mobility in India: an analysis of the COVID-19 Community Mobility Reports, 2020. Child Youth Serv Rev 2020 Sep;116:105160. [doi: 10.1016/j.childyouth.2020.105160] [Medline: 32834269]
31. Sobolevsky S, Szell M, Campari R, Couronné T, Smoreda Z, Ratti C. Delineating geographical regions with networks of human interactions in an extensive set of countries. PLoS One 2013;8(12):e81707. [doi: 10.1371/journal.pone.0081707] [Medline: 24367490]
32. Steinegger B, Granell C, Rapisardi G, et al. Joint analysis of the epidemic evolution and human mobility during the first wave of COVID-19 in Spain: retrospective study. JMIR Public Health Surveill 2023 May 22;9:e40514. [doi: 10.2196/40514] [Medline: 37213190]
33. Colizza V, Vespignani A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: theory and simulations. J Theor Biol 2008 Apr 7;251(3):450-467. [doi: 10.1016/j.jtbi.2007.11.028] [Medline: 18222487]
34. Pullano G, Di Domenico L, Sabbatini CE, et al. Underdetection of cases of COVID-19 in France threatens epidemic control. Nature New Biol 2021 Feb 4;590(7844):134-139. [doi: 10.1038/s41586-020-03095-6]
35. Coston A, Guha N, Ouyang D, Lu L, Chouldechova A, Ho DE. Leveraging administrative data for bias audits: assessing disparate coverage with mobility data for COVID-19 policy. In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency FAccT ’21: Association for Computing Machinery; 2021:173-184. [doi: 10.1145/3442188.3445881]
36. Fox R. “What about bias in your dataset?”: Quantifying sampling bias in SafeGraph Patterns. URL: https://medium.com/safegraph/what-about-bias-in-the-safegraph-dataset-b95b7208893c [accessed 2025-02-10]
37. Xia J, Yin K, Yue Y, et al. Impact of human mobility on COVID-19 transmission according to mobility distance, location, and demographic factors in the Greater Bay Area of China: population-based study. JMIR Public Health Surveill 2023 Apr 26;9:e39588. [doi: 10.2196/39588] [Medline: 36848228]
38. Baker RE, Mahmud AS, Miller IF, et al. Infectious disease in an era of global change. Nat Rev Microbiol 2022 Apr;20(4):193-205. [doi: 10.1038/s41579-021-00639-z] [Medline: 34646006]

Abbreviations

CDC: Centers for Disease Control and Prevention
HHS: Health and Human Services

Edited by A Mavragani; submitted 30.07.24; peer-reviewed by C Argyropoulos, M Mazzoli; revised version received 09.12.24; accepted 25.12.24; published 18.02.25.

Please cite as:
Pullano G, Alvarez-Zuzek LG, Colizza V, Bansal S
Characterizing US Spatial Connectivity and Implications for Geographical Disease Dynamics and Metapopulation Modeling: Longitudinal Observational Study
JMIR Public Health Surveill 2025;11:e64914
URL: https://publichealth.jmir.org/2025/1/e64914
doi:10.2196/64914

Original Paper

An Explainable Artificial Intelligence Text Classifier for Suicidality Prediction in Youth Crisis Text Line Users: Development and Validation Study

Julia Thomas[1,2,3], MA; Antonia Lucht[3], MSci; Jacob Segler[4], MA; Richard Wundrack[3], PhD; Marcel Miché[1], PhD; Roselind Lieb[1], Prof Dr; Lars Kuchinke[5], Prof Dr; Gunther Meinlschmidt[2,6,7,8], Prof Dr

[1]Division of Clinical Psychology and Epidemiology, Faculty of Psychology, University of Basel, Basel, Switzerland

[2]Division of Clinical Psychology and Cognitive Behavioural Therapy, International Psychoanalytic University Berlin, Berlin, Germany

[3]Department of Research, Analytics and Development, krisenchat gGmbH, Berlin, Germany

[4]Division of Child and Adolescent Psychiatry/Psychotherapy, Universitätsklinikum Ulm, Ulm, Germany

[5]Division of Methods and Statistics, International Psychoanalytic University Berlin, Berlin, Germany

[6]Clinical Psychology and Psychotherapy, Methods and Approaches, Department of Psychology, Trier University, Trier, Germany

[7]Department of Digital and Blended Psychosomatics and Psychotherapy, Psychosomatic Medicine, University Hospital and University of Basel, Basel, Switzerland

[8]Department of Psychosomatic Medicine, University Hospital Basel, University of Basel, Basel, Switzerland

Corresponding Author:
Julia Thomas, MA
Division of Clinical Psychology and Epidemiology
Faculty of Psychology
University of Basel
Missionsstrasse 60/62
Basel, 4055
Switzerland
Phone: 49 30 57714627
Email: julia.thomas@krisenchat.de

Abstract

Background: Suicide represents a critical public health concern, and machine learning (ML) models offer the potential for identifying at-risk individuals. Recent studies using benchmark datasets and real-world social media data have demonstrated the capability of pretrained large language models in predicting suicidal ideation and behaviors (SIB) in speech and text.

Objective: This study aimed to (1) develop and implement ML methods for predicting SIBs in a real-world crisis helpline dataset, using transformer-based pretrained models as a foundation; (2) evaluate, cross-validate, and benchmark the model against traditional text classification approaches; and (3) train an explainable model to highlight relevant risk-associated features.

Methods: We analyzed chat protocols from adolescents and young adults (aged 14-25 years) seeking assistance from a German crisis helpline. An ML model was developed using a transformer-based language model architecture with pretrained weights and long short-term memory layers. The model predicted suicidal ideation (SI) and advanced suicidal engagement (ASE), as indicated by composite Columbia-Suicide Severity Rating Scale scores. We compared model performance against a classical word-vector-based ML model. We subsequently computed discrimination, calibration, clinical utility, and explainability information using a Shapley Additive Explanations value-based post hoc estimation model.

Results: The dataset comprised 1348 help-seeking encounters (1011 for training and 337 for testing). The transformer-based classifier achieved a macroaveraged area under the curve (AUC) receiver operating characteristic (ROC) of 0.89 (95% CI 0.81-0.91) and an overall accuracy of 0.79 (95% CI 0.73-0.99). This performance surpassed the word-vector-based baseline model (AUC-ROC=0.77, 95% CI 0.64-0.90; accuracy=0.61, 95% CI 0.61-0.80). The transformer model demonstrated excellent prediction for nonsuicidal sessions (AUC-ROC=0.96, 95% CI 0.96-0.99) and good prediction for SI and ASE, with AUC-ROCs of 0.85 (95% CI 0.97-0.86) and 0.87 (95% CI 0.81-0.88), respectively. The Brier Skill Score indicated a 44% improvement in classification performance over the baseline model. The Shapley Additive Explanations model identified language features predictive of SIBs, including self-reference, negation, expressions of low self-esteem, and absolutist language.

Conclusions: Neural networks using large language model–based transfer learning can accurately identify SI and ASE. The post hoc explainer model revealed language features associated with SI and ASE. Such models may potentially support clinical decision-making in suicide prevention services. Future research should explore multimodal input features and temporal aspects of suicide risk.

(JMIR Public Health Surveill 2025;11:e63809) doi:10.2196/63809

KEYWORDS

deep learning; explainable artificial intelligence (XAI); large language model (LLM); machine learning; neural network; prevention; risk monitoring; suicide; transformer model; suicidality; suicidal ideation; self-murder; self-harm; youth; adolescent; adolescents; public health; language model; language models; chat protocols; crisis helpline; help-seeking behaviors; German; Shapley; decision-making; mental health; health informatics; mobile phone

Introduction

Suicide, the third leading cause of premature mortality among German adolescents [1], encompasses a spectrum from suicidal ideation (SI) to preparatory actions [2-5]. Adolescents' vulnerability to suicidal phenomena stems from a complex interplay of biological, genetic, psychological, and social factors [6-9]. Alarmingly, less than 50% of adolescents who attempt suicide receive appropriate psychiatric intervention [10]. This gap between need and care emphasizes the critical importance of accurately identifying and timely identification of at-risk individuals, challenging mental health care providers and educational institutions.

The digital age, characterized by the frequent use of the internet and smartphones, has transformed youth help-seeking behaviors, with online text-based services becoming the preferred communication mode [11-13]. These platforms align with young people's digital environments [12,14], reducing help-seeking barriers such as limited service availability and stigma [15]. Their immediacy and anonymity are crucial for crisis help-seeking, establishing text-based helplines as critical public health measures for suicide-related issues [15,16]. While presenting challenges for clinicians in rapidly assessing risks and implementing interventions [17], these digitized services also offer opportunities. They enable traditional clinical risk monitoring models [18] to be complemented or potentially superseded by machine learning (ML) approaches.

ML has shown promise in identifying risk factors associated with suicide risk [19-21] and suicidal outcomes, including SI [22], behaviors [21,23], attempts [24], and completed suicides [25,26]. ML models use various input data types, from electronic health records [14,24,27] to textual data [19,28], collected in diverse settings [12,19,24,28,29]. These "theory-free" approaches [30] often outperform traditional detection methods [31].

Natural language processing advancements, particularly pretrained large language models based on transformer architectures, have significantly enhanced language classification tasks [32,33]. These "foundational models" [34], such as BERT, RoBERTa, LLaMA, and GPT, provide probabilistic language representations applicable to diverse tasks, opening new avenues for ML-based prediction in language-based psychological domains [34].

Language models provide rich text embeddings for transfer learning and fine-tuning [35,36], potentially enabling more accurate detection of SI and behaviors, even with limited domain-specific data [37]. Recent research has highlighted ML techniques, particularly foundational language models, in comprehensive suicidality prediction [38,39] and extraction of relevant textual indicators [40,41]. Transformer-based models outperform traditional ML approaches in suicide prediction using textual data, especially in social media datasets [38,42-45], though their real-world clinical applicability has remained unexplored.

The clinical implementation of ML models requires rigorous validation [46]. The "black box" nature of ML decisions necessitates explainable AI to emphasize interpretability for clinicians [30,31]. While traditional rule-based ML approaches such as decision trees offer straightforward explainability, the complexity of deep neural networks in foundational models presents interpretability challenges [47]. Model-agnostic post hoc explanation approaches, such as Shapley Additive Explanations (SHAPs), address this complexity [48-50], potentially elucidating key language features associated with suicide risk.

Deep neural network models have been applied to social media data [29,51-56], and explainable AI studies have been conducted on medical tabular data [57] and electronic health records [58]. However, studies using textual data from clinical populations, particularly helpline data, have not adequately addressed explainability.

Clinical implementations of ML as a diagnostic tool require rigor and transparency beyond interpretability. Ensuring model stability, tuning, and clinical value is crucial when considering the cost-benefit ratio of diagnostic decisions [59,60]. Metrics such as calibration and clinical utility [61-64] can emphasize the reliability and practical value of ML-based decisions in suicide risk prediction, though they remain underused in this context.

Our study aims to bridge the gap between state-of-the-art natural language processing techniques and real-world clinical applications in suicide prevention, addressing the lack of explainable transformer predictions for suicide risk assessment using clinical textual inputs validated for clinical utility. We develop and test a model for real-time crises within crisis text lines, integrating transformer-based models with SHAPs and incorporating clinical utility metrics.

We hypothesize that our transformer-based model will outperform the traditional baseline model in predicting SI and advanced suicidal engagement (ASE). Additionally, the SHAPs application is expected to highlight interpretable language features influencing the classification process, providing clinically relevant insights for decision-making in suicide prevention. This research aims to advance both the technical capabilities and practical applicability of ML in mental health assessment, potentially enhancing early intervention strategies in suicide prevention.

Methods

Overview

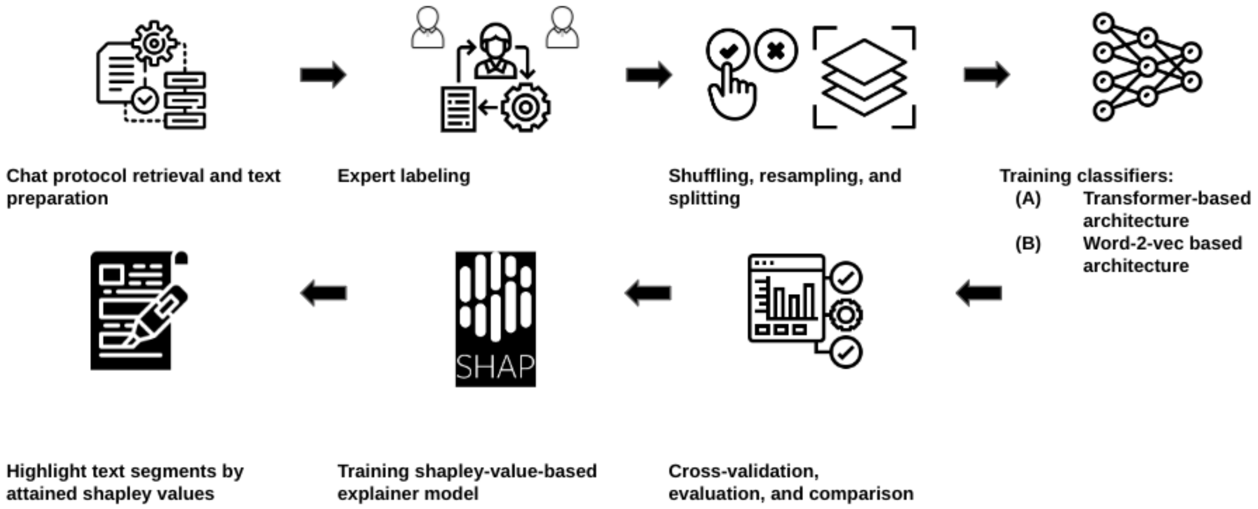
This study used a comprehensive approach to develop and evaluate ML models for identifying SI and behaviors in crisis helpline conversations. Our methodology encompassed several vital components, including data acquisition and preparation, model development and training, model evaluation and comparison, and explainability analysis. The statistical analysis comprised three main stages: preliminary sample analysis, development and training of two neural networks, and comprehensive evaluation of model performance. The following sections detail each component of our methodology, including data collection, preprocessing, model architecture, training procedures, and evaluation metrics.

Study Design

This study consists of several steps (Figure 1 [65-71]). We obtained chat protocols from krisenchat, a German crisis text line for youth. The dataset underwent cleaning, preprocessing, and expert labeling.

Post labeling, the data were randomized, partitioned, resampled, and encoded for model training. We then trained two text classification models and assessed model performance using various metrics. Lastly, we used a SHAPs explainer model to interpret the classification results.

Figure 1. Study design for suicidality prediction model development and evaluation using German youth crisis helpline session transcripts collected between November 11, 2021, and April 30, 2022: data flow, model development, and evaluation process for image sources.



Data Collection and Preprocessing

We obtained chat transcripts from krisenchat, a prominent German preclinical crisis intervention service for individuals up to 25 years of age [72]. The dataset comprised sessions conducted between November 30, 2021, and April 30, 2022. We excluded sessions from individuals younger than 14 years of age and those lacking age information due to informed consent considerations. The final dataset included transcripts from individuals experiencing suicidal thoughts and behaviors, as well as those seeking advice on other topics.

We defined a chat session as all messages from a single individual during a counseling session, typically at most 2 hours. We removed all personal identifiers and preprocessed the text data to enhance readability and reduce variability. This process eliminated extraneous elements such as links, HTML tags, and special characters beyond typical German punctuation.

From sessions marked with SI by krisenchat counselors, we identified relevant conversations addressing SI within this study's period. We excluded sessions too brief to assess suicidal ideation and behaviors (SIBs) or those terminating prematurely. Of 14,073 sessions, we identified 3193 control sessions, with 2886 meeting length and age criteria. We randomly selected 500 nonsuicidal sessions for manual review, replacing any indicating suicidal thoughts or severe self-harm with nonsuicidal cases. We used the imbalanced-learn library to address the class imbalance in our dataset by applying random oversampling of minority classes in the test set after splitting.

Measures and Labeling

We used the Columbia-Suicide Severity Rating Scale (C-SSRS), the gold standard for assessing SIB [73,74]. The C-SSRS categorizes ideation levels from a general desire to die to actual suicide attempts, quantifying severity and intensity. It defines suicidal behavior as any life-ending action, including nonharmful behaviors, interrupted attempts, and preparatory actions.

To address data imbalance and align with recent suicidology theories [75,76], we applied composite categories of SI and

ASE based on C-SSRS levels. This approach reflects the distinct processes of SI development and progression to suicidal behaviors [77]. The probability of acting on suicidal thoughts is influenced by an individual's capacity to endure pain and access to lethal means [78].

We defined SI using C-SSRS items "wish to be dead" and "nonspecific active suicidal thoughts," encompassing thoughts from the desire not to wake up to explicit life-ending statements. ASE comprised C-SSRS items involving specific methods, intent, or attempts, including contemplation of lethal methods, setting definite dates for action, and engaging in preparatory behaviors. The consideration of lethal means marks the transition to suicidal behaviors, defining the ASE category.

Three independent expert raters, extensively trained in C-SSRS, labeled the sessions. Their training involved mutual panel ratings and group discussions across 50 trial conversations not included in this study. Each rater subsequently assigned ratings independently. Raters conducted integrity discussions between rating sessions to maintain consistency and mitigate observer drift, exchanging thoughts and rating standards. Upon completing all session ratings, the raters convened to address discrepancies and reach a consensus on final ratings through comprehensive discussions.

Reporting Standards

This study adheres to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis for artificial intelligence (AI) abstract and reporting guidelines checklist for prediction model development and validation [79] (Tables S1 and S2 in Multimedia Appendix 1) and the CONSORT-AI (Consolidated Standards of Reporting Trials for Artificial Intelligence) checklist [80] (Table S3 in Multimedia Appendix 1).

Ethical Considerations

This study was conducted using ethical standards for human participant research. This study's protocol received approval from the Ethics Committee of the International Psychoanalytic University Berlin (2023_08). Informed consent was obtained through krisenchat's terms of service, which explicitly state that user data may be used for research purposes without direct identification of individuals. All personally identifiable information was removed from the chat transcripts during data preprocessing to protect privacy and confidentiality. Participants were not compensated directly, as this study used existing data from the crisis helpline. No images of participants were used in this study; the source text was solely used for training the classifier and was cleaned of all personally identifiable information. These measures ensure that no identification of individual participants is possible.

Software

All preprocessing, modeling, and evaluation tasks were performed using TensorFlow's [81] official Docker Image to construct a containerized GPU runtime on a Linux machine, using Python (version 3.8.1; Python Software Foundation) and CUDA (version 11.4; NVIDIA Corp). The machine had a 4 GB GPU and 16 GB of RAM. We used NumPy [82] and pandas [83] for preprocessing, while the Imblearn library was used for resampling. Model architectures and weights were sourced from the Hugging Face Transformers [35] library, and we used TensorFlow's Keras [84] module for modeling. Visualization and metrics were facilitated through matplotlib [85], seaborn [86], sci-kit-learn [87], and sklearn [88]. The final script is openly accessible on GitHub [89].

Model Architecture

Overview

One-way ANOVA to evaluate differences in word usage and age across 3 groups: nonsuicidal, SI, and ASE.

We developed two neural network models to classify SI and ASE in crisis interactions. Both models share a similar architecture, comprising three main blocks: an embedding block for the mathematical representation of messages, a long short-term memory (LSTM) block to model temporal dependencies, and a classification block.

Model 1, a transformer-multilayer perceptron (T-MLP), uses a pretrained transformer encoder (XLM-RoBERTa-base) [90], available on Hugging Face [35]. This multilingual model, trained on 2.5 TB of CommonCrawl data in 100 languages, tokenizes and encodes input text into a 768-dimensional embedding. Each message is embedded separately and then attached to an array of embedded messages. The embeddings feed into a time-distributed dense layer, followed by an LSTM encoder with dropout regularization, and finally, a multilayer perceptron for classification.

Model 2, a word2vector-multilayer perceptron (W2V-MLP), uses pretrained German word-vector embeddings (300-dimensional) processed through convolutional filters. These word-wise embeddings, trained using FastText on standard crawl and Wikipedia datasets, form an embedding matrix. This matrix is processed through a convolutional layer, max pooling, dropout, an LSTM encoder, and a multilayer perceptron layer for classification.

Both models output probabilities for three categories: not suicidal, SI, and ASE. Detailed model parameters are provided in Table S1 in Multimedia Appendix 2.

Model Training, Cross-Validation, and Evaluation

We implemented a three-phase evaluation and cross-validation protocol: (1) cross-validation of training, (2) post hoc estimation of overall evaluation metrics, and (3) assessment of calibration and clinical utility metrics.

To evaluate model stability, we performed 5 iterations of repeated cross-validation. We initialized data at random seeds, then shuffled and partitioned it (1011 for training and 337 for testing), maintaining consistent class ratios through stratified shuffling. We applied random oversampling to both sets individually to ensure balanced class ratios.

We limited training to a maximum of 100 epochs, incorporating an early stopping mechanism triggered by test accuracy, with a patience setting of 5 epochs. We selected the best model based on its area under the curve (AUC) performance, derived from receiver operating characteristics (ROCs).

We assessed model performance using standard class-wise metrics, including AUC, precision, recall, and F_1-scores, computing these metrics macroaveraged across classes. We determined class-wise metrics using a one-versus-all approach, binarizing predictions for each class. AUC measures the model's discriminative ability, with scores of 0.7-0.8 considered acceptable and 0.9 indicating high discrimination. We computed ROC curves using the one-versus-all method, quantifying the model's classification ability across all thresholds [80]. We applied bootstrapping with 1000 subsamples from the test dataset to establish 95% CIs [81]. We provided confusion matrices to visualize misclassifications [82].

To demonstrate clinical utility, we conducted decision curve analysis (DCA) for the SI and ASE classes [61,62]. DCA calculates a "net benefit" (NB) value based on the harm-to-benefit ratio of clinical decisions. An NB of 0.2 indicates the model detects 20 true positives without increasing false positives. We set threshold probabilities for SI between 0.3 and 0.5 and for ASE between 0.15 and 0.30 based on our sample's event rates and clinical considerations [63].

We assessed model calibration by calculating the overall Brier score (BS), which measures the mean square error of prediction probabilities against actual outcomes. Lower BS values indicate better calibration accuracy. We decomposed the BS into reliability, resolution, and uncertainty to offer insights beyond AUC-ROC [61]. We also calculated the Brier skill score to compare performance improvements of our transformer model over the word2vec model in calibration.

We visually assessed calibration using plots depicting the relationship between predicted probabilities and observed outcome frequencies [86]. A perfectly calibrated model aligns with the plot's diagonal; deviations indicate under- or overprediction.

Detailed DCA calculations, rationale, interpretations, and BS decomposition analysis are available in Table S3 in Multimedia Appendix 2.

Explainability Analysis

We used a SHAPs model-agnostic post hoc explainer to enhance our understanding of the models' prediction processes. Shapley values, derived from cooperative game theory [91], provide local explanations for each prediction [92]. These values quantify each feature's impact on an instance's prediction relative to its average impact across other feature combinations. The algorithm uses tokenized text vectors as inputs and considers all possible feature combinations to calculate these contributions.

In addition to examining individual results, we determined the overall importance of specific language features in predicting nonsuicidal, SI, or ASE categories. This approach provides a comprehensive understanding of the linguistic markers that influence the model's decision-making process.

Results

Overview

The final dataset comprised 1348 eligible sessions, partitioned into 1011 training and 337 testing cases. After resampling, the training set expanded to 1254 cases, and the test set to 420 cases. We provide dataset characteristics in Table S2 in Multimedia Appendix 2. The mean participant age was 17.9 (95% CI 17.7-18.1) years. We observed significant differences in word count across groups ($F_{2,97}$=48.34, P<.001).

Model Training, Evaluation, and Cross-Validation and Evaluation

The training curves demonstrated that the transformer model learned consistently and stably (Figure S1 in Multimedia Appendix 3). In contrast, the baseline word vector model rapidly reached its learning peak before exhibiting overfitting, halted only by the early stopping mechanism (Figure S1 in Multimedia Appendix 3). The Training Curves section in Multimedia Appendix 3 provides a more detailed interpretation of these learning curves.

The transformer model (T-MLP) achieved superior performance with an overall accuracy of 0.79 (95% CI 0.73-0.99) and a macroaveraged AUC-ROC of 0.89 (95% CI 0.81-0.91). It demonstrated excellent discrimination for nonsuicidal sessions (AUC-ROC=0.96, 95% CI 0.96-0.99) and good discrimination for SI and ASE with AUC-ROCs of 0.85 (95% CI 0.97-0.86) and 0.87 (95% CI 0.81-0.88), respectively. The baseline word2vec model (W2V-MLP) showed lower performance across all metrics (Table 1).

Table 1. Classification performance metrics of transformer-based (T-MLP[a]) and word vector-based (W2V-MLP[b]) models for predicting NS[c], SI[d], and ASE[e] in German youth crisis helpline users (N=337 evaluation set).

Model and class	Precision (positive predictive value)	Recall (sensitivity)	F_1-score (harmonic mean of precision and recall)	AUC[f]-ROC[g] (95% CI)
T-MLP				0.89 (0.81-0.91)
NS	0.97	1	0.98	0.96 (0.96-0.99)
SI	0.71	0.65	0.68	0.85 (0.97-0.86)
ASE	0.69	0.72	0.71	0.87 (0.81-0.88)
W2V-MLP				0.77 (0.64-0.9)
NS	0.61	0.86	0.71	0.89 (0.84-0.9)
SI	0.59	0.69	0.61	0.78 (0.68-0.78)
ASE	0.67	0.28	0.39	0.64 (0.62-0.72)

[a] T-MLP: transformer-multilayer perceptron.

[b] W2V-MLP: word2vector-multilayer perceptron.

[c] NS: not suicidal.

[d] SI: suicidal ideation.

[e] ASE: advanced suicidal engagement.

[f] AUC: area under the curve.

[g] ROC: receiver operating characteristic.

Evaluation and Cross-Validation

Class-wise metrics demonstrate the transformer model's consistent superiority over the baseline model across all classification metrics. Figure 2 illustrates the performance of the T-MLP model. The AUC-ROC analysis (Figure 2A) confirms the model's high discriminative ability across different classes and threshold probabilities. Precision-recall curves (Figure 2B) reveal some uncertainty at lower recall thresholds, suggesting potential for improved calibration. The confusion matrix (Figure 2C) demonstrates the model's effectiveness in distinguishing nonsuicidal cases while showing some inconsistency in classifying SI and ASE.

The T-MLP model exhibits a low false negative rate of 0.01, indicating high accuracy in identifying suicidal sessions. Conversely, the W2V-MLP model (Figure 3) demonstrates inferior performance. Its AUC-ROC analysis (Figure 3A) and precision-recall curves (Figure 3B) show lower discriminative ability. The confusion matrix (Figure 3C) reveals a tendency to over-predict ASE, with a higher false negative rate of 0.17 and a false positive rate of 25% for nonsuicidal outcomes.

These results underscore the T-MLP model's enhanced capability in accurately classifying SIBs compared to the baseline W2V-MLP model while also highlighting areas for potential improvement in model calibration.

DCA demonstrates the T-MLP model's superior clinical utility. For SI prediction, the T-MLP model yields higher NBs than treat-all and treat-none strategies within 0.3-0.5 threshold probabilities, identifying 10-18 true positives without increasing false positives (Figure 4A). The W2V-MLP model, while initially beneficial, exhibits potentially detrimental clinical decisions beyond 0.4 thresholds (Figure 5A).

For ASE, the T-MLP model outperforms alternative strategies within 0.15-0.3 thresholds, achieving an NB of 0.20-0.25 (Figure 4B). Conversely, the W2V-MLP model underperforms relative to the treat-all strategy, potentially leading to suboptimal clinical decisions (Figure 5B). These findings underscore the T-MLP model's enhanced clinical applicability in suicide risk assessment.

We assessed model calibration using Brier and Log loss scores. The T-MLP model achieved lower scores (Brier: 0.10, Log loss: 0.50) compared to the W2V-MLP model (Brier: 0.18, Log loss: 0.92), indicating superior calibration performance. However, both models exhibited resolution challenges in specific classes (Table 2, and Figures 5 and 6).

The Brier Skill Score revealed a 44.4% improvement in classification performance by the T-MLP model over the baseline. Both models demonstrated low variability in predicted probabilities across samples, with an overall uncertainty of 0.22 due to class rebalancing.

Calibration curve analysis (Figure 6) exposed distinct prediction patterns. The T-MLP model consistently overpredicted for the nonsuicidal class across all decision thresholds (Figure 6A) while underpredicting SI and ASE at probabilities above 0.5 and overpredicting at lower thresholds. The W2V-MLP model underpredicted for the nonsuicidal class up to a 0.6 threshold and overpredicted suicidal outcomes at lower thresholds (Figure 6B).

These calibration patterns suggest potential overestimation of clinical utility within predefined thresholds for both models, necessitating cautious interpretation of the DCA results.

Figure 2. Performance analysis of the transformer-based model (T-MLP) for predicting not suicidal, SI, and ASE based on session transcripts of German youth crisis helpline users between November 30, 2021, and April 30, 2022 (N=337 evaluation set). (A) Class-wise one-versus-all AUC-ROC analysis with 95% CIs derived from 1000 bootstrap samples. (B) Class-wise precision-recall analysis with 95% CIs derived from 1000 bootstrap samples. (C) Confusion matrix of the test set: light colors represent low numbers, dark colors represent high numbers; correct classifications are on the diagonal from top-left to bottom-right. ASE: advanced suicidal engagement; AUC: area under the curve; SI: suicidal ideation; ROC: receiver operating characteristic; T-MLP: transformer-multilayer perceptron.

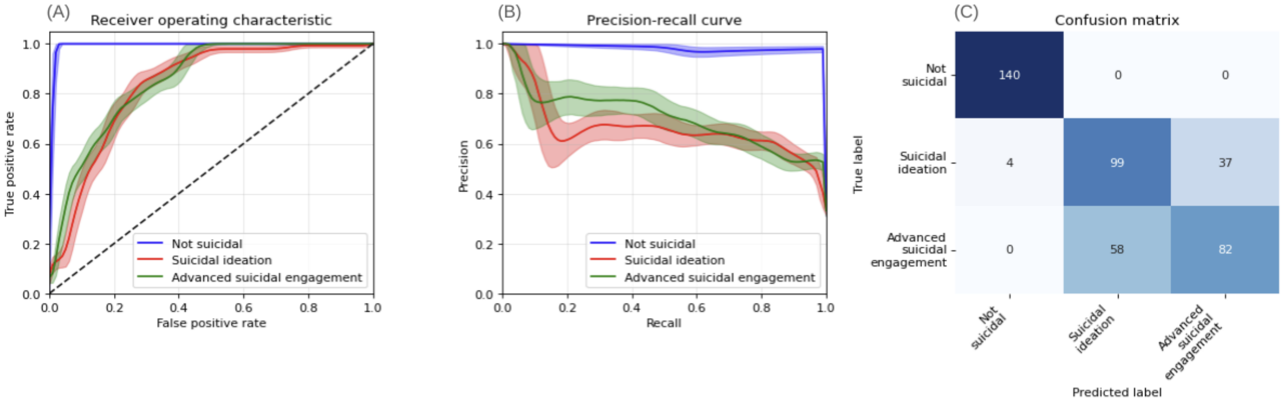


Figure 3. Performance analysis of word vector-based model (W2V-MLP) for predicting not suicidal, SI, and ASE in German youth crisis helpline users (N=337 evaluation set). (A) Class-wise one-versus-all AUC-ROC analysis with 95% CIs derived from 1000 bootstrap samples. (B) Class-wise sensitivity-specificity analysis with 95% CIs derived from 1000 bootstrap samples. (C) Confusion matrix of the test set: light colors represent low numbers, dark colors represent high numbers; correct classifications are on the diagonal from top-left to bottom-right. ASE: advanced suicidal engagement; AUC: area under the curve; SI: suicidal ideation; ROC: receiver operating characteristic; W2V-MLP: word2vector-multilayer perceptron.

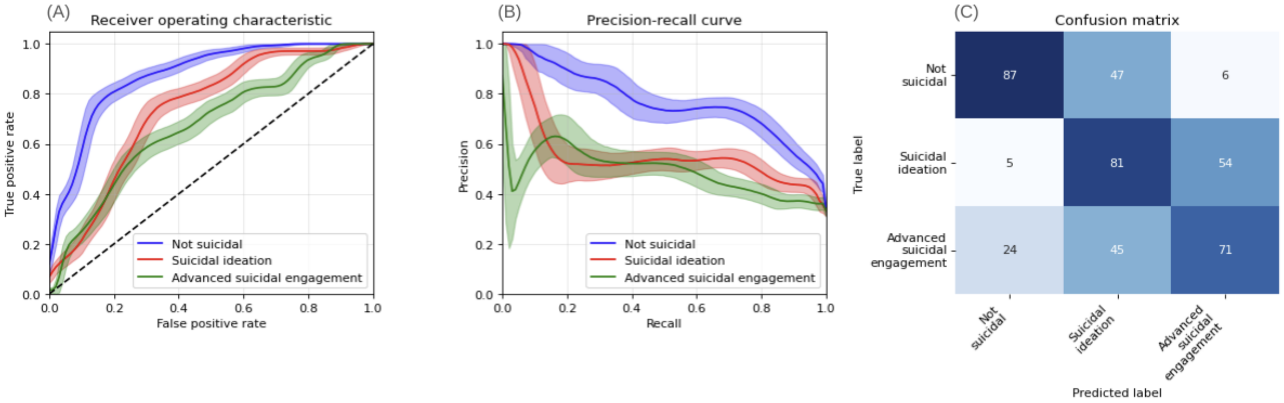


Figure 4. Decision curve analysis of the transformer-based model (T-MLP) for predicting suicidal ideation and advanced suicidal engagement in German youth crisis helpline users (N=337 evaluation set). (A) Class-wise decision curve analysis for suicidal ideation. (B) Class-wise decision curve analysis for advanced suicidal engagement. Color coding: treat-all strategy in gray, treat-none strategy in black, classifier (T-MLP) in blue, theoretical maximal benefit in red. Irrelevant threshold segments are grayed out. T-MLP: transformer-multilayer perceptron.

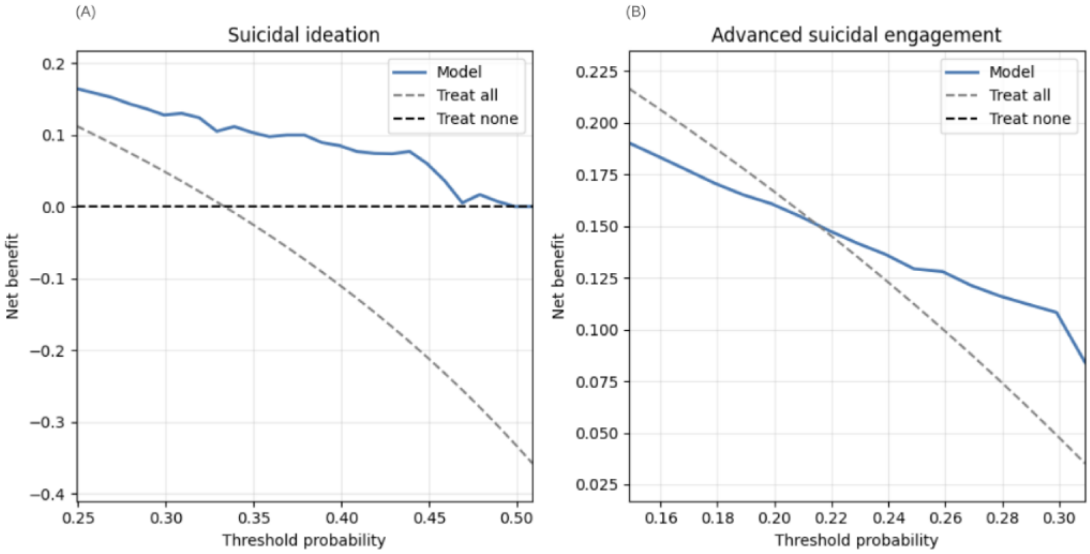


Figure 5. Decision curve analysis of word-vector model (T-MLP) for predicting suicidal ideation and advanced suicidal engagement in German youth crisis helpline users (N=337 evaluation set). (A) Class-wise decision curve analysis for advanced suicidal engagement. (B) Class-wise decision curve analysis for advanced suicidal engagement. Color coding: treat-all strategy in gray, treat-none strategy in black, classifier (T-MLP) in blue, theoretical maximal benefit in red. Irrelevant threshold segments are grayed out. T-MLP: transformer-multilayer perceptron.

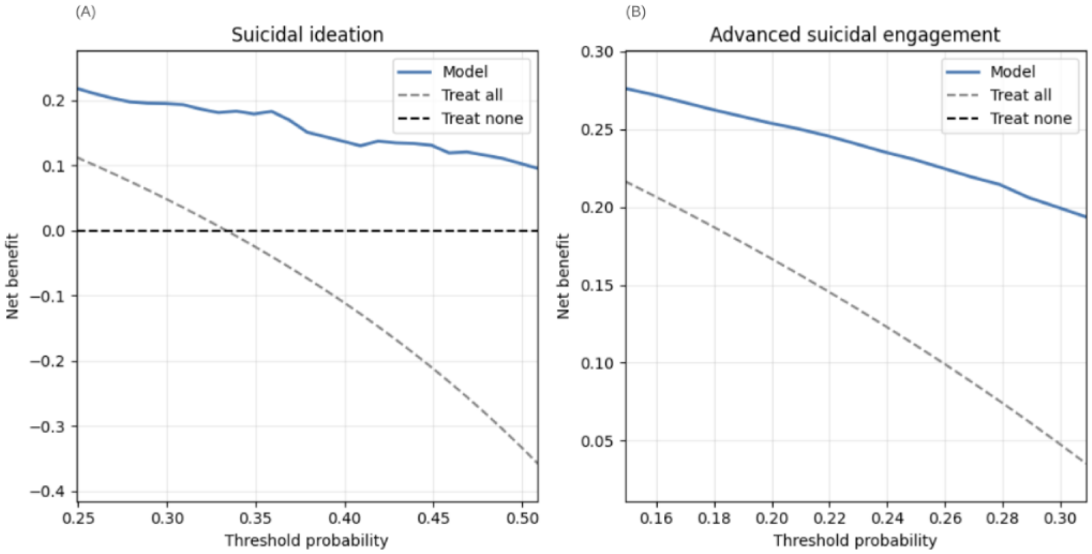


Figure 6. Calibration analysis of transformer-based (T-MLP) and word vector-based (W2V-MLP) models for predicting nonsuicidal, suicidal ideation, and advanced suicidal engagement based on session transcripts of German youth crisis helpline users between November 30, 2021, and April 30, 2022 (N=337 evaluation set). (A) Class-wise calibration curves for the T-MLP model. (B) Class-wise calibration curves for the W2V-MLP model. Color coding: not suicidal (blue), suicidal ideation (orange), and advanced suicidal engagement (green). T-MLP: transformer-multilayer perceptron; W2V-MLP: word2vector-multilayer perceptron.

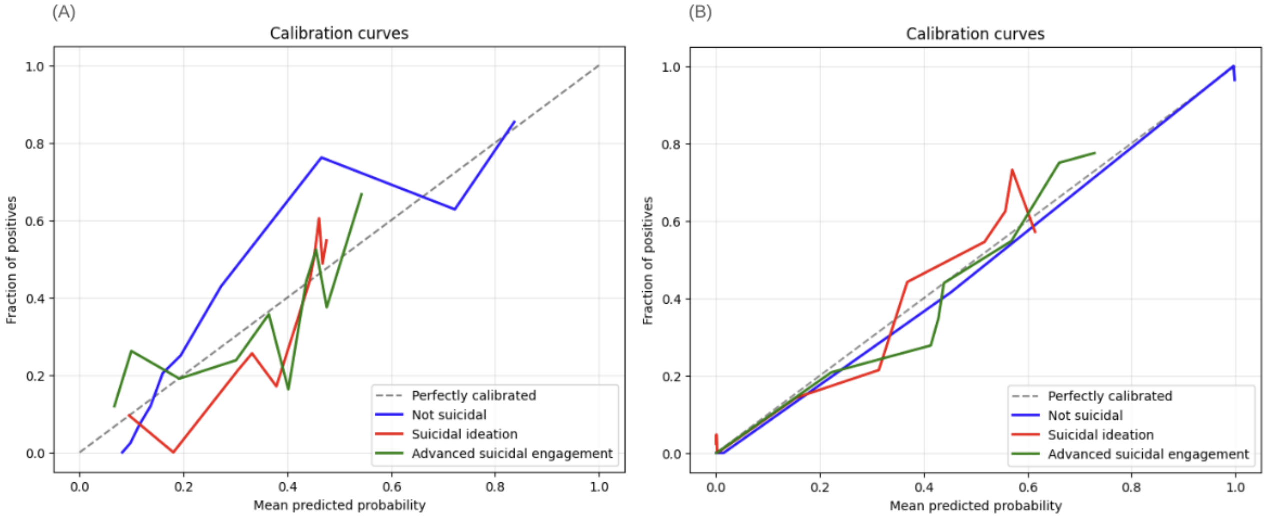


Table 2. Brier score decomposition and model comparison for transformer-based (T-MLP[a]) and word vector-based (W2V-MLP[b]) models in predicting NS[c], SI[d], and ASE[e] based on session transcripts of German youth crisis helpline users between November 30, 2021, and April 30, 2022 (N=337 evaluation set).

	Transformer[f]			word2vec[g]		
Class	NS	SI	ASE	NS	SI	ASE
Reliability	0.002	0.007	0.008	0.001	0.010	0.011
Resolution	0.219	0.863	0.106	0.06	0.015	0.023
Uncertainty	0.222	0.222	0.222	0.222	0.222	0.222

[a] T-MLP: transformer-multilayer perceptron.

[b] W2V-MLP: word2vector-multilayer perceptron.

[c] NS: not suicidal

[d] SIS: suicidal ideation.

[e] ASE: advanced suicidal engagement.

[f] Average Brier loss=0.10; Brier skill score=44.4%.

[g] Average Brier loss=0.18; Brier skill score=44.4%.

Explainability Analysis

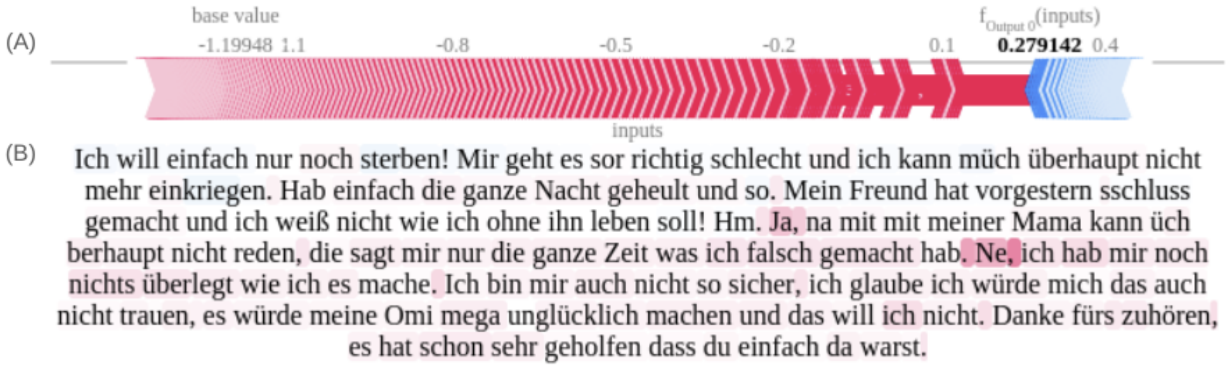
We implemented SHAPs to generate additive text explanations for the transformer model. Due to privacy constraints, we present a vignette illustrating the model's interpretation of feature importance rather than specific outputs.

Global feature importance analysis revealed linguistic traits associated with SI, including self-reference, negation, expressions of low self-esteem, problem emphasis, and articulations of isolation. ASE texts exhibited similar characteristics with additional absolutist language and frequent references to lethal methods and injuries.

Figures 7A and 7B depict a synthetic session vignette predicting the nonsuicidal group, highlighting de-escalating statements in the latter half of the conversation as influential for a positive prediction. Figures 8A, 8B, 9A, and 9B illustrate tokens associated with SI and ASE, respectively. These figures emphasize self-referential statements, expressions indicating a desire to end one's life, and words associated with negative sentiment.

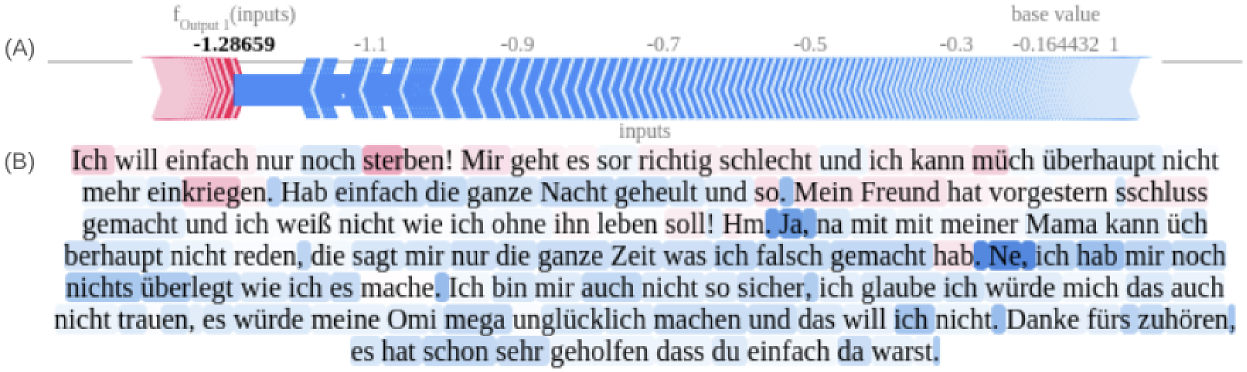
It is noteworthy that in transformer architectures, word tokens do not have fixed impacts on the output due to their context-dependent evaluation. This contextual approach, while representing an advancement over other word-based methods, limits the generalizability of derived Shapley values and plots. Interestingly, the model did not assign high importance to the misspelled word "sschluss gemacht" (intended as "schluss gemacht" meaning "broke up"), demonstrating its robustness to typographical errors.

Figure 7. Shapley values-based text explanation plot for nonsuicidal outcome prediction in German youth crisis helpline users, plot based on synthetic case vignette. (A) Force plot of aggregated Shapley values indicating the predominance of not suicidal features over SI-associated risk markers. (B) The text highlights illustrating content associated with the target class (not suicidal) in red and suicidal content in blue. SI: suicidal ideation.



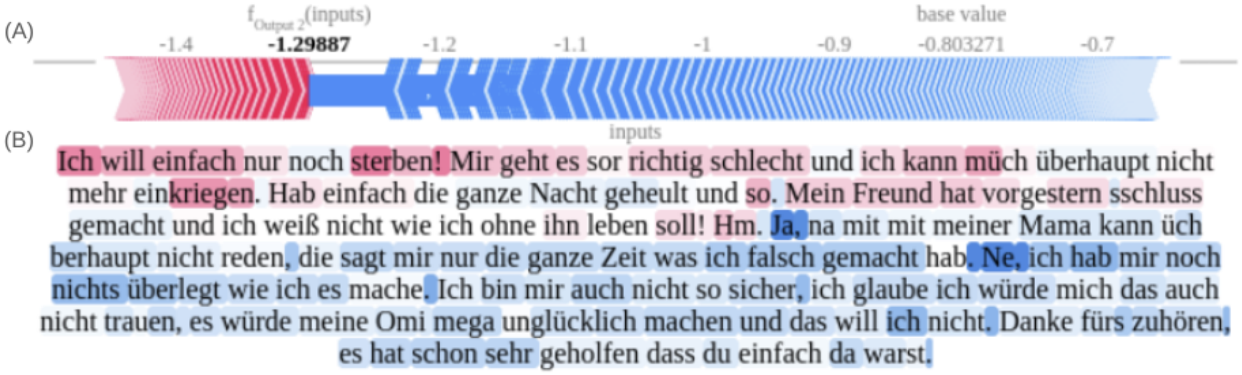
Ich will einfach nur noch sterben! Mir geht es sor richtig schlecht und ich kann müch überhaupt nicht mehr einkriegen. Hab einfach die ganze Nacht geheult und so. Mein Freund hat vorgestern sschluss gemacht und ich weiß nicht wie ich ohne ihn leben soll! Hm. Ja, na mit mit meiner Mama kann üch berhaupt nicht reden, die sagt mir nur die ganze Zeit was ich falsch gemacht hab. Ne, ich hab mir noch nichts überlegt wie ich es mache. Ich bin mir auch nicht so sicher, ich glaube ich würde mich das auch nicht trauen, es würde meine Omi mega unglücklich machen und das will ich nicht. Danke fürs zuhören, es hat schon sehr geholfen dass du einfach da warst.

Figure 8. Shapley values-based text explanation plot for suicidal ideation outcome prediction in German youth crisis helpline users, plot based on synthetic case vignette. (A) Force plot of aggregated Shapley values indicating the balance between suicidal ideation-associated risk markers and nonsuicidal features. (B) Text highlights illustrating content associated with the target class (suicidal ideation) in red and nonsuicidal content in blue.



Ich will einfach nur noch sterben! Mir geht es sor richtig schlecht und ich kann müch überhaupt nicht mehr einkriegen. Hab einfach die ganze Nacht geheult und so. Mein Freund hat vorgestern sschluss gemacht und ich weiß nicht wie ich ohne ihn leben soll! Hm. Ja, na mit mit meiner Mama kann üch berhaupt nicht reden, die sagt mir nur die ganze Zeit was ich falsch gemacht hab. Ne, ich hab mir noch nichts überlegt wie ich es mache. Ich bin mir auch nicht so sicher, ich glaube ich würde mich das auch nicht trauen, es würde meine Omi mega unglücklich machen und das will ich nicht. Danke fürs zuhören, es hat schon sehr geholfen dass du einfach da warst.

Figure 9. Shapley values-based text explanation plot for advanced suicidal engagement outcome prediction in German youth crisis helpline users, plot based on synthetic case vignette. (A) Force plot of aggregated Shapley values indicating the balance between advanced suicidal engagement–associated risk markers and nonsuicidal features. (B) Text highlights illustrate content associated with the target class (advanced suicidal engagement) in red and nonsuicidal content in blue. Translation: “I just want to die right now. I am feeling pretty bad, and I can’t calm myself down. I was crying the whole night and stuff. My boyfriend dumped me the day before yesterday, and I don't know how to live without him!” “Hm. Yeah, I can't talk to my mum about such stuff. She just points out my flaws all the time.” “No, I don't have a specific Plan of how to do it; I am also pretty unsure - I think I would never have the courage to do it, my Granny would be super sad and I don't want that. Thank you for listening to me. Thank you for being there for me. This helped a lot.”.



Ich will einfach nur noch sterben! Mir geht es sor richtig schlecht und ich kann müch überhaupt nicht mehr einkriegen. Hab einfach die ganze Nacht geheult und so. Mein Freund hat vorgestern sschluss gemacht und ich weiß nicht wie ich ohne ihn leben soll! Hm. Ja, na mit mit meiner Mama kann üch berhaupt nicht reden, die sagt mir nur die ganze Zeit was ich falsch gemacht hab. Ne, ich hab mir noch nichts überlegt wie ich es mache. Ich bin mir auch nicht so sicher, ich glaube ich würde mich das auch nicht trauen, es würde meine Omi mega unglücklich machen und das will ich nicht. Danke fürs zuhören, es hat schon sehr geholfen dass du einfach da warst.

Discussion

Overview

We aimed to develop, evaluate, cross-validate, and benchmark a transformer-based prediction model for detecting SI and ASE among adolescents seeking emergency online counseling services in Germany.

The primary findings of this study demonstrate the proficiency of our transformer-based prediction model in identifying SI and ASE among adolescents using online counseling services. Our model outperformed the baseline word-vector approach, exhibiting excellent to very good performance in distinguishing between nonsuicidal sessions, SI, and ASE. Additionally, the Shapley analysis successfully identified word tokens associated with suicidal outcomes, providing valuable insights into language features indicative of SI. These findings highlight the potential of ML models in enhancing suicide prevention efforts.

Principal Results

Our key findings underscore the superior performance of our transformer-based prediction model in identifying SI and ASE using natural language data from adolescents. The model was trained on rigorously collected data evaluated by 3 expert raters using an established and validated scale. We adhered to a strict validation protocol, including repeated cross-validation and clinical utility metrics, to ensure our results are reliable, interpretable, and applicable in clinical settings.

The transformer model showed strong capabilities in differentiating between session transcripts containing nonsuicidal statements and those featuring discussions of SI or ASE, despite some calibration challenges. These findings suggest the model’s potential as a valuable screening tool for identifying individuals who may be at risk for SI and ASE, potentially supporting clinical decision-making and enabling timely interventions. The model excelled at recognizing transcripts with nonsuicidal content, though it faced some difficulties in precisely distinguishing more nuanced levels of

SIB. Notably, our model substantially surpassed the performance of the baseline model that used pretrained word vectors, demonstrating improved overall metrics and a reduced risk of misclassification. While the model achieved high sensitivity and specificity in identifying nonsuicidal cases, its specificity for SI and ASE was comparatively lower.

Analysis of calibration curves indicates a necessity for recalibration before clinical implementation, as current reliability metrics suggest some degree of instability. Although the transformer model exhibits evident clinical utility, the observed calibration issues may lead to an overestimation of its practical performance. Furthermore, while the application of SHAPs analysis effectively identified word tokens associated with suicidal outcomes, it demonstrated reduced effectiveness in interpreting out-of-context tokens. This limitation potentially constrains the model's capacity to comprehensively capture the dynamic nature of youth language patterns.

Overall, the transformer model outperformed the word-vector model, demonstrating superior performance and higher clinical utility, even when trained on a relatively small, domain-specific dataset. This advantage primarily stems from the transformer encoders' advanced ability to represent natural language, partly due to their contextual embedding strategy and training on larger datasets. Additionally, the architecture inherently excels at analyzing deeper levels of language structure.

Our findings offer strong support for the use of ML models, particularly those using transformers and transfer learning, in detecting and classifying levels of SI in clinical psychological and psychiatric research that relies heavily on text and speech data. The transformer model surpasses both the treat-all and treat-none strategies and the baseline model for detecting SI and ASE. Including an explainability component provides clinicians with insights into the specific language features associated with SI or advanced engagement [93].

The SHAPs algorithm generated local explanations for the model's predictions, which we aggregated to gain insight into the global importance of features in the language of suicidal adolescents. The explainer model effectively pinpointed specific language markers that influence the model's predictions. The importance of explainability is underscored by regulations such as the AI Act of the European Union, which emphasizes transparency in AI applications [94].

The language markers identified by the model are consistent with existing research on suicidal language markers, validating our results. Notably, our analysis confirmed that the usage of first-person singular pronouns and absolutist language is linked to SI [95-97].

This study demonstrates the potential of our model to accurately assess and identify suicidal thoughts in adolescents, with significant implications for prevention strategies addressing SIBs. The model's high applicability to the online crisis counseling context facilitates its potential adoption by other online counseling services, potentially aiding in risk mitigation and supporting clinical decision-making processes. The model's capacity to identify key risk markers may enable more precise risk assessments and inform targeted interventions for individuals at elevated risk of ASE. Furthermore, the model's ability to track and detect changes in language usage over time could contribute to earlier identification of individuals experiencing increased risk.

In summary, the developed model can significantly enhance the effectiveness and efficiency of prevention efforts at any level of SI.

Limitations

The primary limitation of this study is the small sample size, which may restrict the generalizability of the model's performance. As evidenced by the learning curves (Figures S1 and S2 in Multimedia Appendix 3), the transformer model converged with minimal overfitting in contrast to the baseline model. The limited sample size might have influenced class learnability, particularly for the word-vector approach, which requires more parameter estimation. Repeated cross-validation revealed training instability and sensitivity to weight initialization, potentially leading to class-specific biases. Calibration metrics indicated significant reliability issues for both models, suggesting the need for recalibration before clinical implementation. Due to these calibration concerns, the clinical utility of the models should be interpreted cautiously, as small sample sizes introduce more noise to neural networks [98]. These issues could be mitigated by using larger datasets, implementing robust weight initialization, or applying additional regularization techniques.

Furthermore, due to metric requirements, we applied resampling techniques to the rarer outcomes, which may lead to overfitting minority classes and lower generalization. Performance improvements over the baseline may be overstated due to less emphasis on feature engineering in the word-vector model. Incorporating n-grams, tf-idf (term frequency–inverse document frequency), or dictionary-based features could address the lack of sequential text structure in word embeddings, which this study did not fully address.

Potential age or gender biases in the expression of suicidality may exist, as the sample was not gender-balanced, potentially skewing results toward female expressions of SIBs. The class imbalance necessitated the use of oversampling techniques, which might introduce biases. Technical limitations led to session truncations, affecting data quality and representativeness.

This study's exclusive focus on textual data, while appropriate for the crisis text line setting, may limit the comprehensive assessment of suicidality. Suicidal behavior is multifaceted, and other data types such as audio or visual cues could provide additional valuable information. The absence of these modalities in our analysis is a limitation inherent to the text-based nature of the crisis line service studied.

Token-based post hoc explanations might not fully capture the sequential nature of speech, possibly missing complex effects detected by transformers and LSTM layers. Future research should explore specific techniques for transformer-based explainability and evaluate the linguistic features of highlighted tokens. Limiting analysis to text tokens may overlook the importance of other communication forms, such as emoticons, which play a significant role in conveying meaning. Future models should consider these aspects.

Additionally, the vignette demonstrated that despite being trained on more extensive data, the transformer encoder still will not fully capture irregular or misspelled language. Adapting to new language variations, often developed among youth, heavily depends on collecting domain-specific datasets for embedding models.

Comparison to Prior Work

The discriminative abilities of our model are comparable, albeit slightly lower, to those reported by Zhang et al [99], who analyzed 659 suicide notes, 431 last statements, and 2000 neutral posts. Their model achieved 95%, 94.9%, and 94.9% in precision, recall, and F_1-score, respectively, although it was limited to binary targets and explicitly focused on suicide notes. Similarly, Broadbent et al [28] found that using sentence-wise embeddings of counseling sessions, a transformer embedding surpassed a baseline tf-idf model in reducing the false-negative rate.

In contrast, Aldhyani et al [100] demonstrated that a convolutional neural network and bidirectional LSTM model excelled over the XGBoost model, achieving 95% accuracy in detecting SI compared to the latter's 91.5%. However, when using validated dictionary features such as Linguistic Inquiry and Word Count, boosted tree algorithms performed better than a neural network architecture combining convolutional and LSTM layers (convolutional neural network and bidirectional LSTM) on binary targets. The key distinction of this study from Aldhyani et al [100] lies in the emphasis on feature engineering and the use of dictionaries, which were leading methods for building text classifiers before the emergence of transfer learning with large pretrained language models [62].

Unique to our study, we also include metrics for calibration and clinical utility, enhancing the practical value of our findings.

Generalizability

The generalizability of our findings varies by domain. Given the small and domain-specific sample, the results may primarily apply to German-speaking adolescents seeking mental health assistance. However, the analysis of language features suggests that findings might extend to a broader demographic group.

Clinical Implications and Implications for Future Research

Our transformer-based model has the potential to assist clinicians in identifying at-risk individuals and improving intervention prioritization. Continuous monitoring using such models is cost-effective and can detect cases that might otherwise be missed, facilitating referral to qualified human review and leading to more accurate and timely interventions. This research significantly contributes to prevention efforts addressing SIBs.

Despite the promise of such models, as indicated in our study, future research could enhance the model's accuracy and utility. Subsequent studies should aim to collect sample sizes that are adequate for the power requirements of ML. If faced with the common problem of class imbalance, researchers could opt for adjusting the cost function of the learner or using synthetic data generation through augmentation to mitigate this issue.

Other key priorities include validating the model across diverse clinical settings such as inpatient facilities, rural clinics, and school-based mental health services to enhance its generalizability. Additionally, evaluating the model's performance across different demographic groups and cultural contexts will be essential to ensure its effectiveness and identify potential biases.

Developing dynamic models that track changes over time and incorporate additional factors such as demographics could improve performance. This could be achieved by using retrained transformers in extracting longitudinal event data [101] or by incorporating explanations into the clinicians' counseling environment [102]. Integrating transformer embeddings into multimodal graph network models might refine precision and bolster prevention efforts by addressing the multifactorial nature of suicide.

Integrating multiple data types could provide a more holistic view of an individual's mental state. However, such approaches would require careful consideration of privacy concerns, data integration challenges, and the development of more complex models capable of processing diverse input types. Future studies could explore the relative contributions of different data modalities to suicide risk assessment and the feasibility of implementing multimodal systems in various clinical settings.

ML shows potential in mental health applications, but it should be considered a complementary tool to clinical expertise and evidence-based decision-making rather than a standalone solution. In suicidology, it is imperative to recognize that suicide risk emerges from complex life circumstances that require comprehensive social interventions, support systems, and targeted public policies. While ML algorithms can serve as diagnostic tools to detect linguistic markers associated with suicide risk in individuals, they cannot independently prevent or resolve the underlying personal challenges. Furthermore, the implementation of ML in this context necessitates careful ethical considerations to ensure that these tools augment human judgment without introducing or amplifying biases. Future research should focus on integrating ML approaches within broader, multifaceted suicide prevention strategies that address the full spectrum of risk factors and protective measures.

Conclusions

Suicide prevention represents a critical public health concern, particularly among adolescents. Timely identification of individuals displaying SIBs is crucial for early intervention. Our study introduces an explainable transformer-based ML model that outperforms a baseline word-vector approach in identifying SI and ASE in adolescents seeking help via a German crisis helpline. This model could be invaluable for clinicians prioritizing intervention cases. However, challenges remain in discerning finer-grained types of SI. This research underscores the potential of ML in detecting SI and represents a step toward more effective suicide prevention. Future work should focus on enhancing model accuracy in longitudinal setups or incorporating multimodal features.

Acknowledgments

The publication was funded and supported by the Open Access Fund of Universität Trier and by the German Research Foundation (DFG).

Data Availability

This study used highly sensitive and confidential crisis helpline chat transcripts from vulnerable individuals, including minors. Due to the nature of these conversations, which frequently contain personal details and sensitive information regarding mental health and SI, the raw data cannot be made publicly available. This restriction is necessary to protect participant privacy and confidentiality and to comply with ethical guidelines and data protection regulations, including the General Data Protection Regulation. Even in an anonymized form, the risk of reidentification cannot be completely mitigated, thus precluding data sharing.

Conflicts of Interest

JT, AL, RW, and JS are or were employed and receive a salary from krisenchat. krisenchat did not influence aspects like data collection, analysis, and reporting of results. GM received funding from the Stanley Thomas Johnson Stiftung and Gottfried und Julia Bangerter-Rhyner-Stiftung (project PC 28/17 and PC 05/18), from Gesundheitsförderung Schweiz (project 18.191/K50001), from the Swiss Heart Foundation (project =FF21101), from the Research Foundation of the International Psychoanalytic University Berlin (project 5087 and 5217), from the Swiss National Science Foundation (SNSF; project 100014_135328), from the German Federal Ministry of Education and Research (budget item 68606) in the context of an evaluation project conducted among others in collaboration with krisenchat, from the Hasler Foundation (project 23004), in the context of a Horizon Europe project from the Swiss State Secretariat for Education, Research and Innovation (SERI; contract 22.00094), and from Wings Health in the context of a proof-of-concept study. GM is a cofounder, member of the board, and shareholder of Therayou AG, active in digital and blended mental health care. GM receives royalties from publishing companies as author, including a book published by Springer, and an honorarium from Lundbeck for speaking at a symposium. Furthermore, GM is compensated for providing psychotherapy to patients, acting as a supervisor, serving as a self-experience facilitator ("Selbsterfahrungsleiter"), and for postgraduate training of psychotherapists and supervisors.

Multimedia Appendix 1
Reporting guideline checklists.
[DOCX File , 37 KB - publichealth_v11i1e63809_app1.docx]

Multimedia Appendix 2
Supplementary tables and figures.
[DOCX File , 52 KB - publichealth_v11i1e63809_app2.docx]

Multimedia Appendix 3
A primer on neural networks and metrics.
[DOCX File , 256 KB - publichealth_v11i1e63809_app3.docx]

References

1. Hink AB, Killings X, Bhatt A, Ridings LE, Andrews AL. Adolescent suicide-understanding unique risks and opportunities for trauma centers to recognize, intervene, and prevent a leading cause of death. Curr Trauma Rep 2022;8(2):41-53 [FREE Full text] [doi: 10.1007/s40719-022-00223-7] [Medline: 35399601]
2. Becker M, Correll CU. Suicidality in Childhood and Adolescence. Dtsch Arztebl Int 2020 Apr 10;117(15):261-267. [doi: 10.3238/arztebl.2020.0261] [Medline: 32449889]
3. Shneidman ES. Definition of Suicide. New Jersey: Wiley; 1985.
4. Harmer B, Lee S, Rizvi A, Saadabadi A. Suicidal ideation. StatPearls. Florida: StatPearls Publishing URL: http://www.ncbi.nlm.nih.gov/books/NBK565877 [accessed 2024-07-29]
5. O'Carroll PW, Berman AL, Maris RW, Moscicki EK, Tanney BL, Silverman MM. Beyond the tower of babel: a nomenclature for suicidology. Suicide Life Threat Behav 1996;26(3):237-252. [doi: 10.1111/j.1943-278x.1996.tb00609.x]
6. Campisi SC, Carducci B, Akseer N, Zasowski C, Szatmari P, Bhutta ZA. Suicidal behaviours among adolescents from 90 countries: a pooled analysis of the global school-based student health survey. BMC Public Health 2020;20(1):1102 [FREE Full text] [doi: 10.1186/s12889-020-09209-z] [Medline: 32772922]

7. Turecki G, Brent DA, Gunnell D, O'Connor RC, Oquendo MA, Pirkis J, et al. Suicide and suicide risk. Nat Rev Dis Primers 2019;5(1):74 [FREE Full text] [doi: 10.1038/s41572-019-0121-0] [Medline: 31649257]
8. Benton TD, Muhrer E, Jones JD, Lewis J. Dysregulation and suicide in children and adolescents. Child Adolesc Psychiatr Clin N Am 2021;30(2):389-399. [doi: 10.1016/j.chc.2020.10.008] [Medline: 33743946]
9. Viner RM, Ozer EM, Denny S, Marmot M, Resnick M, Fatusi A, et al. Adolescence and the social determinants of health. Lancet 2012;379(9826):1641-1652. [doi: 10.1016/S0140-6736(12)60149-4] [Medline: 22538179]
10. Pelkonen M, Marttunen M. Child and adolescent suicide: epidemiology, risk factors, and approaches to prevention. Paediatr Drugs 2003;5(4):243-265. [doi: 10.2165/00128072-200305040-00004] [Medline: 12662120]
11. Richards D, Viganó N. Online counseling: a narrative and critical review of the literature. J Clin Psychol 2013;69(9):994-1011. [doi: 10.1002/jclp.21974] [Medline: 23630010]
12. Grimland M, Benatov J, Yeshayahu H, Izmaylov D, Segal A, Gal K, et al. Predicting suicide risk in real-time crisis hotline chats integrating machine learning with psychological factors: exploring the black box. Suicide Life Threat Behav 2024;54(3):416-424. [doi: 10.1111/sltb.13056] [Medline: 38345174]
13. Cunningham CE, Walker JR, Eastwood JD, Westra H, Rimas H, Chen Y, The Mobilizing Minds Research Group. Modeling mental health information preferences during the early adult years: a discrete choice conjoint experiment. J Health Commun 2014;19(4):413-440 [FREE Full text] [doi: 10.1080/10810730.2013.811324] [Medline: 24266450]
14. Bey R, Cohen A, Trebossen V, Dura B, Geoffroy P, Jean C, et al. Natural language processing of multi-hospital electronic health records for public health surveillance of suicidality. npj Ment Health Res 2024;3(1):6 [FREE Full text] [doi: 10.1038/s44184-023-00046-7] [Medline: 38609541]
15. Brody C, Star A, Tran J. Chat-based hotlines for health promotion: a systematic review. mHealth 2020;6:36 [FREE Full text] [doi: 10.21037/mhealth-2019-di-13] [Medline: 33437833]
16. Hoffberg AS, Stearns-Yoder KA, Brenner LA. The effectiveness of crisis line services: a systematic review. Front Public Health 2020;7:399 [FREE Full text] [doi: 10.3389/fpubh.2019.00399] [Medline: 32010655]
17. Lejeune A, Le Glaz A, Perron PA, Sebti J, Baca-Garcia E, Walter M, et al. Artificial intelligence and suicide prevention: a systematic review. Eur Psychiatry 2022;65(1):1-22 [FREE Full text] [doi: 10.1192/j.eurpsy.2022.8] [Medline: 35166203]
18. Bernert RA, Hilberg AM, Melia R, Kim JP, Shah NH, Abnousi F. Artificial intelligence and suicide prevention: a systematic review of machine learning investigations. Int J Environ Res Public Health 2020;17(16):5929 [FREE Full text] [doi: 10.3390/ijerph17165929] [Medline: 32824149]
19. Cohen J, Wright-Berryman J, Rohlfs L, Trocinski D, Daniel L, Klatt TW. Integration and validation of a natural language processing machine learning suicide risk prediction model based on open-ended interview language in the emergency department. Front Digit Health 2022;4:818705 [FREE Full text] [doi: 10.3389/fdgth.2022.818705] [Medline: 35187527]
20. Kusuma K, Larsen M, Quiroz JC, Gillies M, Burnett A, Qian J, et al. The performance of machine learning models in predicting suicidal ideation, attempts, and deaths: a meta-analysis and systematic review. J Psychiatr Res 2022;155:579-588. [doi: 10.1016/j.jpsychires.2022.09.050] [Medline: 36206602]
21. Nordin N, Zainol Z, Mohd Noor MH, Chan LF. Suicidal behaviour prediction models using machine learning techniques: a systematic review. Artif Intell Med 2022;132:102395. [doi: 10.1016/j.artmed.2022.102395] [Medline: 36207078]
22. Hill RM, Oosterhoff B, Kaplow JB. Prospective identification of adolescent suicide ideation using classification tree analysis: models for community-based screening. J Consult Clin Psychol 2017;85(7):702-711. [doi: 10.1037/ccp0000218] [Medline: 28414489]
23. Hopkins D, Rickwood DJ, Hallford DJ, Watsford C. Structured data vs. unstructured data in machine learning prediction models for suicidal behaviors: a systematic review and meta-analysis. Front Digit Health 2022;4:945006 [FREE Full text] [doi: 10.3389/fdgth.2022.945006] [Medline: 35983407]
24. Nock MK, Millner AJ, Ross EL, Kennedy CJ, Al-Suwaidi M, Barak-Corren Y, et al. Prediction of suicide attempts using clinician assessment, patient self-report, and electronic health records. JAMA Netw Open 2022;5(1):e2144373 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.44373] [Medline: 35084483]
25. Choi SB, Lee W, Yoon J, Won J, Kim DW. Ten-year prediction of suicide death using Cox regression and machine learning in a nationwide retrospective cohort study in South Korea. J Affect Disord 2018;231:8-14. [doi: 10.1016/j.jad.2018.01.019] [Medline: 29408160]
26. Chen Q, Zhang-James Y, Barnett EJ, Lichtenstein P, Jokinen J, D'Onofrio BM, et al. Predicting suicide attempt or suicide death following a visit to psychiatric specialty care: a machine learning study using Swedish national registry data. PLoS Med 2020;17(11):e1003416 [FREE Full text] [doi: 10.1371/journal.pmed.1003416] [Medline: 33156863]
27. Boudreaux ED, Rundensteiner E, Liu F, Wang B, Larkin C, Agu E, et al. Applying machine learning approaches to suicide prediction using healthcare data: overview and future directions. Front Psychiatry 2021;12:707916 [FREE Full text] [doi: 10.3389/fpsyt.2021.707916] [Medline: 34413800]
28. Broadbent M, Medina Grespan M, Axford K, Zhang X, Srikumar V, Kious B, et al. A machine learning approach to identifying suicide risk among text-based crisis counseling encounters. Front Psychiatry 2023;14:1110527 [FREE Full text] [doi: 10.3389/fpsyt.2023.1110527] [Medline: 37032952]

29. Roy K, Gaur M, Soltani M, Rawte V, Kalyan A, Sheth A. ProKnow: process knowledge for safety constrained and explainable question generation for mental health diagnostic assistance. Front Big Data 2023;5:1056728 [FREE Full text] [doi: 10.3389/fdata.2022.1056728] [Medline: 36700134]
30. Eloranta S, Boman M. Predictive models for clinical decision making: deep dives in practical machine learning. J Intern Med 2022;292(2):278-295 [FREE Full text] [doi: 10.1111/joim.13483] [Medline: 35426190]
31. Schafer KM, Kennedy G, Gallyer A, Resnik P. A direct comparison of theory-driven and machine learning prediction of suicide: a meta-analysis. PLoS One 2021;16(4):e0249833. [doi: 10.1371/journal.pone.0249833] [Medline: 33844698]
32. Selva Birunda S, Kanniga Devi R. A review on word embedding techniques for text classification. In: Raj JS, Iliyasu AM, Bestak R, Baig ZA, editors. Innovative Data Communication Technologies and Application. Lecture Notes on Data Engineering and Communications Technologies. Singapore: Springer; 2021:267-281.
33. Apidianaki M. From word types to tokens and back: a survey of approaches to word meaning representation and interpretation. Comput Linguist 2023;49(2):465-523. [doi: 10.1162/coli_a_00474]
34. Rosenbloom B, Hudson DA, Adeli E, Altman R, Arora S, Arx SV, et al. On the opportunities and risks of foundation models. arXiv Preprint posted on October 1, 2019. [doi: 10.5260/chara.21.2.8]
35. Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, Moi A, et al. Transformers: state-of-the-art natural language processing. 2020 Presented at: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations; October 2020; Virtual p. 38-45. [doi: 10.18653/v1/2020.emnlp-demos.6]
36. Malte A, Ratadiya P. Evolution of transfer learning in natural language processing. arXiv. Posted online on October 16, 2019 2019. [doi: 10.48550/arXiv.1910.07370]
37. Haroz EE, Grubin F, Goklish N, Pioche S, Cwik M, Barlow A, et al. Designing a clinical decision support tool that leverages machine learning for suicide risk prediction: development study in partnership with native American care providers. JMIR Public Health Surveill 2021;7(9):e24377 [FREE Full text] [doi: 10.2196/24377] [Medline: 34473065]
38. Ananthakrishnan G, Jayaraman AK, Trueman TE, Mitra SK, Murugappan AA. Suicidal intention detection in tweets using BERT-based transformers. 2022 Presented at: International Conference on Computing, Communication, and Intelligent Systems (ICCCIS); November 4-5, 2022; Greater Noida. [doi: 10.1109/icccis56430.2022.10037677]
39. Lin E, Sun J, Chen H, Mahoor MH. Data quality matters: suicide intention detection on social media posts using a RoBERTa-CNN model. arXiv. Posted online on February 3, 2024 2024. [doi: 10.48550/arXiv.2402.02262]
40. Alhamed F, Ive J, Specia L. Proceedings of the 9th Workshop on Computational Linguistics and Clinical Psycholog. In: Using Large Language Models (LLMs) to extract evidence from pre-annotated social media data. 2024 Presented at: Proceedings of the Eighth Workshop on Computational Linguistics and Clinical Psychology; 2024 July 29; Seattle, USA p. 232-327 URL: https://aclanthology.org/2024.clpsych-1.22 [doi: 10.18653/v1/2022.clpsych-1.23]
41. Tanaka R, Fukazawa Y. Integrating supervised extractive and generative language models for suicide risk evidence summarization. arXiv. Posted online on March 20, 2024 2024.
42. Devika SP, Pooja MR, Arpitha MS, Vinayakumar R. BERT-Based Approach for Suicide Depression Identification. Proc Third Int Conf Adv Comput Eng Commun Syst 2023;612:435-444. [doi: 10.1007/978-981-19-9228-5_36]
43. Sawhney R, Joshi H, Gandhi S, Shah R. A time-aware transformer based model for suicide ideation detection on social media. 2020 Presented at: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP); 2020 November 01; NA p. 7685-7697.
44. Akintoye O, Wei N, Liu Q. Suicide detection in tweets using LSTM and transformers. 2024 Presented at: 4th Asia Conference on Information Engineering (ACIE); 2024 January 26-28; Singapore, Singapore.
45. Chatterjee M, Kumar P, Samanta P, Sarkar D. Suicide ideation detection from online social media: a multi-modal feature based technique. Int J Inf Manage Data Insights 2022;2(2):100103. [doi: 10.1016/j.jjimei.2022.100103]
46. Pigoni A, Delvecchio G, Turtulici N, Madonna D, Pietrini P, Cecchetti L, et al. Machine learning and the prediction of suicide in psychiatric populations: a systematic review. Transl Psychiatry 2024;14(1):140 [FREE Full text] [doi: 10.1038/s41398-024-02852-9] [Medline: 38461283]
47. Goh ATC. Back-propagation neural networks for modeling complex systems. Artif Intell Eng 1995;9(3):143-151. [doi: 10.1016/0954-1810(94)00011-s]
48. Kokalj E, Škrlj B, Lavrač N, Pollak S, Robnik-Šikonja M. BERT meets Shapley: extending SHAP explanations to transformer-based classifiers. 2021 Presented at: Proceedings of the EACL Hackashop on News Media Content Analysis and Automated Report Generation; April 1, 2021; Virtual p. 16-21 URL: https://aclanthology.org/2021.hackashop-1.3
49. Lundberg S, Lee SI. A unified approach to interpreting model predictions. arXiv. Posted online on November 25, 2017 2017;24.
50. Nordin N, Zainol Z, Mohd Noor MH, Chan LF. An explainable predictive model for suicide attempt risk using an ensemble learning and Shapley Additive Explanations (SHAP) approach. Asian J Psychiatr 2023;79:103316. [doi: 10.1016/j.ajp.2022.103316] [Medline: 36395702]
51. Malhotra A, Jindal R. XAI transformer based approach for interpreting depressed and suicidal user behavior on online social networks. Cogn Syst Res 2024;84(1):101186. [doi: 10.1016/j.cogsys.2023.101186]
52. Sandoval AA, Aguilar-Canto F, Jiménez D, Calvo H. Suicide tendency content detection with natural language processing and LIME explainer. Comput. Sci. Res 2023;152(12):143-152.

53. Mumenin N, Basar R, Hossain ABMK, Hossain MD, Hasan ML. Suicidal ideation detection from social media texts using an interpretable hybrid model. 2023 Presented at: 6th International Conference on Electrical Information and Communication Technology (EICT); 2023 December 07-09; Khulna, Bangladesh p. 1-6.
54. Naseem U, Khushi M, Kim J, Dunn AG. Hybrid text representation for explainable suicide risk identification on social media. IEEE Trans Comput Soc Syst 2024;11:4663-4672. [doi: 10.1109/TCSS.2022.3184984]
55. Thushari PD, Aggarwal N, Vajrobol V, Saxena GJ, Singh S, Pundir A. Identifying discernible indications of psychological well-being using ML: explainable AI in reddit social media interactions. Soc Netw Anal Min 2023;13(1):141. [doi: 10.1007/s13278-023-01145-1]
56. Verma A, Harper M, Assi S, Al-Hamid A, Yousif MG, Mustafina J, et al. Suicide ideation detection: a comparative study of sequential and transformer hybrid algorithms. In: Wah YB, Berry MW, Mohamed A, Al-Jumeily D, editors. Data Science and Emerging Technologies. DaSET 2022. Lecture Notes on Data Engineering and Communications Technologies, vol 165. Singapore: Springer; 2023.
57. Tang H, Miri Rekavandi A, Rooprai D, Dwivedi G, Sanfilippo FM, Boussaid F, et al. Analysis and evaluation of explainable artificial intelligence on suicide risk assessment. Sci Rep 2024;14(1):6163 [FREE Full text] [doi: 10.1038/s41598-024-53426-0] [Medline: 38485985]
58. Mahamud AH, Dey AK, Alam ANMS, Alam MGR, Zaman S. Implementation of Explainable AI in Mental Health Informatics: Suicide Data of the United Kingdom. 2022 Presented at: 12th International Conference on Electrical and Computer Engineering (ICECE); December 21-23, 2022; Dhaka, Bangladesh.
59. Kirtley OJ, van Mens K, Hoogendoorn M, Kapur N, de Beurs D. Translating promise into practice: a review of machine learning in suicide research and prevention. Lancet Psychiatry 2022;9(3):243-252. [doi: 10.1016/S2215-0366(21)00254-6] [Medline: 35183281]
60. Luk JW, Pruitt LD, Smolenski DJ, Tucker J, Workman DE, Belsher BE. From everyday life predictions to suicide prevention: clinical and ethical considerations in suicide predictive analytic tools. J Clin Psychol 2022;78(2):137-148. [doi: 10.1002/jclp.23202] [Medline: 34195998]
61. Van Calster B, Vickers AJ. Calibration of risk prediction models: impact on decision-analytic performance. Med Decis Making 2015;35(2):162-169. [doi: 10.1177/0272989X14547233] [Medline: 25155798]
62. Shah ND, Steyerberg EW, Kent DM. Big data and predictive analytics: recalibrating expectations. JAMA 2018;320(1):27-28. [doi: 10.1001/jama.2018.5602] [Medline: 29813156]
63. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J 2014;35(29):1925-1931 [FREE Full text] [doi: 10.1093/eurheartj/ehu207] [Medline: 24898551]
64. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of COVID-19: systematic review and critical appraisal. BMJ 2020;369:m1328 [FREE Full text] [doi: 10.1136/bmj.m1328] [Medline: 32265220]
65. Prozess icons. Flaticon. 2024 Dec 13. URL: https://www.flaticon.com/de/kostenlose-icons/prozess [accessed 2024-12-13]
66. Automatisierung icons. Flaticon. URL: https://www.flaticon.com/de/kostenlose-icons/automatisierung [accessed 2024-12-13]
67. Decision icons. Flaticon. URL: https://www.flaticon.com/free-icons/decision [accessed 2024-12-13]
68. Schicht icons. Flaticon. URL: https://www.flaticon.com/de/kostenlose-icons/schicht [accessed 2024-12-13]
69. Etwas beachten icons. Flaticon. URL: https://www.flaticon.com/de/kostenlose-icons/etwas-beachten [accessed 2024-12-13]
70. Algorithmus icons. Flaticon. URL: https://www.flaticon.com/de/kostenlose-icons/algorithmus [accessed 2024-12-13]
71. Richtigkeit icons. Flaticon. URL: https://www.flaticon.com/de/kostenlose-icons/richtigkeit [accessed 2024-12-13]
72. Eckert M, Efe Z, Guenthner L, Baldofski S, Kuehne K, Wundrack R, et al. Acceptability and feasibility of a messenger-based psychological chat counselling service for children and young adults ("krisenchat"): A cross-sectional study. Internet Interv 2022;27:100508 [FREE Full text] [doi: 10.1016/j.invent.2022.100508] [Medline: 35242589]
73. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-suicide severity rating scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011;168(12):1266-1277 [FREE Full text] [doi: 10.1176/appi.ajp.2011.10111704] [Medline: 22193671]
74. Reeves KW, Vasconez G, Weiss SJ. Characteristics of suicidal ideation: a systematic review. Arch Suicide Res 2022;26(4):1736-1756. [doi: 10.1080/13811118.2021.2022551] [Medline: 35040371]
75. Joiner T. Why People Die by Suicide. U.S: Harvard University Press; 2005.
76. O'Connor RC, Kirtley OJ. The integrated motivational-volitional model of suicidal behaviour. Philos Trans R Soc Lond B Biol Sci 2018;373(1754):20170268 [FREE Full text] [doi: 10.1098/rstb.2017.0268] [Medline: 30012735]
77. Chu C, Buchman-Schmitt JM, Stanley IH, Hom MA, Tucker RP, Hagan CR, et al. The interpersonal theory of suicide: a systematic review and meta-analysis of a decade of cross-national research. Psychol Bull 2017;143(12):1313-1345. [doi: 10.1037/bul0000123] [Medline: 29072480]
78. Klonsky ED, May AM. The three-step theory (3ST): a new theory of suicide rooted in the “Ideation-to-Action” framework. Int J Cogn Ther 2015;8(2):114-129. [doi: 10.1521/ijct.2015.8.2.114]
79. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024;385:e078378 [FREE Full text] [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]

80. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AICONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med 2020;26(9):1364-1374 [FREE Full text] [doi: 10.1038/s41591-020-1034-x] [Medline: 32908283]
81. TensorFlow. 2023. URL: https://zenodo.org/records/10126399 [accessed 2023-12-06]
82. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature 2020;585(7825):357-362 [FREE Full text] [doi: 10.1038/s41586-020-2649-2] [Medline: 32939066]
83. GitHub - pandas-dev/pandas at v2.1.3. URL: https://github.com/pandas-dev/pandas/tree/v2.1.3 [accessed 2024-12-11]
84. GitHub - keras-team/keras: Deep learning for humans. URL: https://github.com/keras-team/keras [accessed 2024-12-11]
85. Matplotlib ? Visualization with Python. URL: https://matplotlib.org/ [accessed 2024-07-31]
86. Waskom M. seaborn: statistical data visualization. J Open Source Softw 2021;6(60):3021. [doi: 10.21105/joss.03021]
87. Andersson JM. JimmyMAndersson/StatKit. URL: https://github.com/JimmyMAndersson/StatKit [accessed 2024-07-31]
88. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. arXiv. Posted online on June 5, 2018 2018.
89. krisenchat/p91-suicide-model: This is a transformer-based text classification model which aims to predict and explain session-propensities for suicide. GitHub. URL: https://github.com/krisenchat/p91-suicide-model [accessed 2024-07-30]
90. Conneau A, Khandelwal K, Goyal N, Chaudhary V, Wenzek G, Guzmán F, et al. Unsupervised cross-lingual representation learning at scale. arXiv. Posted online on April 8, 2020 2020.
91. Elkind E, Rothe J. Cooperative game theory. In: Rothe J, editor. Economics and Computation. Springer Texts in Business and Economics. Berlin: Springer; 2016.
92. Merrick L, Taly A. The explanation game: explaining machine learning models using Shapley values. In: Machine Learning and Knowledge Extraction. CD-MAKE 2020. Lecture Notes in Computer Science. Switzerland: Springer, Cham; 2020.
93. Wysocki O, Davies JK, Vigo M, Armstrong AC, Landers D, Lee R, et al. Assessing the communication gap between AI models and healthcare professionals: explainability, utility and trust in AI-driven clinical decision-making. Artif Intell 2023;316:103839. [doi: 10.1016/j.artint.2022.103839]
94. Key Issue 5: Transparency Obligations - EU AI Act. URL: https://www.euaiact.com/key-issue/5 [accessed 2024-04-27]
95. Sierra G, Andrade-Palos P, Bel-Enguix G, Osornio-Arteaga A, Cabrera-Mora A, García-Nieto L, et al. Suicide risk factors: a language analysis approach in social media. J Lang Soc Psychol 2021;41(3):312-330 [FREE Full text] [doi: 10.1177/0261927X211036171]
96. Lao C, Lane J, Suominen H. Analyzing suicide risk from linguistic features in social media: evaluation study. JMIR Form Res 2022;6(8):e35563 [FREE Full text] [doi: 10.2196/35563] [Medline: 36040781]
97. Homan S, Gabi M, Klee N, Bachmann S, Moser A, Duri' M, et al. Linguistic features of suicidal thoughts and behaviors: a systematic review. Clin Psychol Rev 2022;95:102161 [FREE Full text] [doi: 10.1016/j.cpr.2022.102161] [Medline: 35636131]
98. Warner B, Misra M. Understanding neural networks as statistical tools. Am Stat 1996 Nov;50(4):284-293 [FREE Full text] [doi: 10.1080/00031305.1996.10473554]
99. Zhang Z, Rousson V, Lee WC, Ferdynus C, Chen M, Qian X, written on behalf of AME Big-Data Clinical Trial Collaborative Group. Decision curve analysis: a technical note. Ann Transl Med 2018 Aug;6(15):308 [FREE Full text] [doi: 10.21037/atm.2018.07.02] [Medline: 30211196]
100. Aldhyani THH, Alsubari SN, Alshebami AS, Alkahtani H, Ahmed ZAT. Detecting and analyzing suicidal ideation on social media using deep learning and machine learning models. Int J Environ Res Public Health 2022 Oct 03;19(19):12635 [FREE Full text] [doi: 10.3390/ijerph191912635] [Medline: 36231935]
101. Douglass RW, Gartzke E, Lindsay JR, Gannon JA, Scherer TL. What is escalation? Measuring crisis dynamics in international relations with human and LLM generated event data. arXiv. Posted online on January 18, 2024 2024.
102. Stern W, Goh SJ, Nur N, Aragon PJ, Mercer T, Bhattacharyya S, et al. Natural Language Explanations for Suicide Risk Classification Using Large Language Models. ceur-ws.org 2024;3649:1-10.

Abbreviations

AI: artificial intelligence
ASE: advanced suicidal engagement
AUC: area under the curve
BS: Brier score
C-SSRS: Columbia-Suicide Severity Rating Scale
CONSORT-AI: Consolidated Standards of Reporting Trials for Artificial Intelligence
DCA: decision curve analysis
LSTM: long short-term memory
ML: machine learning
NB: net benefit
ROC: receiver operating characteristic

SHAP: Shapley Additive Explanation
SI: suicidal ideation
SIB: suicidal ideation and behavior
tf-idf: term frequency-inverse document frequency
T-MLP: transformer-multilayer perceptron
W2V-MLP: word2vector-multilayer perceptron

Edited by A Mavragani; submitted 01.07.24; peer-reviewed by EU Kucuksille, GM Bairy, S Mao; comments to author 18.07.24; revised version received 30.08.24; accepted 07.11.24; published 29.01.25.

Please cite as:
Thomas J, Lucht A, Segler J, Wundrack R, Miché M, Lieb R, Kuchinke L, Meinlschmidt G
An Explainable Artificial Intelligence Text Classifier for Suicidality Prediction in Youth Crisis Text Line Users: Development and Validation Study
JMIR Public Health Surveill 2025;11:e63809
URL: https://publichealth.jmir.org/2025/1/e63809
doi:10.2196/63809
PMID:

Epidemiological Features, Clinical Symptoms, and Environmental Risk Factors for Notifiable Japanese Encephalitis in Taiwan From 2008 to 2020: Retrospective Study

Fu-Huang Lin[1], PhD; Yu-Ching Chou[1], PhD; Chi-Jeng Hsieh[2], PhD; Chia-Peng Yu[1], PhD

[1]School of Public Health, National Defense Medical Center, Taipei City, Taiwan

[2]Department of Health Care Administration, Asia Eastern University of Science and Technology, New Taipei City, Taiwan

Corresponding Author:
Chia-Peng Yu, PhD
School of Public Health, National Defense Medical Center, Taipei City, Taiwan

Abstract

Background: Japanese encephalitis (JE) is a zoonotic parasitic disease caused by the Japanese encephalitis virus (JEV), and may cause fever, nausea, headache, or meningitis. It is currently unclear whether the epidemiological characteristics of the JEV have been affected by the extreme climatic conditions that have been observed in recent years.

Objective: This study aimed to examine the epidemiological characteristics, trends, and potential risk factors of JE in Taiwan from 2008 to 2020. Specifically, the study focused on gender, age, season, residential area, clinical manifestations, high-risk areas, and the impact of environmental and climate factors.

Methods: This study reviewed publicly available annual summary data on reported JE cases in the Taiwan Centers for Diseases Control between 2008 and 2020.

Results: This study collected 309 confirmed domestic patients and 4 patients with imported JE. There was an increasing trend in the incidence of JE, 0.69‐1.57 cases per 1,000,000 people, peaking in 2018. Case fatality rate was 7.7% (24/313). Comparing sex, age, season, and place of residence, the incidence rate was highest in males, 40‐ to 59-year-old patients, summer, and the Eastern region, with 1.89, 3.27, 1.25, and 12.2 cases per million people, respectively. The average coverage rate of the JE vaccine for children in Taiwan is 94.9%. Additionally, the major clinical manifestations of the cases included fever, unconsciousness, headache, stiff necks, psychological symptoms, vomiting, and meningitis. The major occurrence places of JE included paddy fields, pig farms, pigeon farms, poultry farms, and ponds. For air pollution factors, linear regression analysis showed that SO_2 (ppb) concentration was positively associated with JE cases (β=2.184, *P*=.02), but O_3 (ppb) concentration was negatively associated with them (β=−0.157, *P*=.01). For climate factors, relative humidity (%) was positively associated with JE cases (β=.380, *P*=.02).

Conclusions: This study is the first to report confirmed cases of JE from the surveillance data of the Taiwan Centers for Diseases Control between 2008 and 2020. It identified residence, season, and age as risk factors for JE in Taiwan. Air pollution and climatic factors also influenced the rise in JE cases. This study confirmed that JE remains a prevalent infectious disease in Taiwan, with its epidemic gradually increasing in severity. These findings empower clinicians and health care providers to make informed decisions, guiding their care and resource allocation for patients with JE, a disease that significantly impacts the health and well-being of the Taiwanese population.

(JMIR Public Health Surveill 2025;11:e63053) doi:10.2196/63053

KEYWORDS

epidemiology; Japanese encephalitis virus; domestic; environmental factor; retrospective study

Introduction

The Japanese encephalitis virus (JEV) is a significant public health concern in Asia, causing an estimated 68,000 clinical cases annually, primarily in East Asia and Southeast Asia [1]. Although symptomatic Japanese encephalitis (JE) occurs infrequently, it has a serious case fatality rate, with up to 30% of encephalitis cases resulting in death. Encephalitis survivors have a high likelihood of developing permanent neurological or psychiatric sequelae, with an estimated occurrence rate of 30%-50% of cases. The endemic transmission of JEV poses a health threat to over 3 billion people in a substantial portion of the World Health Organization Southeast Asia and Western Pacific regions [1]. The disease currently lacks specific antiviral treatment, necessitating supportive care measures like managing symptoms and providing respiratory and nutritional support to help patients overcome the infection [2]. Vaccination stands as the effective strategy to prevent the disease, with World Health Organization recommending the integration of JE vaccination into national immunization programs in all regions where the disease poses a significant public health threat [1,3].

JE is a mosquito-borne viral infection caused by the JEV, a member of the *Flaviviridae* family closely related to dengue, yellow fever, and West Nile viruses [4]. The primary means of transmission of JEV from animals to humans is through the bite of infected mosquitoes, particularly Culex species, with *Culex tritaeniorhynchus* (*Cx. tritaeniorhynchus*) being of particular importance. Pigs and ardeid birds serve as the main amplifying hosts for the virus [5]. The clinical presentation of JE in humans varies widely, encompassing a spectrum ranging from mild or asymptomatic infection to severe encephalitis. This severe form of the disease is characterized by significant mortality, with a high death rate among infected individuals. Additionally, survivors of severe JE often experience sequelae within the central nervous system [6].

In recent years, air pollution has been identified as a significant public health concern in China, with an estimated 1 million deaths annually attributed to its effects. This has underscored the urgency for the implementation of effective prevention strategies. However, future climate change is likely to increase the frequency and duration of severe weather conditions, thereby promoting exposure to air pollution and exacerbating the impacts on human health [7]. A previous study indicated that tropical and subtropical populations of *Cx. tritaeniorhynchus* exhibited higher genetic diversity and that the effects of environmental factors (climatic conditions) provide significant advantages for the establishment and expansion of *Cx. tritaeniorhynchus* [8]. An understanding of the impact of environmental factors on JE can provide valuable insights into its morbidity and mortality, which remain unclear.

JE poses a significant public health threat, particularly in Asian countries, especially in South Asia, Southeast Asia, and East Asia [9]. Several countries in these regions have successfully implemented various intervention measures to reduce JE morbidity, including early diagnosis, prompt treatment, national immunization programs, and effective vector control strategies [10,11]. Despite these efforts, over 3 billion people still reside in JE-endemic countries, with an estimated 67,900 cases occurring annually [8]. The severity of JE is evident in its high mortality rate, with approximately 20% - 30% of cases resulting in fatalities. Additionally, 30% - 50% of JE survivors experience significant long-term neurological sequelae [9,12]. Previous studies have also highlighted the ongoing public health concern posed by JE in Taiwan [13].

Nestled at 23°4' north latitude and 121°0' east longitude, Taiwan enjoys a subtropical climate with average monthly temperatures ranging from a comfortable 16 to 29 °C. Average monthly relative humidity hovers between 75% and 90%. Despite its status as a developed nation with a per capita gross domestic product of US$35,244 [14], Taiwan continues to grapple with the challenges posed by JE. Despite occasional imported cases of JE in recent years [15], local infections persist in Taiwan, indicating limitations in the effectiveness of preventive measures in controlling or eradicating the disease. The number of JE cases in Taiwan ranges from sporadic confirmed cases to cluster epidemics, with the middle-aged population exhibiting the highest rate of JEV. Addressing the disease burden caused by JEV is imperative, and effective public health measures should be adopted in Taiwan. Although some studies provided epidemiological data on JE in Taiwan [16], their study did not delve into the clinical symptoms of JE patients over the past decade nor investigate the potential association between climate and air pollution factors and the number of cases [17]. The objective of this study is to conduct a comprehensive investigation of the epidemiological, clinical, and environmental characteristics of JE in Taiwan. To this end, population surveillance data gathered over a 13-year period will be used.

Methods

Study Design and Population

Study Area

Taiwan is an island in East Asia located between 21°45’ N and 25°56’ N. The Tropic of Cancer (23.5° N) passes through Chiayi City, situated in south-central Taiwan, dividing the entire island into 2 climate zones. Taiwan has a total land area of 35,980 km² and a population of approximately 23 million people, resulting in an average population density of 635 individuals per km². The northern part of Taiwan belongs to the subtropical climate zone, whereas the southern part belongs to the tropical climate zone. Consequently, the weather in Taiwan is relatively warm, with high humidity occurring throughout the year [18,19].

Case Definitions

A clinical case was defined as an individual of any age experiencing an acute onset of fever accompanied by a change in mental status and new onset of seizures (excluding simple febrile seizures) at any time of the year [9,13]. A confirmed case was defined as a clinical case with positive laboratory test results, including the presence of IgM antibodies specific to the JE virus in a single sample of cerebrospinal fluid (CSF) or serum; a 4-fold increase in IgG antibodies; the detection of JE virus antigens in tissue via immunohistochemistry; or the detection of the JE virus genome in serum, plasma, blood, CSF, or tissue samples. Additionally, cases meeting the clinical case definition and epidemiologically linked to a confirmed case were also considered confirmed [20,21].

Data Collection

Surveillance for JE Infection

JE has been categorized as a notifiable infectious disease since 1955. Physicians are required to report all cases that meet the case definition of JE, collect samples, and send them to the Taiwan Centers for Disease Control (TCDC) within one week of the case being reported for examination [20]. We collected data from all JE-confirmed cases reported to the TCDC from January 2008 to December 2020. The reported information included patient age, sex, area of residence, geographic location, seasonal variation of exposure, and date of JE onset.

Laboratory Tests

In 1998, the TCDC developed an E/M-specific capture IgM and IgG enzyme-linked immunosorbent assay (E/M-specific IgM/IgG ELISA) for JE and dengue fever (DF) [13,22,23].

Hemagglutination inhibition (HI) and E/M-specific IgM/IgG ELISA are both employed as screening tools for detecting antibodies against JE and DF [24-26]. This study employed a previously described differential testing algorithm [26]. Briefly, positive samples on the JE ELISA underwent further testing using DF ELISA. The combined results from both assays informed the final interpretation. Positive JE ELISA and negative DF ELISA indicated the presence of JEV antibodies only, while positive results in both tests suggested a false-positive JE result. Long-term evaluation demonstrated high sensitivity and specificity of the E/M-specific IgM/IgG ELISA, effectively differentiating JEV infections from dengue virus infections [26]. Since 2001, real-time polymerase chain reaction (PCR) analysis of acute-phase serum samples collected within 7 days of symptom onset and CSF from individuals diagnosed with JE has been employed for diagnostic purposes [27]. Nevertheless, enzyme-linked immunosorbent assay (ELISA) remains the primary diagnostic testing method, with the E/M-specific IgM/IgG ELISA being the most widely used ELISA method.

The JE laboratory has confirmed that clinical cases meeting one of the following specific laboratory criteria should be defined as confirmed cases of JE: (1) an HI titer of the convalescent serum of 1:160 and at least a 4-fold increase between the HI titers of convalescent and acute sera [24,25]; (2) an HI titer of 1:320 derived from a single serum sample [24,25]; (3) an IgM-positive serum sample, as determined by an ELISA test, or a fourfold increase in IgG levels between paired serum samples [25]; (4) a sample exhibiting a positive real-time PCR [25]; (5) a sample that has been found to be positive for indirect immunofluorescence antibody staining following the isolation of the virus from CSF [25,28]; (6) virus isolation can be accomplished through cell culture techniques using either the mosquito C6/36 cell line or plaque assays with the BHK-21 cell line [26]; and (7) an alternative approach is the extraction of viral RNA from JEV-infected culture medium using the QIAamp Viral RNA Mini Kit (Qiagen, Germantown, MD) [15].

Surveillance of Environmental Factors

The dataset comprised monthly air pollutant data from 2008 to 2020, obtained from the air quality monitoring network of the Environmental Protection Agency. The pollutants analyzed included total suspended particulates, particulate matter (PM) 2.5 and 10 (PM2.5 and PM10), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), carbon monoxide (CO), ozone (O_3), dust fall volume, and lead concentration [29]. Furthermore, monthly weather data from the Meteorological Bureau of the Ministry of Communications for the same period were analyzed. The meteorological variables included temperature, rainfall, relative humidity, atmospheric pressure, the number of days with precipitation, and the duration of sunshine [30]. Statistical analyses and correlation tests were employed to assess the temporal and spatial trends in air pollutants and meteorological factors, as well as their potential association with JE case numbers. The data collected for this study were stored in the supplementary file.

Statistical Analysis

This study presents a retrospective historical analysis of all cases of JE diagnosed in Japan since 2008, including both domestic and imported cases. The number of individuals diagnosed with JE from 2008 to 2020 was confirmed, and the epidemiological characteristics of these cases were examined. These characteristics included sex (male and female), age groups (<20, 20‐39, 40‐59, and ≥60 years old), time of diagnosis (spring, summer, fall, and winter), and area of residence (northern, central, southern, and eastern Taiwan). Disparities and outcomes were analyzed. Our analysis then focused on the variables of sex, age, time of diagnosis, changes in living area, trends, and related outcomes for cases of JE from 2011 to 2020. Descriptive data are presented as means and summaries where appropriate. Categorical variables were compared using chi-squared tests. Odds ratios were calculated using logistic regression, with 95% confidence intervals estimated using parameter estimation. All statistical analyses were performed using SPSS (IBM SPSS version 21; Asia Analytics Taiwan Ltd). All statistical tests were 2-tailed, with an α value of .05. A *P* value less than .05 was considered statistically significant.

Ethical Considerations

The ethical considerations of this study were consistent with those of Holland et al [31] in *JAMA Psychiatry* and Pan et al [32] in *BMC Infectious Diseases*. All case data are accessible via the internet, with information regarding JE cases available on the TCDC website. This study was deemed exempt from ethical approval according to the Communicable Disease Control Act of Taiwan [33] because the datasets used were deidentified and the study did not involve any of the datasets in a way that could potentially lead to the identification of individuals. Along with the absence of personal identifiers and the absence of any potential harm to individuals, the study was conducted in accordance with institutional research guidelines to ensure scientific rigor and ethical standards.

Results

Demographic Data

The figures and tables selected for analysis in this study are representative of an examination of demographic characteristics (incidence rates), patient clinical symptoms, environmental factors, air pollution, and climate factors on JE in Taiwan over the recent 13-year period. The flowchart of this study is shown in Figure 1. There were 313 confirmed cases of local and imported infection, consisting of 309 (98.7%) confirmed local cases, 195 (62.3%) male patients, 182 (58.1%) patients aged 40‐59 years, 245 (78.3%) cases in summer, and 145 (46.3%) cases in the southern region. There were 4 (1.3%) confirmed cases of imported infection (Table 1). There was no statistically significant difference between the number of confirmed cases in 2008 and 2020 and the relationship between different identities (local or imported), sex, age group, season, and place of residence (all *P*>.05) (Table 1). There was a statistically significant difference between the number of confirmed cases in 2008 and 2020 and the relationship between different seasons and place of residence (all *P*<.05) (Table 1). The main environmental factors associated with JE were paddy fields

(160/582, 27.5%), pig farms (158/582, 27.1%), pigeon farms (125/582, 21.5%), poultry farms (65/582, 11.2%), ponds (29/582, 5%), and domestic pets (18/582, 3.1%), as indicated in Table 2.

Figure 1. Flowchart of the study sample selection from the Taiwan Centers for Disease Control Database (CDC) in Taiwan from January 2008 to December 2020.

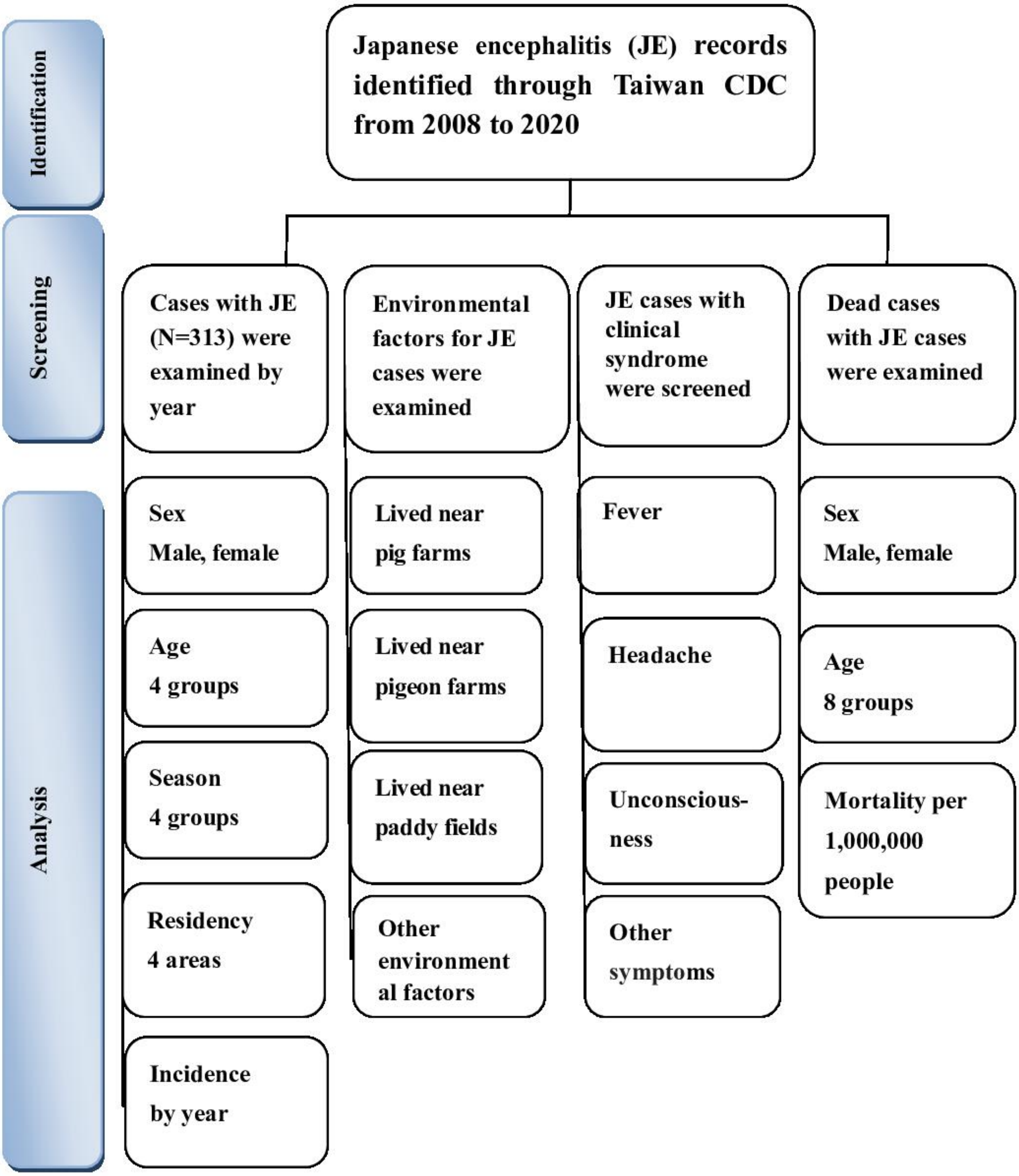


Table . Epidemiological features of Japanese encephalitis cases in Taiwan from 2008 to 2020.

Variable	Total	Year													*P* value
Domestic		2008 (n=17)	2009 (n=18)	2010 (n=33)	2011 (n=22)	2012 (n=32)	2013 (n=16)	2014 (n=18)	2015 (n=30)	2016 (n=23)	2017 (n=25)	2018 (n=37)	2019 (n=21)	2020 (n=21)	
Individual, n															.12
Domestic	309	17	17	32	22	32	16	18	30	23	25	37	19	21	
Imported	4	0	1	1	0	0	0	0	0	0	0	0	2	0	
Sex, n															.90
Male	195	12	13	22	14	17	12	12	19	12	15	20	13	14	
Female	118	5	5	11	8	15	4	6	11	11	10	17	8	7	
Age group, n															.26
<20	10	3	1	1	1	0	1	0	0	1	0	1	0	1	
20 - 39	68	5	3	11	8	9	5	3	7	2	4	6	3	2	
40 - 59	182	9	9	16	9	17	9	10	20	16	14	24	14	15	
≥60	53	0	5	5	4	6	1	5	3	4	7	6	4	3	
Season, n															<.001
Spring	40	1	2	2	0	10	2	0	0	2	3	10	7	1	
Summer	245	15	14	29	20	19	14	15	28	14	19	25	13	20	
Fall	28	1	2	2	2	3	0	3	2	7	3	2	1	0	
Winter	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Residency, n															.008
Northern	78	2	5	10	9	4	5	2	7	4	8	10	5	7	
Central	65	4	2	6	5	5	3	5	4	5	5	9	5	7	
Southern	145	9	4	14	4	21	8	11	18	11	10	17	11	7	
Eastern	25	2	7	3	4	2	0	0	1	3	2	1	0	0	

Table . Environmental factors for Japanese encephalitis cases in Taiwan between 2008 and 2020.

Variable (lived nearby)		Year												
	Total, n	2008, n	2009, n	2010, n	2011, n	2012, n	2013, n	2014, n	2015, n	2016, n	2017, n	2018, n	2019, n	2020, n
Pig farms	158	4	6	8	10	23	10	11	13	11	13	25	13	11
Pigeon farms	125	7	4	14	2	17	5	7	9	11	10	19	8	12
Paddy fields	160	9	6	17	9	16	9	11	14	11	8	26	14	10
Egret nests	6	1	—[a]	1	—	4	—	—	—	—	—	—	—	—
Henhouses	2	1	1	—	—	—	—	—	—	—	—	—	—	—
Pond	29	2	2	3	1	7	3	2	3	1	1	3	1	—
Clam breeding farms	1	—	1	—	—	—	—	—	—	—	—	—	—	—
Poultry farms	65	—	—	9	6	11	3	—	5	6	5	12	4	4
Goat farms	4	—	—	2	—	1	—	—	—	—	1	—	—	—
Domestic pets	18	—	—	1	5	5	3	—	2	1	—	—	—	1
Farmlands	6	—	—	—	3	3	—	—	—	—	—	—	—	—
Orchards	8	—	—	—	3	5	—	—	—	—	—	—	—	—

[a] Not applicable

Clinical Symptoms

The main clinical symptoms of JE were fever (280/866, 32.3%), unconsciousness (162/866, 18.7%), headache (136/866, 15.7%), stiff neck (63/866, 7.3%), psychiatric symptoms (59/866, 6.8%), and vomiting (49/866, 5.7%), as shown in Table 3.

There were 24 deaths from JE, of which 15 (62.5%) were male and 9 (37.5%) were aged 50 - 60 years. The mortality rate (per 1,000,000 people) ranged from 0.04 to 0.17 between 2008 and 2020 (Table 4).

Table . Clinical symptoms of Japanese encephalitis cases in Taiwan from 2008 to 2020.

Variable		Year												
	Total	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Fever, n	280	16	18	30	18	26	15	14	28	21	21	33	20	20
Headache, n	136	13	5	19	13	15	8	7	9	9	8	14	6	10
Stiff necks, n	63	1	6	3	6	6	3	5	3	5	7	11	3	4
Muscle cramps, n	35	7	2	5	—[a]	—	3	2	6	2	3	2	—	3
Vomiting, n	49	4	2	11	4	5	3	1	2	1	3	8	4	1
Speaking difficulty, n	9	—	2	—	3	4	—	—	—	—	—	—	—	—
Coma, n	30	4	5	12	7	2	—	—	—	—	—	—	—	—
Unconscious, n	162	5	6	21	11	9	11	9	13	17	14	20	12	14
Encephalitis, n	3	—	—	—	1	2	—	—	—	—	—	—	—	—
Paralysis, n	1	—	—	—	1	—	—	—	—	—	—	—	—	—
Psychological symptoms, n	59	—	—	—	—	9	—	5	11	5	5	15	6	3
Paralyzed limbs, n	3	—	—	—	—	3	—	—	—	—	—	—	—	—
Diarrhea, n	2	—	—	—	—	2	—	—	—	—	—	—	—	—
Muscular ache, n	2	—	—	—	—	2	—	—	—	—	—	—	—	—
Encephalitic stimulation symptoms, n	10	—	—	—	—	—	—	1	3	2	1	2	—	1
Dystonia, n	16	—	—	—	—	—	—	—	4	4	2	1	2	3
Meningitis symptoms, n	6	—	—	—	—	—	—	—	—	—	—	—	—	3

Variable	Year													
	Total	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Aseptic meningitis, n	—	—	—	—	—	—	—	—	—	—	—	—	—	3

[a] Not applicable.

Table . The mortality rate per 1,000,000 people and the characteristics of deceased cases with Japanese encephalitis from 2008 to 2020[a] in Taiwan.

Variable	2008 (n=2)	2009 (n=2)	2010 (n=2)	2011 (n=1)	2013 (n=1)	2014 (n=3)	2015 (n=1)	2016 (n=4)	2017 (n=3)	2018 (n=2)	2019 (n=3)
Mortality rate per 1,000,000 people	0.09	0.09	0.09	0.04	0.04	0.13	0.04	0.17	0.13	0.08	0.13
Sex, n											
Male	1	2	0	1	0	3	0	3	2	2	1
Female	1	0	2	0	1	0	1	1	1	0	2
Age group, n											
<20	—[b]	—	—	—	—	—	—	—	—	—	—
20 - 39	—	—	1	—	—	1	—	—	—	—	—
40 - 59	1	1	1	—	1	2	1	3	1	—	3
≥60	1	1	—	1	—	—	—	1	2	2	—

[a] There were no reported deaths due to Japanese encephalitis in 2012 and 2020.

[b] Not applicable

Regression Analysis

Linear regression analysis showed that air pollution factors were associated with JE cases: SO_2 concentration (β=2.184, SE 0.922; P=.02) and O_3 concentration (β=−0.157, SE 0.063; P=.01) (Table 5). Linear regression analysis showed that climatic factors were associated with JE cases (R^2=0.344, $F_{6,137}$=11.980, P<.001) (Table 6).

The incidence of confirmed cases of JE per million people was 0.69 - 1.57 (Figure 2A); 1.02 - 1.89 in males (Figure 2B); 1.25 - 3.27 in the 40 - 59 years age group (Figure 2C); 0.55 - 1.25 in summer (Figure 2D); and 0 - 12.2 in the northern region (Figure 2E), during 2008 - 2020. The national coverage of JE vaccine in Taiwan is shown in Figure 3.

Table . Association between air pollution factors and Japanese encephalitis cases examined through multiple linear regression analysis. R^2=0.417; *F* test (*df*) –> 6.824 (9, 86); *P* <.001; n=96.

Variables	Nonstandardization coefficient		*P* value
	β value	Standard error	
TSP[a] (μg/m^3)	−0.028	0.055	.61
PM[b] 2.5 (μg/m^3)	0.022	0.177	.90
PM 10 (μg/m^3)	0.007	0.124	.95
SO_2 (ppb)	2.184	0.922	.02
CO (ppm)	−17.244	19.092	.37
NO_2 (ppb)	−0.342	0.564	.55
O_3 (ppb)	−0.157	0.063	.01
Dustfall volume (tonne/km^2/month)	0.029	0.509	.96
Lead (μg/m^3)	43.751	72.026	.55

[a] TSP: total suspended particulate.

[b] PM: particulate matter.

Table . Relationship between climate factors and Japanese encephalitis cases analyzed using multiple linear regression analysis. R^2=0.344; *F* test (*df*) –> 11.980 (6, 137); *P*<.001; n=144.

Variables	Nonstandardization coefficient		*P* value
	β value	Standard error	
Temperature (℃)	−0.025	0.192	.90
Precipitation (mm)	−0.004	0.004	.30
Relative humidity (%)	0.380	0.163	.02
Mean pressure (hPa)	−0.247	0.155	.12
Number of days with precipitation ≥0.1 mm (day)	−0.017	0.199	.93
Sunshine duration (h)	0.029	0.016	.08

Figure 2. Incidence of confirmed JEV among patients in Taiwan according to (**A**) population, (**B**) gender, (**C**) age, (**D**) season, and (**E**) region of residence by year from 2008 to 2020.

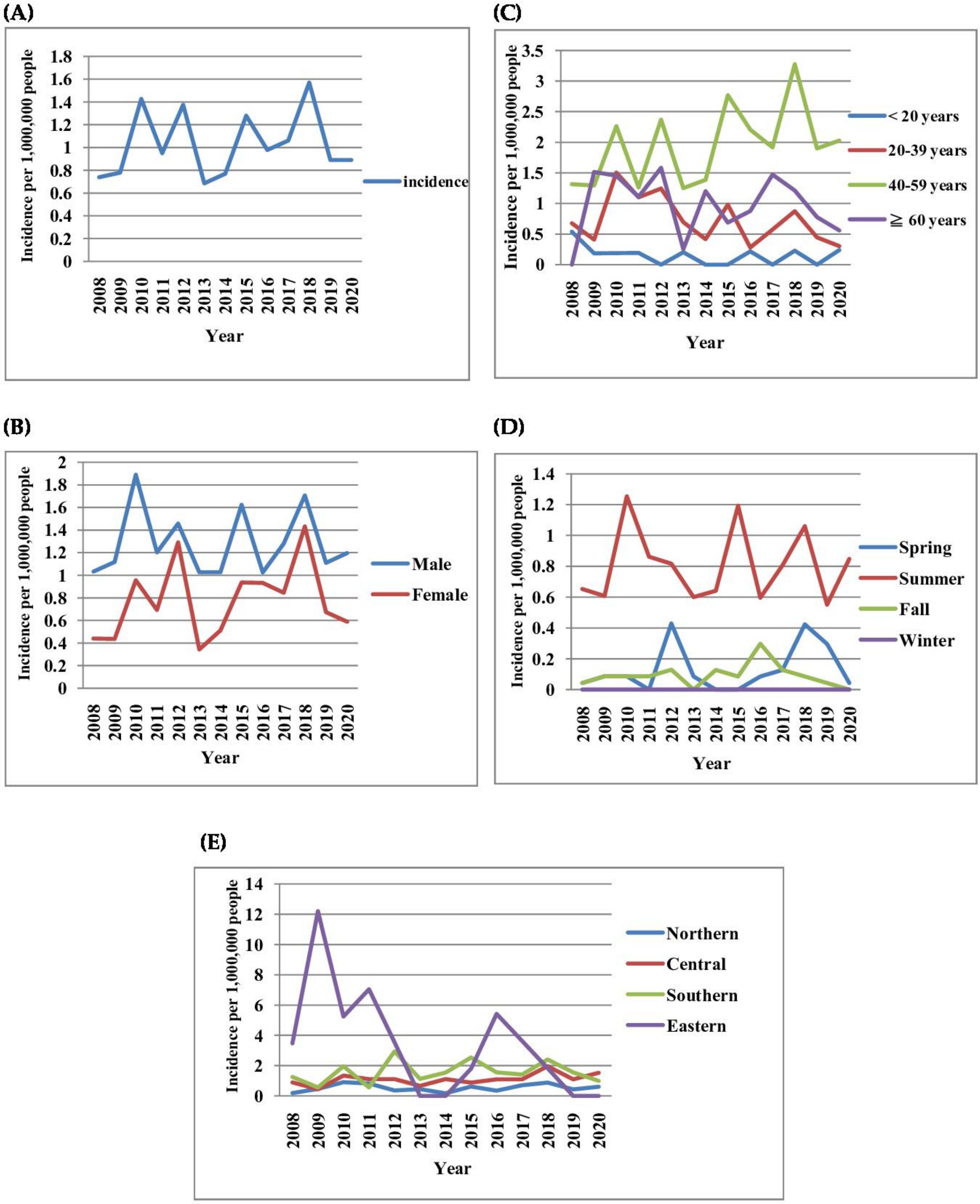
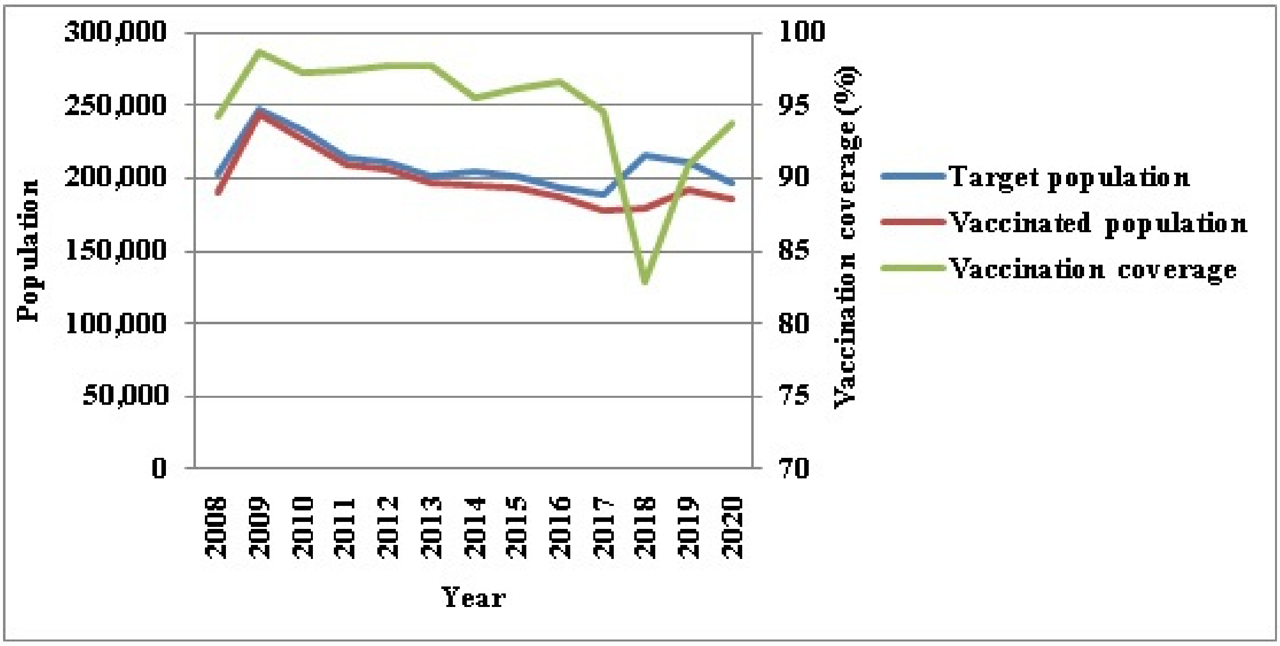


Figure 3. National target population, vaccinated population, and vaccination coverage for Japanese encephalitis in Taiwan from 2008 to 2020.



Discussion

Principal Findings

JE is the most prevalent cause of viral encephalitis in Asia, with transmission occurring through mosquito bites. It is endemic in most of South and Southeast Asia, but the number of cases can vary significantly between regions [34]. The study yielded several significant findings. With regard to epidemiological characteristics, the data indicated that men, middle-aged individuals, those residing in summer, and residents of southern Taiwan are at greater risk. The primary clinical symptoms observed were fever, headache, and unconsciousness. An increase in the concentration of the air pollutant SO_2 and the relative humidity percentage, when considered as climatic factors, was found to correlate with an increase in the number of JE cases.

Epidemiology examines the epidemiological characteristics associated with disease transmission, presentation, and outcome. This discipline has consistently been influenced by policy interventions and disease prevention measures [35,36]. From 2008 to 2020, a total of 313 cases of JE were confirmed in Taiwan. Of these cases, 1.3% were imported, while 98.7% were indigenous, indicating that JE is endemic in Taiwan. The high incidence rate among men is consistent with previous studies [37-39]. It can be postulated that the underlying cause may be differences in personal hygiene habits. Furthermore, most cases occur in individuals aged 40-59 years. Patients in this age group exhibited a higher incidence compared to those younger than 20 years (1.94 vs 0.73), which is consistent with previous studies [13,40,41]. The government of Taiwan can use the similarities and differences in epidemiological characteristics as a foundation for developing and implementing its epidemic prevention policy or strategy.

Among all epidemiological characteristics, seasonal variation is the most susceptible to the distribution of all confirmed cases. In essence, seasonal variation is sufficient to affect the significant increase or decrease of confirmed cases, thereby indicating that the disease in South Asia, Southeast Asian countries, or even Taiwan is endemic [42]. The disease in question typically manifests during the summer months in Taiwan, as evidenced by the findings of previous studies [43-46]. According to the Central Weather Administration of Taiwan, the average summer temperature in Taiwan is 26.4 °C [47]. Consequently, the findings indicate that temperature is a pivotal factor influencing the vector competence and survival of infected mosquitoes. Moreover, southern Taiwan is situated in rural and suburban areas. A previous study indicated that in rural and suburban areas, traditional rice farming and intensive pig breeding provide an ideal environment for both mosquito development and the transmission of JEV among humans. Our findings align with those of previous studies [48-50]. In other words, there are always traces of the JEV in Taiwan that present a threat to public health and increase the clinical medical burden. At this juncture, it is imperative to address these significant challenges head-on, employing strategies to prevent and control them. It is imperative that the Taiwanese government's health department proactively propose local prevention strategies and implement effective surveillance operations to control the epidemic and reduce the number of cases, thereby eliminating the health threat.

JEV is spread by mosquitoes within an enzootic cycle that includes reservoirs among pigs and birds, with humans serving as unintentional dead-end hosts. In the past decade, both serological and genetic evidence from previous studies suggest the presence of JEV in the local fauna [51]. As in the previous study, workplaces of cases in Taiwan are commonly found in pig farms, pigeon farms, and paddy fields [52,53]. This study recommends that Taiwan's government health departments should increase publicity about measures to prevent JE by raising pigs and pigeons in rural areas.

A previous study demonstrated that while the impacts of the COVID-19 lockdown on various sectors, including the economy, research, travel, education, and sports, were clear, the effects on the occurrence of vector-borne diseases (VBDs) were less apparent [54]. The number of cases reported was compared with the number of cases predicted for each year from 2015 to 2020 for the VBDs (JE) under study to infer whether the COVID-19 lockdown had any impact on their prevalence in Taiwan. The actual number of cases was 21, both before and after the lockdown (2019 and 2020). The predicted number of cases for 2020, based on the trend of the previous 5 years (2015 - 2019), was also close to the actual number of cases. This suggests that the JEV has been consistently present in the local area of Taiwan for a long time. During the COVID-19 period, the number of JE cases did not fluctuate significantly. The analysis suggests that the lockdown did not have a significant impact on the incidence of VBDs such as JE.

In the majority of cases, patients infected with JE do not exhibit any clinical symptoms [55,56]. Consequently, the clinical symptoms observed in JE cases in Taiwan are consistent with those previously documented in other studies [57,58]. It is challenging to differentiate between JEV and other pathogens that can cause encephalitis and meningitis, including enterovirus, mumps virus, mycoplasma, herpes virus, and other viruses. This necessitates the use of viral culture or serological examination [59]. Therefore, this study proposes that the current understanding of JE serum complex flavivirus cross-reactivity, which may result in variable clinical outcomes, could inform future preventive and therapeutic interventions.

During the investigation period of this study, Taiwan exhibited a consistently low incidence rate of JE and its associated mortality from 2008 to 2020. The average incidence and mortality rates were 1.03 cases per 1,000,000 people and 0.09 cases per 1,000,000 people, respectively. A total of 10 cases of JE were reported in children. Consequently, the study indicates that the JE vaccination offered a moderate level of protection among children in JE-endemic Taiwan, a finding consistent with previous studies [60]. The implementation of the JE immunization program has been a pivotal factor in controlling the spread of JE. It is necessary to sustain a high vaccination coverage rate for JE and reinforce the disease surveillance system to ensure the efficacious control and eventual eradication of JE [61].

The prospective effects of climate change on public health are a rapidly expanding field of investigation. This encompasses the immediate consequences of more intense heat waves and declining food security, in addition to the indirect influences on the prevalence of infectious diseases [62-66]. Research has demonstrated that the proximity of collection sites to human dwellings (adjusted odds ratio [AOR] 2.02, *P*=.009) and a relative humidity exceeding 80% (AOR 2.40, *P*=.001) are significant independent risk factors for the transmission of JE viruses [67]. This finding is consistent with the present study. The results of this study indicate a positive correlation between relative humidity and JE cases (β=0.380, *P*=.02) through multiple linear regression analysis. The findings of this study indicate that climate change is a significant factor influencing the incidence of infectious diseases. It is therefore recommended that public health and epidemic prevention experts give this issue their attention, with the implementation of early planning measures to reduce the risk of disease.

Air pollution represents one of the most significant challenges of our era, affecting not only climate change but also public and individual health through increased morbidity and mortality [68]. A multitude of pollutants are significant contributors to human disease [69]. A previous study indicated that a comparison of the combined anthropogenic and environmental risk factors of major mosquito-borne diseases, specifically JE in Thailand, revealed that higher SO_2 surface concentrations were negatively associated with disease case counts [70]. However, the results of the linear regression analysis indicated that there was a positive association between SO_2 concentration and JE (β=2.184, *P*=.02). The concentration of O_3 was found to be negatively correlated with JE (β=−0.157, *P*=.01). To the best of our knowledge, this is the first study to demonstrate that the number of cases of JE increased as the concentration of the air pollutant SO_2 increased, and that there was a positive correlation. The study suggests that the high concentration of SO_2 in summer, which is a season prone to JE, may be a contributing factor. It was hypothesized that there was a correlation between these air pollutants and the disease. Consequently, this study proposes that Taiwan’s official policy should implement a rigorous monitoring program to assess local climate factors and air pollution concentration changes. In the event of minor fluctuations or significant fluctuations in environmental data, it is imperative that the media and the public be promptly informed. This will enable individuals to take prompt action to mitigate environmental pressures and the threat to public health.

Limitations

One limitation of the present study is that the data provided by the Taiwan National Infectious Disease Statistics System lack information about the genotypes or strains of the JEV isolated. Consequently, the type of JEV strain that spread to Taiwan and the affinity between virus strains in Taiwan and other countries were not analyzed in this study. Nevertheless, the advantage of this study was the access to the diverse data provided by Taiwan’s public sector on its web-based platform (including the initial version of the platform) and the evaluation of the impact of the COVID-19 pandemic on the epidemiological features of typhoid and paratyphoid. The information from existing public network platforms in Taiwan is both timely and accurate.

Conclusions

This study is the first to report on the epidemiological characteristics, clinical symptomatology, climatic factors, and air pollutants associated with JE cases in Taiwan between 2008 and 2020. In 2018, the incidence rate of JE was the highest (1.57 per million people). In recent years, the risk of local cases has increased at a rapid pace, resulting in a significant burden of disease, public health challenges, and epidemic prevention. The distribution of JE in Taiwan is regional in nature. The risk of JE among patients residing in different areas increases with age. The clinical manifestations of JE are variable, and the primary areas of endemicity are pig farms, pigeon farms, and paddy fields. The JE vaccine coverage rate in Taiwan is high, with an average of 94.9%. The concentration of certain environmental factors (SO_2, O_3) may influence the increase or decrease in JE cases. The identification of crucial data will facilitate future surveillance and research activities in Taiwan.

Conflicts of Interest

None declared.

References

1. Key facts of japanese encephalitis. WHO. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/japanese-encephalitis [accessed 2025-01-10]
2. Japanese encephalitis: symptoms, diagnosis, and treatment. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/japanese-encephalitis/symptoms-diagnosis-treatment/?CDC_AAref_Val=https://www.cdc.gov/japaneseencephalitis/symptoms/index.html [accessed 2025-01-10]
3. Wilder-Smith A, Halstead SB. Japanese encephalitis: update on vaccines and vaccine recommendations. Curr Opin Infect Dis 2010 Oct;23(5):426-431. [doi: 10.1097/QCO.0b013e32833c1d01] [Medline: 20581670]
4. Leinikki P. Japanese encephalitis. In: Zuckerman AJ, Banatvala JE, Pattison JR, editors. Principles and Practice of Clinical Virology, 5th edition: Wiley; 2004:459-466.
5. Lindenbach B, Rice C. Flaviviridae: the viruses and their replication. In: Fields BN, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, et al, editors. Fields Virology, 4th edition: Lippincott Williams & Wilkins; 2001:991-1041.
6. Ghosh D, Basu A. Japanese encephalitis-a pathological and clinical perspective. PLoS Negl Trop Dis 2009 Sep 29;3(9):e437. [doi: 10.1371/journal.pntd.0000437] [Medline: 19787040]
7. Hong C, Zhang Q, Zhang Y, et al. Impacts of climate change on future air quality and human health in China. Proc Natl Acad Sci USA 2019 Aug 27;116(35):17193-17200. [doi: 10.1073/pnas.1812881116] [Medline: 31405979]
8. Zhang Y, Wang H, Du J, et al. Population genetic structure of Culex tritaeniorhynchus in different types of climatic zones in China. BMC Genomics 2024;25(1):673. [doi: 10.1186/s12864-024-10589-4]
9. Centers for Disease Control and Prevention (CDC). Japanese encephalitis surveillance and immunization—Asia and Western pacific, 2012. MMWR Morb Mortal Wkly Rep 2013;62:658-662.
10. Heffelfinger JD, Li X, Batmunkh N, et al. Japanese encephalitis surveillance and immunization—Asia and Western Pacific Regions, 2016. Wkly Epidemiol Rep 2017;92:321-332. [doi: 10.15585/mmwr.mm6622a3]
11. Le Flohic G, Porphyre V, Barbazan P, Gonzalez JP. Review of climate, landscape, and viral genetics as drivers of the Japanese encephalitis virus ecology. PLoS Negl Trop Dis 2013;7(9):e2208. [doi: 10.1371/journal.pntd.0002208] [Medline: 24069463]
12. Fischer M, Hills S, Staples E, Johnson B, Yaich M, Solomon T. Japanese encephalitis prevention and control: advances, challenges, and new initiatives. In: Scheld WM, Hammer SM, Hughes JM, editors. Emerging Infections 8: ASM Press; 2008:93-124.
13. Chang YK, Chang HL, Wu HS, Chen KT. Epidemiological features of Japanese encephalitis in Taiwan from 2000 to 2014. Am J Trop Med Hyg 2017 Feb 8;96(2):382-388. [doi: 10.4269/ajtmh.16-0330] [Medline: 27821699]
14. Average GDP per person (USD) [Article in Mandarin]. National Statistics, ROC (Taiwan). 2022. URL: https://www.stat.gov.tw/point.asp?index=1 [accessed 2025-1-10]
15. Su CL, Yang CF, Teng HJ, et al. Molecular epidemiology of Japanese encephalitis virus in mosquitoes in Taiwan during 2005-2012. PLoS Negl Trop Dis 2014 Oct;8(10):e3122. [doi: 10.1371/journal.pntd.0003122] [Medline: 25275652]
16. Wu YH, Chen JM, Su CY. Occurrence of old age Japanese encephalitis: current situation in Taiwan. J Formos Med Assoc 2017 Dec;116(12):915-916. [doi: 10.1016/j.jfma.2017.04.010] [Medline: 28528139]
17. Hsu JY, Hung CC, Tsou TP, Chen WC. Epidemiology and risk factors of Japanese encephalitis in Taiwan, 2010-2022. PLoS Negl Trop Dis 2023 Oct;17(10):e0011421. [doi: 10.1371/journal.pntd.0011421] [Medline: 37782654]
18. Lin CL, Chang HL, Lin CY, Chen KT. Seasonal patterns of Japanese encephalitis and associated meteorological factors in Taiwan. Int J Environ Res Public Health 2017 Oct 29;14(11):1317. [doi: 10.3390/ijerph14111317] [Medline: 29109371]
19. Ho YC, Su BH, Su HJ, et al. The association between the incidence of mumps and meteorological parameters in Taiwan. Hum Vaccin Immunother 2015;11(6):1406-1412. [doi: 10.1080/21645515.2015.1029687] [Medline: 25891825]
20. Taiwan National Infectious Disease Statistics System [Website in Mandarin]. URL: http://nidss.cdc.gov.tw [accessed 2025-01-10]
21. Solomon T, Thao TT, Lewthwaite P, et al. A cohort study to assess the new WHO Japanese encephalitis surveillance standards. Bull World Health Organ 2008 Mar;86(3):178-186. [doi: 10.2471/blt.07.043307] [Medline: 18368204]
22. Hsu LC, Chen YJ, Hsu FK, et al. The incidence of Japanese encephalitis in Taiwan--a population-based study. PLoS Negl Trop Dis 2014 Jul;8(7):e3030. [doi: 10.1371/journal.pntd.0003030] [Medline: 25058573]

23. Ravi V, Robinson JS, Russell BJ, et al. Evaluation of IgM antibody capture enzyme-linked immunosorbent assay kits for detection of IgM against Japanese encephalitis virus in cerebrospinal fluid samples. Am J Trop Med Hyg 2009 Dec;81(6):1144-1150. [doi: 10.4269/ajtmh.2009.09-0144] [Medline: 19996450]
24. Clark CH, Casals J. Techniques for hemagglutination inhibition with arthropod viruses. Am J Trop Med Hyg 1958;7:561-573. [doi: 10.4269/ajtmh.1958.7.561]
25. Manual for the Laboratory Diagnosis of Japanese Encephalitis Virus Infection: World Health Organization; 2007.
26. Johnson BW, Goodman CH, Jee Y, Featherstone DA. Differential diagnosis of Japanese encephalitis virus infections with the Inbios JE Detect™ and DEN Detect™ MAC-ELISA Kits. Am Soc Trop Med Hyg 2016;94(4):820-828. [doi: 10.4269/ajtmh.15-0631]
27. Sapkal GN, Wairagkar NS, Ayachit VM, Bondre VP, Gore MM. Detection and isolation of Japanese encephalitis virus from blood clots collected during the acute phase of infection. Am J Trop Med Hyg 2007 Dec;77(6):1139-1145. [Medline: 18165537]
28. Mathur A, Kumar R, Sharma S, Kulshreshtha R, Kumar A, Chaturvedi UC. Rapid diagnosis of Japanese encephalitis by immunofluorescent examination of cerebrospinal fluid. Indian J Med Res 1990 Jan;91:1-4. [Medline: 2160913]
29. Environmental statistics inquiry network [Article in Mandarin]. Ministry of Environment. URL: https://statis.moenv.gov.tw/epanet [accessed 2025-01-13]
30. Monthly weather [Article in Mandarin]. Central Weather Administration. URL: https://www.cwa.gov.tw/V8/C/C/Statistics/monthlydata.html [accessed 2025-01-13]
31. Holland KM, Jones C, Vivolo-Kantor AM, et al. Trends in US Emergency Department visits for mental health, overdose, and violence outcomes before and during the COVID-19 pandemic. JAMA Psychiatry 2021 Apr 1;78(4):372-379. [doi: 10.1001/jamapsychiatry.2020.4402] [Medline: 33533876]
32. Pan CY, Liu WL, Su MP, et al. Epidemiological analysis of the Kaohsiung city strategy for dengue fever quarantine and epidemic prevention. BMC Infect Dis 2020 May 15;20(1):347. [doi: 10.1186/s12879-020-4942-y] [Medline: 32414340]
33. Communicable Disease Control Act. Laws & Regulations Database of The Republic of China (Taiwan). 2023. URL: https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=L0050001 [accessed 2025-01-10]
34. Lopalco PL, Biasio LR. Japanese encephalitis can be devastating. Ann Ig 2024;36(3):370-375. [doi: 10.7416/ai.2024.2616] [Medline: 38436079]
35. Gordis L. The dynamics of disease transmission. In: Gordis L, editor. Epidemiology, 3rd edition 2004:1934-2015.
36. Stanley SL Jr. Amoebiasis. Lancet 2003 Mar 22;361(9362):1025-1034. [doi: 10.1016/S0140-6736(03)12830-9] [Medline: 12660071]
37. Srivastava N, Deval H, Mittal M, et al. Extent of disability among paediatric Japanese encephalitis survivors and predictors of poor outcome: a retrospective cohort study in North India. BMJ Open 2022 Oct 31;12(10):e060795. [doi: 10.1136/bmjopen-2022-060795] [Medline: 36316071]
38. Pommier JD, Gorman C, Crabol Y, et al. Childhood encephalitis in the Greater Mekong region (the SouthEast Asia Encephalitis Project): a multicentre prospective study. Lancet Glob Health 2022 Jul;10(7):e989-e1002. [doi: 10.1016/S2214-109X(22)00174-7] [Medline: 35714649]
39. Udeze AO, Odebisi-Omokanye MB. Sero-evidence of silent Japanese encephalitis virus infection among inhabitants of Ilorin, North-central Nigeria: a call for active surveillance. J Immunoassay Immunochem 2022 May 4;43(3):250-258. [doi: 10.1080/15321819.2021.1993897] [Medline: 34809529]
40. Wu D, Chen X, Liu W, et al. Emergence of Japanese encephalitis among adults 40 years of age or older in northern China: epidemiological and clinical characteristics. Transbound Emerg Dis 2021 Nov;68(6):3415-3423. [doi: 10.1111/tbed.13945] [Medline: 33283432]
41. Kwak BO, Hong YJ, Kim DH. Changes in age-specific seroprevalence of Japanese encephalitis virus and impact of Japanese encephalitis vaccine in Korea. Clin Exp Pediatr 2022 Mar;65(3):108-114. [doi: 10.3345/cep.2020.01984] [Medline: 34592804]
42. Lin FH, Chen BC, Chou YC, et al. The epidemiology of Entamoeba histolytica infection and its associated risk factors among domestic and imported patients in Taiwan during the 2011-2020 period. Medicina (Kaunas) 2022 Jun 17;58(6):820. [doi: 10.3390/medicina58060820] [Medline: 35744083]
43. Lee PI, Huang YC, Hwang KP, et al. Recommendations for the use of Japanese encephalitis vaccines. Pediatr Neonatol 2020 Feb;61(1):3-8. [doi: 10.1016/j.pedneo.2019.11.009] [Medline: 31870559]
44. Griffith MM, Fukusumi M, Kobayashi Y, et al. Epidemiology of vaccine-preventable diseases in Japan: considerations for pre-travel advice for the 2019 Rugby World Cup and 2020 Summer Olympic and Paralympic Games. Western Pac Surveill Response J 2018;9(2):26-33. [doi: 10.5365/wpsar.2017.8.4.002] [Medline: 30057855]
45. Folly AJ, Dorey-Robinson D, Hernández-Triana LM, et al. Temperate conditions restrict Japanese encephalitis virus infection to the mid-gut and prevents systemic dissemination in Culex pipiens mosquitoes. Sci Rep 2021 Mar 17;11(1):6133. [doi: 10.1038/s41598-021-85411-2] [Medline: 33731761]
46. Lo SH, Tang HJ, Lee SSJ, et al. Determining the clinical characteristics and prognostic factors for the outcomes of Japanese encephalitis in adults: a multicenter study from southern Taiwan. J Microbiol Immunol Infect 2019 Dec;52(6):893-901. [doi: 10.1016/j.jmii.2019.08.010] [Medline: 31628089]

47. Monthly average climate [Article in Mandarin]. Central Weather Administration. URL: https://www.cwa.gov.tw/V8/C/C/Statistics/monthlymean.html [accessed 2025-01-10]
48. Tian HY, Bi P, Cazelles B, et al. How environmental conditions impact mosquito ecology and Japanese encephalitis: an eco-epidemiological approach. Environ Int 2015 Jun;79:17-24. [doi: 10.1016/j.envint.2015.03.002] [Medline: 25771078]
49. Di Francesco J, Choeung R, Peng B, et al. Comparison of the dynamics of Japanese encephalitis virus circulation in sentinel pigs between a rural and a peri-urban setting in Cambodia. PLoS Negl Trop Dis 2018 Aug;12(8):e0006644. [doi: 10.1371/journal.pntd.0006644] [Medline: 30138381]
50. Lindahl JF, Ståhl K, Chirico J, Boqvist S, Thu HTV, Magnusson U. Circulation of Japanese encephalitis virus in pigs and mosquito vectors within Can Tho city, Vietnam. PLoS Negl Trop Dis 2013;7(4):e2153. [doi: 10.1371/journal.pntd.0002153] [Medline: 23593520]
51. Lim MJ, Loh ZY, Yeo HL, et al. Isolation and genetic characterization of Japanese encephalitis virus two decades after its elimination in Singapore. Viruses 2022 Nov 28;14(12):2662. [doi: 10.3390/v14122662] [Medline: 36560666]
52. Williams CR, Webb CE, Higgs S, van den Hurk AF. Japanese encephalitis virus emergence in Australia: public health importance and implications for future surveillance. Vector Borne Zoonotic Dis 2022 Nov;22(11):529-534. [doi: 10.1089/vbz.2022.0037] [Medline: 36354964]
53. Walsh MG, Pattanaik A, Vyas N, et al. High-risk landscapes of Japanese encephalitis virus outbreaks in India converge on wetlands, rain-fed agriculture, wild Ardeidae, and domestic pigs and chickens. Int J Epidemiol 2022 Oct 13;51(5):1408-1418. [doi: 10.1093/ije/dyac050] [Medline: 35355081]
54. Mayilsamy M, Vijayakumar A, Veeramanoharan R, Rajaiah P, Balakrishnan V, Kumar A. Impact of COVID-19 lockdown during 2020 on the occurrence of vector-borne diseases in India. J Vector Borne Dis 2023;60(2):207. [doi: 10.4103/0972-9062.364762]
55. Sahu RC, Suthar T, Pathak A, Jain K. Interventions for the prevention and treatment of Japanese encephalitis. Curr Infect Dis Rep 2022 Nov;24(11):189-204. [doi: 10.1007/s11908-022-00786-1]
56. Turtle L, Solomon T. Japanese encephalitis - the prospects for new treatments. Nat Rev Neurol 2018 Apr 26;14(5):298-313. [doi: 10.1038/nrneurol.2018.30] [Medline: 29697099]
57. Pham D, Howard-Jones AR, Hueston L, et al. Emergence of Japanese encephalitis in Australia: a diagnostic perspective. Pathology (Phila) 2022 Oct;54(6):669-677. [doi: 10.1016/j.pathol.2022.07.001] [Medline: 35995617]
58. Guo H, Sun L, Shen X, Hu W. A retrospective study of the clinical characteristics of Japanese encephalitis in adults. J Integr Neurosci 2022 Jul 18;21(5):125. [doi: 10.31083/j.jin2105125] [Medline: 36137966]
59. Wang LP, Yuan Y, Liu YL, et al. Etiological and epidemiological features of acute meningitis or encephalitis in China: a nationwide active surveillance study. Lancet Reg Health West Pac 2022 Mar;20:100361. [doi: 10.1016/j.lanwpc.2021.100361] [Medline: 35036977]
60. Khare B, Kuhn RJ. The Japanese encephalitis antigenic complex viruses: from structure to immunity. Viruses 2022 Oct 8;14(10):2213. [doi: 10.3390/v14102213] [Medline: 36298768]
61. Tandale BV, Khude PM, Deshmukh PS, et al. Effectiveness of Japanese encephalitis vaccination among children in central India. J Med Virol 2023 Jan;95(1):e28399. [doi: 10.1002/jmv.28399] [Medline: 36512338]
62. Zhang WX, Zhao S, Pan C, et al. Mass immunisation to eradicate Japanese encephalitis: real-world evidence from Guizhou Province in 2005-2021. J Virus Erad 2024 Mar;10(1):100366. [doi: 10.1016/j.jve.2024.100366] [Medline: 38586471]
63. Mora C, Dousset B, Caldwell IR, et al. Global risk of deadly heat. Nat Clim Chang 2017 Jul;7(7):501-506. [doi: 10.1038/nclimate3322]
64. Mazdiyasni O, AghaKouchak A, Davis SJ, et al. Increasing probability of mortality during Indian heat waves. Sci Adv 2017 Jun;3(6):e1700066. [doi: 10.1126/sciadv.1700066] [Medline: 28630921]
65. Schmidhuber J, Tubiello FN. Global food security under climate change. Proc Natl Acad Sci USA 2007 Dec 11;104(50):19703-19708. [doi: 10.1073/pnas.0701976104] [Medline: 18077404]
66. Springmann M, Mason-D'Croz D, Robinson S, et al. Global and regional health effects of future food production under climate change: a modelling study. Lancet 2016 May 7;387(10031):1937-1946. [doi: 10.1016/S0140-6736(15)01156-3] [Medline: 26947322]
67. Wu X, Lu Y, Zhou S, Chen L, Xu B. Impact of climate change on human infectious diseases: empirical evidence and human adaptation. Environ Int 2016 Jan;86:14-23. [doi: 10.1016/j.envint.2015.09.007] [Medline: 26479830]
68. Diptyanusa A, Herini ES, Indarjulianto S, Satoto TBT. Estimation of Japanese encephalitis virus infection prevalence in mosquitoes and bats through nationwide sentinel surveillance in Indonesia. PLoS One 2022;17(10):e0275647. [doi: 10.1371/journal.pone.0275647] [Medline: 36223381]
69. Manisalidis I, Stavropoulou E, Stavropoulos A, Bezirtzoglou E. Environmental and health impacts of air pollution: a review. Front Public Health 2020;8:14. [doi: 10.3389/fpubh.2020.00014] [Medline: 32154200]
70. Tewari P, Ma P, Gan G, et al. Non-linear associations between meteorological factors, ambient air pollutants and major mosquito-borne diseases in Thailand. PLoS Negl Trop Dis 2023 Dec;17(12):e0011763. [doi: 10.1371/journal.pntd.0011763] [Medline: 38150471]

Abbreviations

AOR: adjusted odds ratio
CSF: cerebrospinal fluid
DF: dengue fever
HI: hemagglutination inhibition
JE: Japanese encephalitis
JEV: Japanese encephalitis virus
PM: particulate matter
TCDC: Taiwan Centers for Diseases Control
VBD: vector-borne disease

Edited by C Argyropoulos; submitted 11.06.24; peer-reviewed by D Valvi, RJPS Guimaraes, YL Cheong; revised version received 21.11.24; accepted 23.11.24; published 28.01.25.

Please cite as:
Lin FH, Chou YC, Hsieh CJ, Yu CP
Epidemiological Features, Clinical Symptoms, and Environmental Risk Factors for Notifiable Japanese Encephalitis in Taiwan From 2008 to 2020: Retrospective Study
JMIR Public Health Surveill 2025;11:e63053
URL: https://publichealth.jmir.org/2025/1/e63053
doi:10.2196/63053

Original Paper

Mindfulness Intervention for Health Information Avoidance in Older Adults: Mixed Methods Study

Chenyu Gu[1], PhD; Liquan Qian[2], PhD; Xiaojie Zhuo[2], PhD

[1]School of Journalism and Communication, Minjiang University, Fuzhou, China
[2]School of Arts and Media, Wuhan College, Wuhan, China

Corresponding Author:
Xiaojie Zhuo, PhD
School of Arts and Media
Wuhan College
No. 333 Huangjiahu Avenue
Jiangxia District, Hubei Province
Wuhan, 430212
China
Phone: 86 180 5922 1673
Email: 3074@mju.edu.cn

Abstract

Background: The global aging population and rapid development of digital technology have made health management among older adults an urgent public health issue. The complexity of online health information often leads to psychological challenges, such as cyberchondria, exacerbating health information avoidance behaviors. These behaviors hinder effective health management; yet, little research examines their mechanisms or intervention strategies.

Objective: This study investigates the mechanisms influencing health information avoidance among older adults, emphasizing the mediating role of cyberchondria. In addition, it evaluates the effectiveness of mindfulness meditation as an intervention strategy to mitigate these behaviors.

Methods: A mixed methods approach was used, combining quantitative and qualitative methodologies. Substudy 1 developed a theoretical model based on self-determination theory to explore internal (positive metacognition and health self-efficacy) and external (subjective norms and health information similarity) factors influencing health information avoidance, with cyberchondria as a mediator. A cross-sectional survey (N=236) was conducted to test the proposed model. Substudy 2 involved a 4-week mindfulness meditation intervention (N=94) to assess its impact on reducing health information avoidance behaviors.

Results: Study 1 showed that positive metacognition (β=.26, *P*=.002), health self-efficacy (β=.25, *P*<.001), and health information similarity (β=.29, *P*<.001) significantly predicted health information avoidance among older adults. Cyberchondria mediated these effects: positive metacognition (effect=0.106, 95% CI 0.035-0.189), health self-efficacy (effect=0.103, 95% CI 0.043-0.185), and health information similarity (effect=0.120, 95% CI 0.063-0.191). Subjective norms did not significantly predict health information avoidance (β= .11, *P*=.13), and cyberchondria did not mediate this relationship (effect= 0.045, 95% CI 0.102 to 0.016). Study 2 found that after the 4-week mindfulness intervention, the intervention group (group 1: n=46) exhibited significantly higher mindfulness levels than the control group (group 2: n=48; M_{group1}=4.122, M_{group2}=3.606, *P*<.001) and higher levels compared with preintervention (M_{t2}=4.122, M_{t1}=3.502, *P*<.001, where t1=preintervention and t2=postintervention). However, cyberchondria levels did not change significantly (M_{t1}=2.848, M_{t2}=2.685, *P*=.18). Nevertheless, the results revealed a significant interaction effect between mindfulness and cyberchondria on health information avoidance (effect= 0.357, *P*=.002, 95% CI 0.580 to 0.131), suggesting that mindfulness intervention effectively inhibited the transformation of cyberchondria into health information avoidance behavior.

Conclusions: This study reveals the role of cyberchondria in health information avoidance and validates mindfulness meditation as an effective intervention for mitigating such behaviors. Findings offer practical recommendations for improving digital health information delivery and health management strategies for older adults.

(JMIR Public Health Surveill 2025;11:e69554) doi:10.2196/69554

KEYWORDS

health information avoidance; cyberchondria; self-determination theory; mindfulness; elderly

Introduction

Background

With the global aging population and the rise of digital health information, the health information behavior of older adults has become a critical focus [1]. The adoption of internet and age-friendly technologies has enabled older adults to seek health information online. However, the overwhelming volume and inconsistent quality of information, combined with lower information literacy among older adults, often hinder efficient access to reliable health information, posing challenges to their health management.

Health information avoidance is a growing concern [2]. It involves individuals deliberately avoiding health-related information, often driven by anxiety and uncertainty, which can lead to severe health consequences [3]. Given their greater need for health information, avoidance behaviors in older adults can hinder effective health management and delay timely diagnosis or treatment [4]. However, research on its causes and interventions remains limited. Cyberchondria, characterized by excessive health information searching, is another critical issue. It disrupts daily activities and exacerbates health anxiety, despite being initially aimed at reducing uncertainty [5]. While it has been linked to problems like internet addiction and sleep disorders, its connection to health information avoidance remains underexplored [6,7]. This study posits that cyberchondria amplifies anxiety, which in turn motivates information avoidance as a self-protection mechanism [8]. Viewing cyberchondria as a psychological driver of health information avoidance offers valuable insights into this behavior.

This study introduces mindfulness meditation as a potential intervention. Rooted in Buddhist traditions, mindfulness has been shown to reduce anxiety and improve emotional regulation and cognitive flexibility [9,10]. By applying mindfulness meditation with older adults, this study evaluates its effects on cyberchondria and health information avoidance, providing a cost-effective health management strategy.

In summary, this study aims to explore the psychological mechanisms underlying health information avoidance among older adults, with a particular focus on cyberchondria as a mediating factor. Drawing from self-determination theory, we propose a model that incorporates both internal (positive metacognition and health self-efficacy) and external (subjective norms and health information similarity) factors to explain how they contribute to health information avoidance. Furthermore, we aim to evaluate the effectiveness of mindfulness meditation as a practical intervention to mitigate the impact of cyberchondria on information avoidance behaviors. This study seeks to address the following two research questions:

1. What are the formation pathways of health information avoidance among older adults?
2. Can mindfulness meditation effectively intervene in cyberchondria and reduce health information avoidance behaviors among older adults?

By addressing these questions, this research intends to contribute theoretical and practical insights to enhance older adults' health management capabilities in the digital era.

Theory Background and Hypothesis Development

Health Information Avoidance Among Older Adults

In the context of global digitalization and an aging population, the health information behavior of older adults has become a critical topic in health management [11]. Health information avoidance, the deliberate decision to avoid health-related information, is particularly prevalent among older adults due to lower information literacy, anxiety about health issues, and fear of disease severity [12]. This behavior poses significant risks, such as missed medical advice or delayed disease prevention and treatment, as older adults often experience poorer health than younger individuals [1].

Although avoidance may temporarily reduce stress, it often exacerbates anxiety by failing to address underlying health concerns and increasing uncertainty [13,14]. It also undermines proactive health management, reducing engagement in health-promoting behaviors and increasing the risk of chronic diseases [15]. Despite growing research, key gaps remain, particularly in understanding how external environments interact with psychological factors and in developing targeted interventions for older adults. Cyberchondria, defined as repeated health searches that heighten anxiety, offers a valuable framework for understanding avoidance behaviors [16]. It highlights the uncertainty and negative emotions triggered by health information, which align with the drivers of avoidance. Incorporating this framework can help explain the anxiety-avoidance cycle in older adults and support the development of targeted interventions to break this cycle.

Cyberchondria

Cyberchondria refers to compulsive and excessive searches for health-related information, often intended for self-reassurance. However, these searches usually provide only temporary relief, with anxiety often worsening during and after the process. Despite its negative consequences, such as harm to mental health, overmedicalization, and susceptibility to medical scams, cyberchondria has become a significant public health concern with the rise of mobile internet use [17,18].

Research on cyberchondria highlights 2 main areas. First, contributing factors include low self-esteem, anxiety sensitivity, low tolerance for uncertainty, and high neuroticism [19-21]. Second, its negative impacts include reduced trust in doctors, increased self-treatment behaviors, medical errors, lower quality of life, and wasted health care resources [22,23]. It has also been linked to problematic internet use [24]. Despite these findings, little research has examined the link between cyberchondria and health information avoidance. This study aims to fill this gap by exploring how cyberchondria contributes to avoidance behaviors, providing new insights into health behaviors [25,26]. Understanding this connection is crucial for designing interventions to help vulnerable populations better manage health information and reduce the negative effects of cyberchondria.

Self-Determination Theory

Self-determination theory (SDT), proposed by Deci and Ryan [27], explains human motivation and behavior through the fulfillment of 3 psychological needs: competence, relatedness, and autonomy. Behavior is influenced by internal motivation, driven by psychological cognition, and external motivation, shaped by external pressures or rewards [28]. This dual perspective has been widely applied in health psychology, effectively explaining behaviors such as the adoption of personal health record systems driven by autonomous traits (internal factors) and compulsive social media use influenced by social interaction (external factors) [29-31].

Although extensively applied, SDT has not been used to explain cyberchondria. By integrating internal motivations and external environments, SDT provides a valuable framework for understanding its development and guiding effective health interventions [32]. Responding to calls for further exploration of cyberchondria's antecedents and consequences [33], this study uses SDT to analyze the factors influencing cyberchondria and its connection to health information avoidance.

Cyberchondria and Health Information Avoidance

Cyberchondria is characterized by compulsive and frequent internet-based searching for health-related information, which intensifies individuals' anxiety and concerns about their health. This behavior often leads to cognitive and emotional responses, such as excessive worry about illness and overinterpretation of symptoms, ultimately impacting health behaviors [34]. While cyberchondria has been linked to behaviors like internet addiction, evidence suggests that internet addiction may eventually result in information avoidance. Individuals with cyberchondria often experience heightened anxiety and negative emotional responses when exposed to health-related information [35]. These negative emotions reinforce their fear of health issues, leading to a hypersensitivity toward disease-related information. Based on protection motivation theory, individuals may perceive excessive health information as a source of emotional distress and adopt avoidance behaviors to reduce anxiety and stress [36]. Furthermore, individuals with cyberchondria often interpret ambiguous or uncertain information as a potential threat. This heightened threat perception can result in strong psychological resistance, causing them to proactively reduce exposure to health information to avoid discomfort. Over time, this behavior manifests as health information avoidance [37]. In light of these mechanisms, the following hypotheses are proposed in this study:

H1: Cyberchondria positively influences the health information avoidance.

Internal Factors: Positive Metacognition and Health Self-Efficacy

Next, we explore the factors that lead to cyberchondria in older adults. According to the SDT mentioned earlier, an individual's behavioral intention is influenced by internal cognitive factors. This study selects positive metacognition as one of the key internal factors influencing cyberchondria. Metacognition is commonly defined as "thinking about thinking" and encompasses the beliefs, strategies, and methods individuals use to regulate their internal cognitive processes [38]. Metacognitive beliefs can be categorized as either positive or negative. Positive metacognition refers to the belief that focusing on health issues and extensively thinking about them enhances one's sense of safety and control. This construct is widely used to explain health anxiety [39].

Research has shown that positive metacognition effectively predicts attentional biases toward health-related information. Individuals with strong positive metacognitive tendencies are more vigilant about their health and more likely to seek health-related information [40]. Such individuals often believe that worrying about their health and gaining more knowledge about illnesses provide a sense of security. This biased thinking pattern drives compulsive health information-seeking behaviors, which are hallmarks of cyberchondria [41]. However, excessive information-seeking often leads to heightened anxiety, cognitive overload, and feelings of inadequacy in managing one's health, which can result in health information avoidance as a coping mechanism. Therefore, this study hypothesizes that positive metacognition positively influences an individual's likelihood of developing cyberchondria, which, in turn, leads to health information avoidance. Based on this logic, the following research hypotheses are proposed:

H2a: Positive metacognition positively influences cyberchondria.

H2b: Cyberchondria mediates the effect of positive metacognition on health information avoidance.

The concept of self-efficacy was first introduced by Bandura [42] in 1977, referring to an individual's confidence and belief in their ability to achieve specific behavioral goals. Ajzen [43] later incorporated it into the theory of planned behavior, which suggests that self-efficacy is a significant predictor of an individual's behavioral intention. Numerous studies have since confirmed that self-efficacy is an important component of the internal cognitive system, significantly affecting personal behavior and decision-making [44]. Self-efficacy has been widely studied in the field of health behaviors [45], where it is often referred to as health self-efficacy. According to Lee et al [46], health self-efficacy refers to an individual's confidence in their ability to manage their health. Individuals with higher levels of health self-efficacy are more likely to feel capable of addressing health issues independently and have a stronger inclination toward self-healing. This self-healing tendency reduces communication with health care professionals and encourages individuals to search for health and treatment-related information for self-treatment [47]. Furthermore, the tendency toward self-healing has been shown to be significantly associated with cyberchondria [22]. Therefore, this study speculates that individuals with higher levels of health self-efficacy are more inclined to gather health-related information through internet-based channels, which may, in turn, lead to a tendency toward cyberchondria, causing these users to become trapped in a health information bubble. Based on this reasoning, the following research hypotheses are proposed:

H3a: Health self-efficacy positively influences cyberchondria.

H3b: Cyberchondria mediates the effect of health self-efficacy on health information avoidance.

External Factors: Health Information Similarity and Subjective Norms

Health information similarity refers to the consistency or resemblance in the content and structure of health information encountered on digital platforms, often driven by personalized recommendation algorithms tailored to users' preferences and behaviors [48,49]. These algorithms frequently expose users to overlapping health content, which can create cognitive confusion and negative emotional responses. For instance, similar symptom descriptions across diseases make it difficult to differentiate conditions, increasing uncertainty and anxiety [50]. This anxiety often triggers more frequent searches, leading to information overload and reinforcing a cycle of heightened anxiety.

While previous studies link information similarity to health anxiety, its impact on cyberchondria remains underexplored. Repeated exposure to overlapping health content can cause individuals to misinterpret ordinary symptoms such as severe illnesses, further amplifying uncertainty and intensifying cyberchondria [51]. This exacerbates excessive information searching and, in turn, contributes to health information avoidance. Based on this reasoning, the following hypotheses are proposed:

H4a: Health information similarity positively influences cyberchondria.

H4b: Cyberchondria mediates the effect of health information similarity on health information avoidance.

Subjective norms refer to the perceived social pressure or expectations from others, reflecting the social acceptance of a particular behavior. Both the risk information seeking and processing model and the theory of planned behavior emphasize that social factors (subjective norms) significantly influence an individual's intention to seek health information [52]. In this study, subjective norms specifically refer to the expectations from important others, such as family and friends, regarding an individual's health information-seeking behavior. When individuals perceive strong interest or encouragement from their social circle, subjective norms can create pressure to continue searching for health information. This social pressure often leads individuals to feel a heightened sense of responsibility, causing them to persist in health information searches despite experiencing anxiety and discomfort. Such pressure can exacerbate anxiety, trapping individuals in a cycle of frequent searches and information overload [53]. Furthermore, subjective norms can make it difficult for individuals to exercise self-restraint, increasing doubts about their health and ultimately contributing to cyberchondria.

Subjective norms may also worsen cyberchondria by influencing self-perception. When individuals believe that those around them place great importance on health information searching, they may see not engaging in such behavior as irresponsible, further intensifying their psychological burden and health anxiety [54]. To alleviate this emotional distress, individuals may choose to avoid health information altogether, attempting to reduce the psychological burden caused by social pressure. Based on this reasoning, the following hypotheses are proposed:

H5a: Subjective norms positively influence cyberchondria.

H5b: Cyberchondria mediates the effect of subjective norms on health information avoidance.

Mindfulness as a Proactive Psychological Intervention Method

The resource model of self-control and ego depletion theory posits that self-control tasks, such as compulsive health information searches, deplete cognitive resources ("ego depletion"), impairing subsequent performance and leading to avoidance of related activities like processing additional health information [55,56]. However, self-control can be improved through consistent training. Mindfulness, rooted in Buddhist traditions, is defined as present-focused attention and nonjudgmental acceptance [57]. It includes 4 elements: being present, attention, awareness, and acceptance [58]. Mindfulness training enhances self-control, reduces cognitive depletion, and improves emotional regulation, making it effective for addressing anxiety, depression, and stress [59,60].

In the context of cyberchondria, individuals experience intense anxiety due to excessive focus on health information, leading to repeated searches that amplify uncertainty and anxiety. This cycle depletes cognitive resources, increasing health information avoidance. Mindfulness training helps break this cycle by enhancing awareness and acceptance, thereby reducing anxiety and distress [61]. Improved emotional regulation allows individuals to manage anxiety without excessive searching, disrupting the "anxiety–information searching–greater anxiety–information avoidance" loop. In addition, mindfulness promotes cognitive flexibility, helping individuals identify and resist cognitive biases common in cyberchondria, such as misinterpreting ordinary symptoms as severe illnesses [62]. By encouraging observation over reaction, mindfulness reduces health anxiety and prevents avoidance behaviors stemming from excessive concern about health information.

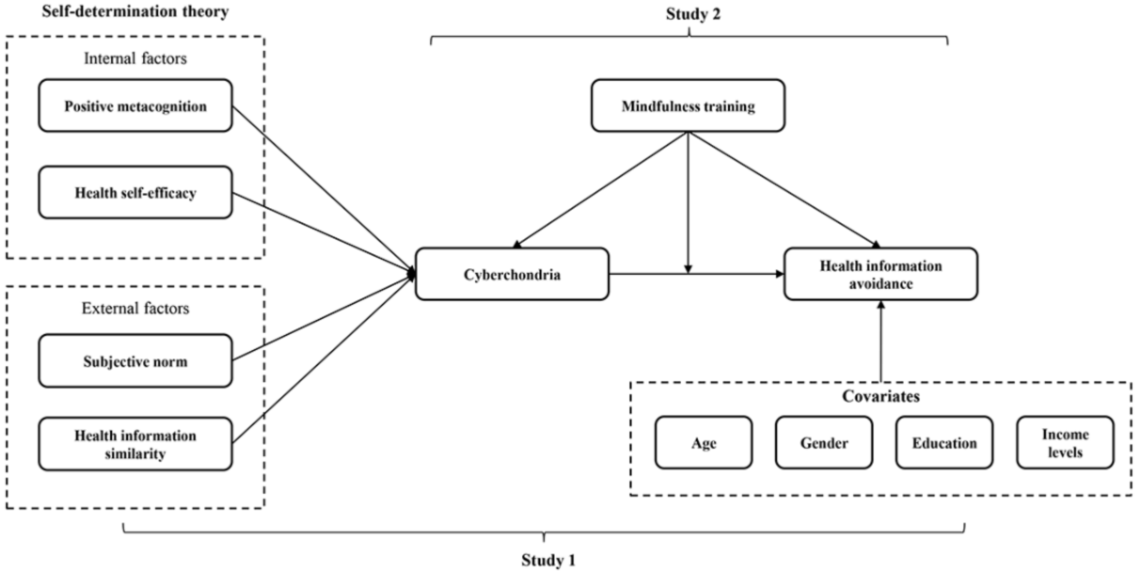
H6a: Mindfulness training significantly reduces individuals' levels of cyberchondria.

H6b: Mindfulness training negatively moderates the positive effect of cyberchondria on health information avoidance.

Conceptual Model

In summary, this study consists of 2 substudies. Study 1 aims to explore the causes of cyberchondria among older adults in the digital age and the impact model of health information avoidance behavior among older adults. Study 2 aims to validate the intervention effects of mindfulness training on cyberchondria and health information avoidance behavior in older adults. The specific research models are shown in Figure 1.

Figure 1. Conceptual model of health information avoidance among older adults.



Methods

Study 1

Participants and Procedure

Participants for study 1 were primarily recruited through health communities and social media platforms commonly used by older adults, such as WeChat Moments (Tencent), Douyin (the Chinese version of TikTok, ByteDance), and Jinri Toutiao (Today's Headlines). The target population included individuals aged 60 years and older who were capable of independently completing internet-based questionnaires. To enhance sample diversity and ensure an adequate sample size, additional outreach was conducted through local community networks, such as senior activity centers.

Data for study 1 was collected through internet-based survey platform. Before the survey began, the researchers clearly informed all participants about the purpose of the study and assured them that all responses would be strictly confidential and used solely for academic purposes. The survey was conducted anonymously, with no personal information being disclosed. Participants were also informed of the importance of providing truthful responses and were reminded of their right to withdraw from the survey at any time without any conditions. Informed consent was obtained from all participants.

Initially, 328 questionnaires were collected, with 92 invalid responses excluded, resulting in 236 valid questionnaires (mean age 63.93 years; 91, 38.56% males; 145, 61.44% females). Demographic details are provided in the Multimedia Appendix 1. The response rate of valid questionnaires was 71.95%. The criteria for selecting valid questionnaires were correctly answering screening questions, a total response time of more than 1 minute, and not selecting the same response option for 9 consecutive items.

Measurements

Study 1 measured the following latent variables: health information avoidance [63], cyberchondria [64], positive metacognition [65], health self-efficacy [66], subjective norms, and health information similarity [67,68]. The detailed measurement scales can be found in Multimedia Appendix 1. To ensure validity, all measurement tools were adapted from established research and adjusted to fit the context of health information dissemination. The average score of the items in each scale was used to represent each latent variable, and all questionnaires were measured using a 5-point Likert scale. A pretest was also conducted with 10 volunteers to ensure that all items were clearly stated and accurately understood.

Measurement of Model

Since this study is exploratory in nature, the research model, although built based on theoretical and logical reasoning, does not originate from an existing model. Therefore, it is appropriate to use "Smart PLS" for partial least squares (PLS) path analysis () to test the research model in this context [69].

Reliability and validity testing. As shown in Table 1, the factor loadings for all measurement items ranged from 0.728 to 0.950, indicating that all items met the retention criteria. In addition, the Cronbach α values for each latent variable ranged from 0.870 to 0.915, and the composite reliability (CR) values for all latent variables were greater than the acceptable threshold of 0.7, indicating that the reliability of the scales was satisfactory. Furthermore, the average variance extracted (AVE) values for all variables exceeded the standard acceptance value of 0.5, demonstrating adequate convergent validity. The variance inflation factor (VIF) values for all factors were below 10, indicating that there were no multicollinearity issues in the scales used in this study [70,71].

Table 1. Reliability and convergent validity results of latent variables.

Latent variable and items	Factor loadings	VIF[a]	Cronbach α values	CR[b]	AVE[c]
Health information avoidance			.912	.925	.739
HIA[d] 1	.841	2.608			
HIA 2	.903	3.805			
HIA 3	.803	2.056			
HIA 4	.854	2.461			
HIA 5	.894	2.929			
Cyberchondria			.915	.916	.796
CC[e] 1	.904	3.421			
CC 2	.912	3.329			
CC 3	.898	3.127			
CC 4	.854	2.202			
Positive metacognition			.898	.884	.710
PM[f] 1	.847	2.528			
PM 2	.854	2.408			
PM 3	.797	1.955			
PM 4	.869	2.744			
PM 5	.844	2.373			
Health self-efficacy			.870	.903	.718
HS[g] 1	.873	2.145			
HS 2	.823	2.053			
HS 3	.851	2.377			
HS 4	.843	2.100			
Subjective norm			.912	.936	.850
SN[h] 1	.944	3.727			
SN 2	.882	2.521			
SN 3	.940	3.908			
Health information similarity			.880	.927	.731
IS[i] 1	.728	2.246			
IS 2	.877	2.815			
IS 3	.850	2.239			
IS 4	.950	4.355			

[a] VIF: variance inflation factor.

[b] CR: composite reliability.

[c] AVE: average variance extracted.

[d] HIA: health information avoidance.

[e] CC: cyberchondria.

[f] PM: positive metacognition.

[g] HS: health self-efficacy.

[h] SN: subjective norm.

[i] IS: health information similarity.

Next, discriminant validity among the variables was verified, and the specific results are shown in Table 2. The square root of the AVE for all variables (on the diagonal) was greater than the Pearson correlation coefficients between the variables, indicating that the scales had satisfactory discriminant validity. A common method bias test was also conducted using the single-factor test. The variance explained by the first unrotated factor was 31.999%, which is below the critical threshold of 40%, indicating that common method bias is not an issue in this study [72].

Table 2. Discriminant validity analysis of constructs in study 1.

Constructs	HIA[a]	CC[b]	HS[c]	PM[d]	SN[e]	IS[f]
HIA	.860[g]	—[h]	—	—	—	—
CC	.410	.892[g]	—	—	—	—
HS	.346	.382	.848[g]	—	—	—
PM	.212	.343	.184	.843[g]	—	—
SN	.158	.137	.019	.651	.922[g]	—
IS	.339	.435	.291	.371	.246	.855[g]

[a] HIA: health information avoidance.

[b] CC: cyberchondria.

[c] HS: Health self-efficacy.

[d] PM: Positive metacognition.

[e] SN: subjective norm.

[f] IS: health information similarity.

[g] square roots of the average variance extracted.

[h] Not applicable.

Model Testing

The PLS algorithm was used to calculate the R^2 values for each variable, all of which were greater than the standard acceptance value of 0.1, indicating good predictive accuracy of the model [73]. Blindfolding was then conducted, and the results showed that the Stone-Geisser Q^2 values for all variables were greater than 0, indicating that the model effectively predicts the relationships between variables [74]. In addition, the SRMR value was 0.054, which is below the standard threshold of 0.08, the NFI value was 0.931, and the RMS θ value was 0.103. These results demonstrate that the model has good fit [75].

Study 2

Participants and Procedure

The required sample size for the intervention study was estimated using GPower, with an effect size greater than 0.8 considered appropriate for the experimental study. Setting α=.05, 1-β=.8, and effect size=0.4, the minimum required sample size was calculated to be 44 [76].

The mindfulness training required participants to invest a significant amount of time, making it challenging to recruit a sufficient number of participants. To address this difficulty, our research team collaborated with students from 3 courses at 3 universities, encouraging them to invite their older family members (older than 60 years old and capable of independently using social media) to participate in the mindfulness training program. To motivate the students, the researchers offered course credits and a reward of approximately US $20. To improve the quality of the mindfulness training and the responses to the survey, the researchers informed participants before the experiment began that their responses would be kept strictly confidential and used only for academic purposes. Participants were also informed of the importance of honest responses and were made aware of their right to withdraw from the study at any time without conditions. Informed consent was obtained from all participants. The research was conducted in accordance with the standards of the Declaration of Helsinki.

We initially recruited 101 participants and divided them into 2 groups: group 1 (mindfulness training group, n=51) and group 2 (control group, n=50). The research process was conducted in 3 stages.

Stage 1 (t1)

At the start of the study, a survey measured participants' levels of cyberchondria, health information avoidance, state mindfulness (Brown and Ryan [77]), and demographic variables (gender, age, income, and education). An independent samples *t* test showed no significant differences between the 2 groups for cyberchondria (M_{group1}=2.848, M_{group2}=3.110, *P*=.12, t_{99}=–1.557), health information avoidance (M_{group1}=3.063, M_{group2}=3.384, *P*=.11, t_{99}=–1.631), or state mindfulness (M_{group1}=3.502, M_{group2}=3.524, *P*=.77, t_{99}=–.297).

Stage 2

Mindfulness training started 1 week after the initial survey to avoid bias from survey content. Group 1 underwent a short-term intervention based on the mindfulness-based stress reduction (MBSR) program (Kabat-Zinn [78]), while group 2 received no intervention. The training included five 10-minute audio sessions focusing on breathing, bodily sensations, thoughts, and emotions. Participants were instructed to complete a daily session at noon with randomly assigned recordings and record their practice (refer to Multimedia Appendix 1 for audio excerpts).

Stage 3 (t2)

Following 4 weeks of training, consistent with previous findings on mindfulness duration (Rooks et al [79]), 5 participants from group 1 and 2 from group 2 withdrew or failed to complete the study, leaving a final valid sample of 46 in group 1 and 48 in group 2. Cyberchondria, health information avoidance, and state mindfulness levels were reassessed. Reanalysis of initial data showed no significant differences between retained participants, ensuring data integrity.

Ethical Considerations

This study strictly adhered to the ethical principles outlined in the Declaration of Helsinki and was approved by the Ethics Committee of the School of Journalism and Communication at Minjiang University (Ref: MJUCER20240107). Both substudies were conducted in strict accordance with relevant ethical guidelines to protect participants' rights and ensure data confidentiality.

All participants provided informed consent before participating in the study. They were fully informed about the study's objectives, procedures, data usage policies, and their right to withdraw from the study at any time without penalty. As the data were collected through questionnaires, the consent forms were obtained electronically.

To protect privacy and data security, all participant data were anonymized and deidentified during both the data collection and analysis processes. No personally identifiable information was recorded during the study, and all data were securely stored in an encrypted system to prevent unauthorized access.

Participants received appropriate compensation based on their involvement in the study. In substudy 1, a cross-sectional survey, participants were rewarded US $2 upon completing the questionnaire. In substudy 2, a 4-week longitudinal survey, participants received a gift valued at approximately US $20 after completing the study. These compensations were intended to acknowledge participants' time and effort while avoiding undue influence on their decision to participate. This manuscript and its supplementary materials do not include any images or materials that could identify participants.

Results

Study 1

There were no missing values in the data. A bootstrapping test with a sample size of 5000 was conducted on the collected data to explore the path coefficients and their significance within the model. The final test results are presented in Tables 3 and 4.

The results of the data analysis indicate that cyberchondria exacerbates health information avoidance among older adults (β=.41, *P*<.001), supporting H1. From the perspective of internal factors, both positive metacognition and health self-efficacy have a positive influence on cyberchondria (β=.26, *P*=.002; β=.25, *P*<.001), which in turn further increases health information avoidance behavior (95% Boot CI 0.035-0.189; 95% Boot CI 0.043-0.185). Thus, H2a, H2b, H3a, and H3b are all supported. Regarding external environmental factors, the similarity of health information also significantly positively affects cyberchondria (β=.29, *P*<.001) and leads to health information avoidance through cyberchondria (95% Boot CI 0.063-0.191), supporting H4a and H4b. However, subjective norms did not show a significant direct relationship with cyberchondria (β=–.11, *P*=.13), and cyberchondria did not mediate the relationship between subjective norms and health information avoidance (95% Boot CI –0.102 to 0.016), leading to the rejection of H5a and H5b.

The results of study 1 indicate that cyberchondria significantly exacerbates health information avoidance among older adults, influenced by both internal factors (positive metacognition and health self-efficacy) and external factors (health information similarity). However, subjective norms did not have a significant effect on cyberchondria, suggesting that external social pressure is not a major driver of health information avoidance behavior in older adults. Based on these findings, the next study will focus on validating whether mindfulness training can effectively intervene in cyberchondria and the resulting health information avoidance behavior. The specific findings of study 1 are shown in Figure 2.

Table 3. Hypothesis testing results of study 1 (direct effect).

Direct effect	β values	*t* test (*df*)	*P* values
H1: CC[a]→HIA[b]	.410	7.016 (226)	<.001[c]
H2a: PM[d]→CC	.259	3.035 (226)	.002[e]
H3a: HS[f]→CC	.251	3.453 (227)	.001[e]
H4a: IS[g]→CC	.293	4.344 (227)	<.001[c]
H5a: SN[h]→CC	-.109	–1.507 (228)	.13

[a] CC: cyberchondria.

[b] HIA: health information avoidance.

[c] *P*<.001.

[d] PM: positive metacognition.

[e] *P*<.01.

[f] HS: health self-efficacy.

[g] IS: health information similarity.

[h] SN: subjective norm.

Table 4. Hypothesis testing results of study 1 (indirect effect).

Indirect effect	Effect value	*t* values (*df*)	95% CI	*P* values
H2b: PM[a]→CC[b]→HIA[c]	0.106	2.697 (220)	0.035 to 0.189	.007[d]
H3b: HS[e]→CC→HIA	0.103	2.821 (221)	0.043 to 0.185	.005[d]
H4b: IS[f]→CC→HIA	0.120	3.691 (221)	0.063 to 0.191	<.001[g]
H5b: SN[h]→CC→HIA	–0.045	1.459 (222)	–0.102 to 0.016	.15

[a] PM: positive metacognition.

[b] CC: cyberchondria.

[c] HIA: health information avoidance.

[d] *P*<.01.

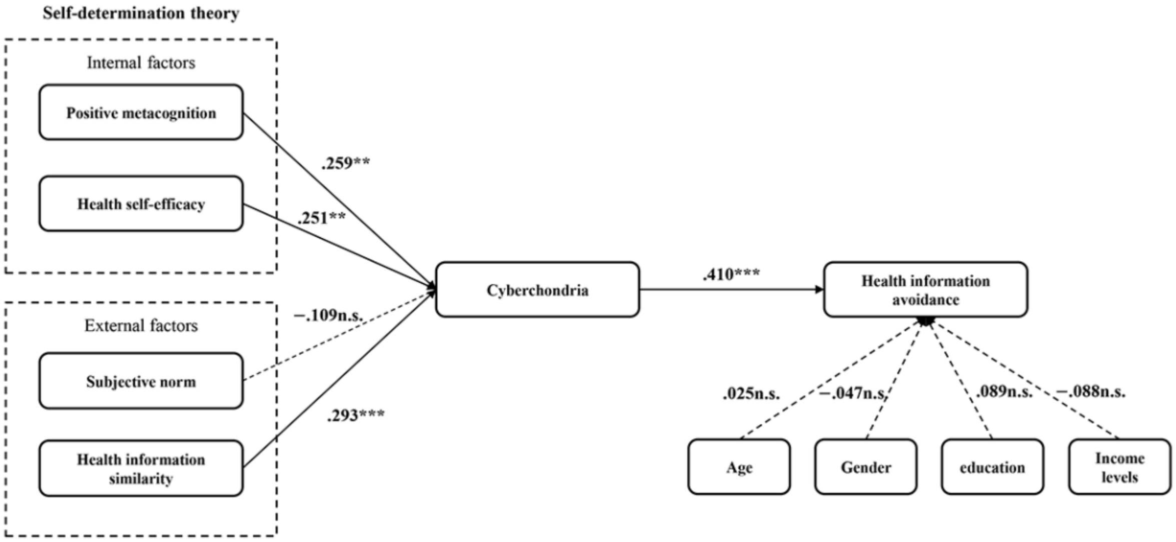
[e] HS: health self-efficacy.

[f] IS: health information similarity.

[g] *P*<.001.

[h] SN: subjective norm.

Figure 2. Result model of study 1: path analysis.***P*<.01, ****P*<.001, n.s. *P*>.05.



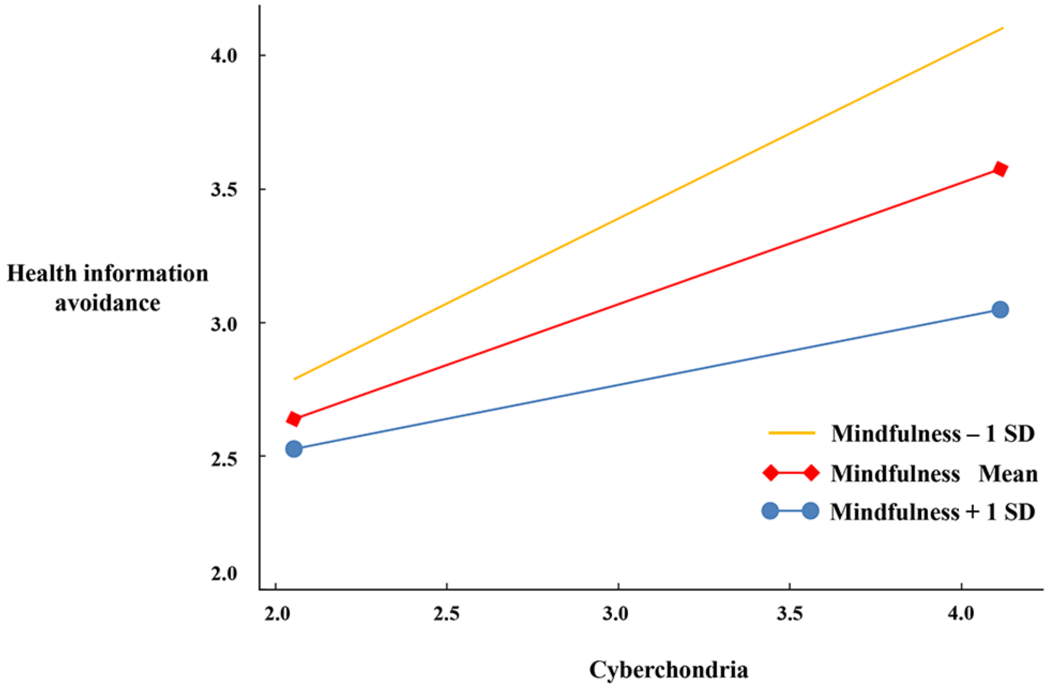
Study 2

Before hypothesis testing, the effectiveness of the mindfulness intervention was first examined. To control for the effects of covariates such as age, gender, and education level, an ANCOVA (analysis of covariance) was conducted to test the manipulation of state mindfulness. The analysis showed that the state mindfulness score of the intervention group was significantly higher than that of the control group (M_{group1}=4.122, M_{group2}=3.606, $F_{1,93}$=16.866, P<.001). Subsequently, the within-group differences in t1 and t2 for both the intervention and control groups were compared. The results showed that the state mindfulness level of the intervention group at t2 was significantly higher than at t1 (M_{t2}=4.122, M_{t1}=3.502, $F_{1,96}$=85.024, P<.001), while there was no significant difference between t1 and t2 in the control group (M_{t2}=3.524, M_{t1}=3.583, $F_{1,97}$=.222, P=.64). In summary, the manipulation of state mindfulness was successful.

Next, ANCOVA was conducted to test H6a, with age, gender, education level, and income level as covariates. The results showed that there was no significant difference in the level of cyberchondria between t1 and t2 in the intervention group (M_{t1}=2.848, M_{t2}=2.685, $F_{1,96}$=1.835, P=.18), indicating that H6a was not supported, meaning that mindfulness training did not improve the level of cyberchondria. Subsequently, H6b was tested using Hayes' suggested Bootstrap method with SPSS (IBM Corp) process model 1 to examine the interaction effect, with 5000 bootstrap samples and a 95% CI. The data analysis showed that the interaction effect of mindfulness level and cyberchondria on health information avoidance was significant (effect= 0.357, P=.002, 95% CI 0.580 to 0.131). Thus, H6b was supported.

In conclusion, the results of study 2 supported H6b but rejected H6a. Although mindfulness training did not directly reduce the level of cyberchondria, it effectively inhibited the transformation of cyberchondria into health information avoidance behavior (Figure 3).

Figure 3. The Interaction effect of mindfulness and cyberchondria on health information avoidance.



Discussion

Principal Findings

Cyberchondria and health information avoidance have become increasingly pressing public health issues in the digital age, posing significant threats to the health of older adults. Therefore, exploring their underlying mechanisms and potential intervention strategies is of great theoretical and practical significance. Based on the self-determination theory, this study conducted 2 substudies to explore the causes of cyberchondria and health information avoidance among older adults, as well as the intervention effects of mindfulness training.

Internal and External Factors Leading to Health Information Avoidance: Cyberchondria as the Mediator

The results reveal that internal factors (positive metacognition and health self-efficacy) and external factors (health information similarity) significantly predict cyberchondria among older adults, leading to health information avoidance.Positive metacognition fosters overthinking and health-related anxiety, as excessive focus on health concerns amplifies uncertainty. Previous research links positive metacognition to problematic behaviors [80], and this study confirms its role in exacerbating cyberchondria and health information avoidance. Health self-efficacy also influences cyberchondria. Although high self-efficacy promotes confidence in managing health, ambiguous health information undermines this confidence, driving excessive searches and heightening anxiety. This aligns with self-efficacy theory, where unresolved issues lead to repetitive checking and stress [81].

Health information similarity indirectly contributes to avoidance by increasing anxiety and cognitive overload. Repeated exposure to similar information, often driven by algorithms, makes it difficult to differentiate between diseases, heightening uncertainty and driving avoidance behaviors through cyberchondria [82]. However, subjective norms did not significantly affect cyberchondria, suggesting that older adults rely more on personal judgment and experience rather than external social influences [83,84]. This finding underscores the need for interventions targeting individual-level factors rather than external pressures.

Mindfulness Training as an Effective Intervention for Cyberchondria and Health Information Avoidance

The results also indicate that while mindfulness training did not directly reduce the level of cyberchondria, it significantly mitigated the transition from cyberchondria to health information avoidance behavior. Mindfulness meditation, by enhancing emotional regulation and cognitive flexibility, helps older adults avoid being dominated by anxiety when faced with health information, even though it may not directly prevent the onset of cyberchondria. This effectively reduces their tendency to avoid health information. These findings are consistent with previous studies on the effects of mindfulness, supporting the conclusion that mindfulness training can help individuals better accept current emotions and reduce fear of future uncertainties [85]. Furthermore, our findings emphasize the role of mindfulness training in regulating health information behavior, demonstrating that it not only improves the emotional state of older adults but also increases their acceptance of health information, thereby reducing avoidance behaviors. This highlights the potential of mindfulness training to effectively break the vicious cycle of "anxiety–information search–greater anxiety–information avoidance." Mindfulness can enable older adults to maintain a more balanced perspective when encountering health-related information, which prevents anxiety from escalating into avoidant behavior.

Implications and Future Directions

This Study Makes Three Theoretical Contributions

First, it makes significant contributions to SDT by deepening the understanding of how internal and external factors influence health behaviors through psychological mechanisms. It validates the mediating role of cyberchondria in health information avoidance among older adults and demonstrates how positive metacognition, health self-efficacy, and health information similarity interact within this process. This fills a critical gap in previous research by uncovering the mediating mechanisms of health information avoidance and provides new insights into the cognitive-emotional processes of older adults, an information-vulnerable population. Furthermore, this study extends the applicability of SDT by highlighting the dynamic balance between intrinsic motivation and external support in older adults navigating complex digital health environments. In addition, this study systematically explores the role of mindfulness interventions in managing health information behaviors. While mindfulness did not directly reduce cyberchondria, it significantly inhibited its progression to information avoidance. By enhancing older adults' capacity to cope with uncertainty and anxiety, mindfulness fosters a more open attitude toward health information. This expands mindfulness beyond emotional regulation to include its impact on complex health decision-making, providing theoretical support for its application in public health interventions. Finally, this study challenges the universal applicability of subjective norms in shaping health behaviors, as proposed in the theory of planned behavior [86]. It finds that older adults rely more on autonomous judgment than societal expectations due to their life experience and independence, emphasizing the importance of considering demographic differences and group heterogeneity in health behavior research.

Practically, This Study Offers Three Recommendations

First, for internal factors (eg, positive metacognition and excessive health self-efficacy), psychological education and cognitive behavioral therapy can help older adults reduce overfocus on health risks and avoid unnecessary health information searches [87]. Second, optimizing recommendation algorithms on health information platforms to reduce repetitive content and providing personalized information tailored to older adults' health conditions can alleviate anxiety caused by information overload and similarity. Finally, enhancing digital health literacy through health seminars and digital skills training, combined with mindfulness programs, can offer a comprehensive intervention strategy to address health information avoidance. These measures collectively aim to reduce health anxiety, improve engagement with health information, and enable older adults to achieve proactive health management.

Limitations

This study provides valuable insights into the causes of cyberchondria and health information avoidance among older adults and evaluates mindfulness meditation as an intervention. However, several limitations should be noted. First, the 4-week mindfulness intervention, while effective in the short term, does not address the sustainability of its effects. Future research should explore longer interventions and conduct follow-ups to assess lasting impacts. Combining mindfulness with approaches such as health education or social support may yield stronger and more enduring outcomes. In addition, while positive metacognition and health self-efficacy emerged as critical internal factors, specific intervention strategies for addressing these were not explored. Future research should investigate techniques such as cognitive behavioral therapy and health coaching to mitigate their negative impacts and further reduce health information avoidance. Expanding on these areas will help refine interventions and improve health management strategies for older adults.

Conclusion

This study used 2 substudies to explore the causes of cyberchondria and health information avoidance among older adults, as well as the intervention effects of mindfulness meditation. Study 1 constructed an influence model based on self-determination theory, using a cross-sectional survey to verify that internal factors such as positive metacognition (β=.26, *P*=.002) and health self-efficacy (β=.25, *P*<.001), as well as external factors such as health information similarity (β=.29, *P*<.001), were significant predictors of cyberchondria among older adults, which in turn led to health information avoidance behavior (β=.41, *P*<.001). However, subjective norms did not show a significant association with cyberchondria among older adults (β=–.11, *P*=.13). Study 2 conducted a 4-week mindfulness intervention, demonstrating the effectiveness of mindfulness in inhibiting the transition from cyberchondria to health information avoidance (effect=–0.357, *P*=.002, 95% CI –0.580 to –0.131]). However, it was also found that mindfulness did not directly reduce the occurrence of cyberchondria (*P*=.18).

Acknowledgments

The authors thank all the participants in this study. This research was funded by Minjiang University Research Start-up Funding (grant 324-32404314). The authors confirm that no generative AI tools, including ChatGPT or similar models, were used in the writing, editing, or preparation of this manuscript. All content was developed independently by the authors, based on original research and analysis.

Data Availability

The dataset of this study is accessible through science data bank website. In compliance with ethical guidelines and to protect participant privacy, all personally identifiable information has been removed from the dataset. Researchers are welcome to access the data through the provided link. For additional inquiries or specific questions about the dataset, please contact the corresponding author.

Authors' Contributions

CG contributed to conceptualization, methodology, software, formal analysis, data curation, writing-original draft preparation, writing-review and editing, visualization, and project administration. XZ handled methodology, validation, formal analysis, and resources. LQ managed validation, investigation, resources, visualization, and project administration.

Conflicts of Interest

None declared.

Multimedia Appendix 1

A: Demographic information B: Measures C: Study 2 Stimulus materials D: Checklist for Reporting Results of Internet E-Surveys (CHERRIES).

[DOCX File , 26 KB - publichealth_v11i1e69554_app1.docx]

References

1. Zhao YC, Zhao M, Song S. Online health information seeking behaviors among older adults: systematic scoping review. J Med Internet Res 2022;24(2):e34790 [FREE Full text] [doi: 10.2196/34790] [Medline: 35171099]
2. Cheng X, Huang X, Yang B, Xiao Y. Unveiling the paradox of technostress: Impacts of technology-driven stressors on the elderly's avoidance behaviors. Inf. Manag 2023;60(8):103875. [doi: 10.1016/j.im.2023.103875]
3. Barbour JB, Rintamaki LS, Ramsey JA, Brashers DE. Avoiding health information. J Health Commun 2012;17(2):212-229. [doi: 10.1080/10810730.2011.585691] [Medline: 22004015]
4. Sun Y, Feng Y, Shen X, Guo X. Fear appeal, coping appeal and mobile health technology persuasion: a two-stage scenario-based survey of the elderly. Inf. Technol. People 2022;36(1):362-386. [doi: 10.1108/itp-07-2021-0519]
5. Starcevic V, Berle D. Cyberchondria: towards a better understanding of excessive health-related Internet use. Expert Rev Neurother 2013;13(2):205-213. [doi: 10.1586/ern.12.162] [Medline: 23368807]
6. Vismara M, Varinelli A, Pellegrini L, Enara A, Fineberg NA. New challenges in facing cyberchondria during the coronavirus disease pandemic. Curr Opin Behav Sci 2022;46:101156 [FREE Full text] [doi: 10.1016/j.cobeha.2022.101156] [Medline: 35581995]
7. Zheng H, Chen X, Jiang S, Sun L. How does health information seeking from different online sources trigger cyberchondria? The roles of online information overload and information trust. Inf. Process. Manag 2023;60(4):103364. [doi: 10.1016/j.ipm.2023.103364]
8. Chae J. Who avoids cancer information? Examining a psychological process leading to cancer information avoidance. J Health Commun 2016;21(7):837-844. [doi: 10.1080/10810730.2016.1177144] [Medline: 27337343]
9. Linardon J. Rates of attrition and engagement in randomized controlled trials of mindfulness apps: Systematic review and meta-analysis. Behav Res Ther 2023;170:104421 [FREE Full text] [doi: 10.1016/j.brat.2023.104421] [Medline: 37862854]

10. Creswell JD. Mindfulness Interventions. Annu Rev Psychol 2017;68:491-516. [doi: 10.1146/annurev-psych-042716-051139] [Medline: 27687118]
11. Davodi SR, Zendehtalab H, Zare M, Behnam Vashani H. Effect of health promotion interventions in active aging in the elderly: a randomized controlled trial. Int J Community Based Nurs Midwifery 2023;11(1):34-43 [FREE Full text] [doi: 10.30476/IJCBNM.2022.96246.2117] [Medline: 36650846]
12. Howell JL, Shepperd JA. Reducing health-information avoidance through contemplation. Psychol Sci 2013;24(9):1696-1703. [doi: 10.1177/0956797613478616] [Medline: 23842956]
13. Leyva B, Taber JM, Trivedi AN. Medical care avoidance among older adults. J Appl Gerontol 2020;39(1):74-85. [doi: 10.1177/0733464817747415] [Medline: 29249189]
14. Case D, Andrews JE, Johnson JD, Allard SL. Avoiding versus seeking: the relationship of information seeking to avoidance, blunting, coping, dissonance, and related concepts. J Med Libr Assoc 2005;93(3):353-362 [FREE Full text] [Medline: 16059425]
15. Howell JL, Lipsey NP, Shepperd JA. Health information avoidance. The Wiley Encyclopedia of health psychology 2020:279-286. [doi: 10.1002/9781119057840.ch77]
16. Schenkel SK, Jungmann SM, Gropalis M, Witthöft M. Conceptualizations of cyberchondria and relations to the anxiety spectrum: systematic review and meta-analysis. J Med Internet Res 2021;23(11):e27835 [FREE Full text] [doi: 10.2196/27835] [Medline: 34792473]
17. Vismara M, Caricasole V, Starcevic V, Cinosi E, Dell'Osso B, Martinotti G, et al. Is cyberchondria a new transdiagnostic digital compulsive syndrome? A systematic review of the evidence. Compr Psychiatry 2020;99:152167 [FREE Full text] [doi: 10.1016/j.comppsych.2020.152167] [Medline: 32146315]
18. Tyrer P, Cooper S, Tyrer H, Wang D, Bassett P. Increase in the prevalence of health anxiety in medical clinics: possible cyberchondria. Int J Soc Psychiatry 2019;65(7-8):566-569 [FREE Full text] [doi: 10.1177/0020764019866231] [Medline: 31379243]
19. Bajcar B, Babiak J. Self-esteem and cyberchondria: The mediation effects of health anxiety and obsessive–compulsive symptoms in a community sample. Curr Psychol 2019;40(6):2820-2831. [doi: 10.1007/s12144-019-00216-x]
20. Fergus TA. Anxiety sensitivity and intolerance of uncertainty as potential risk factors for cyberchondria: a replication and extension examining dimensions of each construct. J Affect Disord 2015;184:305-309. [doi: 10.1016/j.jad.2015.06.017] [Medline: 26141807]
21. Bajcar B, Babiak J. Neuroticism and cyberchondria: The mediating role of intolerance of uncertainty and defensive pessimism. Pers. Individ. Differ 2020;162:110006. [doi: 10.1016/j.paid.2020.110006]
22. Khan AW, Pandey J. Dark side consequences of cyberchondria: an empirical investigation. Aslib J. Inf. Manag 2022;74(5):801-817. [doi: 10.1108/ajim-08-2021-0222]
23. Mathes BM, Norr AM, Allan NP, Albanese BJ, Schmidt NB. Cyberchondria: Overlap with health anxiety and unique relations with impairment, quality of life, and service utilization. Psychiatry Res 2018;261:204-211. [doi: 10.1016/j.psychres.2018.01.002] [Medline: 29324396]
24. Fergus TA, Spada MM. Cyberchondria: Examining relations with problematic Internet use and metacognitive beliefs. Clin Psychol Psychother 2017;24(6):1322-1330. [doi: 10.1002/cpp.2102] [Medline: 28621035]
25. Brown RJ, Skelly N, Chew-Graham CA. Online health research and health anxiety: A systematic review and conceptual integration. Clin. Psychol.: Sci. Pract 2020;27(2):e12299. [doi: 10.1111/cpsp.12299]
26. Starcevic V. Keeping Dr. Google under control: how to prevent and manage cyberchondria. World Psychiatry 2023;22(2):233-234 [FREE Full text] [doi: 10.1002/wps.21076] [Medline: 37159368]
27. Deci EL, Ryan RM. Self-determination theory. Handbook of theories of social psychology 2012;1(20):416-436. [doi: 10.4135/9781446249215.n21]
28. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
29. Gillison FB, Rouse P, Standage M, Sebire SJ, Ryan RM. A meta-analysis of techniques to promote motivation for health behaviour change from a self-determination theory perspective. Health Psychol Rev 2019;13(1):110-130. [doi: 10.1080/17437199.2018.1534071] [Medline: 30295176]
30. Assadi V, Hassanein K. Consumer adoption of personal health record systems: a self-determination theory perspective. J Med Internet Res 2017;19(7):e270 [FREE Full text] [doi: 10.2196/jmir.7721] [Medline: 28751301]
31. Ali F, Zhang Q, Tauni MZ, Butt AH, Ahsan T. Contingent self-esteem, social interaction fears, and compulsive WeChat usage. Behav. Inf. Technol 2022;42(6):1-12. [doi: 10.1080/0144929x.2022.2039768]
32. Li Y, Wang X. Seeking health information on social media: a perspective of trust, self-determination, and social support. j organ end user com 2018;30(1):1-22. [doi: 10.4018/JOEUC.2018010101]
33. Laato S, Islam AKMN, Islam MN, Whelan E. What drives unverified information sharing and cyberchondria during the COVID-19 pandemic? Eur. J. Inf. Syst 2020;29(3):288-305. [doi: 10.1080/0960085x.2020.1770632]
34. Yalçın İ, Boysan M, Eşkisu M, Çam Z. Health anxiety model of cyberchondria, fears, obsessions, sleep quality, and negative affect during COVID-19. Curr Psychol 2024;43(9):8502-8519 [FREE Full text] [doi: 10.1007/s12144-022-02987-2]

35. Zhu X, Zheng T, Ding L, Zhang X. Exploring associations between eHealth literacy, cyberchondria, online health information seeking and sleep quality among university students: A cross-section study. Heliyon 2023;9(6):e17521 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e17521] [Medline: 37408886]
36. Sweeny K, Melnyk D, Miller W, Shepperd JA. Information avoidance: who, what, when, and why. Rev. Gen. Psychol 2010;14(4):340-353. [doi: 10.1037/a0021288]
37. McMullan RD, Berle D, Arnáez S, Starcevic V. The relationships between health anxiety, online health information seeking, and cyberchondria: Systematic review and meta-analysis. J Affect Disord 2019;245:270-278. [doi: 10.1016/j.jad.2018.11.037] [Medline: 30419526]
38. Spada MM, Nikcević AV, Moneta GB, Wells A. Metacognition as a mediator of the relationship between emotion and smoking dependence. Addict Behav 2007;32(10):2120-2129. [doi: 10.1016/j.addbeh.2007.01.012] [Medline: 17307299]
39. Nadeem F, Malik N, Atta M, Ullah I, Martinotti G, Pettorruso M, et al. Relationship between health-anxiety and cyberchondria: role of metacognitive beliefs. J Clin Med 2022;11(9):2590 [FREE Full text] [doi: 10.3390/jcm11092590] [Medline: 35566713]
40. Kaur A, Butow PN, Thewes B. Do metacognitions predict attentional bias in health anxiety? Cogn Ther Res 2011;35(6):575-580. [doi: 10.1007/s10608-011-9387-6]
41. Starcevic V, Baggio S, Berle D, Khazaal Y, Viswasam K. Cyberchondria and its relationships with related constructs: a network analysis. Psychiatr Q 2019;90(3):491-505 [FREE Full text] [doi: 10.1007/s11126-019-09640-5] [Medline: 31098922]
42. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
43. Ajzen I. The theory of planned behavior. Organ. Behav. Hum. Decis. Process 1991;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
44. Aminuddin HB, Jiao N, Jiang Y, Hong J, Wang W. Effectiveness of smartphone-based self-management interventions on self-efficacy, self-care activities, health-related quality of life and clinical outcomes in patients with type 2 diabetes: a systematic review and meta-analysis. Int J Nurs Stud 2021;116:103286. [doi: 10.1016/j.ijnurstu.2019.02.003] [Medline: 30827741]
45. Simonetti V, Durante A, Ambrosca R, Arcadi P, Graziano G, Pucciarelli G, et al. Anxiety, sleep disorders and self-efficacy among nurses during COVID-19 pandemic: a large cross-sectional study. J Clin Nurs 2021;30(9-10):1360-1371 [FREE Full text] [doi: 10.1111/jocn.15685] [Medline: 33534934]
46. Lee SY, Hwang H, Hawkins R, Pingree S. Interplay of negative emotion and health self-efficacy on the use of health information and its outcomes. Commun. Res 2008;35(3):358-381. [doi: 10.1177/0093650208315962]
47. Essery R, Geraghty AWA, Kirby S, Yardley L. Predictors of adherence to home-based physical therapies: a systematic review. Disabil Rehabil 2017;39(6):519-534. [doi: 10.3109/09638288.2016.1153160] [Medline: 27097761]
48. Liang H, Fu K. Information overload, similarity, and redundancy: unsubscribing information sources on Twitter. J Comput-Mediat Comm 2016;22(1):1-17. [doi: 10.1111/jcc4.12178]
49. Cinelli M, De Francisci Morales G, Galeazzi A, Quattrociocchi W, Starnini M. The echo chamber effect on social media. Proc Natl Acad Sci U S A 2021;118(9):e2023301118 [FREE Full text] [doi: 10.1073/pnas.2023301118] [Medline: 33622786]
50. Malloch YZ, Zhang J. Seeing others receive support online: effects of self-disclosure and similarity on perceived similarity and health behavior intention. J Health Commun 2019;24(3):217-225. [doi: 10.1080/10810730.2019.1595226] [Medline: 30912708]
51. Deng Z, Deng Z, Fan G, Wang B, Fan W, Liu S. More is better? Understanding the effects of online interactions on patients health anxiety. Asso for Info Science & Tech 2023;74(11):1243-1264. [doi: 10.1002/asi.24822]
52. Griffin RJ, Dunwoody S, Neuwirth K. Proposed model of the relationship of risk information seeking and processing to the development of preventive behaviors. Environ Res 1999;80(2 Pt 2):S230-S245. [doi: 10.1006/enrs.1998.3940] [Medline: 10092438]
53. Zheng H, Kyung Kim H, Joanna Sin S, Theng Y. A theoretical model of cyberchondria development: Antecedents and intermediate processes. Telemat Inform 2021;63:101659. [doi: 10.1016/j.tele.2021.101659]
54. Honora A, Wang K, Chih W. How does information overload about COVID-19 vaccines influence individuals' vaccination intentions? The roles of cyberchondria, perceived risk, and vaccine skepticism. Comput Human Behav 2022;130:107176 [FREE Full text] [doi: 10.1016/j.chb.2021.107176] [Medline: 35013641]
55. Inzlicht M, Schmeichel BJ. What is ego depletion? Toward a mechanistic revision of the resource model of self-control. Perspect Psychol Sci 2012;7(5):450-463. [doi: 10.1177/1745691612454134] [Medline: 26168503]
56. Reinecke L, Hartmann T, Eden A. The guilty couch potato: the role of ego depletion in reducing recovery through media use. J Commun 2014;64(4):569-589. [doi: 10.1111/jcom.12107]
57. Ma J, Zhao D, Xu N, Yang J. The effectiveness of immersive virtual reality (VR) based mindfulness training on improvement mental-health in adults: a narrative systematic review. Explore (NY) 2023;19(3):310-318 [FREE Full text] [doi: 10.1016/j.explore.2022.08.001] [Medline: 36002363]
58. Shapiro SL, Carlson LE, Astin JA, Freedman B. Mechanisms of mindfulness. J Clin Psychol 2006;62(3):373-386. [doi: 10.1002/jclp.20237] [Medline: 16385481]

59. Cullen M. Mindfulness-based interventions: an emerging phenomenon. Mindfulness 2011;2(3):186-193. [doi: 10.1007/s12671-011-0058-1]
60. Hathaisaard C, Wannarit K, Pattanaseri K. Mindfulness-based interventions reducing and preventing stress and burnout in medical students: a systematic review and meta-analysis. Asian J Psychiatr 2022;69:102997. [doi: 10.1016/j.ajp.2021.102997] [Medline: 34995839]
61. Toivonen KI, Zernicke K, Carlson LE. Web-based mindfulness interventions for people with physical health conditions: systematic review. J Med Internet Res 2017;19(8):e303 [FREE Full text] [doi: 10.2196/jmir.7487] [Medline: 28860106]
62. Yogeswaran V, El Morr C. Effectiveness of online mindfulness interventions on medical students' mental health: a systematic review. BMC Public Health 2021;21(1):2293 [FREE Full text] [doi: 10.1186/s12889-021-12341-z] [Medline: 34920715]
63. Howell JL, Shepperd JA. Establishing an information avoidance scale. Psychol Assess 2016;28(12):1695-1708. [doi: 10.1037/pas0000315] [Medline: 27099980]
64. Jokić-Begić N, Mikac U, Čuržik D, Sangster Jokić C. The development and validation of the short Cyberchondria Scale (SCS). J Psychopathol Behav Assess 2019;41(4):662-676. [doi: 10.1007/s10862-019-09744-z]
65. Bailey R, Wells A. Development and initial validation of a measure of metacognitive beliefs in health anxiety: The MCQ-HA. Psychiatry Res 2015;230(3):871-877. [doi: 10.1016/j.psychres.2015.11.035] [Medline: 26626951]
66. Sheer VC. A meta-synthesis of health-related self-efficacy instrumentation: problems and suggestions. J Nurs Meas 2014;22(1):77-93. [doi: 10.1891/1061-3749.22.1.77] [Medline: 24851665]
67. Ajzen I. Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. J Applied Social Pyschol 2006;32(4):665-683. [doi: 10.1111/j.1559-1816.2002.tb00236.x]
68. Walsh G, Mitchell V. Consumer vulnerability to perceived product similarity problems: scale development and identification. J. Macromarketing 2005;25(2):140-152. [doi: 10.1177/0276146705280613]
69. Hair JF, Sarstedt M, Ringle CM, Mena JA. An assessment of the use of partial least squares structural equation modeling in marketing research. J. of the Acad. Mark. Sci 2011;40(3):414-433. [doi: 10.1007/s11747-011-0261-6]
70. Hair JF, Ringle CM, Gudergan SP, Fischer A, Nitzl C, Menictas C. Partial least squares structural equation modeling-based discrete choice modeling: an illustration in modeling retailer choice. Bus Res 2018;12(1):115-142. [doi: 10.1007/s40685-018-0072-4]
71. Fornell C, Larcker DF. Structural equation models with unobservable variables and measurement error: algebra and statistics. J. Mark. Res 1981;18(3):382-388. [doi: 10.2307/3150980]
72. Podsakoff PM, MacKenzie SB, Lee J, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
73. Hair JF, Ringle CM, Sarstedt M. PLS-SEM: Indeed a silver bullet. J. Mark. Theory Pract 2014;19(2):139-152. [doi: 10.2753/MTP1069-6679190202]
74. Dijkstra TK, Henseler J. Consistent and asymptotically normal PLS estimators for linear structural equations. Comput Stat Data Anal 2015;81:10-23. [doi: 10.1016/j.csda.2014.07.008]
75. Hair JF, Risher JR, Sarstedt M, Ringle CM. When to use and how to report the results of PLS-SEM. Eur. Bus. Rev 2019;31(1):2-24. [doi: 10.1108/ebr-11-2018-0203]
76. Cohen J. Statistical power analysis for the behavioral science. Cambridge: Academic Press; 2013.
77. Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. J Pers Soc Psychol 2003;84(4):822-848. [doi: 10.1037/0022-3514.84.4.822] [Medline: 12703651]
78. Kabat-Zinn J. Mindfulness-based interventions in context: past, present, and future. Clinical Psychology: science and practice 2003;10(2):144-156. [doi: 10.1093/clipsy.bpg016]
79. Rooks JD, Morrison AB, Goolsarran M, Rogers SL, Jha AP. “We Are Talking About Practice”: the influence of mindfulness vs. relaxation training on athletes‘ attention and well-being over high-demand intervals. J Cogn Enhanc 2017;1(2):141-153. [doi: 10.1007/s41465-017-0016-5]
80. Casale S, Rugai L, Fioravanti G. Exploring the role of positive metacognitions in explaining the association between the fear of missing out and social media addiction. Addict Behav 2018;85:83-87. [doi: 10.1016/j.addbeh.2018.05.020] [Medline: 29864680]
81. Muris P. Relationships between self-efficacy and symptoms of anxiety disorders and depression in a normal adolescent sample. Pers. Individ. Differ 2002;32(2):337-348. [doi: 10.1016/S0191-8869(01)00027-7]
82. Ahmed A, Aziz S, Toro CT, Alzubaidi M, Irshaidat S, Serhan HA, et al. Machine learning models to detect anxiety and depression through social media: a scoping review. Comput Methods Programs Biomed Update 2022;2:100066 [FREE Full text] [doi: 10.1016/j.cmpbup.2022.100066] [Medline: 36105318]
83. Ahn J, Kahlor LA. No regrets when it comes to your health: anticipated regret, subjective norms, information insufficiency and intent to seek health information from multiple sources. Health Commun 2020;35(10):1295-1302. [doi: 10.1080/10410236.2019.1626535] [Medline: 31177829]
84. Heid AR, Zarit SH, Fingerman KL. "My Parent is so stubborn!"-perceptions of aging parents' persistence, insistence, and resistance. J Gerontol B Psychol Sci Soc Sci 2016;71(4):602-612 [FREE Full text] [doi: 10.1093/geronb/gbu177] [Medline: 25573152]

85. Tudor K, Maloney S, Raja A, Baer R, Blakemore S, Byford S, MYRIAD Team, et al. Universal mindfulness training in schools for adolescents: a scoping review and conceptual model of moderators, mediators, and implementation factors. Prev Sci 2022;23(6):934-953 [FREE Full text] [doi: 10.1007/s11121-022-01361-9] [Medline: 35267177]
86. Korn L, Haynie DL, Luk JW, Sita K, Simons-Morton BG. Attitudes, subjective norms, and perceived behavioral control associated with age of first use of cannabis among adolescents. J Sch Health 2021;91(1):50-58. [doi: 10.1111/josh.12977] [Medline: 33216363]
87. Podina IR, Cosmoiu A, Rusu P, Chivu AC. Positive thinking is not adaptive thinking: a cognitive-behavioral take on interpretation bias modification for social anxiety. J Rat-Emo Cognitive-Behav Ther 2020;38(3):424-444. [doi: 10.1007/s10942-020-00344-5]

Abbreviations

ANCOVA: analysis of covariance
AVE: average variance extracted
CR: composite reliability
MBSR: mindfulness-based stress reduction
PLS: partial least squares
SDT: self-determination theory
VIF: variance inflation factor

Edited by A Mavragani; submitted 02.12.24; peer-reviewed by S You, Q Duan; comments to author 25.12.24; revised version received 29.12.24; accepted 30.12.24; published 28.01.25.

Please cite as:
Gu C, Qian L, Zhuo X
Mindfulness Intervention for Health Information Avoidance in Older Adults: Mixed Methods Study
JMIR Public Health Surveill 2025;11:e69554
URL: https://publichealth.jmir.org/2025/1/e69554
doi:10.2196/69554
PMID:

Original Paper

Alternative Presentations of Overall and Statistical Uncertainty for Adults' Understanding of the Results of a Randomized Trial of a Public Health Intervention: Parallel Web-Based Randomized Trials

Christine Holst[1,2], PhD; Steven Woloshin[2,3], MD, MS; Andrew D Oxman[1,2], MD; Christopher Rose[1,4], PhD; Sarah Rosenbaum[1,2], PhD; Heather Menzies Munthe-Kaas[1,2], PhD

[1]Centre for Epidemic Interventions Research, Norwegian Institute of Public Health, Oslo, Norway

[2]Lisa Schwartz Foundation for Truth in Medicine, Norwich, VT, United States

[3]Dartmouth Institute for Health Policy and Clinical Practice, Geisel Medical School, West Lebanon, NH, United States

[4]Cluster for Reviews and Health Technology Assessments, Division for Health Services, Norwegian Institute of Public Health, Oslo, Norway

Corresponding Author:
Christine Holst, PhD
Centre for Epidemic Interventions Research
Norwegian Institute of Public Health
Myrens Verksted 3L
Oslo, 0213
Norway
Phone: 47 48234044
Email: Christine.Holst@fhi.no

Abstract

Background: Well-designed public health messages can help people make informed choices, while poorly designed messages or persuasive messages can confuse, lead to poorly informed decisions, and diminish trust in health authorities and research. Communicating uncertainties to the public about the results of health research is challenging, necessitating research on effective ways to disseminate this important aspect of randomized trials.

Objective: This study aimed to evaluate people's understanding of overall and statistical uncertainty when presented with alternative ways of expressing randomized trial results.

Methods: Two parallel, web-based, individually randomized trials (3×2 factorial designs) were conducted in the United States and Norway. Participants were randomized to 1 of 6 versions of a text (summary) communicating results from a study examining the effects of wearing glasses to prevent COVID-19 infection. The summaries varied in how *overall uncertainty* ("Grading of Recommendations Assessment, Development and Evaluation [GRADE] language," "plain language," or "no explicit language") and *statistical uncertainty* (whether a margin of error was shown or not) were presented. Participants completed a web-based questionnaire exploring 4 coprimary outcomes: 3 to measure understanding of overall uncertainty (benefits, harms, and sufficiency of evidence), and one to measure statistical uncertainty. Participants were adults who do not wear glasses recruited from web-based research panels in the United States and Norway. Results of the trials were analyzed separately and combined in a meta-analysis.

Results: In the US and Norwegian trials, 730 and 497 individuals were randomized, respectively; data for 543 (74.4%) and 452 (90.9%) were analyzed. More participants had a correct understanding of uncertainty when presented with plain language (United States: 37/99, 37% and Norway: 40/76, 53%) than no explicit language (United States: 18/86, 21% and Norway: 34/80, 42%). Similar positive effect was seen for the GRADE language in the United States (26/79, 33%) but not in Norway (30/71, 42%). There were only small differences between groups for understanding the uncertainty of harms. Plain language improved correct understanding of evidence sufficiency (odds ratio 2.05, 95% CI 1.17-3.57), compared to no explicit language. The effect of GRADE language was inconclusive (odds ratio 1.34, 95% CI 0.79-2.28). The understanding of statistical uncertainty was improved when the participants were shown the margin of error compared to not being shown: Norway: 16/75, 21% to 24/71, 34% vs 1/71, 1% to 2/76, 3% and the United States: 21/101, 21% to 32/90, 36% vs 0/86, 0% to 3/79, 4%).

Conclusions: Plain language, but not GRADE language, was better than no explicit language in helping people understand overall uncertainty of benefits and harms. Reporting margin of error improved understanding of statistical uncertainty around the effect of wearing glasses, but only for a minority of participants.

Trial Registration: ClinicalTrials.gov NCT05642754; https://tinyurl.com/4mhjsm7s

(JMIR Public Health Surveill 2025;11:e62828) doi:10.2196/62828

KEYWORDS

communication; Grading of Recommendations Assessment, Development, and Evaluation language; GRADE language; statistical uncertainty; overall uncertainty; randomized trial

Introduction

Background

Public health messaging matters—it shapes how people understand important health risks and what can be done to mitigate them. Ideally, such messages are based on findings from robust research, but even important messages based on solid research can fail to properly reach the target audience when poorly communicated, as was seen during the COVID-19 pandemic [1]. Decision scientists have articulated basic principles for effective health communication, such as using simple and familiar wording, using clear visual design, presenting structured comparisons of alternatives, and careful testing in the target audiences [2-5].

A recent randomized trial assessed the effect of these principles on communication effectiveness in the context of COVID-19 home test kit instructions using a real example [6]. The trial showed that individuals randomized to instructions that did not follow best decision science principles (ie, the actual instructions authorized by the Food and Drug Administration), compared to those that did (ie, carefully pretested intervention instructions), were more likely to fail to quarantine when quarantine was the right choice (33% vs 14% failed; 95% CI for the 19% difference being 6%-31%).

Evidence from randomized trials documents the importance of both the format and content of health messaging. Formatting examples include how the use of percentages (eg, 10%) versus frequency formats (eg, 10 in 100) can improve comprehension [7] and how absolute versus relative risk measures for communicating treatment effects are better understood [3,8] and help people make decisions more consistent with their values [9].

Content examples include the importance of presenting both benefits and harms when describing interventions [10] and highlighting study limitations, such as how simple nondirective explanations about surrogate outcomes and newly approved drugs enhance evidence-based decision-making about prescription drugs [11,12].

Furthermore, there is evidence supporting the importance of communication about the uncertainty of research findings in both research summaries aimed at professionals [13-16] and plain language summaries for the public [17]. This includes both *statistical uncertainty* (ie, imprecision) [16] and the *overall uncertainty*, due to the risk of bias, inconsistency, indirectness, and publication bias [13,15]. However, there are still open questions about the best formats and language for presenting both kinds of uncertainty, and how people understand and react to such information.

A recent pair of trials [18] found that including “quality cues” in communications tempered the publics’ tendency to assume that the quality of evidence presented without such cues is high (when it was not), and reporting that evidence quality was low decreased trust, perception of intervention efficacy, and the likelihood of adopting it.

The certainty of the evidence can affect the decisions that people make. If the purpose of a message is to inform people rather than to persuade them [19], it is necessary to include information about the degree and source of uncertainty related to the effect estimate. Not doing so can be misleading.

This study is the first of several planned studies to evaluate strategies to improve communication of research findings (Message Lab). The goal of Message Lab is to develop and promote best practices in message development (ie, attending to the foregoing communication principles and evidence); facilitate user testing; and conduct randomized trials assessing the effects of public health messages on the public’s understanding of the messages, beliefs, decisions, and behaviors. Therefore, this study is also designed as a proof-of-concept exercise to develop and test a method to efficiently and effectively evaluate communication strategies intended to summarize the results of randomized trials using web-based trial platforms.

Objectives

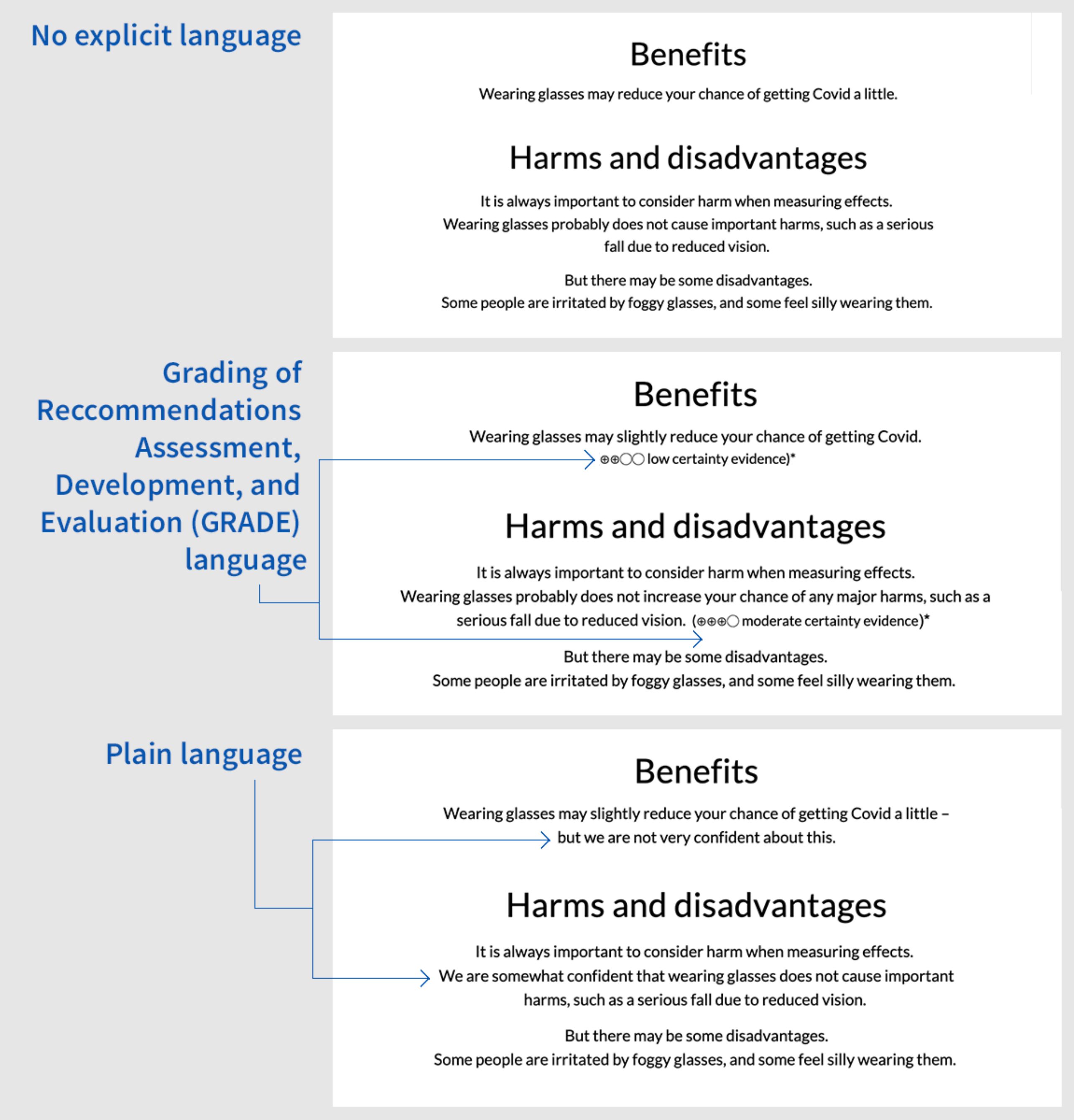
This study aimed to evaluate the effect of alternative formats for communicating overall and statistical uncertainty on the public’s understanding of uncertainty and the sufficiency of evidence. The objectives of the study were (1) to compare the effects of 3 ways of communicating the overall uncertainty of the effect of wearing glasses on reducing the chance of acquiring COVID-19, and (2) to compare the effects of including the margin of error (MOE; also called CI) compared to not including it.

Methods

Design

We designed a web-based, parallel-group, individually randomized, pragmatic trial to compare the effects of different ways of communicating uncertainty when reporting the results of a randomized trial to the public. The trial was prespecified, registered at ClinicalTrials.gov (NCT05642754) Zenodo (7428981), and, except as noted, conducted following a published protocol [20]. We used a published trial assessing the effect of wearing glasses on the risk of being infected with COVID-19 as our example [21]. We used a 3×2 factorial trial design because we were interested in the effects of presenting overall uncertainty in each of 3 ways (ie, Grading of Recommendations Assessment, Development and Evaluation [GRADE] language, plain language, or no explicit language) combined with the effects of presenting or not presenting statistical uncertainty (in the form of an MOE). This resulted in 6 (3×2) groups differing in how the COVID-19 study was summarized (Figures 1 and 2).

Figure 1. Three alternative versions presented to participants in the US trial communicating overall uncertainty about the possible benefits and harms of wearing glasses to reduce the chances of acquiring COVID-19.



No explicit language

Benefits

Wearing glasses may reduce your chance of getting Covid a little.

Harms and disadvantages

It is always important to consider harm when measuring effects.
Wearing glasses probably does not cause important harms, such as a serious fall due to reduced vision.

But there may be some disadvantages.
Some people are irritated by foggy glasses, and some feel silly wearing them.

Grading of Reccommendations Assessment, Development, and Evaluation (GRADE) language

Benefits

Wearing glasses may slightly reduce your chance of getting Covid.
⊕⊕◯◯ low certainty evidence)*

Harms and disadvantages

It is always important to consider harm when measuring effects.
Wearing glasses probably does not increase your chance of any major harms, such as a serious fall due to reduced vision. (⊕⊕⊕◯ moderate certainty evidence)*

But there may be some disadvantages.
Some people are irritated by foggy glasses, and some feel silly wearing them.

Plain language

Benefits

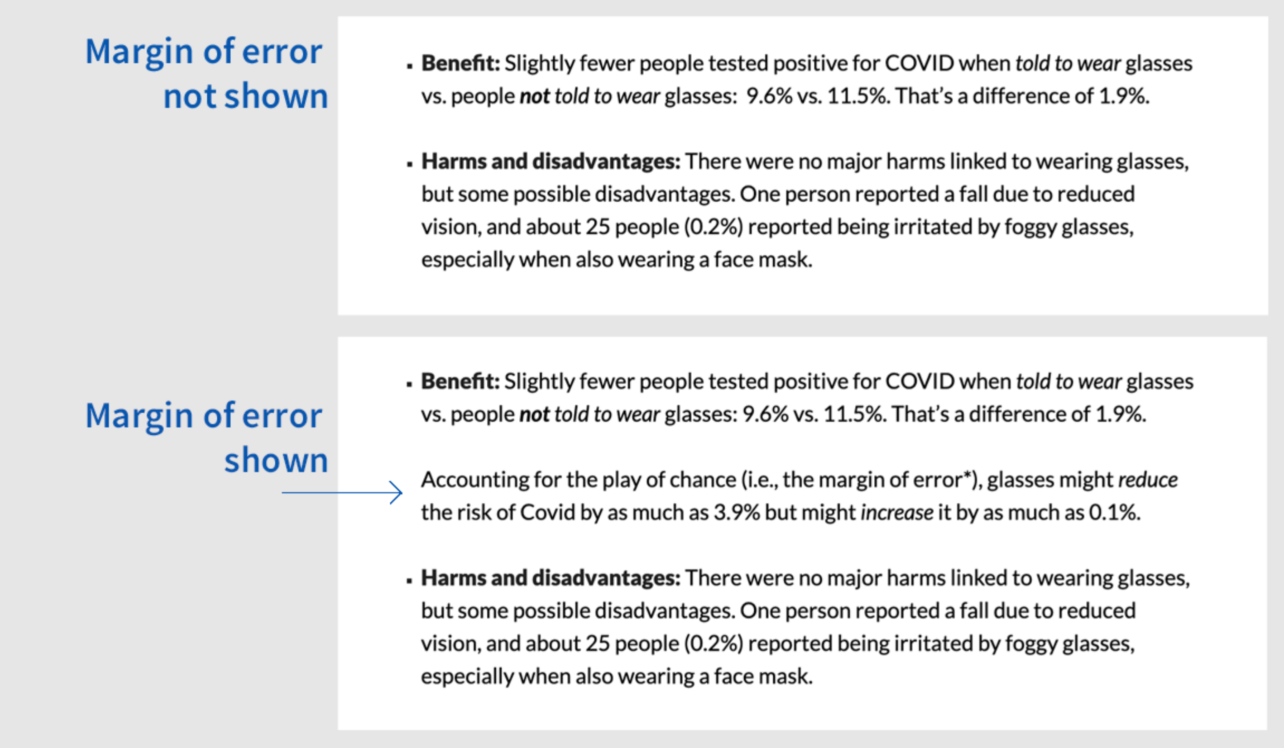
Wearing glasses may slightly reduce your chance of getting Covid a little – but we are not very confident about this.

Harms and disadvantages

It is always important to consider harm when measuring effects.
We are somewhat confident that wearing glasses does not cause important harms, such as a serious fall due to reduced vision.

But there may be some disadvantages.
Some people are irritated by foggy glasses, and some feel silly wearing them.

Figure 2. Two alternative versions presented to participants in the US trial communicating statistical uncertainty (margin of error).



Margin of error not shown

- **Benefit:** Slightly fewer people tested positive for COVID when *told to wear* glasses vs. people ***not*** *told to wear* glasses: 9.6% vs. 11.5%. That's a difference of 1.9%.
- **Harms and disadvantages:** There were no major harms linked to wearing glasses, but some possible disadvantages. One person reported a fall due to reduced vision, and about 25 people (0.2%) reported being irritated by foggy glasses, especially when also wearing a face mask.

Margin of error shown

- **Benefit:** Slightly fewer people tested positive for COVID when *told to wear* glasses vs. people ***not*** *told to wear* glasses: 9.6% vs. 11.5%. That's a difference of 1.9%.

  Accounting for the play of chance (i.e., the margin of error*), glasses might *reduce* the risk of Covid by as much as 3.9% but might *increase* it by as much as 0.1%.
- **Harms and disadvantages:** There were no major harms linked to wearing glasses, but some possible disadvantages. One person reported a fall due to reduced vision, and about 25 people (0.2%) reported being irritated by foggy glasses, especially when also wearing a face mask.

Recruitment of Participants

We implemented the design in 2 nearly identical trials in Norway and the United States in April 2023 and May 2023. For each trial, we used quota sampling from a web-based platform (Prolific; Prolific Academic Ltd) of volunteer research participants in the US trial [22] and from an independent commercial research agency (Opinion) [23] with a panel of 120,000 people living in Norway. All participants were thus part of established research panels, which provide recruitment and management of participants for online research. Participants were eligible if they were aged ≥18 years and said that they did not regularly wear glasses (based on a prescreen feature in Prolific and a screener question in Opinion). In Norway, 2667 persons received an invitation from Opinion to participate, while the study was open to 34,242 eligible persons in the Prolific platform (see the Results section for more details). Participants were literate in English (used in the US version) or Norwegian (used in the Norwegian version). Both platforms applied processes to prevent bots and fraudulent participants [24], and to further increase data quality, we built in attention and comprehension checks [25].

Randomization

Eligible participants were randomized in a 1:1:1 ratio to 1 of the 6 comparison groups (described above). In the United States, participants were randomized via the web-based questionnaire after clicking on the link to the summary. In Norway, the participants were randomized to each summary by the research agency.

Interventions and Control

Overview

The alternative versions of the presentation of overall and statistical uncertainty included in the 6 summaries are provided in Figures 1 and 2, and the full summaries are provided in Multimedia Appendix 1. Each version included one of the alternatives for communicating overall uncertainty about the benefits and harms (ie, GRADE, plain, or no explicit language) and either included or did not include a presentation of the MOE:

- Version 1: no explicit language and MOE not shown (reference or control version)
- Version 2: no explicit language and MOE shown
- Version 3: plain language and MOE not shown
- Version 4: plain language and MOE shown
- Version 5: GRADE language and MOE not shown
- Version 6: GRADE language and MOE shown

Development and User Testing of the Interventions

The summaries were drafted in English by 4 researchers (SW, AO, HMK, SR) and translated to Norwegian by 2 researchers (IHE, CH). Note that the summaries underwent multiple rounds of user testing and modifications as needed before the trial. We used human-centered design methods to develop the summaries to present to the participants [26]. Initially, we invited 4 native English speakers and 3 native Norwegian speakers to unmoderated user testing using Loop11, a digital user experience platform [27]. In total, 7 user test participants (n=5, 71% female and n=2, 29% male; n=2, 29% with secondary school education and n=5, 71% with postsecondary education; n=2, 29% living in Canada and n=5, 71% living in Norway; and n=3, 43% aged >55 years and n=4, 57% aged between 30 and 45 years) were introduced to the project and asked to imagine themselves in the following scenario: "Imagine you hear that glasses may reduce your chance of becoming infected with COVID. You go online to find out more information and find a website that says... [insert COVID trial summary version 1, 2 or 3].”

They were then asked to read 5 pieces of text: 3 versions of the summaries with progressively more information (eg, the first version was the shortest summary with no text related to uncertainty or the MOE, the second version had text related to uncertainty but not the MOE, and the third version had text related to uncertainty and the MOE); the trial questions; and the text that would be available via hyperlinks and hover text in the summaries. For each piece of text, the user test participants w asked for their first set of impressions and then a series of follow-up questions related to content, font, format, language, and anything else.

We revised the summaries according to feedback. The translations were reviewed by a third researcher. The English summaries were presented to 1 more participant in English (moderated, in-person user experience interview) using the same format as the first 4 interviews. The summaries were subsequently revised, and changes were also made to the Norwegian summaries.

We gathered feedback on the Norwegian summaries from colleagues and conducted 2 in-person user experience interviews with 2 native Norwegian speakers using the same question guide as the English user experience interviews. Revisions were made according to the feedback, and where appropriate, these revisions were back translated into the English versions of the COVID-19 trial summaries.

Furthermore, we gathered feedback from the user experience participants on other materials related to the process of participating in the randomized trial (eg, from the invitation to participate to the text sent after participants completed the questionnaire).

Outcomes

We defined 3 binary coprimary outcomes to measure understanding of overall uncertainty and 1 coprimary outcome to measure understanding of statistical uncertainty (precision of the effect estimate for the benefit). Each of the 3 outcomes for overall uncertainty were measured by comparing participants’ answers to questions (Table 1) about uncertainty to expected (correct) answers based on the size of the effects and the certainty of the evidence (the full questionnaire is provided in the Multimedia Appendix 2).

The 3 *coprimary outcomes* were as follows (all used 4-point ordinal response sets):

- Understanding of the uncertainty of the benefit (question 9)
- Understanding of the sufficiency of the evidence (question 15)
- Understanding of the uncertainty of important harms (question 12)

An additional coprimary outcome was included for the 3 versions reporting the MOE, assessing the understanding of statistical uncertainty (ie, the precision of the effect estimates: choose which among the 4 statements was most consistent with the information provided, eg, question 10: “wearing glasses may *reduce* the chance of COVID a little, but might *reduce it a lot*”).

Secondary outcomes included questions (Table 1) regarding the following:

- The perceived benefit and harms of wearing glasses to reduce the chance of acquiring COVID-19 infections (questions 8 and 11)
- Intended behavior (whether participants would wear glasses to reduce the chance of acquiring COVID-19 in areas with high and low COVID-19 infection rates; questions 3 and 7)
- Perceptions of the information; trustworthiness (question 13), sufficiency (question 14), clarity (questions 17 and 18), and helpfulness (question 19); and the likelihood of sharing it with others (question 20)
- Decisional conflict (questions 4 and 5) [28]

Table 1. Outcomes and the corresponding questionnaire items.

Outcome	Questionnaire
Primary outcomes	
Understanding uncertainty	- Question 9. How sure are you about the effect of wearing glasses on your chance of getting COVID?
Understanding uncertainty of important harms	- Question 12. How sure are you about whether wearing glasses to reduce COVID can cause important harms?
Understanding sufficiency	- Question 15. Not enough is known to be sure about the effects of wearing glasses to reduce the chance of getting COVID.
Understanding statistical uncertainty	- Question 10. Which of the following statements is most consistent with the information provided?
Secondary outcomes	
Perceived benefit and harm	- Question 8. What is the possible effect of wearing glasses on your chance of getting COVID? - Question 11. How likely do you think it is that wearing glasses to reduce the chance of getting COVID can cause any important harms?
Intended behavior	- Question 3. If there were a surge of COVID-19 cases in your area, how likely would you be to wear glasses or recommend wearing glasses to reduce the chance of getting COVID? - Question 7. If there were very few COVID cases in your area, how likely would you be to wear glasses or recommend wearing glasses to reduce the chance of getting COVID?
Perceptions of the information	- Question 13. This information seems like a trustworthy summary of what is known about the effects of wearing glasses to reduce the chance of getting COVID. - Question 14. The summary gives me enough information to understand what is known about the effects of wearing glasses to reduce the chance of getting COVID. - Question 17. I think the information about whether wearing glasses affects the chance of getting COVID was... - Question 18. I think the information about whether wearing glasses to prevent COVID has important harms was... - Question 19. If you were making a decision about wearing glasses to prevent COVID, would you find the information we showed you helpful? - Question 20. Say you knew someone who heard that wearing glasses might affect your chance of getting COVID. How likely would you be to share the information you just saw with them?
Decisional conflict	- Question 4. The answer about wearing glasses if there were a surge of COVID was hard for me to give. - Question 5. The information in the summary helped me make an informed decision about wearing glasses if there were a surge of COVID.

Data Collection

In the United States, we collected data using SurveyMonkey (SurveyMonkey Inc). Participants were directed to the platform after agreeing to participate in the study. They were asked to enter their personal identifying number (as a panel member on the relevant recruiting platform). In Norway, Opinion collected the data in their own panel via an email invitation and a link to the survey on their own platform. After reading initial information, the participants were randomized to read one of the 6 summaries; they were asked to answer 18 questions about the summary; 4 questions about themselves (eg, whether they wore glasses, what was their highest level of education, and what was their concern regarding COVID-19) to assess saliency of the scenario; and 3 questions to assess their numeracy (Multimedia Appendix 2). They submitted their responses electronically, using the "submit" button (refer to Statistical Analysis section).

Statistical Analysis

The US and Norwegian trials were analyzed in the same way, except as noted. We excluded participants who completed the survey in <3 minutes, reported regularly wearing glasses, or failed the attention checks. The attention checks (questions 6 and 21) included 2 true or false questions to verify the consistency of participants' responses, ensuring that they were paying attention and not just randomly answering. Data were duplicated for some US participants who appeared to have submitted the same responses multiple times in a short period. We assumed that these participants clicked the submit button several times or refreshed their browsers, so we analyzed only the first data submitted by these participants. All analyses were performed before unblinding as prespecified according to the intention-to-treat principle—all randomized participants meeting the inclusion criteria were included and analyzed in the arms to which they were randomized [20].

All outcomes were binomial. We used logistic regression to estimate odds ratios (ORs) for the treatments and their interactions. Model fit was assessed using the Hosmer-Lemeshow test. To aid in interpretation, we reexpressed ORs as risk differences, accounting for statistical uncertainty on baseline odds, main effects, and interactions. No data were missing for any of the participants meeting the inclusion criteria.

We used fixed effects meta-analysis to pool estimates across the trials, obtain overall estimates of effect, and assess the country as a potential effect modifier. We performed prespecified subgroup analyses for both trials to explore differences in treatment effect with respect to numeracy and, in the US trial, saliency. Numeracy was defined as scoring 3 (vs <3) on a validated instrument [29]. Saliency was defined as being very or extremely worried about acquiring COVID-19 and considering it very or extremely important to take action to reduce the chance of acquiring COVID-19. It was not possible to perform the analysis for saliency for the Norwegian trial because only 20 participants met the saliency criteria.

Furthermore, we performed nonprespecified analyses for potential effect modification. In the US trial, we explored effect modification by response time (responding within 7 min vs >7 min). The choice of 7 minutes was data driven, chosen to be the whole number of minutes closest to the median response time. We also explored effect modification by education level (having graduated vs not having graduated from a university).

We presented 2-sided 95% CIs and *P* values, wherever applicable, throughout the study. Meta- and subgroup analyses were presented using forest plots, with *P* values testing null hypotheses of homogeneity (there was no difference between the estimates for the 2 trials or no effect modification). All analyses were performed using Stata 18 (StataCorp LLC).

Ethical Considerations

This study was considered for ethics approval by the Regional Committee for Medical and Health Research Ethics in Norway and was found not to require ethics approval because it falls outside the committee’s mandate under the Health Research Act (reference 557972).

It was also deemed exempt from further review by the Committee for the Protection of Human Subjects (STUDY00032615) and the Dartmouth College Institutional Review Board.

By clicking on the study link in Prolific and Opinions platforms, participants implicitly consented to take part in the study. Everyone invited to participate were given information about who was conducting the research (Dartmouth University was emphasized in the US trial, while Norwegian Institute of Public Health was emphasized in the Norwegian trial), the aim of the research, what they would be asked to do, and how long it would take (Multimedia Appendix 3). It was clarified that participation was voluntary and that participants could discontinue participation at any time. We did not collect personal data from the participants (as per policy of the recruitment agencies), all collected data were anonymous, and we could not trace any data back to the participants. We did not inform the participants how the data would be stored or for how long; however, we did provide contact information to participants to ask any questions.

However, the final process of obtaining informed consent did not completely align with the procedures described in the study protocol [20]. During user testing, we received feedback that the introductory text was too long and cumbersome. Thus, we attempted to shorten it. Given that the study was completely anonymous (no personal or identifying information was gathered from the participants), posed minimal risk, and involved consenting members of the web-based panel platforms we used, we decided that we did not need to provide information regarding how data would be stored. In addition, we could remove a specific sentence asking for informed consent, as implied consent via clicking on the link to participate in the study was deemed sufficient. However, considering emerging best practice guidelines for web-based trials, we will provide more information and seek explicit consent in future trials.

We paid participants 100 Norwegian Kroner (US $8.40) and US $4.30 (50 Norwegian Kroner) in the Norwegian and US trials, respectively.

Patient and Public Involvement

We involved members of the public in user testing during the development of the survey and COVID-19 trial summaries.

Results

Overview

We invited 2667 Norwegians to participate in the trial, of which 1782 (66.82%) did not respond, 388 (14.55%) were not eligible, and 497 (18.63%) were randomized to one of the 6 intervention groups (Figure 3).

Of the 497 participants, 45 (9%) were excluded from the analysis for the reasons provided in Figure 3.

The study was open for 34,242 people in the United States to participate in the trial, of which 33,512 (97.88%) did not respond in time and 730 (2.13%; required sample size) were randomized to one of the 6 intervention groups (Figure 4).

Of the 730 participants, 187 (25.6%) were later excluded from the analysis for the reasons provided in Figure 4. The average age of participants was 45 (SD 16.8) years in the Norwegian trial and 38 (SD 13.2) years in the US trial (Table 2).

In total, 50% (226/452) of the Norwegians and 48.3% (262/543) of the US participants were female. The educational level was higher in the US participants (284/543, 52.3% had at least a college degree) than in the Norwegian participants (197/452, 43.6% had at least a college degree). Altogether, approximately half (517/995, 52%) of participants in both trials failed to answer all 3 numeracy questions correctly (Table 2).

There were only minor differences among the comparison groups in both the Norwegian and US trials (Tables S1 and S2 in Multimedia Appendix 4).

Figure 3. Modified CONSORT (Consolidated Standards of Reporting Trials) flow diagram of the participant inclusion process in April 2023. The flow diagram shows the eligible, recruited, and allocated participants from the Opinion panel in the Norwegian randomized controlled trial. *A less formal expression of the overall uncertainty language used in ordinary or familiar conversation, corresponding to the same Grading of Recommendations Assessment, Development and Evaluation (GRADE) assessment of the certainty of the evidence. **On the basis of the Cochrane Effective Practice and Organization of Care Group's guidance for communicating the certainty of evidence based on the GRADE approach to assessing the certainty of evidence [30].

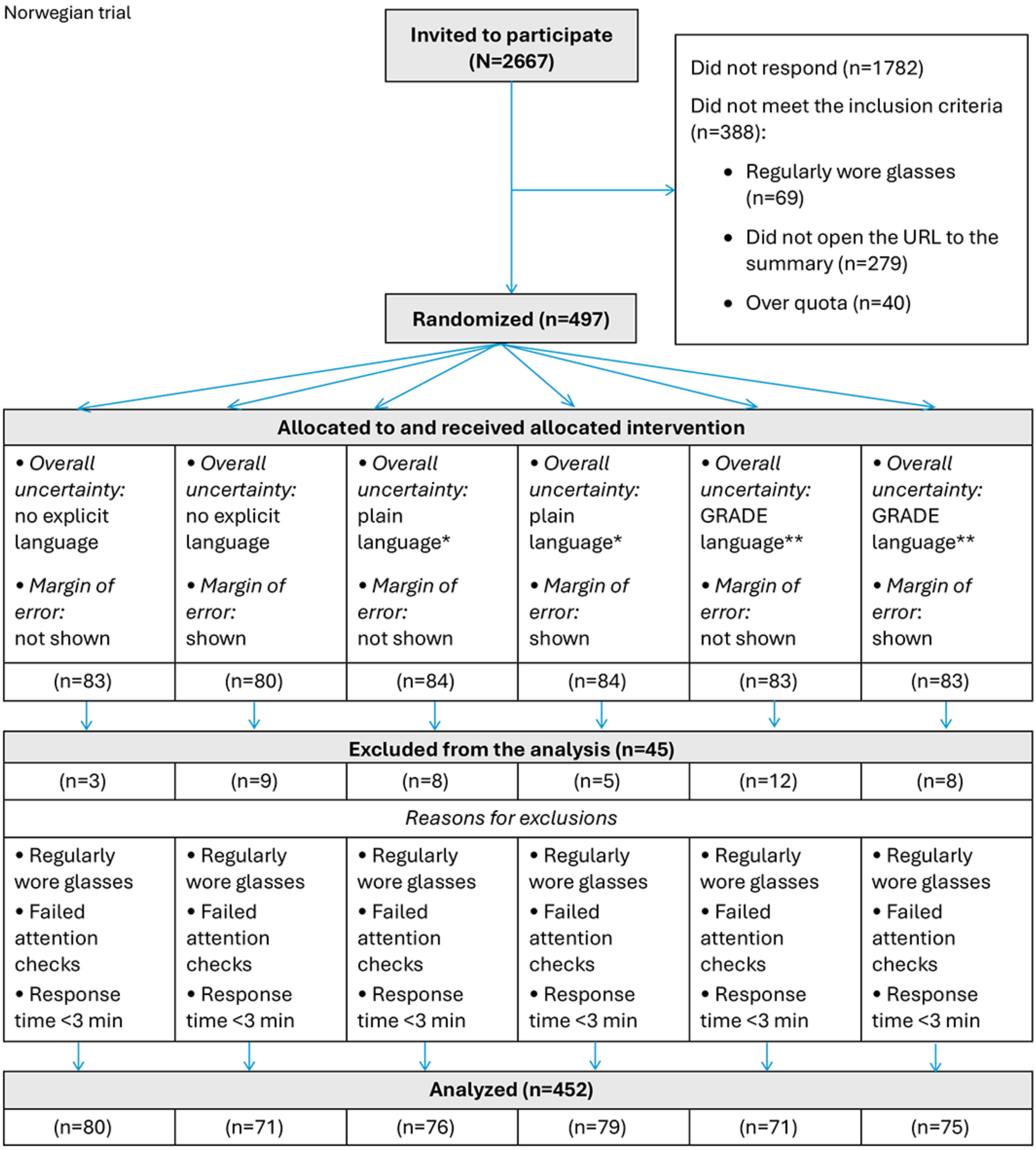


Figure 4. Modified CONSORT (Consolidated Standards of Reporting Trials) flow diagram of the participant inclusion process in April 2023. The flow diagram shows the eligible, recruited, and allocated participants from the Prolific panel in the US randomized controlled trial. *A less formal expression of the overall uncertainty language used in ordinary or familiar conversation, corresponding to the same Grading of Recommendations Assessment, Development and Evaluation (GRADE) assessment of the certainty of the evidence. **On the basis of the Cochrane Effective Practice and Organization of Care Group's guidance for communicating the certainty of evidence based on the GRADE approach to assessing the certainty of evidence [30].

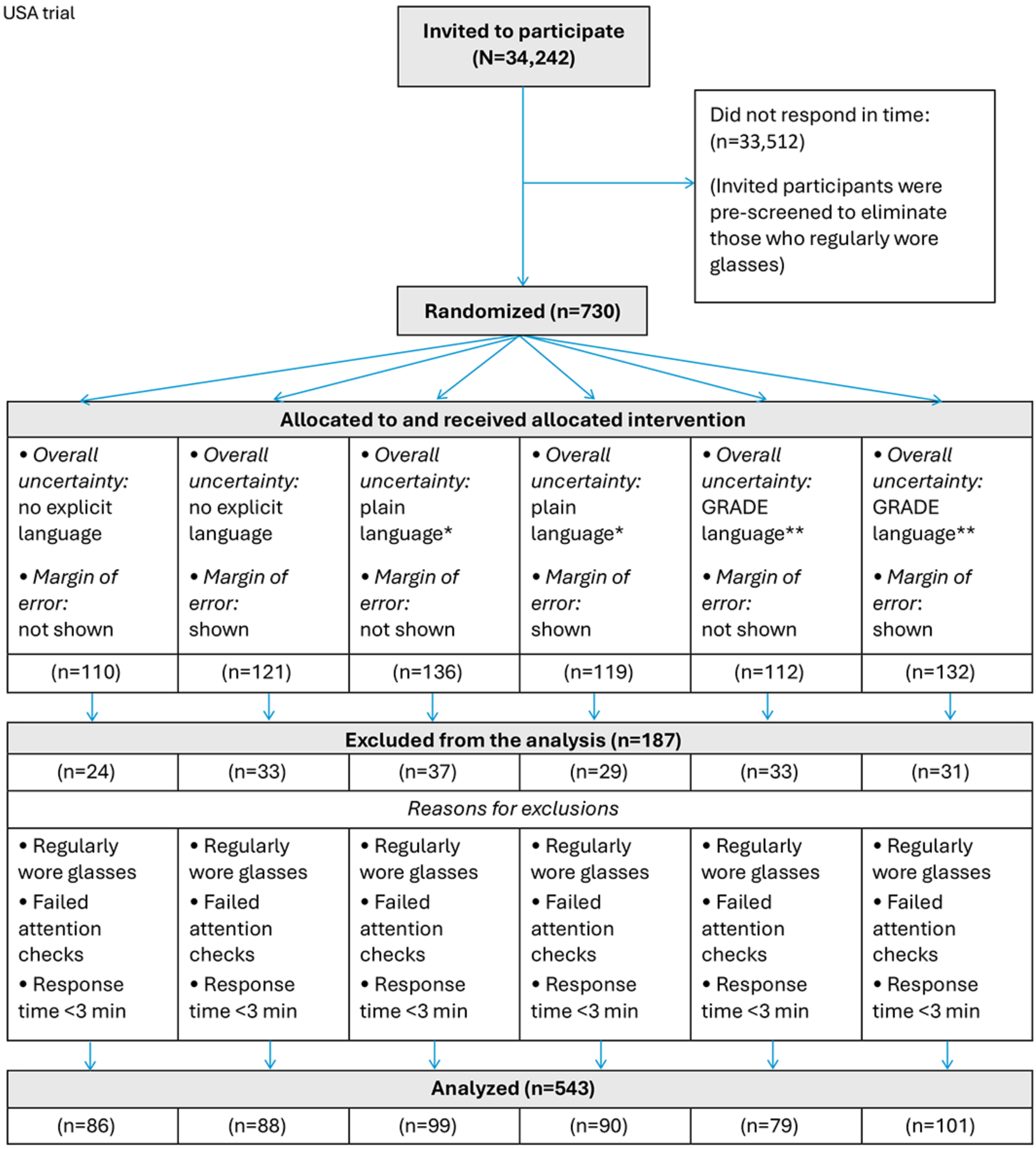


Table 2. Characteristics of the Norwegian and American participants from the web-based research panels taking part in the randomized controlled trial on evaluating the people's understanding of overall and statistical uncertainty when presented with alternative expressions in April 2023.

	Norwegian trial (n=452)	US trial (n=543)	Total (N=995)
Age (y), mean (SD)	45 (16.8)	38 (13.2)	41 (15.4)
Age group (y), n (%)			
18 to 29	109 (24.1)	155 (28.5)	264 (26.5)
30 to 39	74 (16.4)	195 (35.9)	269 (27)
40 to 49	81 (17.9)	98 (18)	179 (18)
50 to 59	88 (19.5)	27 (5)	115 (11.6)
60 to 69	59 (13.1)	58 (10.7)	117 (11.8)
≥70	41 (9.1)	10 (1.8)	51 (5.1)
Sex, n (%)			
Female	226 (50)	262 (48.3)	488 (49)
Male	226 (50)	281 (51.7)	507 (51)
Employment status, n (%)			
Full time	209 (46.2)	213 (39.2)	422 (42.4)
Part time	34 (7.5)	65 (12)	99 (9.9)
Unpaid[a]	—[b]	55 (10.1)	55 (5.5)
Unemployed or seeking work	—	51 (9.4)	51 (5.1)
Welfare[c]	114 (25.2)	—	114 (11.5)
Self-employed	22 (4.9)	—	22 (2.2)
Missing	17 (3.8)	129 (23.8)	146 (14.7)
Student	49 (10.8)	—	49 (4.9)
Other	7 (1.5)	30 (5.5)	37 (3.7)
Highest education, n (%)			
<High school	33 (7.3)	25 (4.6)	58 (5.8)
High school degree	131 (29)	92 (16.9)	223 (22.4)
Some college	91 (20.1)	142 (26.2)	233 (23.4)
College degree	124 (27.4)	204 (37.6)	328 (33)
Graduate or professional school	73 (16.2)	80 (14.7)	153 (15.4)
Numeracy, n (%)			
<3	235 (52)	282 (51.9)	517 (52)
3	217 (48)	261 (48.1)	478 (48)

[a] Includes participants who are homemakers, retired, or disabled.

[b] Not applicable.

[c] Includes participants in any welfare program, maternity leave, pensioner, or unemployment benefits.

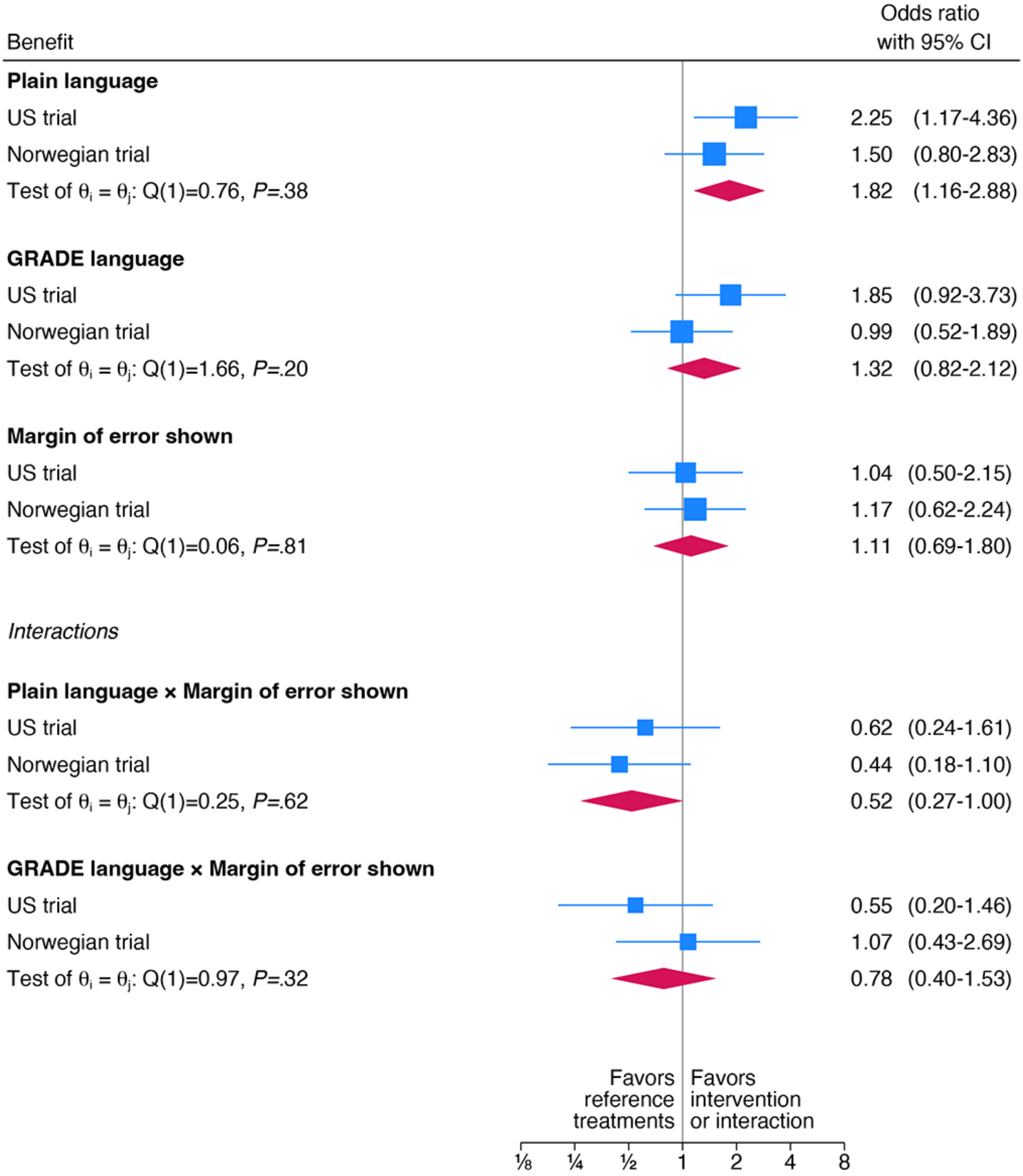
Understanding of the Uncertainty of the Benefit

Overall

Overall, plain (but not GRADE) language, compared to no explicit language, improved correct understanding of the overall uncertainty of the effect of wearing glasses on the chance of acquiring COVID-19 (ie, "mixed but more unsure than sure;" OR 1.82, 95% CI 1.16-2.88; Figure 5).

When the MOE was shown together with plain language, the effect on understanding of the overall uncertainty was substantially reduced (OR 0.52, 95% CI 0.27-1.00; Figure 5): participants were more likely to perceive the evidence as "very unsure" rather than "mixed but more unsure than sure" (the correct response) (Table S3 and S4 in Multimedia Appendix 4).

Figure 5. Meta-analysis of the understanding of the uncertainty of the benefit of wearing glasses. Odds ratios for answering “Mixed but more unsure than sure” to the question “How sure are you about the effect of wearing glasses on your chance of getting COVID?” in the Norwegian and US randomized controlled trials on evaluating people’s understanding of overall and statistical uncertainty when presented with alternative expressions. GRADE: Grading of Recommendations Assessment, Development and Evaluation.



Individual Trials

Overall, on average, across all 6 comparison groups, there were 25.2% (114/452) of Norwegian participants who responded that the evidence was very unsure compared to US participants, of whom 16.4% (89/543) reported the same (Table S3 and S4 in Multimedia Appendix 4).

The risk difference for the Norwegian trial was less certain (risk difference of 10.1%, 95% CI –5.5 to 25.7; Table 3) than in the US trial.

In the US trial, using plain language increased correct understanding of how sure we can be about the effect of wearing glasses on the chance of acquiring COVID-19, from 21% (18/86) to 37% (37/99; risk difference 16.4%; 95% CI 3.6-29.3; Table 4).

Table 3. Risk differences for the understanding of the overall uncertainty of the benefit of wearing glasses in the Norwegian randomized controlled trial on evaluating people's understanding of overall and statistical uncertainty when presented with alternative expressions for the question "How sure are you about the effect of wearing glasses on your chance of getting COVID?" with responses including very sure, mixed but more sure than unsure, mixed but more unsure than sure (correct), and very unsure.

Uncertainty	Margin of error	Participants[a], n (%)	Odds ratio[b] (95% CI)	Risk difference[c] (%), (95% CI)	*P* value
No explicit language (n=80)	Not shown	34 (42.5)	1	0	—[d]
No explicit language (n=71)	Shown	33 (46.5)	1.17 (0.62 to 2.24)	4.0 (–11.89 to 19.85)	.62
Plain language (n=76)	Not shown	40 (52.6)	1.50 (0.80 to 2.83)	10.1 (–5.47 to 25.73)	.21
Plain language (n=79)	Shown	29 (36.7)	0.78 (0.42 to 1.48)	–5.8 (–20.97 to 9.39)	.46
GRADE[e] language (n=71)	Not shown	30 (42.3)	0.99 (0.52 to 1.89)	–0.3 (–16.04 to 15.54)	.98
GRADE language (n=75)	Shown	36 (48)	1.25 (0.66 to 2.35)	5.5 (–10.16 to 21.16)	.49

[a] Participants randomized to the intervention answering correctly or as anticipated.

[b] Odds ratios include the main and interaction effects.

[c] Risk differences account for uncertainty on the baseline odds.

[d] Not applicable.

[e] GRADE: Grading of Recommendations Assessment, Development and Evaluation.

Table 4. Risk differences for the understanding of the overall uncertainty of the benefit of wearing glasses in the US randomized controlled trial on evaluating people's understanding of overall and statistical uncertainty when presented with alternative expressions for the question "How sure are you about the effect of wearing glasses on your chance of getting COVID?" with responses including very sure, mixed but more sure than unsure, mixed but more unsure than sure (correct), and very unsure.

Overall uncertainty	Margin of error	Participants[a], n (%)	Odds ratio (95% CI)[b]	Risk difference[c] (%), (95% CI)	*P* value
No explicit language (n=86)	Not shown	18 (20.9)	1	0	—[d]
No explicit language (n=88)	Shown	19 (21.6)	1.04 (0.50 to 2.15)	0.7 (–11.5 to 12.8]	.92
Plain language (n=99)	Not shown	37 (37.4)	2.25 (1.17 to 4.36)	16.4 (3.6 to 29.3)	.02
Plain language (n=90)	Shown	25 (27.8)	1.45 (0.73 to 2.91)	6.8 (–5.8 to 19.5)	.29
GRADE[e] language (n=79)	Not shown	26 (32.9)	1.85 (0.92 to 3.73)	12.0 (–1.5 to 25.4)	.08
GRADE language (n=101)	Shown	22 (21.8)	1.05 (0.52 to 2.12)	0.9 (–10.9 to 12.6)	.89

[a] Participants randomized to the intervention answering correctly or as anticipated.

[b] Odds ratios include the main and interaction effects.

[c] Risk differences account for uncertainty on the baseline odds.

[d] Not applicable.

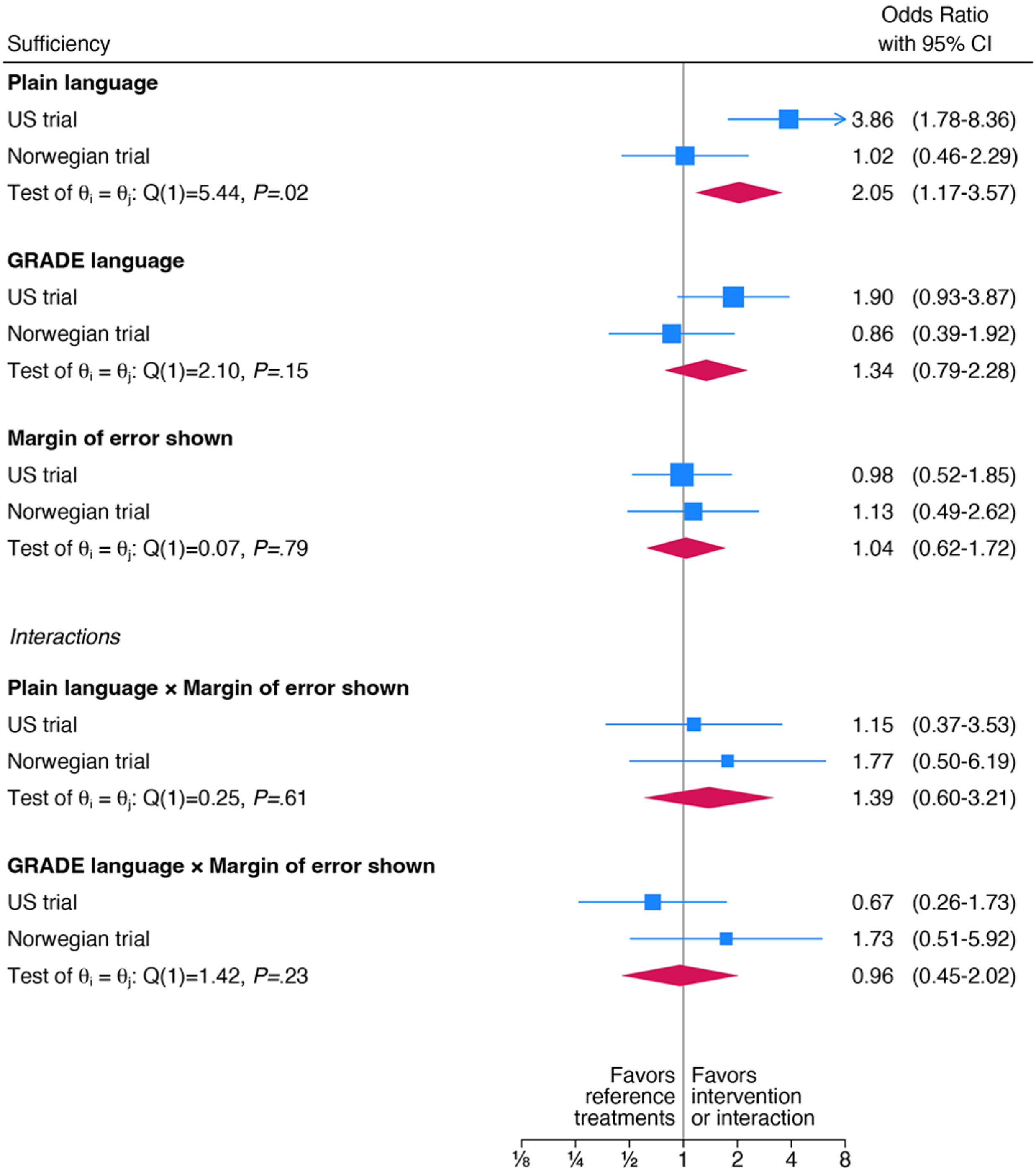
[e] GRADE: Grading of Recommendations Assessment, Development and Evaluation.

Understanding of the Sufficiency of the Evidence

Overall

Overall, plain, compared to no explicit language, improved correct understanding of the sufficiency of the evidence (OR 2.05, 95% CI 1.17-3.57; Figure 6). However, the results of the 2 trials were heterogeneous (Cochran Q_1=5.44; *P*=.02). The OR for the Norwegian trial was 1.02 (95% CI 0.46-2.29), whereas the OR for the US trial was 3.86 (95% CI 1.78-8.36). The effect of GRADE language, compared to no explicit language, was inconclusive on understanding of the sufficiency of the evidence (OR 1.34, 95% CI 0.79-2.28; Figure 6).

Figure 6. Meta-analysis presenting understanding the sufficiency of the evidence. Odds ratios for agreeing or strongly agreeing with “Not enough is known to be sure about the effects of wearing glasses to reduce the chance of getting COVID” in the Norwegian and US randomized controlled trials on evaluating people’s understanding of overall and statistical uncertainty when presented with alternative expressions. GRADE: Grading of Recommendations Assessment, Development and Evaluation.



Individual Trials

The results of the Norwegian trial for this outcome likely reflected a ceiling effect. In total, 81% (65/80) of the participants, who were shown no explicit language for overall uncertainty and no MOE (Table 5), agreed that not enough was known to be sure about the effects of wearing glasses to reduce the chance of acquiring COVID-19.

On the basis of the results of the US trial, using plain language increased the proportion of participants who agreed or strongly agreed that not enough was known to be sure about the effects of wearing glasses to reduce the chance of getting COVID-19, from 67% (58/86) to 89% (88/99); risk difference 21.4%, 95% CI 9.8-33.1 (Table 6). The effect of using the GRADE language was less certain.

Participants randomized to the intervention answering correctly or as anticipated.

Table 5. Risk differences for sufficiency of the evidence in the Norwegian randomized controlled trial on evaluating people's understanding of overall and statistical uncertainty when presented with alternative expressions for the question "not enough is known to be sure about the effects of wearing glasses to reduce the chance of getting COVID," with responses including strongly agree, agree (correct), disagree, and strongly disagree.

Overall uncertainty	Margin of error	Participants[a], n (%)	Odds ratio (95% CI)[b]	Risk difference[c] (%; 95% CI)	*P* value
No explicit language (n=80)	Not shown	65 (81.2)	1	0	—[d]
No explicit language (n=71)	Shown	59 (83.1)	1.13 (0.49 to 2.62)	1.9 (–10.36 to 14.06)	.77
Plain language (n=76)	Not shown	62 (81.6)	1.02 (0.46 to 2.29)	0.3 (–11.88 to 12.54)	.96
Plain language (n=79)	Shown	71 (89.9)	2.05 (0.81 to 5.15)	8.6 (–2.21 to 19.46)	.13
GRADE[e] language (n=71)	Not shown	56 (78.9)	0.86 (0.39 to 1.92)	–2.4 (–15.16 to 10.40)	.72
GRADE language (n=75)	Shown	66 (88)	1.69 (0.69 to 4.14)	6.8 (–4.53 to 18.03)	.25

[a] Participants randomized to the intervention answering correctly or as anticipated.

[b] Odds ratios include the main and interaction effects.

[c] Risk differences account for uncertainty on the baseline odds.

[d] Not applicable.

[e] GRADE: Grading of Recommendations Assessment, Development and Evaluation.

Table 6. Risk differences for sufficiency of the evidence in the US randomized controlled trial on evaluating people's understanding of overall and statistical uncertainty when presented with alternative expressions for the question "not enough is known to be sure about the effects of wearing glasses to reduce the chance of getting COVID," with responses including strongly agree, agree (correct), disagree, and strongly disagree.

Overall uncertainty	Margin of error	Participants[a], n (%)	Odds ratio (95% CI)[b]	Risk difference[c] (%; 95% CI)	*P* value
No explicit language (n=86)	Not shown	58 (67.4)	1	0	—[d]
No explicit language (n=88)	Shown	59 (67)	0.98 (0.52 to 1.85)	–0.4 (–14.34 to 13.55)	.96
Plain language (n=99)	Not shown	88 (88.9)	3.86 (1.78 to 8.36)	21.5 (9.77 to 33.13)	<.001
Plain language (n=90)	Shown	81 (90)	4.34 (1.91 to 9.90)	22.6 (10.87 to 34.24)	<.001
GRADE[e] language (n=79)	Not shown	63 (79.7)	1.90 (0.93 to 3.87)	12.3 (–0.98 to 25.59)	.08
GRADE language (n=101)	Shown	73 (72.3)	1.26 (0.67 to 2.36)	4.8 (–8.37 to 18.04)	.47

[b] Odds ratios include the main and interaction effects.

[c] Risk differences account for uncertainty on the baseline odds.

[d] Not applicable.

[e] GRADE: Grading of Recommendations Assessment, Development and Evaluation.

Understanding of the Uncertainty of Important Harms

Overall

There was no difference between plain and GRADE languages, compared to no explicit language, on the ability to correctly understand the uncertainty of important harms (OR 0.98, 95% CI 0.68-1.63 and OR 1.23, 95% CI 0.85-1.78; see Figure S1 in Multimedia Appendix 4).

Individual Trials

Understanding of the uncertainty of important harms varied across the trials. In total, 14.2% (64/452) of the participants in the Norwegian trial and 31.1% (169/543) in the US trial correctly understood what we aimed to communicate about the certainty of the evidence for important harms (Tables S5 and S6 in Multimedia Appendix 4). In both trials, there were only small differences between the comparison groups that could have occurred by chance alone.

Understanding Statistical Uncertainty (MOE)

Overall

In both countries, showing, compared to not showing, the MOE increased the proportion of people who chose the answer most consistent with the information provided (may reduce the chance of COVID a little but might increase it a little).

Individual Trials

In the Norwegian trial, 27.1% (61/225 and between 21%-34% in each of the 3 groups) of the participants answered correctly when the MOE was shown. Only 1.8% (4/227 and between 1%-3% in each of the 3 groups) answered correctly when MOE was not shown. In the US trial, 27.6% (77/279 and between 21%-36% in each of the 3 groups)) answered correctly when MOE was shown, while 1.9% (5/264 and between 0%-4% in each of the 3 groups) answered correctly when the MOE was not shown (Tables S7 and S8 in Multimedia Appendix 4).

The majority of the total number of participants shown the MOE (366/504, 72.6%) from both trials (and between 64% and 79% in each of the 3 groups) *failed* to correctly understand the MOE.

As expected, very few people correctly guessed the MOE when it was not shown. A correct answer when not shown the MOE would have been “Don’t know.” The proportion of participants who responded “Don’t know” when the MOE was shown was 4.8% (24/504) compared to 4.7% (23/491) when MOE was not shown.

Secondary Outcomes

Interest in Wearing Glasses to Reduce COVID-19 Risk During a Surge in Cases

Interest in wearing glasses to reduce COVID-19 risk during a surge was consistently lower in the Norwegian versus US trials (Figure S2 in Multimedia Appendix 4). In the Norwegian trial, plain language with or without the MOE and the GRADE language without the MOE reduced the proportion of participants who responded that they would be likely or very likely to wear or recommend wearing glasses to reduce the chance of acquiring COVID-19 if there were a surge of COVID-19 cases (risk difference: –13.6%, 95% CI –24.7 to –2.5; –18.6%, 95% CI –28.7 to –8.4; and −15.5%, 95% CI –26.4 to –4.5, respectively; Table S9 in Multimedia Appendix 4).

In the US trial, the combination of using either plain or GRADE languages to communicate the overall uncertainty of the benefit of wearing glasses and showing the MOE reduced the proportion of participants who responded that they would be likely or very likely to wear or recommend wearing glasses to reduce the chance of acquiring COVID-19 if there were a surge of COVID-19 cases (risk difference: –20.5%, 95% CI –33.1 to –7.9 and 21.2%, 95% CI –33.5 to –8.9, respectively; Table S10 in Multimedia Appendix 4).

The difference between the 2 trials is in part due to the lower proportion of participants in the Norwegian reference group (18/80, 23%) compared to the US reference group (31/86, 36%), who responded that they would be likely or very likely to wear or recommend wearing glasses to reduce the chance of acquiring COVID-19 if there were a surge of COVID-19 cases. This might be due to the translation from English to Norwegian.

Interest in Wearing Glasses to Reduce COVID-19 Risk if No Surge in Cases

We did not find an effect of plain or GRADE languages, compared to no explicit language, on the likelihood of wearing glasses or recommending wearing glasses to reduce the chance of acquiring COVID-19 if there were very few COVID-19 cases, the perceived benefit of wearing glasses, or the perceived chance of important harms (Figure S3 in Multimedia Appendix 4).

Perceptions of the Information Provided

In both trials, perceptions of the information provided differed little across plain and GRADE languages compared to no explicit language, with few exceptions; in both trials (US trial more so than Norway), perceptions of *helpfulness* were somewhat lower with plain language, with or without the MOE (Tables S11 and S12 in Multimedia Appendix 4).

In the Norwegian trial, with plain or GRADE languages, reporting the MOE reduced perceptions of the *trustworthiness* of the information (Table S13 and S14 in Multimedia Appendix 4).

In the Norwegian trial, the GRADE language without the MOE reduced the perception that information about whether wearing glasses affects the chance of acquiring COVID-19 was *sufficient* (Table S15 and S16 in Multimedia Appendix 4).

In the Norwegian trial, plain language without the MOE and the GRADE language with it reduced perceptions that the information about the benefit of glasses was *clear.* A similar effect was seen in the US trial for plain language with the MOE (Tables S17 and S18 in Multimedia Appendix 4).

In the Norwegian trial, plain language with the MOE reduced perceptions that information about whether glasses had important harms was *clear*. Similarly, this effect was seen in the US trial when the MOE was shown with no explicit language or plain language (Tables S19 and S20 in Multimedia Appendix 4).

In the Norwegian trial, plain language with the MOE reduced the proportion of participants who responded that they definitely or probably would *share the information* with someone who heard that wearing glasses might affect the chance of acquiring COVID-19 (Table S21 and S22 in Multimedia Appendix 4).

Decisional Conflict

Plain language without the MOE increased the feeling that the decision about wearing glasses if there were a surge of COVID-19 cases was hard to make (Table S23 and S24 in Multimedia Appendix 4). Most participants felt that they made an informed decision about wearing glasses if there were a surge of COVID-19 (264/452, 58.4%; between 52.6% and 67.6% in the Norwegian trial and 406/543, 74.8%; between 70.9% and 83% in the US trial; Table S25 and S26 in Multimedia Appendix 4).

Potential Modifying Factors

We did not find credible evidence of effect modification for numeracy, education, salience, or the time taken to complete the questionnaire.

Discussion

Principal Findings

The US and Norwegian trials comparing different ways of communicating the overall and statistical uncertainty of research results to the public generated mixed results. Plain language improved readers’ understanding of the overall uncertainty of the benefit but only to a modest extent. The effect of the GRADE language was uncertain but, at best, had a modest effect. Furthermore, reporting the MOE reduced understanding by making the evidence seem more uncertain than it actually was. Reporting the MOE did improve understanding of statistical uncertainty around the effect of glasses but only for a minority group of people. A more detailed discussion of the most important findings follows in the subsequent sections.

Communicating Overall Uncertainty of the Benefit

Plain language is probably more accessible than the GRADE language (eg, the phrase “not very confident” seems easier to understand than “moderate uncertainty”) for communicating the overall uncertainty of benefit. Given the modest findings, more work is needed on the language; other ways of communicating uncertainty, such as visualizations; and easily accessible explanations.

It should be noted that we displayed the GRADE symbol together with the GRADE language. The GRADE symbols for uncertainty are similar to widely used symbols used for ranking the quality of, for example, hotels, restaurants, and consumer products [30]. Nonetheless, they may be unfamiliar in this context and not easily understood without further explanation.

In another small trial [31], the GRADE symbols were compared to letters to convey the quality of evidence, and both letters and symbols were well understood.

Both the plain language and the GRADE language summaries included text explaining why we were “not very confident” (it was “uncertain” that wearing glasses may slightly reduce the chance of acquiring COVID-19). These explanations were available under a tab labeled “Keep in mind.” Neither the number of participants who read this text nor the effect of the text on their understanding of the uncertainty of the evidence is known. More Norwegians responded that the evidence was very unsure compared to US participants. One reason for the differences between the Norwegian and US results may be that “may reduce” was translated to “kan muligens redusere” (“can possibly reduce”) in Norwegian (based on the approved translation used by the Norwegian Institute of Public Health) [32]. The word “kan” in Norwegian can mean either “may” or “can,” and these 2 terms may differ with regards to their emphasis on the degree of uncertainty, which is why the moderating adverb “muligens” (directly translating to “possibly”) was included in the Norwegian version. Although we user-tested the summaries before the trials, further exploration of the extent to which people find this text helpful and how it impacts their understanding of the uncertainty of the evidence is warranted. Furthermore, cultural understandings and comfort with terms related to or describing the concept of “uncertainty” could be further explored to communicate uncertainty more accurately in languages other than English.

Communicating Overall Uncertainty of Important Harm

Neither plain language nor the GRADE language improved the understanding of the overall uncertainty of important harm. This could be because we did not explain why we said that there was moderate certainty evidence and prima facie, it seemed implausible to participants that wearing glasses could cause important harm. In fact, about one-third of the Norwegian participants were very sure, and this might be because they considered it implausible that wearing glasses could cause important harm. On the other hand, the other participants (about half) might have been somewhat or very unsure because there was so little information. The difference between the Norwegian and US trials might have been due to translation. The English version for the control group (No explicit language) stated that “wearing glasses probably does not cause important harms, such as serious fall due to reduced vision” while the Norwegian version, if back translated stated “Using glasses probably doesn’t cause serious injury, for example after falling” but does not refer to “reduced vision.” Roughly twice as many participants in the US trial, compared to participants in the Norwegian trial, correctly understood the overall uncertainty of the evidence for important harm. The most likely reason for this is the difference between the English summaries and the Norwegian translations. In this case, the explanation for why important harms (serious injury after falling is plausible because of reduced vision) was not mentioned in the Norwegian translation. This highlights the need for more extensive user testing and the use of back translation or other means of ensuring that translations are correctly understood.

Effect of MOE on Overall Uncertainty

Showing the MOE did help people understand statistical uncertainty (ie, glasses may reduce the chance of acquiring COVID-19 a little *but might* increase it a little). While this is encouraging, the modest effect suggests more work is needed to ensure that people understand what MOE implies.

In the group of participants who were shown plain language, showing the MOE (for benefit) decreased the proportion of people who answered this question correctly. This probably happened because the effect of glasses was small, and the MOE really highlighted that adding uncertainty to uncertainty in the context of a small effect, to begin with, made people feel the effect was less certain than it was. This finding underscores the need for more work to help people calibrate their sense of uncertainty. It would be interesting to see what would happen in another example where the intervention effect was bigger than the effect of glasses. Another reason for this may be that participants who were shown the MOE were also shown 2 reasons for rating the evidence as low certainty (wide MOE *and* important study limitations), whereas those who were not shown the MOE were only presented with 1 source of uncertainty (important study limitations). Findings from a 2020 study [33] looking at the effect of communicating uncertainty found that participants who were shown 3 sources of uncertainty were more likely to report a weaker perception of the effectiveness of the intervention (drug) than those who were presented with only 2 sources of uncertainty. More research needs to be conducted to explore this hypothesis.

Effect of MOE on Intended Behavior

Reporting overall uncertainty using plain language or the GRADE language or reporting the MOE decreased the likelihood that participants would wear glasses to protect themselves against COVID-19 if there was a surge in cases. This finding is consistent with other research findings that suggest reporting uncertainty decreased the likelihood that people would use eye protection to reduce their chance of acquiring COVID-19 [18]. It is also logical that the less certain one is about the benefits of doing something, the less likely it is to be done. These results highlight the importance of effectively communicating uncertainty if the intention is to inform people rather than to persuade them [19].

Effect of MOE on Perceptions of Information

In Norwegian trial, showing the MOE in combination with plain language reduced the perceived trustworthiness of the information (the results were similar, although smaller, and not statistically significant for the other versions). This is consistent with the findings of Schneider et al [18], which found that including a clue that evidence of the effect of eye protection was low quality or certainty reduced the perceived trustworthiness of an infographic. In contrast, in our US trial, the reduction in perceived trustworthiness when showing the MOE was substantially smaller and not statistically significant in any version. This suggests that we cannot conclude that communicating uncertainty necessarily reduces the perceived trustworthiness of information about the effects of interventions. Indeed, more candid communication over time might make changes in recommendations seem less arbitrary and help preserve people’s trust in health authorities [14].

There is some evidence that plain language with or without the MOE and the GRADE language with the MOE reduced the perception that the summary was sufficient. This may be due to the participants confusing the sufficiency of the evidence (ie, a small, uncertain effect) with the sufficiency of the summary, which was meant to communicate the effect. While the question we asked specifically aimed at the latter (ie, whether the information was a sufficient summary of what is known about the effects of wearing glasses to reduce the chance of acquiring COVID-19), future qualitative work should be done to help distinguish these 2 kinds of sufficiency. Similarly, showing the MOE also reduced the perception that the summaries were clear, perhaps for the same reasons. In the Norwegian trial, the likely ceiling effect may reflect the mentioned translation of “may” to “can possibly” in the Norwegian summary.

Nevertheless, most participants found the decision about wearing glasses to reduce the chance of acquiring COVID-19 hard to make, with or without the MOE, regardless of which summary they were shown.

Limitations

Study limitations include the weak, uncertain effect that we were trying to summarize (from the Glasses trial), which magnified the communication challenges, and several language translation issues, which may limit the cross-country comparisons. Furthermore, our results may have been influenced by a lack of saliency, that is, COVID-19 infection rates were relatively low when we conducted the study, and participants might have responded differently in a more realistic or pressing scenario. Furthermore, we need to explore the effect of paying participants in web-based trials on the quality of the responses [24].

As this study was the research team’s first attempt at conducting a web-based trial, we encountered and reflected on a number of challenges and opportunities. First, while we were satisfied with the web-based platforms we eventually chose, it was time-consuming and challenging to identify a proper platform to facilitate the conduct of the trial, and we eventually needed to incorporate 2 platforms to meet our specific needs. Another challenge was related to obtaining informed consent—it was unclear what elements of the consent were covered by the web-based platform or our invitation text. Since the conduct of our trial, more guidance has been provided on the platforms and other academic websites offering guidance on informed consent in web-based trials.

In the next Message Lab trial for assessing different ways of communicating evidence, we would like to set up Google Analytics (Google LLC), if using Google Sites to present the health messages to the participants. This will allow us to investigate further on how much time people spend on the different sections and how they navigate the texts presented.

Conclusions

Our study has several strengths, including the randomized factorial design, which let us explore the interaction between uncertainty language (GRADE, plain, or no explicit) and the MOE on understanding the benefits, harms, and the corresponding uncertainties and allowed replication in 2 distinct populations. Our study shows that explicitly reporting uncertainty affects peoples’ understanding, perceptions, and intended actions. We found that plain language was better than no explicit language in helping people understand the overall uncertainty of the evidence, although most participants still did not correctly understand how sure they could be. Reporting MOE reduced understanding of the overall uncertainty by making people feel that the evidence was even less certain. Reporting MOE improved the interpretation of statistical uncertainty around the effect of glasses, but only for a minority group of participants.

This study underscores how much more work needs to be done to develop effective ways to communicate overall uncertainty and just how unclear the numbers are (statistical uncertainty). If this communication is done poorly, it may simply add to confusion and lead to poor decisions. If done properly, effective communication around uncertainty can help people to make the best decisions they can, given the evidence that is known.

Acknowledgments

This research is funded by the Norwegian Institute of Public Health. The authors would like to thank Ingeborg Hess Elgersma and the Centre for Epidemic Interventions Research team for their valuable feedback on the translations and questionnaire development.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
All summaries in both English and Norwegian.
[PDF File (Adobe PDF File), 4235 KB - publichealth_v11i1e62828_app1.pdf]

Multimedia Appendix 2
Questionnaires in English and Norwegian.
[PDF File (Adobe PDF File), 432 KB - publichealth_v11i1e62828_app2.pdf]

Multimedia Appendix 3
Information provided to the participants.
[PDF File (Adobe PDF File), 561 KB - publichealth_v11i1e62828_app3.pdf]

Multimedia Appendix 4
Additional result tables and figures.
[PDF File (Adobe PDF File), 893 KB - publichealth_v11i1e62828_app4.pdf]

Multimedia Appendix 5
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1253 KB - publichealth_v11i1e62828_app5.pdf]

References

1. Fischhoff B. The COVID communication breakdown: how to fix public health messaging. Foreign Affairs. 2021 Oct 4. URL: https://www.foreignaffairs.com/articles/united-states/2021-10-04/covid-communication-breakdown [accessed 2025-03-10]
2. Fischhoff B. The sciences of science communication. Proc Natl Acad Sci U S A 2013 Aug 20;110 Suppl 3(Suppl 3):14033-14039 [FREE Full text] [doi: 10.1073/pnas.1213273110] [Medline: 23942125]
3. Fischhoff B. Risk perception and communication. In: Detels R, Karim QA, Baum F, Li L, Leyland AH, editors. Oxford Textbook of Global Public Health, Seventh Edition. Oxford, UK: Oxford University Press; 2021.
4. Oxman AD, Glenton C, Flottorp S, Lewin S, Rosenbaum S, Fretheim A. Development of a checklist for people communicating evidence-based information about the effects of healthcare interventions: a mixed methods study. BMJ Open 2020 Jul 21;10(7):e036348 [FREE Full text] [doi: 10.1136/bmjopen-2019-036348] [Medline: 32699132]
5. Woloshin S, Yang Y, Fischhoff B. Communicating health information with visual displays. Nat Med 2023 May 08;29(5):1085-1091. [doi: 10.1038/s41591-023-02328-1] [Medline: 37156935]
6. Woloshin S, Dewitt B, Krishnamurti T, Fischhoff B. Assessing how consumers interpret and act on results from at-home COVID-19 self-test kits: a randomized clinical trial. JAMA Intern Med 2022 Mar 01;182(3):332-341 [FREE Full text] [doi: 10.1001/jamainternmed.2021.8075] [Medline: 35099501]
7. Woloshin S, Schwartz LM. Communicating data about the benefits and harms of treatment: a randomized trial. Ann Intern Med 2011 Jul 19;155(2):87-96 [FREE Full text] [doi: 10.7326/0003-4819-155-2-201107190-00004] [Medline: 21768582]
8. Akl EA, Oxman AD, Herrin J, Vist GE, Terrenato I, Sperati F, et al. Using alternative statistical formats for presenting risks and risk reductions. Cochrane Database Syst Rev 2011 Mar 16;2011(3):CD006776 [FREE Full text] [doi: 10.1002/14651858.CD006776.pub2] [Medline: 21412897]
9. Carling CL, Kristoffersen DT, Montori VM, Herrin J, Schünemann HJ, Treweek S, et al. The effect of alternative summary statistics for communicating risk reduction on decisions about taking statins: a randomized trial. PLoS Med 2009 Aug 25;6(8):e1000134 [FREE Full text] [doi: 10.1371/journal.pmed.1000134] [Medline: 19707575]
10. Schwartz LM, Woloshin S. The drug facts box: improving the communication of prescription drug information. Proc Natl Acad Sci U S A 2013 Aug 20;110 Suppl 3(Suppl 3):14069-14074 [FREE Full text] [doi: 10.1073/pnas.1214646110] [Medline: 23942130]
11. Blalock SJ, Sage A, Bitonti M, Patel P, Dickinson R, Knapp P. Communicating information concerning potential medication harms and benefits: what gist do numbers convey? Patient Educ Couns 2016 Dec;99(12):1964-1970 [FREE Full text] [doi: 10.1016/j.pec.2016.07.022] [Medline: 27444232]
12. Schwartz LM, Woloshin S. Communicating uncertainties about prescription drugs to the public: a national randomized trial. Arch Intern Med 2011 Sep 12;171(16):1463-1468. [doi: 10.1001/archinternmed.2011.396] [Medline: 21911629]
13. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011 Apr;64(4):401-406. [doi: 10.1016/j.jclinepi.2010.07.015] [Medline: 21208779]

14. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol 2011 Dec;64(12):1283-1293. [doi: 10.1016/j.jclinepi.2011.01.012] [Medline: 21839614]
15. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ. What is "quality of evidence" and why is it important to clinicians? BMJ 2008 May 01;336(7651):995-998. [doi: 10.1136/bmj.39490.551019.be]
16. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). J Clin Epidemiol 2011 Apr;64(4):407-415. [doi: 10.1016/j.jclinepi.2010.07.017] [Medline: 21247734]
17. Santesso N, Rader T, Nilsen ES, Glenton C, Rosenbaum S, Ciapponi A, et al. A summary to communicate evidence from systematic reviews to the public improved understanding and accessibility of information: a randomized controlled trial. J Clin Epidemiol 2015 Feb;68(2):182-190 [FREE Full text] [doi: 10.1016/j.jclinepi.2014.04.009] [Medline: 25034199]
18. Schneider CR, Freeman AL, Spiegelhalter D, van der Linden S. The effects of quality of evidence communication on perception of public health information about COVID-19: two randomised controlled trials. PLoS One 2021 Nov 17;16(11):e0259048 [FREE Full text] [doi: 10.1371/journal.pone.0259048] [Medline: 34788299]
19. Oxman AD, Fretheim A, Lewin S, Flottorp S, Glenton C, Helleve A, et al. Health communication in and out of public health emergencies: to persuade or to inform? Health Res Policy Syst 2022 Mar 05;20(1):28 [FREE Full text] [doi: 10.1186/s12961-022-00828-z] [Medline: 35248064]
20. Woloshin S, Oxman AD, Rose CJ, Rosenbaum SE, Holst C, Diaz-Hemard L, et al. Communicating uncertainty about the effects of wearing glasses to reduce the chance of getting COVID: protocol for a randomized trial. Zenodo. 2022 Nov 18. URL: https://zenodo.org/records/7428981 [accessed 2025-03-10]
21. Fretheim A, Elgersma IH, Helleve A, Elstrøm P, Kacelnik O, Hemkens LG. Effect of wearing glasses on risk of infection with SARS-CoV-2 in the community: a randomized clinical trial. JAMA Netw Open 2022 Dec 01;5(12):e2244495 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.44495] [Medline: 36454571]
22. Prolific homepage. Prolific. URL: https://www.prolific.com/ [accessed 2024-03-15]
23. Vi måler nordmenns meninger, og leverer innsikt som bringer deg videre. Opinion. URL: https://www.opinion.no/ [accessed 2024-03-15]
24. Bradley P. Bots and data quality on crowdsourcing platforms. Prolific. 2018 Aug 10. URL: https://www.prolific.com/resources/bots-and-data-quality-on-crowdsourcing-platforms [accessed 2024-02-20]
25. Douglas BD, Ewell PJ, Brauer M. Data quality in online human-subjects research: comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. PLoS One 2023 Mar 14;18(3):e0279720 [FREE Full text] [doi: 10.1371/journal.pone.0279720] [Medline: 36917576]
26. Ergonomics of human-system interaction part 210: human-centred design for interactive systems. International Organization for Standardization. 2019. URL: https://www.iso.org/standard/77520.html [accessed 2025-03-10]
27. Loop11 homepage. Loop11. URL: https://www.loop11.com/ [accessed 2024-03-15]
28. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making 1995 Feb 01;15(1):25-30. [doi: 10.1177/0272989x9501500105]
29. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med 1997 Dec 01;127(11):966-972. [doi: 10.7326/0003-4819-127-11-199712010-00003] [Medline: 9412301]
30. Schünemann HJ, Best D, Vist G, Oxman AD, GRADE Working Group. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. CMAJ 2003 Sep 30;169(7):677-680 [FREE Full text] [Medline: 14517128]
31. Akl EA, Maroun N, Guyatt G, Oxman AD, Alonso-Coello P, Vist GE, et al. Symbols were superior to numbers for presenting strength of recommendations to health care consumers: a randomized trial. J Clin Epidemiol 2007 Dec;60(12):1298-1305. [doi: 10.1016/j.jclinepi.2007.03.011] [Medline: 17998085]
32. Formuleringer i GRADE for å uttrykke resultater fra kunnskapsoppsummeringer om effekt. Folkehelseinstituttet (Norwegian Institute of Public Health). URL: https://www.fhi.no/ku/kunnskaps-og-beslutningsstotte/metodeboka/framgangsmate/hente-ut-data-sammenfatte-og-gradere/?term= [accessed 2024-03-15]
33. Büchter RB, Betsch C, Ehrlich M, Fechtelpeter D, Grouven U, Keller S, et al. Communicating uncertainty in written consumer health information to the public: parallel-group, web-based randomized controlled trial. J Med Internet Res 2020 Aug 10;22(8):e15899 [FREE Full text] [doi: 10.2196/15899] [Medline: 32773375]

Abbreviations

GRADE: Grading of Recommendations Assessment, Development and Evaluation
MOE: margin of error
OR: odds ratio

Edited by A Mavragani; submitted 04.06.24; peer-reviewed by R Büchter, H Pilabré; comments to author 11.12.24; revised version received 30.01.25; accepted 06.02.25; published 18.03.25.

Please cite as:
Holst C, Woloshin S, Oxman AD, Rose C, Rosenbaum S, Munthe-Kaas HM
Alternative Presentations of Overall and Statistical Uncertainty for Adults' Understanding of the Results of a Randomized Trial of a Public Health Intervention: Parallel Web-Based Randomized Trials
JMIR Public Health Surveill 2025;11:e62828
URL: https://publichealth.jmir.org/2025/1/e62828
doi:10.2196/62828
PMID:

HIV Incidence and Associated Risk Factors Among Young Men Who Have Sex With Men in Tianjin, China: Retrospective Cohort Study

Yan Guo[1*], BMed; Hui Gong[1*], MSc; Xiangyu Yan[2], PhD; Xinying Zhang[3], BMed; Tielin Ning[1], BMed; Zhongquan Liu[1], BMed; Long Li[1], MSc; Jie Yang[4]; Liu Li[1], BMed; Changxu Ma[1], BMed; Zhuang Cui[5], PhD; Maohe Yu[1], MSc

[1]Tianjin Centers for Disease Control and Prevention, 6 Hua Yue Road, Hedong District, Tianjin, China
[2]School of Disaster and Emergency Medicine, Tianjin University, Tianjin, China
[3]Tianjin Hongqiao Center for Disease Control and Prevention, Tianjin, China
[4]Shenlan Public Health Consulting Service Center, Tianjin, China
[5]Department of Epidemiology and Biostatistics, Tianjin Medical University, Tianjin, China
[*]these authors contributed equally

Corresponding Author:
Maohe Yu, MSc
Tianjin Centers for Disease Control and Prevention, 6 Hua Yue Road, Hedong District, Tianjin, China

Abstract

Background: Young men who have sex with men (YMSM) have a higher risk of HIV infection. However, evidence of HIV incidence from large cohort studies is limited in this key population, particularly among Chinese YMSM.

Objective: This study aimed to investigate the HIV incidence and associated risk factors among YMSM aged 16 - 24 years in Tianjin, China.

Methods: We conducted a retrospective cohort study among men who have sex with men aged 16 - 24 years from October 2017 to December 2022 through the largest local nongovernmental organization serving men who have sex with men in Tianjin. Participants who responded to the investigations at least twice during the study period were included. They completed questionnaires on demographic information, sexual behaviors, psychoactive substance use, and sexually transmitted infection status. Simultaneously, their blood samples were collected for HIV testing. HIV incidence was calculated by dividing the sum of observed HIV seroconversions by the observed person-years (PYs). A Cox proportional hazards regression model was used to identify risk factors associated with HIV incidence.

Results: A total of 1367 HIV-negative YMSM were included in the cohort, among whom 62 seroconversions occurred, contributing 2384.2 observed PYs; the total incidence was 2.6 (95% CI 2.0 - 3.2) per 100 PYs. The segmented HIV incidence rates were 3 (95% CI 1.5 - 4.5), 2.4 (95% CI 1.5 - 3.3), and 2.7 (95% CI 1.5 - 3.9) per 100 PYs for 2017 - 2018, 2019 - 2020, and 2021 - 2022, respectively. Compared to YMSM who had been followed up fewer than three times, those followed up three or more times had a relatively lower risk of HIV infection (Adjusted hazard ratio [AHR] 0.553, 95% CI 0.325 - 0.941). YMSM who preferred finding sexual partners offline had a higher risk of HIV infection compared to those who preferred finding sexual partners online (AHR 2.207, 95% CI 1.198 - 4.066). Compared to YMSM without syphilis, those infected with syphilis had an increased risk of HIV infection (AHR 2.234, 95% CI 1.137 - 4.391). Additionally, YMSM who used psychoactive substances had a higher risk of HIV infection compared to those who did not use such substances (AHR 2.467, 95% CI 1.408 - 4.321).

Conclusions: Our study complements data on HIV incidence among YMSM in large cities across China. Syphilis infection and the use of psychoactive substances were risk factors associated with HIV occurrence, demonstrating an urgent need for tailored prevention and control interventions for this key population.

(JMIR Public Health Surveill 2025;11:e66487) doi:10.2196/66487

KEYWORDS

young men who have sex with men; YMSM; incidence; HIV; cohort study; risk factor

Introduction

The Global AIDS report released by the Joint United Nations Programme on HIV/AIDS (UNAIDS) [1] estimated that there were 350,000 new HIV infections among people aged 15 - 24 years in 2022. Young people accounted for 26% of HIV infections in the Asia-Pacific region in 2020; among them, 52% were men who have sex with men (MSM), making them a key population that requires focused attention for HIV prevention. Meanwhile, the prevalence of HIV among young men who have sex with men (YMSM) has more than doubled in Indonesia (ie, from 6% in 2011 to 13% in 2019) and nearly tripled in Malaysia (ie, 6% in 2012 to 15% in 2022), and Vietnam (ie, 3% in 2011 to 11% in 2020) [2]. Previous studies suggest that young people from key populations experience high vulnerability and challenges in staying healthy and safe in situations marked by stigma, discrimination and harassment, punitive laws, and social taboos [2]. Moreover, young people living with HIV have the lowest rate of awareness about their infection status [3].

Meanwhile, in recent years, the HIV epidemic in China has risen rapidly among young people. The number of new HIV cases reported yearly among people aged 15 - 24 years increased from 9373 in 2010 to 15,790 in 2019; a total of more than 140,000 HIV cases have been reported in this age group [4]. HIV has become a major infectious disease that endangers the health of young people in China, particularly in some regions where the proportions of reported cases have increased [4]. The spread of HIV among young people is concerning, posing a serious threat to their lives and health. Due to them being in a stage of rapid physical and mental development, lack of sexual health knowledge, and the increase in high-risk behaviors, they are more vulnerable to HIV.

Given that YMSM are susceptible to HIV infection [5,6], studies have shown that HIV infection rates among YMSM are higher than those in older populations [5,6,7,8]. YMSM under 25 years of age represent a demographic with increasing numbers of new HIV infections even as HIV incidence is declining globally [9]. Although there were some data on HIV prevalence among YMSM [9,10], the epidemic remains poorly defined among in this population [11], particularly the data of incidence is scarcely reported. Understanding HIV incidence—especially among subgroups of key populations—is essential for effectively targeting strategies to prevent new HIV infections [12].

Tianjin, located in the Bohai Gulf, is the second most populous city in northern China and one of the four major municipalities, with 16 districts and a resident population of 13.73 million as of 2021 [13,14]. The majority of young HIV cases in Tianjin were attributed to homosexual transmission, and the prevalence of HIV and syphilis among YMSM in the city is high [15]. However, the incidence of HIV among YMSM in Tianjin remains unknown.

This study aimed to (1) identify the HIV-related behavioral characteristics of YMSM, (2) understand the trends of HIV incidence among YMSM, and (3) investigate the association between related factors and HIV seroconversion among YMSM. The findings of this study will provide policymakers with evidence to inform prevention strategies aimed at controlling the HIV epidemic among YMSM.

Methods

Study Design

A retrospective cohort study on HIV incidence among YMSM was conducted in Tianjin from 2017 to 2022. First, to enhance geographical coverage and participant accessibility, seven voluntary HIV counselling and testing sites were set up for on-site surveys. Second, the on-site surveys were conducted at Shenlan, the largest local nongovernmental organization (NGO) serving MSM. The surveys were implemented by professional staff at the NGO, who also comprised MSM. To ensure the quality of the surveys, all professional staff at the 7 voluntary HIV counselling and testing sites received systematic and standard training; furthermore, the surveys were carried out through a unified method. Third, to guarantee the representativeness of the participants, two recruitment approaches were adopted. One approach involved professional staff recruiting YMSM through various strategies, such as offline hotspots (eg, gay baths, gay bars, public toilets) and online platforms (eg, WeChat, QQ, Blued). The second approach involved mobilizing the YMSM recruited through the first approach to introduce new YMSM. Fourth, all participants were required to participate in a baseline survey, followed by a follow-up survey every 6 months.

During the study, the fingerprints of the participants' right index fingers were collected, and a unique fingerprint code was assigned. A verification code was sent to the participants' cell phones to verify their phone numbers. After successful verification, training investigators conducted face-to-face questionnaire interviews. After the completion of the questionnaires, professional blood collectors obtained venous blood samples from the participants. The blood was sent to a professional laboratory for HIV testing. HIV test and questionnaire responses were linked using the participant's fingerprint code, which was unique and served as a proxy identifier to match the HIV testing results to each participant. Participants' fingerprint codes were recorded into an unidentifiable ID number to make the dataset anonymous. The Tianjin Centers for Disease Control and Prevention questionnaire was developed for the target population and presurvey.

Participant's Eligibility Criteria

The target population of this study were YMSM who participated in the survey and resided in Tianjin during the study period. Eligible criteria for study enrollment included: (1) born a male biologically, (2) aged 16 - 24 years old, (3) who engaged in penetrative oral or anal sexual intercourse with males in the past year, (4) voluntarily consented to questionnaires, fingerprint registration, blood sample collection, and related testing, and (5) self-reported negative HIV tests or were unaware of their HIV infection status before participation in the survey.

Questionnaire Interview

The questionnaire included 4 parts. The first part collected information on sociodemographic characteristics such as age, residential district, occupation, education, marital status, and duration of residence. The second part collected information on sexual orientation and sexual behaviors in the last 6 months, including heterosexual behaviors, homosexual behaviors, commercial homosexual partners, primary venue to find sexual partners, and condom use. The third part collected information on the use of psychoactive substances (eg, Rush and Capsule Zero), injection drug use in the past year, syphilis infection status, and peer education received in the past year. As smoking, alcohol consumption, and psychoactive substance use are addictive in nature, and previous research by Guo et al [16] showed a positive relationship between smoking and psychoactive substance use among YMSM, this study further explored these relationships. In this study, drinking was defined as consuming alcohol at least twice within the past 30 days, while smoking was defined as continuous or cumulative smoking for at least 6 months before the survey and smoking within the past 30 days.

HIV Laboratory Test Procedure

Venous blood samples were collected from the study participants. First, rapid HIV test reagents were used; those who tested reactive were retested using ELISA (enzyme-linked immunosorbent assay). If reactive, they were subjected to an HIV supplemental test. The supplemental test was first confirmed using the Western Blot method. If the Western Blot yielded indeterminate or negative results in the confirmatory test, the participants underwent a quantitative nucleic acid test using viral load detection reagents. Participants with nonreactive findings in the initial rapid test or the enzyme immunoassay were further tested using pooled nucleic acid tests.

Statistical Analysis

Based on the baseline data, frequency analyses stratified by HIV seroconversion status were performed. χ^2 tests were used to evaluate the differences in characteristics between participants with and without HIV seroconversion. For HIV seroconversions, the midpoint between the last negative and the first positive test dates was assumed to be the estimated onset date of HIV infection. For participants who did not seroconvert (ie, test positive for HIV), the date of their last survey was used as the endpoint for our analysis. Individual observation time was calculated as the interval between the participant’s baseline survey and their latest HIV test during the study period. HIV incidence was calculated by dividing the number of observed HIV seroconversions by the total observed person-time. The 95% CI for HIV incidence were calculated using person-years (PYs) over the observed time as the denominator. The proportional hazards assumptions for the Cox models were tested by examining the interactions between variables included in the model and the logarithm of the follow-up time; no violations were observed. Therefore, Cox proportional hazards models were used to estimate the associations between various factors and HIV incidence. A multivariate analysis was conducted based on variables identified in univariate analysis; marginally significant variables ($P \leq .10$) in univariate analysis and factors that may be associated with new HIV infections were included in a multivariate Cox proportional hazards regression model to identify factors associated with HIV incidence. All analyses were performed using SPSS software (version 24.0; IBM Corp). A two-tailed $P<.05$ was considered statistically significant.

Ethical Considerations

This study was approved by the Ethics Committee of Tianjin Centers for Disease Control and Prevention (approval number: TJCDC-R-2023 - 020) and was carried out in accordance with the Declaration of Helsinki. Ethical guidelines were followed during all the stages of this study. Comprehensive details about the study were provided to the participants during their first survey, and participation in the study was entirely voluntary and all participants provided written informed consent. Participants had the right to opt out of the study at any time. When accessing the dataset of this study, it was entirely deidentifed, and there was no effective way to link the data to the participants; further, the data were completely anonymized during data processing and statistical analyses. We did not pay any monetary compensation to the participants; however, we provided free HIV counseling and testing services.

Results

The Cohort

A total of 2919 YMSM completed the baseline survey in this study, of which 138 YMSM yielded positive results for HIV and were excluded from the initial investigation, with 2781 HIV-negative YMSM remaining after the first test. An additional 1403 YMSM who completed only the baseline survey were further excluded from the cohort. A total of 1378 YMSM then participated in two or more surveys during the study. Among these, 6 YMSM who underwent follow-up and baseline investigations on the same day were excluded. Additionally, 5 YMSM were excluded due to positive seroconversion within the window period (7d) after the baseline investigation. Lastly, a total of 1367 HIV-negative YMSM aged 16 - 24 years, each with at least two HIV test records were identified and included in our cohort. Among the 1367 HIV-negative YMSM who were enrolled in our retrospective cohort study, 62 YMSM were newly infected with HIV during the study period (Figure 1). All participants in the cohort were followed up and tested simultaneously.

A total of 1403 YMSM were lost to follow-up in our study. Compared to those retained in the cohort YMSM, lost to follow-up YMSM had a higher proportion of out-of-school individuals, a higher frequency of anal intercourse in the last week, lower condom use during their last anal sex experience, and were less likely to use psychoactive substances (Table 1).

Figure 1. Retrospective cohort study screening process among YMSM in Tianjin, China,2017 - 2022. YMSM: young men who have sex with men.

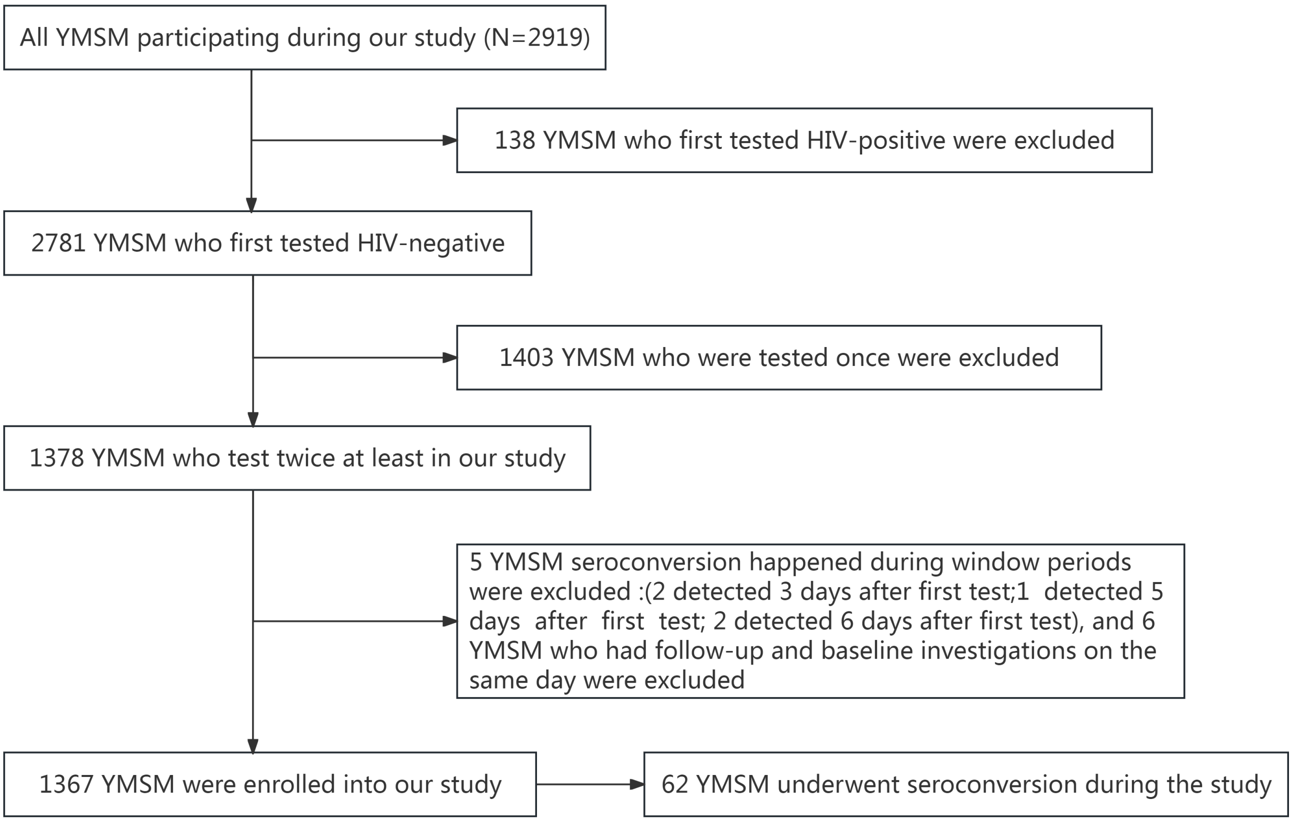


Table . Comparisons of characteristics between the cohort YMSM[a] and the lost to follow-up YMSM in Tianjin, China, 2017 - 2022.

Variables	The cohort YMSM (N=1367), n (%)	The lost to follow-up YMSM (n=1403), n (%)	*P* value
Education level			.67
High school or below	266 (19.5)	282 (20.1)	
College and above	1101 (80.5)	1121 (79.9)	
Duration of residence in Tianjin (years)			.10
<2	304 (22.2)	349 (24.9)	
≧2	1063 (77.8)	1054 (75.1)	
Occupation			.03
Out-of-school YMSM	742 (54.3)	818 (58.3)	
Students YMSM	625 (45.7)	585 (41.7)	
Sexual orientation			.57
Homosexual	1283 (93.9)	1324 (94.4)	
Bisexual/heterosexual	84 (6.1)	79 (5.6)	
Primary venue to find sexual partners			.37
Online	1193 (87.3)	1208 (86.1)	
Offline	174 (12.7)	195 (13.9)	
Number of anal intercourse last week			.02
<2	1127 (82.4)	1109 (79)	
≧2	240 (17.6)	294 (21)	
Used condoms for anal sex last time			.002
No	451 (33)	541 (38.6)	
Yes	916 (67)	862 (61.4)	
Psychoactive substance use			.001
No	699 (51.1)	802 (57.2)	
Yes	668 (48.9)	601 (42.8)	

[a] YMSM: young men who have sex with men.

Demographic Characteristics

The majority (1101/1367, 80.5%) of participants completed college or university education. Most participants (n=1063, 77.8%) lived in Tianjin for over two years. A total of 45.7% (n=625) participants were students. The cohort participants' education level and duration of residence in Tianjin were statistical significant between different HIV seroconversion statuses (*P*<.05), whereas occupation and number of follow-up visits did not significantly differ by different statuses of HIV seroconversion (*P*>.05) (Table 2).

Table . Demographic characteristics of the cohort YMSM[a] in Tianjin,China,2017 - 2022.

Variables	Participants (N=1367), n (%)	HIV seroconversion		*P* value
		Yes, n (%)	No, n (%)	
Education level				.02
High school or below	266 (19.5)	19 (30.6)	247 (18.9)	
College and above	1101 (80.5)	43 (69.4)	1058 (81.1)	
Duration of residence in Tianjin (years)				.03
<2	304 (22.2)	7 (11.3)	297 (22.8)	
≧2	1063 (77.8)	55 (88.7)	1008 (77.2)	
Occupation				.26
Out-of-school YMSM	742 (54.3)	38 (61.3)	704 (53.9)	
Student YMSM	625 (45.7)	24 (38.7)	601 (46.1)	
Times of follow-up in the study				.47
<3	875 (64)	37 (59.7)	838 (64.2)	
≥3	492 (36)	25 (40.3)	467 (35.8)	

[a] YMSM: young men who have sex with men.

Sexual Behaviors, Syphilis Infection, and Psychoactive Substance Use

Among participants, 93.9% (1283/1367) identified themselves as homosexual, and 87.3% (n=1193) mainly found sexual partners online. In the past 6 months, 97.4% (n=1331) had anal intercourse, 17.6% (n=240) had more than two instances of anal intercourse in the last week, and 33.0%(n=451) did not use condoms during their last anal sex experience. Most (n=999, 73.1%) participants had their first homosexual intercourse between the ages of 16 - 19 years. Additionally, 7.3% (n=100) were male sex workers and 7.8% (n=107) participants were infected with syphilis. A total of 48.9% (n=668) used psychoactive substances, 25.2% (n=345) were smokers, and 41.6% (n=569) consumed alcohol (Table 3). The rate of psychoactive drug use was higher among smokers (209/345, 60.6%) than nonsmokers (459/1022, 44.9%*;* *P*<.001). Similarly, the rate of psychoactive drug use was more common among alcohol drinkers (310/569, 54.5%) than nondrinkers at (358/798, 44.9%; *P*<.001).

Table . Sexual behaviors, syphilis infection, and psychoactive substances use among the cohort YMSM[a] in Tianjin, China, 2017‑2022.

Variables	Participants (N=1367), n (%)	HIV seroconversion		*P* value
		Yes, n (%)	No, n (%)	
Sexual orientation				.48
Homosexual	1283 (93.9)	60 (96.8)	1223 (93.7)	
Bisexual/heterosexual	84 (6.1)	2 (3.2)	82 (6.3)	
Primary venue to find sexual partners				.02
Online	1193 (87.3)	48 (77.4)	1145 (87.7)	
Offline	174 (12.7)	14 (22.6)	160 (12.3)	
Number of anal intercourse last week				.32
<2	1127 (82.4)	54 (87.1)	1073 (82.2)	
≧2	240 (17.6)	8 (12.9)	232 (17.8)	
Use a condom for anal sex last time				.50
No	451 (33.0)	18 (29)	433 (33.2)	
Yes	916 (67.0)	44 (71)	872 (66.8)	
Age of first homosexual intercourse (years)				.02
<16	70 (5.1)	8 (12.9)	62 (4.8)	
16‑19	999 (73.1)	42 (67.7)	957 (73.3)	
>19	298 (21.8)	12 (19.4)	286 (21.9)	
Syphilis infection				.006
No	1260 (92.2)	51 (82.3)	1209 (92.6)	
Yes	107 (7.8)	11 (17.7)	96 (7.4)	
Male sex worker				.13
No	1267 (92.7)	61 (98.4)	1206 (92.4)	
Yes	100 (7.3)	1 (1.6)	99 (7.6)	
Psychoactive substance use				.001
No	699 (51.1)	19 (30.6)	680 (52.1)	
Yes	668 (48.9)	43 (69.4)	625 (47.9)	
Smoking				.69
No	1022 (74.8)	45 (72.6)	977 (74.9)	
Yes	345 (25.2)	17 (27.4)	328 (25.1)	
Alcohol consumption				.17
No	798 (58.4)	31 (50)	767 (58.8)	
Yes	569 (41.6)	31 (50)	538 (41.2)	
Know about PrEP[bc]				.03
No	513 (63.3)	26 (81.3)	487 (62.6)	
Yes	297 (36.7)	6 (18.7)	291 (37.4)	

[a] YMSM: young men who have sex with other men.

[b] There is missing data.

[c] PrEP: pre-exposure prophylaxis.

The primary venue for finding sexual partners, age of first homosexual intercourse, syphilis infection, and psychoactive substance use were statistically significant between different HIV seroconversion statuses (all *Ps*<.05). However, the number of instances of anal intercourse in the last week, condom use during last anal sex, male sex worker status, smoking status, and alcohol consumption status were not significantly associated with HIV seroconversion (all *Ps*>.05) (Table 3).

A total of 95.8% (n=599) student MSM considered themselves homosexual men, which was higher than the proportion among out-of-school YMSM; additionally, 93.0%(n=581) of student YMSM mainly found sexual partners online, which was higher than among out-of-school YMSM. Student YMSM also had lower rates of using condoms for anal sex, syphilis infection, psychoactive substance use, smoking, and alcohol consumption compared to out-of-school YMSM (Table 4).

Table . Comparisons of characteristics between the out-of-school YMSM[a] and the student YMSM among the YMSM cohort in Tianjin, China, 2017 - 2022.

Variables	Out-of-school YMSM, n (%)	Students YMSM, n (%)	*P* value
Sexual orientation			.005
Homosexual	684 (92.2)	599 (95.8)	
Bisexual/heterosexual	58 (7.8)	26 (4.2)	
Primary venue to find sexual partners			<.001
Online	612 (82.5)	581 (93.0)	
Offline	130 (17.5)	44 (7.0)	
Frequency of anal intercourse last week			<.001
<2	575 (77.5)	552 (88.3)	
≧2	167 (22.5)	73 (11.7)	
Use of a condom for anal sex during last sexual encounter			<.001
No	189 (25.5)	262 (41.9)	
Yes	553 (74.5)	363 (58.1)	
Age at first homosexual intercourse (years)			<.001
<16	42 (5.6)	28 (4.5)	
16 - 19	497 (67.0)	502 (80.3)	
>19	203 (27.4)	95 (15.2)	
Syphilis infection			<.001
No	657 (88.5)	603 (96.5)	
Yes	85 (11.5)	22 (3.5)	
Psychoactive substance use			<.001
No	321 (43.3)	378 (60.5)	
Yes	421 (56.7)	247 (39.5)	
Smoking			<.001
No	485 (65.4)	537 (85.9)	
Yes	257 (34.6)	88 (14.1)	
Alcohol consumption			<.001
No	383 (51.6)	415 (66.4)	
Yes	359 (48.4)	210 (33.6)	

[a] YMSM: young men who have sex with men.

HIV Incidence and the Associated Factors

Follow-up encounters ranged from 1 to 20, with the follow-up intervals ranging from 0.4 to 60.3 months with a median value of 16.5 (IQR 8.3 - 31.5) months among participants. The total person-time observed was 2384.2 PYs. The overall HIV incidence was 2.6 per 100 PYs (95% CI 2.0 - 3.2). To reduce the impact of random errors on the incidence rate and provide a more comprehensive perspective on observing incidence trends, we used a 2-year incidence rate. The segmented HIV incidence rates were 3.0 (95% CI 1.5 - 4.5), 2.4 (95% CI 1.5 - 3.3), and 2.7 (95% CI 1.5 - 3.9) per 100 PYs in 2017 - 2018, 2019 - 2020, and 2021 - 2022, respectively (Figure 2).

Figure 2. Trends in HIV incidence and 95% CI among YMSM in Tianjin, China, 2017 - 2022. PY: person-year. YMSM: young men who have sex with men.

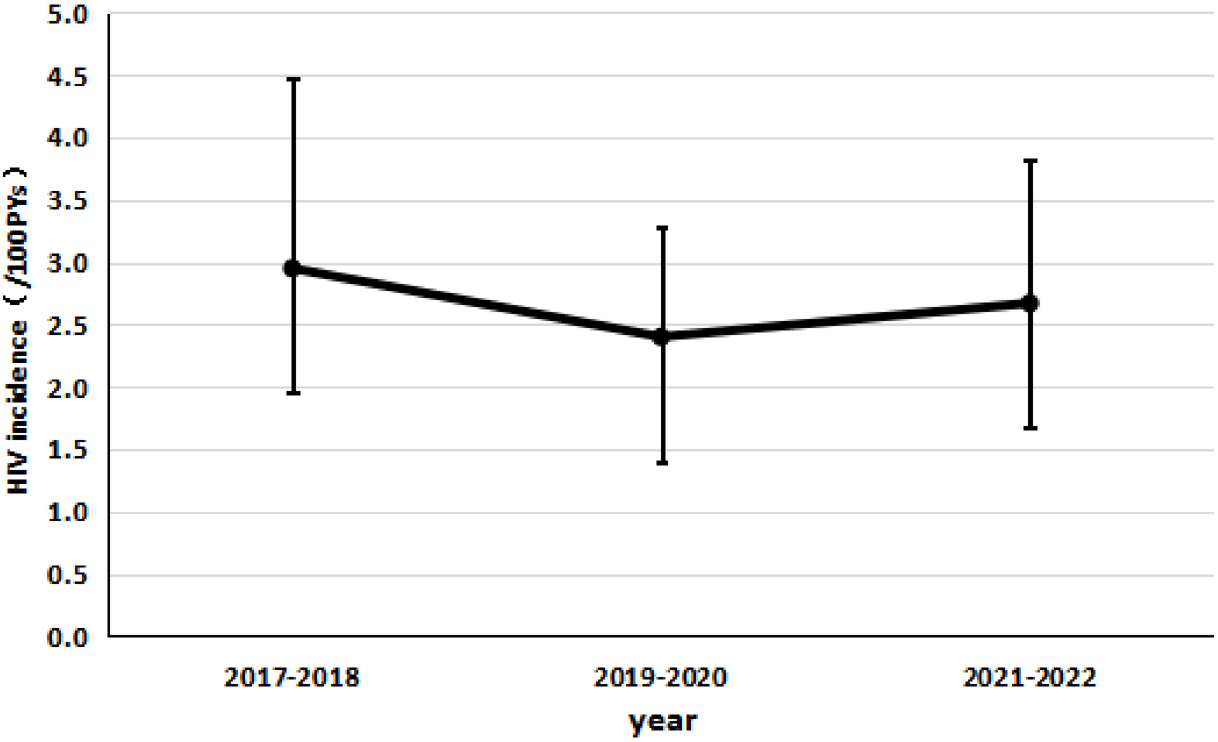


Table 5 shows that YMSM who were followed-up at least three times in the study had an incidence of 2.0 per 100 PYs. and were at a lower risk of HIV infection compared to those followed-up fewer than three times during the study (Adjusted hazard ratio [AHR] 0.553, 95% CI 0.325-0.941). The members who found sexual partners offline had an HIV incidence of 4.7 per 100 PYs. These individuals were at more than twice the risk of HIV infection compared to individuals who find sexual partners online (AHR 2.207, 95% CI 1.198-4.066). People infected with syphilis had an HIV incidence of 5.6 per 100 PYs and had a 2.234-fold higher risk of HIV infection compared to those without infected syphilis (AHR 2.234, 95% CI 1.137-4.391). The YMSM who used psychoactive substances had an HIV incidence of 3.7 per 100 PYs and a 2.467-fold greater risk of HIV infection compared to those who did not use psychoactive substances (AHR 2.467, 95% CI 1.408-4.321).

Table . Factors associated with HIV incidence identified using a proportional hazards regression model among YMSM[a] in Tianjin, China, 2017 - 2022.

Variables	Incidence (per 100 PYs)[b]	Univariate analysis		Multivariate analysis	
		HR[c] value (95% CI)	*P* value	HR value (95% CI)	*P* value
Education level			.08		.08
High school or below	3.7	1.000		1.000	
College and above	2.3	0.616 (0.359-1.057)		0.593 (0.328-1.073)	
Duration of residence in Tianjin (years)			.07		.13
<2	1.4	1.000		1.000	
≧2	2.9	2.061 (0.938-4.527)		1.887 (0.837-4.252)	
Occupation			.63		.53
Nonstudents	2.7	1.000		1.000	
Students	2.4	0.882 (0.528-1.472)		1.201 (0.678-2.126)	
Times of follow-up in the study			.11		.03
<3	3.2	1.000		1.000	
≧3	2.0	0.658 (0.392-1.104)		0.553 (0.325-0.941)	
Sexual orientation			.31		.65
Homosexual	2.7	1.000		1.000	
Bisexual/heterosexual	1.3	0.484 (0.118-1.978)		0.716 (0.170-3.010)	
Primary venue to find sexual partners			.02		.01
Online	2.3	1.000		1.000	
Offline	4.7	2.054 (1.132-3.725)		2.207 (1.198-4.066)	
Number of anal intercourse last week			.48		.55
<2	2.7	1.000		1.000	
≧2	2.1	0.764 (0.364-1.606)		0.791 (0.369-1.695)	
Use a condom for anal sex last time			.65		.88
No	2.9	1.000		1.000	
Yes	2.5	0.881 (0.508-1.530)		0.956 (0.542-1.687)	
Age of first homosexual intercourse (years)					
<16	6.2	1.000	—[d]	1.000	—[d]
16 - 19	2.4	0.383 (0.180-0.815)	.01	0.490 (0.226-1.059)	.07
>19	2.3	0.376 (0.153-0.919)	.03	0.547 (0.217-1.380)	.20
Syphilis Infection			.007		.02
No	2.3	1.000		1.000	
Yes	5.6	2.440 (1.272-4.683)		2.234 (1.137-4.391)	
Male sex worker			.12		.13
No	2.8	1.000		1.000	
Yes	0.8	0.207 (0.029-1.496)		0.209 (0.027-1.614)	
Psychoactive substance use			.002		.002
No	1.6	1.000		1.000	
Yes	3.7	2.369 (1.380-4.065)		2.467 (1.408-4.321)	

[a] YMSM: young men who have sex with men.

[b] PY: person-year.

[c] HR: hazard ratio.

[d]not applicable.

Discussion

Principal Findings

Our study is the first long-term cohort study to assess the incidence of HIV and its associated factors among YMSM in Tianjin. The HIV incidence identified in this study was similar to the incidence among MSM under 24 years old (2.63/100PYs) in Shanghai, published in 2019 [17]. This study found that the incidence of HIV among YMSM in Tianjin showed a stable trend from 2017 to 2022 and was maintained at an elevated level. This steady trend in HIV incidence suggests that risk factors for HIV infection continue to exist; additionally, as a group whose capacity for self-regulation has not fully matured [18], YMSM are more likely to engage in risky behaviors [19]. In our study, approximately 80% of the cohort members had their first homosexual encounter before the age of 20. Individuals who had a sexual debut at an earlier age are more likely to be cognitively immature and involved in high-risk sexual behaviors. Earlier sexual debut is associated with an increased HIV transmission risk [20]. Furthermore, a previous study identified YMSM as potentially more vulnerable to HIV infection than older MSM, reinforcing that YMSM in Tianjin are still at high risk of HIV infection [21].

Our study found a decreased risk of HIV incidence among YMSM who found sexual partners online. In contrast, for depth analysis, among those who found sexual partners offline, the primary venue (135/174, 77.6%) was gay bathhouses, where YMSM are more likely to engage in risky sexual behaviors [22]. Additionally, a study demonstrated that MSM who find partners online have more sexual encounters and multiple partners; however, this behavior was not associated with a greater risk of HIV infection [23]. This may be due to increased distrust and suspicion of partners met online.

As reported in a previous study [24], syphilis infection is a known risk factor associated with HIV incidence, as both HIV and syphilis share the same transmission route [24]. Therefore, we should conduct routine screening and treatment for syphilis infection among YMSM as a measure to prevent HIV infection [25].

Consistent with our findings, previous studies have shown that YMSM who participate in follow-up visits more frequently have a lower HIV incidence [26]. This could be attributed to these individuals being more concerned about HIV infection, which could reduce their sexual risk behaviors, such as decreasing the likelihood of condomless anal sex with a man whose HIV status was unknown. Although data on knowledge of HIV pre-exposure prophylaxis (PrEP) was partially missing in our study, YMSM who experienced HIV seroconversion had a lower awareness rate of HIV PrEP than those who did not. Therefore, we should take effective prevention measures such as PrEP, in addition to routine interventions, to decrease new HIV infections among YMSM.

The use of psychoactive substances was identified as a risk factor for HIV incidence in our study. Additionally, 48.9% of YMSM reported using psychoactive substances, which is higher than that among all MSM in the same region [27]. Psychoactive substances are mostly sold online, and 90% of YMSM in our study mainly found sexual partners online, a higher proportion than among the general MSM population [27]. Furthermore, with the development of network technology and improved internet access, more young people seek health services, particularly HIV-related services, through the internet [28]. Previous studies revealed that MSM who use gay social networking apps reported a higher number of sexual partners, suggesting that we should explore effective intervention strategies using the internet [29,30].

Another finding in our study was that YMSM who smoked or consumed alcohol had a higher prevalence of psychoactive substance use, which is a risk factor for HIV infection. This suggests that strengthening health education initiatives for reducing smoking and alcohol consumption among YMSM may help lower the incidence of HIV.

YMSM in Tianjin remain a vulnerable subpopulation for HIV infection and should be given priority for HIV intervention programs. Ongoing HIV control efforts among young students in schools in Tianjin, such as lectures on HIV, large-scale publicity and education, knowledge contests on HIV prevention, and debate competitions. Given the ongoing HIV epidemic among young students in Tianjin, we designed posters for student YMSM and displayed them in all universities, which yielded a good effect. However, performing effective work is challenging for the out-of-school YMSM due to the absence of a fixed setting, such as a school for intervention. In addition, our study found that, compared to the student YMSM, out-of-school YMSM were more inclined to engage in high-risk behaviors and had a higher proportion of bisexual or heterosexual orientation. Consequently, out-of-school YMSM are more likely to serve as a bridge population for HIV transmission from high-risk MSM to the general population. Therefore, additional prevention options and efforts should be made available to this group.

There are several limitations to this study. First, it was impossible to follow up on the HIV infection status of the individuals who did not enter the cohort; we aim to conduct the follow-up study in the future. Second, as we conducted our study through an NGO, the target population who did not seek HIV testing through the NGO were excluded, potentially leading to selection bias. Finally, given the sensitive nature of the topics, such as sexual behavior, reporting bias may exist.

Conclusions

HIV incidence among YMSM in Tianjin remains high and has not changed over time. Syphilis infection and the use of psychoactive substances continue to be risk factors associated with HIV incidence. Additionally, risky behaviors such as the use of psychoactive substances are widely present among YMSM; therefore, tailored interventions for this population should be strengthened. Our study supplements the data and evidence for research on new HIV infections among YMSM, both in and out-of-school in China's megacities and has significant practical implications.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (72374153), Tianjin Health Research Project (grant No: TJWJ2023QN092), and Tianjin Medical Key Discipline Project (grant number: TJYXZDXK-50A). We would like to express our gratitude to all the staff and study participants.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to privacy restrictions but are available from the corresponding author for reasonable request.

Authors' Contributions

Conceptualization: HG, MY, YG
Supervision: HG, MY, YG
Writing – original draft – HG, YG
Writing – review & editing – CM, JL, LL[1], LL[2],TN, XY, XZ, ZC, ZL.

Conflicts of Interest

None declared.

References

1. The path that ends AIDS: UNAIDS global AIDS update 2023. UNAIDS. 2023. URL: https://idpc.net/publications/2023/07/the-path-that-ends-aids-2023-unads-global-aids-update [accessed 2024-07-13]
2. UNAIDS. Putting young key populations first: HIV and young people from key populations in the asia and pacific region 2022. Geneva: Joint United Nations Programme on HIV/AIDS. 2022. URL: https://www.unaids.org/en/resources/documents/2022/2022-HIV-young-people-key-populations-asia-pacific [accessed 2023-07-05]
3. Baggaley R, Armstrong A, Dodd Z, Ngoksin E, Krug A. Young key populations and HIV: a special emphasis and consideration in the new WHO Consolidated Guidelines on HIV prevention, diagnosis, treatment and care for key populations. J Int AIDS Soc 2015 Feb;18(2S1). [doi: 10.7448/IAS.18.2.19438] [Medline: 25724509]
4. Zhao H, Liu H, Wang L, et al. Epidemiological characteristics of newly- reported HIV cases among youth aged 15-24 years-China, 2010 to 2019. China CDC Weekly 2020;2(48):913. [doi: 10.46234/ccdcw2020.249]
5. Garofalo R, Hotton AL, Kuhns LM, Gratzer B, Mustanski B. Incidence of HIV infection and sexually transmitted infections and related risk factors among very young men who have sex with men. J Acquir Immune Defic Syndr 2016 May 1;72(1):79-86. [doi: 10.1097/QAI.0000000000000933] [Medline: 26745827]
6. Dong Y, Wang L, Burgner DP, et al. Infectious diseases in children and adolescents in China: analysis of national surveillance data from 2008 to 2017. BMJ 2020 Apr 2;369:m1043. [doi: 10.1136/bmj.m1043] [Medline: 32241761]
7. UNAIDS. Young people and HIV. 2021. URL: https://www.unaids.org/en/resources/documents/2021/young-people-and-hiv [accessed 2021-05-07]
8. Mao X, Wang Z, Hu Q, et al. HIV incidence is rapidly increasing with age among young men who have sex with men in China: a multicentre cross-sectional survey. HIV Med 2018 Jun 19;19(8):513-522. [doi: 10.1111/hiv.12623] [Medline: 29923304]
9. Zhang H, Chen CZ, Lin SY, Xu HH, Xue DC, He ZG, et al. HIV/AIDS surveillance in men who have sex with men aged 15-24 based on internet in Fuzhou, 2016-2021. Chin J Epidemiol 2022;43(11):1761-1767. [doi: 10.3760/cma.j.cn112338-20220430-00365]
10. Coelho LE, Torres TS, Veloso VG, et al. The prevalence of HIV among men who have sex with men (MSM) and young MSM in Latin America and the Caribbean: a systematic review. AIDS Behav 2021 Oct;25(10):3223-3237. [doi: 10.1007/s10461-021-03180-5] [Medline: 33587242]
11. Yan G, Fangfang C, Hui G, et al. Trends in Incidence of HIV and Syphilis Among Men Who Have Sex with Men — Tianjin Municipality, China, 2013–2022. China CDC Wkly 2024;6(27):651-657. [doi: 10.46234/ccdcw2024.121]
12. Mwaniki SW, Mugo PM, Phillips TP. Project BESPOKE (Integrated bio-behavioral assessment of HIV and STI among young tertiary student men who have sex with men in Nairobi, Kenya): a respondent-driven sampling survey protocol. Front Public Health 2021 Oct 11;9:619694. [doi: 10.3389/fpubh.2021.619694] [Medline: 34708012]
13. State Council Seventh National Census Leading Group Office.: China Statistics Press; 2020. URL: https://www.stats.gov.cn/sj/ndsj/2020/indexch.htm [accessed 2024-07-13]
14. Tianjin bureau of statistics. Tianjin Statistical Yearbook. 2022. URL: https://stats.tj.gov.cn/nianjian/2022nj/zk/indexch.htm [accessed 2024-07-13]

15. Guo Y, Jie Y, Jianyun B. Study on the new incidence of HIV infection among young men who have sex with men using psychoactive substances in Tianjin. Chin J Viral Dis 2025;15(1):42-48. [doi: 10.16505/j.2095-0136.2024.0062]
16. Guo Y, Yang J, Bai JY, et al. Analysis on psychoactive substance use in young men who have sex with men in Tianjin[J]. Chin J Epidemiol 2024;45(5):687-693. [doi: 10.3760/cma.j.cn112338-20211104-00851]
17. Yin FL, Zhong PS, Zhang Y, Wang JX, Lu M, et al. HIV incidence,retention rate and related factors in a cohort study among men who have sex with men in Shanghai suburb. Chin J AIDS STD 2019 Sep 19;25(9):947-950. [doi: 10.13419/j.cnki.aids.2019.09.19]
18. Mi GD, Zhang BB, Yu F, et al. HIV incidence among men who have sex with men using geosocial networking smartphone application in Beijing, China: an open cohort study. Infect Dis Poverty 2021 Apr 2;10(1):27. [doi: 10.1186/s40249-021-00814-7] [Medline: 33795011]
19. Steinberg L. Risk taking in adolescence: what changes, and why? Ann N Y Acad Sci 2004 Jun;1021(1):51-58. [doi: 10.1196/annals.1308.005] [Medline: 15251873]
20. Balaji AB, An Q, Smith JC, et al. High human immunodeficiency virus incidence and prevalence and associated factors among adolescent sexual minority males-3 cities, 2015. Clin Infect Dis 2018 Mar 5;66(6):936-944. [doi: 10.1093/cid/cix902] [Medline: 29069298]
21. Wei L, Tian J, Guo M, et al. Trajectories of sexual risk behaviors and the associated factors among young men who have sex with men in China. Front Public Health 2022;10:854616. [doi: 10.3389/fpubh.2022.854616] [Medline: 35387185]
22. Wu Z, Xu J, Liu E, et al. HIV and syphilis prevalence among men who have sex with men: a cross-sectional survey of 61 cities in China. Clin Infect Dis 2013 Jul;57(2):298-309. [doi: 10.1093/cid/cit210] [Medline: 23580732]
23. Hong H, Xu J, McGoogan J, Dong H, Xu G, Wu Z. Relationship between the use of gay mobile phone applications and HIV infection among men who have sex with men in Ningbo, China: a cross-sectional study. Int J STD AIDS 2018 Apr;29(5):491-497. [doi: 10.1177/0956462417738468] [Medline: 29099328]
24. Macapagal K, Moskowitz DA, Li DH, et al. Hookup app use, sexual behavior, and sexual health among adolescent men who have sex with men in the United States. J Adolesc Health 2018 Jun;62(6):708-715. [doi: 10.1016/j.jadohealth.2018.01.001] [Medline: 29784114]
25. US Preventive Services Task Force, Mangione CM, Barry MJ, et al. Screening for syphilis infection in nonpregnant adolescents and adults: US Preventive Services Task Force Reaffirmation Recommendation Statement. JAMA 2022 Sep 27;328(12):1243-1249. [doi: 10.1001/jama.2022.15322] [Medline: 36166020]
26. Hui S, Chen F, Li Y, et al. Factors associated with newly HIV infection and transmitted drug resistance among men who have sex with men in Harbin, P.R. China. Front Public Health 2022;10:860171. [doi: 10.3389/fpubh.2022.860171] [Medline: 35719611]
27. Guo Y, Yang J, Bai JY, et al. Analysis on psychoactive substances use in men who have sex with men in Tianjin. Chin J Epidemiol 2022;43(5):686-691. [doi: 10.3760/cma.j.cn112338-20211104-00851] [Medline: 33041319]
28. You X, Gilmour S, Cao W, et al. HIV incidence and sexual behavioral correlates among 4578 men who have sex with men (MSM) in Chengdu, China: a retrospective cohort study. BMC Public Health 2021 Apr 26;21(1):802. [doi: 10.1186/s12889-021-10835-4] [Medline: 33902499]
29. Buhi ER, Daley EM, Oberne A, Smith SA, Schneider T, Fuhrmann HJ. Quality and accuracy of sexual health information web sites visited by young people. J Adolesc Health 2010 Aug;47(2):206-208. [doi: 10.1016/j.jadohealth.2010.01.002] [Medline: 20638014]
30. Zou H, Fan S. Characteristics of men who have sex with men who use smartphone geosocial networking applications and implications for HIV interventions: a systematic review and meta-analysis. Arch Sex Behav 2017 May;46(4):885-894. [doi: 10.1007/s10508-016-0709-3] [Medline: 27043837]

Abbreviations

AHR: adjusted hazard ratio
MSM: men who have sex with men
NGO: nongovernmental organization
PrEP: pre-exposure prophylaxis
PY: person-years
YMSM: young men who have sex with men

Edited by A Mavragani; submitted 20.09.24; peer-reviewed by J Li, K Jiao; revised version received 05.02.25; accepted 06.02.25; published 19.03.25.

Please cite as:
Guo Y, Gong H, Yan X, Zhang X, Ning T, Liu Z, Li L, Yang J, Li L, Ma C, Cui Z, Yu M
HIV Incidence and Associated Risk Factors Among Young Men Who Have Sex With Men in Tianjin, China: Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e66487
URL: https://publichealth.jmir.org/2025/1/e66487
doi:10.2196/66487

Awareness of Sexual Partner’s HIV Status Among Men Who Have Sex With Men in China: Cross-Sec. tional Survey Study

Qian Ma[1], MA; Tingting Jiang[2], MA; Wanjun Chen[2], MA; Shaoqiang Jiang[3], BA; Jinlei Zheng[2], MA; Hui Wang[2], MA; Lin He[2], MA

[1]Department of Rheumatology Immunology & Allergy, Zhejiang University School of Medicine, Hangzhou, China
[2]Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, China
[3]Coastal Service Center, Qiantang District, Hangzhou, Zhejiang, China

Corresponding Author:
Lin He, MA
Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, China

Abstract

Background: Men who have sex with men (MSM) constitute a significant proportion of individuals living with human immunodeficiency virus. Over the past few years, China has implemented various strategies aimed at increasing the rate of HIV testing and reducing HIV transmission among MSM. Among these, the disclosure of HIV serostatus is an effective prevention strategy.

Objective: This study aimed to assess HIV serostatus disclosure and identify factors associated with awareness of sexual partners’ HIV status among MSM to provide a scientific basis for promoting HIV testing and reducing HIV transmission.

Methods: A cross-sectional study based on a large-scale web-based survey was conducted among MSM in Zhejiang province, China, between July and December 2023. MSM who were HIV-negative or had an unknown HIV status were recruited from the Sunshine Test, a web-based platform that uses location-based services to provide HIV prevention services. Participants were required to complete a questionnaire on demographic characteristics, sexual behavior, rush popper use, awareness of sexual partners’ HIV status, and knowledge of pre-exposure prophylaxis (PrEP) and postexposure prophylaxis (PEP). A multinomial regression model was used to identify the factors associated with awareness of sexual partners’ HIV status.

Results: A total of 7629 MSM participated in the study, with 45.2% (n=3451) being aware, 35.4% (n=2701) being partially aware, and 19.4% (n=1477) being unaware of their sexual partner’s HIV status. The multinomial logistic regression analysis revealed the following results. Compared to those who were unaware of their sexual partner’s HIV status, participants who were students (adjusted odds ratio [aOR] 1.43, 95% CI 1.09 - 1.86), had a monthly income of more than US $1400 (aOR 1.36, 95% CI 1.03 - 1.80), had insertive anal sex (aOR 1.35, 95% CI 1.12 - 1.63), had only male sexual partners (aOR 1.53, 95% CI 1.28 - 1.82), had 1 sexual partner in the past 3 months (aOR 2.36, 95% CI 2.01 - 2.77), had used condoms for the past 3 months (aOR 1.72, 95% CI 1.33 - 2.22), had frequently used rush poppers in the past 3 months (aOR 2.27, 95% CI 1.81 - 2.86), were aware of HIV PrEP (aOR 2.04, 95% CI 1.68 - 2.48), were aware of HIV PEP (aOR 1.69, 95% CI 1.39 - 2.06), used mail reagent self-testing (aOR 1.19, 95% CI 1.04 - 1.36), and had previously undergone HIV testing (aOR 1.40, 95% CI 1.16 - 1.69) were associated with increased odds of being aware of their sexual partner’s HIV status.

Conclusions: Overall, 45.2% of MSM who were HIV-negative or had an unknown status were aware of their sexual partners’ HIV status in China. We suggest strengthening targeted interventions through web-based platforms and gay apps to promote the disclosure of HIV serostatus and reduce HIV transmission among MSM.

(JMIR Public Health Surveill 2025;11:e66139) doi:10.2196/66139

KEYWORDS

human immunodeficiency virus; HIV; serostatus disclosure; pre-exposure prophylaxis; PrEP; men who have sex with men; MSM; web-based survey; HIV awareness

Introduction

Men who have sex with men (MSM) constitute a significant proportion of individuals living with human immunodeficiency virus. The overall HIV prevalence among MSM in China has remained at 8% for the past several years [1]. The annual number of newly diagnosed HIV infections is >100,000, with 25.7% of new cases resulting from homosexual transmission in 2023 [2]. The Joint United Nations Program on HIV/AIDS (UNAIDS) proposed a 95-95-95 target by 2030 [3]. The target was to achieve 95% of patients infected with HIV diagnosed, 95% of diagnosed patients treated, and 95% of treated patients achieving viral suppression. In China, the second and third 95% targets (95.1% and 97.3% in 2023, respectively) were achieved with the implementation of a free treatment policy and lifelong follow-up, but only 84.3% of patients have been diagnosed [2]. There are challenges to achieving the first 95% target in China. Scaling up HIV screening for MSM is an important means to achieving the first 95% target, but there are still major challenges. Data from 2023 showed that only 88% of MSM had ever been tested for HIV [4]. Various testing strategies, including regular testing, self-testing, web-based mail testing, and serostatus disclosure, have been promoted to expand HIV testing among MSM [5].

The serostatus disclosure strategy is an effective measure for reducing and controlling HIV transmission. Obermeyer et al [6] have defined HIV serostatus disclosure as “the process of revealing a person’s HIV status, whether positive or negative.” Serostatus disclosure requires a person to ask or know their partner’s HIV status before having sex and to only have sex if they have the same HIV status or use protective measures. Self-disclosure is the most common form of disclosure, which requires a person to share information about his or her HIV status directly with another person. In recent years, serostatus disclosure strategies have been the focus of intervention efforts among people with a high risk of HIV infection and have played an important role in HIV transmission risk reduction. One study showed that disclosure was associated with a 45% reduction in HIV transmission [7]. Empirical studies have shown that disclosure of HIV serostatus by people living with HIV can reduce the risk of HIV transmission by 17.9%‑40.6% compared to no disclosure [8]. Disclosure to sexual partners could also increase the HIV testing rate and reduce transmission risk behaviors [9]. A benefit of disclosing HIV status is that it empowers sexual partners to make shared decisions about how to protect their health while engaging in an active sexual life [10]. Prevalence of HIV disclosure among MSM living with HIV ranges from 12% to 53% [11]. Accordingly, most HIV infections occur when there is a lack of knowledge of the HIV-positive status of the other partner, coupled with a low uptake of preventive measures and inconsistent condom use [12]. Common barriers to disclosure among MSM were a lower perceived risk of HIV infection, history of sexually transmitted infections, engagement in receptive sex, and having sex with casual partners [13-15]. Patients with HIV viral suppression have a lower disclosure rate [16], as it is now established that there is no risk of HIV transmission when the viral load is undetectable [17,18]. Other studies have shown that being unmarried, living in rural formal areas [15], being aged 50 years and over [19], lacking family support [20], poverty, having multiple sexual partners, and stigma [21,22] lead to a lower rate of HIV serostatus disclosure. In addition, substance use prior to sexual encounters is less likely in lower-risk locations compared to highest-risk locations (ie, bathhouse and public sex environment) [23].

In China, individuals infected with HIV are required by law to disclose an HIV infection to their spouses. Furthermore, more than 95% of patients’ spouses know their HIV status, and more than 92% of spouses have been tested for HIV yearly [24]. Therefore, China’s HIV serostatus disclosure strategy is predominantly promoted among high-risk populations susceptible to HIV infections, including MSM. This is achieved by implementing an HIV serostatus disclosure strategy with the aim of expanding HIV testing coverage, enabling the early detection of individuals infected with HIV, and subsequently reducing HIV transmission among MSM. However, there are barriers to implementing this strategy, including the prevalence of multiple sexual partners in the MSM population, fear of being rejected by sexual partners, and the personal privacy of one’s infection status. In addition, the implementation of the “undetectable equals untransmittable” strategy means that many people living with HIV either do not disclose or do not ask about the HIV status of their sexual partners. Therefore, this could be a factor in a partner’s decision to use a condom during sexual intercourse [10]. When implementing the serostatus disclosure strategy, MSM who are HIV-negative or unaware of their infection status are advised to ask their partners about their HIV infection status before sex, to self-test for HIV, and to take appropriate safety measures according to their HIV status. It has also been shown that HIV-negative MSM become more aware of their partners’ serostatuses over time [25].

With the development of the digital economy, an increasing number of MSM are looking for sexual partners through the internet and dating apps. This study shows that the level of HIV knowledge among MSM can be improved and HIV testing can be expanded by relying on the internet. Consequently, web-based interventions and HIV testing among MSM populations, including web-based counseling and mailing of HIV self-test kits, are becoming increasingly common. A previous meta-analysis showed that using mobile health technology could increase the linkage to care among MSM using HIV self-testing [26].

Zhejiang province is economically developed, with an active digital economy. Approximately 66,000 HIV tests were performed among MSM in 2023, accounting for 10% of the national proportion [2], and over 70% of HIV testing among MSM in Zhejiang province relied on the internet. The promotion of pre-exposure prophylaxis (PrEP) and postexposure prophylaxis (PEP) prevention services in China began in 2020. To work toward achieving the target of 95% of HIV infections diagnosed globally and in China, a large-scale, web-based cross-sectional study was conducted. This study aimed to assess the awareness of a sexual partner’s HIV status before engaging in sexual intercourse among MSM, especially for those who were not infected with HIV or did not know their infection status, which could be causing an information gap among MSM in China [13]. In the context of PrEP and PEP prevention services and the implementation of the “undetectable equals untransmittable” strategy, this study aimed to assess HIV serostatus disclosure and identify factors associated with awareness of sexual partners’ HIV status among MSM to provide a scientific basis for promoting HIV testing and reducing HIV transmission.

Methods

Sampling

A cross-sectional survey was conducted between July and December 2023. The survey was based on a large web-based survey of MSM. To determine the required sample size, we used the following formula:

$$n=Z\alpha/2\ P(1-P)$$

The α level was .05, *P* was estimated as 20% ‑ 30%, and δ was 0.05 resulting in a minimum sample size of 1329.

Study Participants

The following inclusion criteria were used to recruit MSM participants: (1) at least 16 years of age, (2) having had sex with men in the past year, (3) HIV-negative or unknown HIV status, and (4) living in Zhejiang province. Participants were excluded if they had previously received 2 or more HIV tests.

Participant Recruitment and Data Collection

Sunshine Coast Public Welfare is a social service organization that uses the internet to provide HIV outreach and testing services to the MSM population. Through the use of location-based services, Sunshine Coastal Public Welfare established a digital HIV prevention service. MSM in Zhejiang province could apply for free HIV-testing services by visiting the official account of the Sunshine Test on WeChat (a popular communication software in China). There were two ways for MSM to access HIV testing services: they could choose to have the test reagent sent by courier, or they could choose to self-test after receiving the reagent. Alternatively, they could choose offline services, where volunteers in their neighborhood provided HIV testing services, or they could visit a nearby voluntary counseling and testing site. MSM who applied for HIV testing services through the Sunshine Test were required to complete a questionnaire that covered demographic characteristics such as age, marital status, and education level, as well as sexual behavior characteristics such as sexual roles, number of sexual partners, use of rush poppers, sexual history, PrEP use, PEP use, and HIV testing history. All the questionnaires were completed using the Sunshine Test [4].

Statistical Analysis

For descriptive analyses, categorical variables were presented as frequencies and proportions, whereas continuous variables were presented as medians and IQRs or means and SDs. Differences in general demographic characteristics of the study population were compared using a χ^2 test. The main outcome was awareness (participants knowing all of their sexual partners' HIV statuses) and partial awareness (participants only knowing some of their partners' HIV statuses) of their sexual partners' HIV status, with those who were unaware of their sexual partner's HIV status being used as the reference group. Multinomial logistic regression was used to identify the factors associated with the awareness and partial awareness of a partners HIV status compared to the unaware group. Factors from the univariate analysis were included in the multinomial regression models. The α level was .05 and β was .1. All statistical analyses were performed using SPSS version 19.0 (IBM Corp).

Ethical Considerations

This study was approved by the Zhejiang Provincial Center for Disease Control and Prevention (approval number 2022-011-01). Written informed consent was obtained from all participants prior to survey completion. Participants received free HIV testing and health counseling, and all study procedures were conducted in accordance with approved guidelines and regulations. Participants could dop out during any stage of the survey. All the data was anonymized no amount of compensation was provided.

Results

A total of 7629 MSM participated in the study (Figure 1), and 64.9% (4956/7629) were aged less than 30 years. Among them, 50.7% (3869/7629) had a bachelor's degree or higher, 16.4% (1254/7629) were students, and 39.1% (2985/7629) had a monthly income of less than US $700. The majority (78.2%, 5964/7629) had only male sexual partners, while 33.3% (2543/7629) had more than 2 sexual partners in the past 3 months. Additionally, 31.2% (2380/7629) reported using rush poppers during sexual activity. The awareness rates of PrEP and PEP were 76.2% (5816/7629) and 78.1% (5960/7629), respectively. Furthermore, 54.8% (4183/7629) had HIV testing using a mail reagent self-test, and 12.9% (982/7629) were undergoing HIV testing for the first time (Table 1).

Figure 1. Study design.

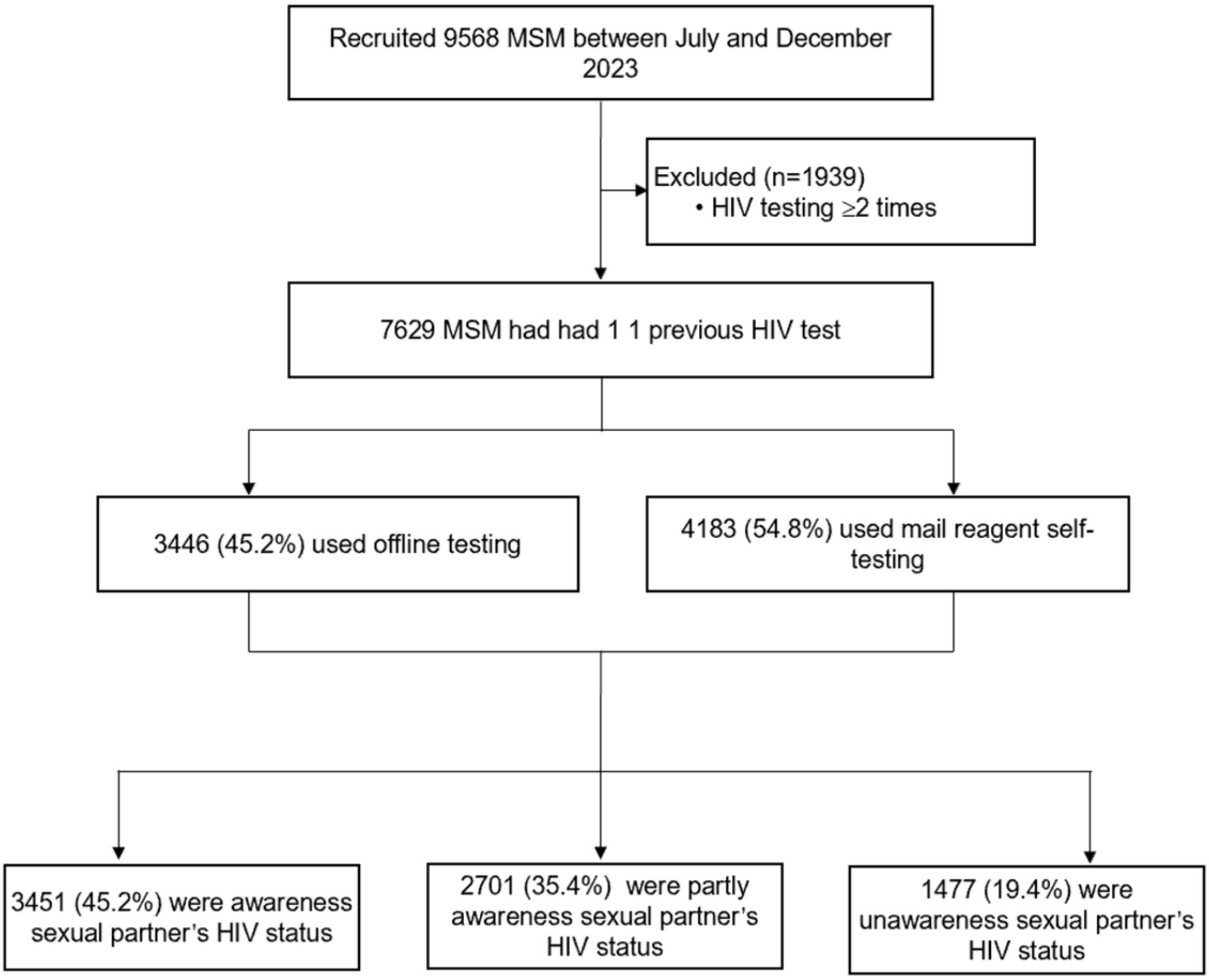


Table . Social demographic and behavioral characteristics among men who have sex with men.

Variables	Total (n=7629), n (%)	Awareness of sexual partner’s HIV status			Chi-square (*df*)	*P* value
		Yes (n=3451)	Partially (n=2701)	No (n=1477)		
Age[a] (years), n (%)					33.458 (4)	<.001
<20	430 (5.6)	185 (43)	136 (31.6)	109 (25.3)		
20-29	4526 (59.3)	2143 (47.3)	1528 (33.8)	855 (18.9)		
≥30	2673 (35)	1123 (42)	1037 (38.8)	513 (19.2)		
Marital status, n (%)					35.719 (4)	<.001
Single	5997 (78.6)	2804 (46.8)	2032 (33.9)	1161 (19.4)		
Married	1388 (18.2)	558 (40.2)	571 (41.1)	259 (18.7)		
Divorced/separated	244 (3.2)	89 (36.5)	98 (40.2)	57 (23.4)		
Education, n (%)					77.958 (6)	<.001
High school and below	1851 (24.3)	683 (36.9)	736 (39.8)	432 (23.3)		
College	1909 (25)	898 (47)	639 (33.5)	372 (19.5)		
Bachelor’s degree	3145 (41.2)	1511 (48)	1099 (34.9)	535 (17)		
Master’s degree or above	724 (9.5)	359 (49.6)	227 (31.4)	138 (19.1)		
Occupation, n (%)					54.237 (6)	<.001
Student	1254 (16.4)	608 (48.5)	375 (29.9)	271 (21.6)		
Company employee	2843 (37.3)	1370 (48.2)	1009 (35.5)	464 (16.3)		
Freelance	1022 (13.4)	426 (41.7)	383 (37.5)	213 (20.8)		
Others	2510 (32.9)	1047 (41.7)	934 (37.2)	529 (21.1)		
Monthly income (US$), n (%)					38.718 (6)	<.001
<$350	1320 (17.3)	577 (43.7)	419 (31.7)	324 (24.5)		
$350-$699	1665 (21.8)	721 (43.3)	633 (38)	311 (18.7)		
$700-$1399	3259 (42.7)	1477 (45.3)	1178 (36.1)	604 (18.5)		
≥$1400	1385 (18.2)	676 (48.8)	471 (34)	238 (17.2)		
Sex roles, n (%)					131.721 (4)	<.001
Receptive anal sex	1750 (22.9)	796 (45.5)	675 (38.6)	279 (15.9)		
Insertive anal sex	2652 (34.8)	1319 (49.7)	949 (35.8)	384 (14.5)		
Both	3227 (42.3)	1336 (41.4)	1077 (33.4)	814 (25.2)		
Sex of partner, n (%)					284.131 (2)	<.001
Male only	5964 (78.2)	2852 (47.8)	2197 (36.8)	915 (15.3)		
Male and female	1665 (21.8)	599 (36)	504 (30.3)	562 (33.8)		
Number of sexual partners in the past 3 months, n (%)					748.589 (4)	<.001
0	1277 (16.7)	475 (37.2)	296 (23.2)	506 (39.6)		
1	3809 (49.9)	2124 (55.8)	1145 (30.1)	540 (14.2)		

Variables	Total (n=7629), n (%)	Awareness of sexual partner’s HIV status			Chi-square (*df*)	*P* value
		Yes (n=3451)	Partially (n=2701)	No (n=1477)		
≥2	2543 (33.3)	852 (33.5)	1260 (49.5)	431 (16.9)		
Condom use in the past 3 months, n (%)					681.686 (6)	<.001
Every time	4266 (55.9)	2171 (50.9)	1445 (33.9)	650 (15.2)		
Sometimes	2005 (26.3)	751 (37.5)	958 (47.8)	296 (14.8)		
No sex	874 (11.5)	288 (33)	168 (19.2)	418 (47.8)		
Never	484 (6.3)	241 (49.8)	130 (26.9)	113 (23.3)		
Ever used rush poppers during sexual behavior, n (%)					169.468 (4)	<.001
Never	5249 (68.8)	2355 (44.9)	1704 (32.5)	1190 (22.7)		
Occasionally	1312 (17.2)	536 (40.9)	601 (45.8)	175 (13.3)		
Often	1068 (14)	560 (52.4)	396 (37.1)	112 (10.5)		
History of STIs[b], n (%)					7.553 (2)	.02
No	7428 (97.4)	3376 (45.4)	2612 (35.2)	1440 (19.4)		
Yes	201 (2.6)	75 (37.3)	89 (44.3)	37 (18.4)		
Awareness of HIV PrEP[c], n (%)					336.193 (2)	<.001
No	1813 (23.8)	551 (30.4)	669 (36.9)	593 (32.7)		
Yes	5816 (76.2)	2900 (49.9)	2032 (34.9)	884 (15.2)		
Knowledge of HIV PEP[d], n (%)					313.701 (2)	<.001
No	1669 (21.9)	528 (31.6)	579 (34.7)	562 (33.7)		
Yes	5960 (78.1)	2923 (49)	2122 (35.6)	915 (15.4)		
HIV testing pathway, n (%)					26.881 (2)	<.001
Offline testing	3446 (45.2)	1526 (44.3)	1165 (33.8)	755 (21.9)		
Mail reagent self-testing	4183 (54.8)	1925 (46)	1536 (36.7)	722 (17.3)		
HIV status, n (%)					7.114 (2)	.03
Positive	82 (1.1)	27 (32.9)	31 (37.8)	24 (29.3)		
Negative	7547 (98.9)	3424 (45.4)	2670 (35.4)	1453 (19.3)		
First time undergoing HIV testing, n (%)					59.378 (2)	<.001
No	6647 (87.1)	3024 (45.5)	2421 (36.4)	1202 (18.1)		
Yes	982 (12.9)	427 (43.5)	280 (28.5)	275 (28)		

[a] The median age was 27 (IQR 23-33).

[b] STI: sexually transmitted infection.

[c] PrEP: pre-exposure prophylaxis.

[d] PEP: postexposure prophylaxis.

Of the 7629 participants, 45.2% (3451/7629) were aware, 35.4% (2701/7629) were partially aware, and 19.4% (1477/7629) were unaware of their sexual partner’s HIV status. Factors associated with high awareness of their sexual partner’s HIV status included being 20 - 29 years old (47.3%, 2143/4526), being single (46.8%, 2804/5997), having a high education level (49.6%, 359/724), having a high monthly income (48.8%, 676/1385), performing insertive anal sex (49.7%, 1319/2652), having one sexual partner in the past 3 months (55.8%, 2124/3809), using condoms in the past 3 months (50.9%, 2171/4266), frequently using rush poppers (52.4%, 560/1068), being aware of HIV PrEP (49.9%, 2900/5816) and PEP (49.0%, 2923/5960), and being HIV-negative (45.4%, 3424/7547). The χ^2 test showed that all factors were associated with awareness of sexual partners' HIV status (Table 1).

As presented in Table 2, multinomial logistic regression analysis showed that, compared to those who were unaware of their sexual partner's HIV status, participants who were students (adjusted odds ratio [aOR] 1.43, 95% CI 1.09 - 1.86), had a monthly income of more than US $1400 (aOR 1.36, 95% CI 1.03 - 1.80), had insertive anal sex (aOR 1.35, 95% CI 1.12 - 1.63), had only male sexual partners (aOR 1.53, 95% CI 1.28 - 1.82), had 1 sexual partner in the past 3 months (aOR 2.36, 95% CI 2.01 - 2.77), used condoms in the past 3 months (aOR 1.72, 95% CI 1.33 - 2.22), frequently used rush poppers in the past 3 months (aOR 2.27, 95% CI 1.81 - 2.86), were aware of HIV PrEP (aOR 2.04, 95% CI 1.68 - 2.48), were aware of HIV PEP (aOR 1.69, 95% CI 1.39 - 2.06), used mail reagent self-testing (aOR 1.19, 95% CI 1.04 - 1.36), and were undergoing HIV testing for the first time (aOR 1.40, 95% CI 1.16 - 1.69) were associated with increased odds of awareness of their sexual partner's HIV status.

Table . Factors associated with awareness of sexual partner’s HIV status among men who have sex with men.

Variables	Partial awareness		Awareness	
	aOR[a] (95% CI)	*P* value	aOR (95% CI)	*P* value
Occupation				
Student	1.05 (0.80 - 1.39)	.71	1.43 (1.09 - 1.86)	.009
Company employee	1.14 (0.96 - 1.36)	.14	1.17 (0.99 - 1.38)	.07
Freelance	1.01 (0.82 - 1.25)	.93	1.01 (0.82 - 1.25)	.91
Others	1.00 (reference)	—[b]	1.00 (reference)	—
Monthly income (US$)				
<$350	1.00 (reference)	—	1.00 (reference)	—
$350-$699	1.30 (1.01 - 1.67)	.045	1.36 (1.03 - 1.80)	.03
$700-$1399	1.11 (0.86 - 1.41)	.43	1.36 (1.07 - 1.73)	.01
≥$1400	0.97 (0.73 - 1.30)	.85	1.36 (1.03 - 1.80)	.03
Sex roles				
Receptive anal sex	1.00 (reference)	—	1.00 (reference)	—
Insertive anal sex	1.08 (0.89 - 1.31)	.35	1.35 (1.12 - 1.63)	.002
Both	0.92 (0.76 - 1.10)	.45	1.02 (0.85 - 1.23)	.82
Sex of partner				
Male only	1.39 (1.16 - 1.66)	<.001	1.53 (1.28 - 1.82)	<.001
Male and female	1.00 (reference)	—	1.00 (reference)	—
Number of sexual partners in the past 3 months				
0	0.84 (0.61 - 1.14)	.27	1.73 (1.27 - 2.32)	<.001
1	0.91 (0.78 - 1.07)	.25	2.36 (2.01 - 2.77)	<.001
≥2	1.00 (reference)	—	1.00 (reference)	—
Condom use in the past 3 moths				
Every time	1.91 (1.45 - 2.52)	<.001	1.72 (1.33 - 2.22)	<.001
Sometimes	2.57 (1.92 - 3.45)	<.001	1.42 (1.08 - 1.88)	.01
No sex	0.64 (0.42 - 0.97)	.04	0.65 (0.44 - 0.95)	.03
Never	1.00 (reference)	—	1.00 (reference)	—
Ever used rush poppers during sex behavior				
Never	1.00 (reference)	—	1.00 (reference)	—
Occasionally	1.63 (1.34 - 1.98)	<.001	1.35 (1.10 - 1.64)	.003
Often	1.77 (1.40 - 2.24)	<.001	2.27 (1.81 - 2.86)	<.001
Awareness of HIV PrEP[c]				
No	1.00 (reference)	—	1.00 (reference)	—
Yes	1.22 (1.00 - 1.49)	.048	2.04 (1.68 - 2.48)	<.001
Awareness of HIV PEP[d]				
No	1.00 (reference)	—	1.00 (reference)	—
Yes	1.56 (1.28 - 1.91)	<.001	1.69 (1.39 - 2.06)	<.001
HIV testing pathway				
Offline testing	1.00 (reference)	—	1.00 (reference)	—

Variables	Partial awareness		Awareness	
	aOR[a] (95% CI)	*P* value	aOR (95% CI)	*P* value
Mail reagent self-testing	1.20 (1.04 - 1.38)	.01	1.19 (1.04 - 1.36)	.01
First time undergoing HIV testing				
No	1.59 (1.30 - 1.93)	<.001	1.40 (1.16 - 1.69)	<.001
Yes	1.00 (reference)	—	1.00 (reference)	—

[a] aOR: adjusted odds ratio.

[b] Statistical analysis was not done for the reference group.

[c] PrEP: pre-exposure prophylaxis.

[d] PEP: postexposure prophylaxis.

Multinomial logistic regression analysis showed that, compared to those who were unaware of their sexual partner's HIV status, participants who had a monthly income of US $350-$699 (aOR 1.30, 95% CI 1.01 - 1.67), only had male sexual partners (aOR 1.39, 95% CI 1.16 - 1.66), used condoms in the past 3 months (aOR 1.91, 95% CI 1.45 - 2.52), frequently used rush poppers in the past 3 months (aOR 1.77, 95% CI 1.40 - 2.24), were aware of HIV PrEP (aOR 1.22, 95% CI 1.00 - 1.49), were aware of HIV PEP (aOR 1.56, 95% CI 1.28 - 1.91), used mail reagent self-testing (aOR 1.20, 95% CI 1.04 - 1.38), and were undergoing HIV testing for the first time (aOR 1.59, 95% CI 1.30 - 1.93) were associated with increased odds of partial awareness of their sexual partner's HIV status (Table 2).

Discussion

Principal Findings

This study described the awareness of sexual partners' HIV status and its associated factors among MSM with an HIV-negative or unknown status in Zhejiang province, China. Based on data from 7629 participants, we found 45.2% (n=3451) of participants were aware of their sexual partner's HIV status and 19.4% (1477) were unaware of their partner's HIV status. Our results show higher percentages compared with studies from Guangzhou, China, which revealed that 36% of MSM were aware of the HIV serostatus of all their sexual partners [27], 37.8% of MSM had received their partner's HIV testing report before engaging in sex [28], and 30.4% of MSM asked all partners about their HIV status [13]. A previous study from the United States found that 55% of participants consistently disclosed their HIV status to their sexual partners, 30% inconsistently disclosed, and 15% did not disclose [29], and these rates were lower than the disclosure rates of 61.2% in Korea [30] and 64.3% in Thailand [31]. The HIV serostatus disclosure strategy has proven effective in preventing and reducing HIV transmission. HIV prevalence among the MSM population has remained at 8% and only 84.3% of HIV-positive people have been diagnosed. We recommend the continued promotion and implementation of serostatus disclosure strategies among MSM, especially proactively asking about their sexual partners' HIV status, taking appropriate protective measures based on HIV status, reducing HIV transmission, and undergoing early diagnosis testing for HIV.

Consistent with previous studies [28,32], we found that MSM who had undergone HIV testing were more likely to obtain their sexual partners' HIV status. MSM who engage in HIV testing and seek their partners' HIV status may reflect a heightened awareness and proactive approach toward HIV management. Individuals who have already undergone HIV testing may be better informed about the importance of knowing their partners' HIV status, thereby reducing HIV transmission risks and promoting safer sexual practices. We found that MSM who used mail reagent self-testing had higher awareness of their partner's HIV status. This suggests that HIV testing through web-based platforms or gay apps could increase HIV serostatus disclosure among MSM [33]. These findings suggest that interventions aimed at increasing HIV testing rates among MSM could indirectly enhance partner disclosure and engagement in HIV-prevention strategies. It supports the need for continued efforts to promote regular HIV testing and web-based self-testing to increase the HIV serostatus disclosure rate. Our study provides evidence that understanding and engaging in HIV testing behaviors may contribute to HIV serostatus disclosure among MSM.

MSM who were aware of HIV PrEP and PEP were more likely to obtain their sexual partners' HIV status. A previous study showed that awareness of and access to PrEP might reduce the need and motivation to discuss HIV status with partners because they may assume that it is the partner's responsibility to use PrEP to protect themselves from HIV infection [13]. PrEP users expressed significantly greater openness to serodifferent partnering than participants who had never used PrEP [34], and MSM were more likely to have sex with a person who disclosed being on PrEP [35]. Of those who had not disclosed their status to their partners, 71% stated PrEP would encourage them to have sex [36]. However, another study showed that PrEP awareness was not associated with HIV serostatus disclosure among Chinese MSM [13]. Knowledge of PrEP and PEP may drive more thorough discussions on HIV status, aligning with a proactive approach to reducing transmission risks. Against the background of a 78% awareness of PrEP and PEP in China, we recommend targeted interventions to increase awareness of PrEP and PEP, thereby increasing the rate of partners' HIV testing and disclosure.

Consistent with a previous study [37], we found that MSM who used condoms and had fewer sexual partners were more likely to obtain the HIV status from their partners. MSM who adhere to consistent condom use and have fewer sexual partners often exhibit a heightened awareness of HIV risks and a commitment to preventing transmission. Findings showed that HIV status disclosure was significantly associated with lower rates of condomless sex compared to nondisclosure [33,38,39]. The probability of HIV exposure through condomless anal intercourse was substantially lower after serostatus disclosure than after nondisclosure [22]. One study showed that MSM engaging in receptive sex were less likely to ask their partners about their HIV status [13]. MSM with an HIV-negative or unknown status have a lower proportion of sexual events with a partner with an unknown HIV status [25]. Studies have demonstrated that MSM who disclose their HIV status are more likely to practice safer sexual behavior with HIV-infected partners than with partners with an HIV-negative or unknown status [23]. These findings highlight the importance of promoting consistent condom use and supporting MSM in maintaining low partner numbers to enhance the disclosure of HIV serostatus.

Moreover, we found that MSM who often used rush poppers had high awareness of their partners’ HIV status. Similar to other studies, drug use was independently associated with disclosure [40]. The possible reasons were as follows. First, MSM who used rush poppers had a higher rates of previous HIV testing, which was associated with greater awareness of their partner’s HIV status [4]. Second, due to the high use of rush poppers among MSM, rush popper use increases the risk of unprotected sexual behavior [39] and HIV transmission [41]. Targeted interventions were conducted for rush users among MSM [42], including reducing rush popper use, condom promotion, HIV serostatus disclosure, and other measures, such as knowing the HIV status of sexual partners before sex. We suggest that targeted interventions should be continued among MSM who use rush poppers, and this approach could help further improve the proportion of HIV serostatus disclosure.

Limitations

This study has some limitations. First, the study used routine surveillance questionnaires, which may not have captured all factors associated with HIV serostatus disclosure in previous studies. For example, the questionnaire only investigated the awareness of PrEP but did not investigate whether they had used it, and previous studies showed that the use of PrEP could also improve the rates of HIV serostatus disclosure [35]. Second, the study focused on participants actively asking about their sexual partners’ HIV status and did not investigate whether the participants disclosed their HIV status to their sexual partners. Finally, the study recruited MSM from web-based applications, 64.9% of whom were under 30 years old. It did not include a few MSM who were over 50 years of age. Previous studies have shown that this population has lower HIV serostatus disclosure [19]. Therefore, the results of this study apply only to MSM who rely on the internet.

Conclusion

The rate of awareness of sexual partners’ HIV status is low among MSM with an HIV-negative or unknown status in China. We suggest strengthening targeted interventions through web-based platforms and dating apps to promote HIV serostatus disclosure among MSM. Future studies could explore strategies to actively ask for a sexual partner’s HIV status among underrepresented demographics, such as older MSM and those who used PrEP, ultimately helping to improve the HIV disclosure rate and reduce HIV transmission.

Acknowledgments

We express our gratitude to the participants and Sunshine Coast Public Welfare for their contributions to the study. The study was funded by a grant from the Health Science and Technology Project in Zhejiang Province Social (2023KY632 and 2021RC048).

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

LH conceived the study design, coordinated the conduct of this research in the field, and performed the statistical analysis. QM drafted the paper. TJ, WC, SJ, JZ, and HW designed the research and played a major role in the survey. All the authors read and approved the final paper.

Conflicts of Interest

None declared.

References

1. Dong MJ, Peng B, Liu ZF, et al. The prevalence of HIV among MSM in China: a large-scale systematic analysis. BMC Infect Dis 2019 Nov 27;19(1):1000. [doi: 10.1186/s12879-019-4559-1] [Medline: 31775654]
2. National center for AIDS STD control and prevention CCFD: HIV prevention and control progress report. : China Center for Disease Control and Prevention; 2023 URL: https://www.chinaaids.cn/ [accessed 2025-01-06]

3. Understanding Fast-Track. UNAIDS. 2015 Jul 7. URL: https://www.unaids.org/en/resources/documents/2015/201506_JC2743_Understanding_FastTrack [accessed 2025-01-06]
4. He L, Jiang T, Chen W, et al. Examining HIV testing coverage and factors influencing first-time testing among men who have sex with men in Zhejiang province, China: cross-sectional study based on a large internet survey. JMIR Public Health Surveill 2024 Jun 14;10:e56906. [doi: 10.2196/56906] [Medline: 38875001]
5. Chen S, Zhang Q, Chan CK, et al. Evaluating an innovative HIV self-testing service with web-based, real-time counseling provided by an artificial intelligence chatbot (HIVST-Chatbot) in increasing HIV self-testing use among chinese men who have sex with men: protocol for a noninferiority randomized controlled trial. JMIR Res Protoc 2023 Jun 30;12:e48447. [doi: 10.2196/48447] [Medline: 37389935]
6. Obermeyer CM, Baijal P, Pegurri E. Facilitating HIV disclosure across diverse settings: a review. Am J Public Health 2011 Jun;101(6):1011-1023. [doi: 10.2105/AJPH.2010.300102] [Medline: 21493947]
7. O’Connell AA, Reed SJ, Serovich JA. The efficacy of serostatus disclosure for HIV transmission risk reduction. AIDS Behav 2015 Feb;19(2):283-290. [doi: 10.1007/s10461-014-0848-2] [Medline: 25164375]
8. Pinkerton SD, Galletly CL. Reducing HIV transmission risk by increasing serostatus disclosure: a mathematical modeling analysis. AIDS Behav 2007 Sep;11(5):698-705. [doi: 10.1007/s10461-006-9187-2] [Medline: 17082982]
9. Simoni JM, Pantalone DW. Secrets and safety in the age of AIDS: does HIV disclosure lead to safer sex? Top HIV Med 2004;12(4):109-118. [Medline: 15516708]
10. Irungu E, Chersich MF, Sanon C, et al. Changes in sexual behaviour among HIV-infected women in West and East Africa in the first 24 months after delivery. AIDS 2012 May 15;26(8):997-1007. [doi: 10.1097/QAD.0b013e3283524ca1] [Medline: 22343965]
11. Wei C, Lim SH, Guadamuz TE, Koe S. HIV disclosure and sexual transmission behaviors among an Internet sample of HIV-positive men who have sex with men in Asia: implications for prevention with positives. AIDS Behav 2012 Oct;16(7):1970-1978. [doi: 10.1007/s10461-011-0105-x] [Medline: 22198313]
12. Eyawo O, de Walque D, Ford N, Gakii G, Lester RT, Mills EJ. HIV status in discordant couples in sub-Saharan Africa: a systematic review and meta-analysis. Lancet Infect Dis 2010 Nov;10(11):770-777. [doi: 10.1016/S1473-3099(10)70189-4] [Medline: 20926347]
13. Cao W, Li J, Sun S, et al. HIV serostatus disclosure among men who have sex with men in China in the era of U=U and PrEP. AIDS Behav 2022 May;26(5):1477-1488. [doi: 10.1007/s10461-021-03502-7] [Medline: 34697704]
14. Murphy DA, de Wit JBF, Donohoe S, Adam PCG. The need to know: HIV status disclosure expectations and practices among non-HIV-positive gay and bisexual men in Australia. AIDS Care 2015;27(Sup1):90-98. [doi: 10.1080/09540121.2015.1062077] [Medline: 26616130]
15. Simbayi LC, Zungu N, Evans M, et al. HIV serostatus disclosure to sexual partners among sexually active people living with HIV in South Africa: results from the 2012 national population-based household survey. AIDS Behav 2017 Jan;21(1):82-92. [doi: 10.1007/s10461-015-1278-5] [Medline: 26767538]
16. Daskalopoulou M, Rodger AJ, Phillips AN, et al. Attitudes to disclosure of HIV-serostatus to new sexual partners and sexual behaviours among HIV-diagnosed gay, bisexual and other men who have sex with men in the UK. AIDS Care 2020 Oct;32(10):1323-1332. [doi: 10.1080/09540121.2020.1728218] [Medline: 32114800]
17. Eisinger RW, Dieffenbach CW, Fauci AS. HIV viral load and transmissibility of HIV infection: undetectable equals untransmittable. JAMA 2019 Feb 5;321(5):451-452. [doi: 10.1001/jama.2018.21167] [Medline: 30629090]
18. Rodger AJ, Cambiano V, Bruun T, et al. Risk of HIV transmission through condomless sex in serodifferent gay couples with the HIV-positive PARTNER taking suppressive antiretroviral therapy (PARTNER): final results of a multicentre, prospective, observational study. Lancet 2019 Jun 15;393(10189):2428-2438. [doi: 10.1016/S0140-6736(19)30418-0] [Medline: 31056293]
19. Lian J, Su XY, Chen XY, et al. Receiving human immunodeficiency virus serostatus disclosure from male sexual partners and related factors among men who have sex with men aged 50 and above. Zhong Yi Xue Ke Xue Yuan Xue Bao 2022 Apr;44(2):221-226. [doi: 10.3881/j.issn.1000-503X.14171] [Medline: 35538756]
20. Mutambara J, Vandirayi S, January J, Zirima H. Factors that influence HIV status disclosure among the elderly: perspectives from Shurugwi, Zimbabwe. AIDS Care 2022 Jan;34(1):112-117. [doi: 10.1080/09540121.2021.1930997] [Medline: 34039191]
21. Smith R, Rossetto K, Peterson BL. A meta-analysis of disclosure of one’s HIV-positive status, stigma and social support. AIDS Care 2008 Nov;20(10):1266-1275. [doi: 10.1080/09540120801926977] [Medline: 18608080]
22. Kavuma D, Kirwana VB, Taani M. Factors associated with HIV positive serostatus disclosure to sexual partners among sexually active young people on anti-retroviral therapy in Central Uganda. HIV AIDS (Auckl) 2023;15:293-311. [doi: 10.2147/HIV.S407535] [Medline: 37312814]
23. Serovich JM, Reed SJ, O’Connell AA, Laschober TC. Relationship between serostatus disclosure and categories of HIV transmission risk in men who have sex with men living with HIV. Int J STD AIDS 2018 Jul;29(8):744-750. [doi: 10.1177/0956462417752267] [Medline: 29431024]
24. Han M. Analysis of the epidemic situation of AIDS in China and prospects for its prevention and control. Chin J AIDS STD 2023;29(3):247-250. [doi: 10.13419/j.cnki.aids.2023.03.01]

25. Sang JM, Wang L, Moore DM, et al. Undetectable or unknown? A longitudinal event-level analysis of disclosure of HIV serostatus and undetectability among gay, bisexual, and other men who have sex with men (gbMSM) in metro Vancouver. AIDS Behav 2021 Aug;25(8):2630-2643. [doi: 10.1007/s10461-021-03224-w] [Medline: 33751314]
26. Chen S, Fang Y, Chan PSF, Kawuki J, Mo P, Wang Z. Counseling supporting HIV self-testing and linkage to care among men who have sex with men: systematic review and meta-analysis. JMIR Public Health Surveill 2024 Jan 24;10:e45647. [doi: 10.2196/45647] [Medline: 38265866]
27. Zeng W, Luo Y, Liu F, et al. Impact of awareness of sexual partners' HIV serostatus on the HIV acquisition among men who have sex with men in Guangzhou, China: a nested case-control study. Arch Sex Behav 2024 Oct;53(9):3655-3662. [doi: 10.1007/s10508-024-02971-w] [Medline: 39152320]
28. Sun HT, Fan XR, Gu YZ, et al. WeChat-based HIV e-report, a new approach for HIV serostatus requests and disclosures among men who have sex with men: prospective subgroup analysis of a randomized controlled trial. JMIR Mhealth Uhealth 2023 May 8;11:e44513. [doi: 10.2196/44513] [Medline: 37155223]
29. Cook CL, Staras SAS, Zhou Z, Chichetto N, Cook RL. Disclosure of HIV serostatus and condomless sex among men living with HIV/AIDS in Florida. PLoS ONE 2018;13(12):e0207838. [doi: 10.1371/journal.pone.0207838] [Medline: 30557362]
30. Kim K, Woo J. Predictors associated with HIV status non-disclosure in Korea. J Korean Med Sci 2022 Apr 25;37(16):e125. [doi: 10.3346/jkms.2022.37.e125] [Medline: 35470599]
31. Morkphrom E, Ratanasuwan W, Sittironnarit G, Rattanaumpawan P. Non-disclosure of HIV serostatus to sexual partners: prevalence, risk factors and clinical impact in patients with HIV. HIV Med 2021 Mar;22(3):194-200. [doi: 10.1111/hiv.13005] [Medline: 33145968]
32. Tang W, Liu C, Cao B, et al. Receiving HIV serostatus disclosure from partners before sex: results from an online survey of Chinese men who have sex with men. AIDS Behav 2018 Dec;22(12):3826-3835. [doi: 10.1007/s10461-018-2062-0] [Medline: 29470809]
33. Konkor I, Antabe R, Mkandawire P, et al. Knowledge of sexual partner's HIV serostatus and the practice of safer sex among heterosexual men of African descent in London, Ontario. Ethn Health 2022 Feb;27(2):375-387. [doi: 10.1080/13557858.2020.1734776] [Medline: 32115995]
34. Kalwicz DA, Rao S, Modrakovic DX, et al. The implications of PrEP use, condom use, and partner viral load status for openness to serodifferent partnering among US sexual minority men (SMM). AIDS Behav 2024 Feb;28(2):524-534. [doi: 10.1007/s10461-023-04241-7] [Medline: 38329558]
35. Medina MM, Crowley C, Montgomery MC, et al. Disclosure of HIV serostatus and pre-exposure prophylaxis use on internet hookup sites among men who have sex with men. AIDS Behav 2019 Jul;23(7):1681-1688. [doi: 10.1007/s10461-018-2286-z] [Medline: 30267365]
36. Bracho-Sanchez E, Stonbraker S, Halpern M, Wood S, Lowenthal E. HIV pre-exposure prophylaxis: assessment of acceptability among members of HIV serodiscordant/serostatus unknown couples. Int J STD AIDS 2022 Jun;33(7):687-693. [doi: 10.1177/09564624221097753] [Medline: 35488451]
37. Sullivan MC, Cruess DG, Huedo-Medina TB, Kalichman SC. Substance use, HIV serostatus disclosure, and sexual risk behavior in people living with HIV: an event-level analysis. Arch Sex Behav 2020 Aug;49(6):2005-2018. [doi: 10.1007/s10508-019-01531-x] [Medline: 31863314]
38. Okafor CN, Li MJ, Hucks-Ortiz C, Mayer KH, Shoptaw S. Disclosure of HIV status and HIV sexual transmission behaviors among HIV-positive Black men who have sex with men in the BROTHERS (HPTN 061) study. J Urban Health 2020 Oct;97(5):692-703. [doi: 10.1007/s11524-020-00419-9] [Medline: 32020466]
39. Fisher MP, Ramchand R, Bana S, Iguchi MY. Risk behaviors among HIV-positive gay and bisexual men at party-oriented vacations. J Stud Alcohol Drugs 2013 Jan;74(1):158-167. [doi: 10.15288/jsad.2013.74.158] [Medline: 23200162]
40. Daskalopoulou M, Rodger A, Phillips AN, et al. Recreational drug use, polydrug use, and sexual behaviour in HIV-diagnosed men who have sex with men in the UK: results from the cross-sectional ASTRA study. Lancet HIV 2014 Oct;1(1):e22-e31. [doi: 10.1016/S2352-3018(14)70001-3] [Medline: 26423813]
41. He L, Pan X, Wang N, et al. New types of drug use and risks of drug use among men who have sex with men: a cross-sectional study in Hangzhou, China. BMC Infect Dis 2018 Apr 17;18(1):182. [doi: 10.1186/s12879-018-3091-z] [Medline: 29665785]
42. Huang PX, Duan Q, Liao MZ, et al. Analysis of rush poppers use and related factors in men who have sex with men in Jinan and Qingdao, 2016. Zhonghua Yu Fang Yi Xue Za Zhi 2020 Aug 6;54(8):861-866. [doi: 10.3760/cma.j.cn112150-20190815-00659] [Medline: 32842316]

Abbreviations

MSM: men who have sex with men
STI: sexually transmitted infections
PEP: postexposure prophylaxis
PrEP: pre-exposure prophylaxis
aOR: adjusted odds ratio
UNAIDS: Joint United Nations Program on HIV/AIDS (UNAIDS)

Edited by A Mavragani; submitted 05.09.24; peer-reviewed by J Jiang, S Chen; revised version received 18.11.24; accepted 19.11.24; published 13.01.25.

Please cite as:
Ma Q, Jiang T, Chen W, Jiang S, Zheng J, Wang H, He L
Awareness of Sexual Partner's HIV Status Among Men Who Have Sex With Men in China: Cross-Sec. tional Survey Study
JMIR Public Health Surveill 2025;11:e66139
URL: https://publichealth.jmir.org/2025/1/e66139
doi:10.2196/66139

Original Paper

The Effects of MyChoices and LYNX Mobile Apps on HIV Testing and Pre-Exposure Prophylaxis Use by Young US Sexual Minority Men: Results From a National Randomized Controlled Trial

Katie B Biello[1,2], MPH, PhD; Kenneth H Mayer[2,3], MD; Hyman Scott[4,5], MD; Pablo K Valente[6], MPH, PhD; Jonathan Hill-Rorie[2], MPH; Susan Buchbinder[4,5], MD; Lucinda Ackah-Toffey[7], MPH; Patrick S Sullivan[7], DVM, PhD; Lisa Hightow-Weidman[8], MPH, PhD; Albert Y Liu[4,5], MPH, MD

[1]Department of Epidemiology, School of Public Health, Brown University, Providence, RI, United States
[2]The Fenway Institute, Fenway Health, Boston, MA, United States
[3]Department of Medicine, Beth Israel Deaconess Medical Center/Harvard Medical School, Boston, MA, United States
[4]Bridge HIV, San Francisco Department of Public Health, San Francisco, CA, United States
[5]Department of Medicine, University of California, San Francisco, San Francisco, CA, United States
[6]University of Connecticut, Waterbury, CT, United States
[7]Rollins School of Global Public Health, Emory University, Atlanta, GA, United States
[8]Institute on Digital Health and Innovation, College of Nursing, Florida State University, Tallahassee, FL, United States

Corresponding Author:
Katie B Biello, MPH, PhD
Department of Epidemiology
School of Public Health
Brown University
121 South Main Street
Box G-S121-2
Providence, RI, 02912
United States
Phone: 1 4018633082
Email: katie_biello@brown.edu

Abstract

Background: Young sexual minority men have among the highest rates of HIV in the United States; yet, the use of evidence-based prevention strategies, including routine HIV testing and pre-exposure prophylaxis (PrEP), remains low. Mobile apps have enormous potential to increase HIV testing and PrEP use among young sexual minority men.

Objective: This study aims to assess the efficacy of 2 theory- and community-informed mobile apps—LYNX (APT Mobility) and MyChoices (Keymind)—to improve HIV testing and PrEP initiation among young sexual minority men.

Methods: Between October 2019 and May 2022, we implemented a 3-arm, parallel randomized controlled trial in 9 US cities to test the efficacy of the LYNX and MyChoices apps against standard of care (SOC) among young sexual minority men (aged 15-29 years) reporting anal sex with cisgender male or transgender female in the last 12 months. Randomization was 1:1:1 and was stratified by site and participant age; there was no masking. The co-primary outcomes were self-reported HIV testing and PrEP initiation over 6 months of follow-up.

Results: A total of 381 young sexual minority men were randomized. The mean age was 22 (SD 3.2) years. Nearly one-fifth were Black, non-Hispanic (n=67, 18%), Hispanic or Latino men (n=67, 18%), and 60% identified as gay (n=228). In total, 200 (53%) participants resided in the Southern United States. At baseline, participants self-reported the following: 29% (n=110) had never had an HIV test and 85% (n=324) had never used PrEP. Sociodemographic and behavioral characteristics did not differ by study arm. Compared to SOC (n=72, 59%), participants randomized to MyChoices (n=87, 74%; *P*=.01) were more likely to have received at least 1 HIV test over 6 months of follow-up; those randomized to LYNX also had a higher proportion of testing (n=80, 70%) but it did not reach the a priori threshold for statistical significance (*P*=.08). Participants in both MyChoices (n=23, 21%) and LYNX (n=21, 20%) arms had higher rates of starting PrEP compared to SOC (n=19, 16%), yet these differences were not statistically significant (*P*=.52).

Conclusions: In addition to facilitating earlier treatment among those who become aware of their HIV status, given the ubiquity of mobile apps and modest resources required to scale this intervention, a 25% relative increase in HIV testing among young sexual minority men, as seen in this study, could meaningfully reduce HIV incidence in the United States.

Trial Registration: ClinicalTrials.gov NCT03965221; https://clinicaltrials.gov/study/NCT03965221

(JMIR Public Health Surveill 2025;11:e63428) doi:10.2196/63428

KEYWORDS

HIV testing; adolescents; sexual minority men; mobile apps; pre-exposure prophylaxis; youths; randomized controlled trial; sexual minority; United States; efficacy; LYNX; MyChoices; sociodemographic; behavioral health; HIV prevention; HIV infection; HIV care; transmission; public health; mHealth; mobile phones

Introduction

Sexual minority men (eg, gay, bisexual, and other men who have sex with men) account for nearly two-thirds of all people living with HIV in the United States while comprising less than 10% of the population. Sexually active young sexual minority men, aged 13-25 and 25-34 years, have among the highest rates of HIV [1,2]. Young sexual minority men are also the least likely to be aware of their HIV status, further enhancing HIV transmission risks and delaying HIV care [2].

The Centers for Disease Control and Prevention (CDC) recommends all sexually active sexual minority men get tested for HIV at least annually, and consider testing more often (ie, every 3 to 6 months) for sexual minority men who are sexually active [3]. HIV testing is essential for early detection, early care engagement, and early treatment—resulting in a lower likelihood of onward transmission and improved quality of life [3]. For those who test negative, it is also an opportunity to engage in pre-exposure prophylaxis (PrEP) care. The US Preventive Services Task Force expanded the recommendation for PrEP in August 2023 to include adolescents, in addition to adults, who report anal or vaginal sexual activity and disclose one of the following criteria: sexual partner with HIV, a recent bacterial sexually transmitted infection (STI), or inconsistent or no condom use with sexual partners [4]. Despite the efficacy of and need for these evidence-based prevention strategies—routine HIV testing and PrEP—use remains suboptimal among young sexual minority men [5-7].

Reasons for a lack of routine HIV testing and limited PrEP use among young sexual minority men are multilevel and complex [8]. For example, at the individual level, young sexual minority men tend to underestimate their personal risk of acquiring HIV [9,10]; at the interpersonal level, young sexual minority men may be concerned that their partner will feel threatened if they test for HIV or initiate PrEP [11,12]; at the clinical level, young sexual minority men may not be comfortable disclosing their sexual activity to providers [13-15]; at the structural level, young sexual minority men may not know where to access HIV testing or PrEP, or have the autonomy or ability to access a testing site or PrEP clinic on their own [11,16]. Any intervention or program to increase routine HIV testing and PrEP use will need to address a plethora of multilevel challenges [8].

Mobile apps—accessed via smartphones which have nearly complete coverage in the United States, particularly among youths—have enormous potential to increase routine HIV testing and PrEP use among young sexual minority men [17-19]. Ninety-five percent of youth report having a smartphone [20], and youth commonly use their mobile devices for a range of activities, including downloading health-related apps and accessing health information [21]. Implemented through mobile apps, mobile health (mHealth) interventions may be a promising tool for promoting behaviors such as HIV testing and PrEP use among young sexual minority men who have been poorly reached through standard clinic services [17-19].

In prior work, 2 distinct apps, MyChoices (Keymind) and LYNX (APT Mobility), each designed to increase HIV testing and PrEP Initiation among young sexual minority men, were developed by separate research teams using different theoretical models to effect behavior change and different approaches for engagement (Multimedia Appendix 1). The process for development and pilot testing of these apps has been described previously [22-25]. In brief, both apps were developed with extensive formative research with young sexual minority men, including interviews and focus groups, beta testing, and pilot randomized controlled trials (RCTs). In parallel pilot RCTs conducted as part of the National Institutes of Health’s Adolescent Medicine Trials Network (ATN) for HIV Interventions, both MyChoices and LYNX were found to be feasible and acceptable, meeting predetermined go or no-go criteria for this current full-scale efficacy trial.

This study presents the results of a 3-arm, RCT conducted in 9 US cities in order to assess the efficacy of the LYNX and MyChoices apps to improve HIV testing and PrEP initiation among young sexual minority men.

Methods

Trial Design

This was a 3-arm, RCT with a 1:1:1 allocation ratio conducted as part of the ATN and the University of North Carolina/Emory Center for Innovative Technology (iTech), and implemented between October 2019 and May 2022 [26,27]. Consented participants completed a baseline assessment and follow-up assessments at 3- and 6-months post baseline. Some participants also completed an additional follow-up assessment between 6 and 12 months post baseline (see Adjustments in Trial Design in Response to the COVID-19 Pandemic section for more details).

Participants

To be eligible participants had to (1) identify as cisgender men, (2) be aged 15-29 years, (3) self-report being HIV serostatus negative or HIV serostatus unknown at screening, (4) not have had an HIV test in the past 3 months, (5) not be currently taking PrEP, (6) self-report at least 1 episode of anal intercourse with a male or transgender female partner during the last 12 months, (7) had to own or have regular and ongoing access to an iOS or Android smartphone, (8) be willing and able to download the MyChoices or LYNX app, and (9) be able to understand, read, and speak English.

Participants were recruited through social media, flyers, and direct outreach at local venues and clinical sites across 9 sites: Atlanta, Georgia (study site: PRISM Health); Boston, Massachusetts (study site: Fenway Health); Bronx, New York City (study site: the Adolescent AIDS Program at Montefiore); Chapel Hill, North Carolina (study site: University of North Carolina, Chapel Hill); Charlotte, North Carolina (study site: RAIN); Chicago, Illinois (study site: AYAR at CORE Center); Houston, Texas (study site: Texas Children’s Hospital); Philadelphia, Pennsylvania (study site: Children’s Hospital of Philadelphia); and Tampa, Florida (study site: University of South Florida).

Ethical Considerations

The study procedures were reviewed and approved by the University of North Carolina institutional review board (IRB) as a single IRB-of-record (study 19-0260). IRB authorization agreements with all participating research entities were enacted. A waiver of parental consent was obtained for individuals aged 15-17 years. All enrolled participants completed a site-specific consent or assent, which contained the essential elements, including a description of the purpose, study procedures, and risks and benefits. No social harms or adverse events were reported. Compensation for completing the survey was set by each site in accordance with norms, cost of living, and input from their IRBs. Participants were not compensated for using the app. The protocol for ATN 143: COMPARE (Comparison of Men’s Prevention Apps to Research Efficacy) is registered at ClinicalTrials.gov (NCT03965221).

Randomization and Masking

Randomization was computer generated and was equally distributed 1:1:1 for study arms. Randomization was stratified by site and participant age (15-18 and 19-29 years). This trial did not use masking. Following randomization, study staff supported participants randomized to the experimental arms through downloading the app, and then provided a brief, structured demonstration of the main components of the app.

Interventions

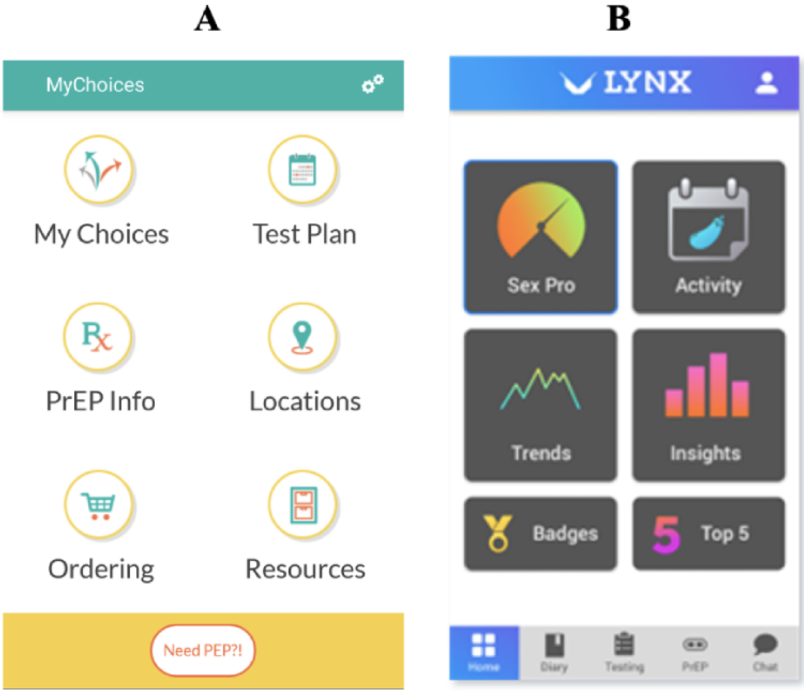
Standard of Care

Participants in all study conditions received CDC fact sheets on HIV testing and PrEP Initiation and had access to any services available as standard of care (SOC) at each study site [28,29].

Experimental Condition: MyChoices Mobile App

As described and detailed previously [23,24], MyChoices was adapted, through extensive formative research and pilot testing, from an app previously demonstrated to increase HIV testing and PrEP Initiation in adult sexual minority men [22-24,30-32]. The MyChoices app is informed by the social cognitive theory [33] and includes the following primary components: sexual health information using a variety of formats (eg, infographics, videos, weblinks to educational resources); access to free home HIV and STI testing kits as well as condoms and lubricants; individualized, tailored HIV prevention recommendations and HIV testing plans based on monthly check-ins; reminders for HIV testing with geolocation-based notifications when near HIV testing sites and GPS-enabled maps with local HIV and STI testing locations and PrEP providers (Figure 1A).

Figure 1. (A) MyChoices and (B) LYNX app landing pages.



Experimental Condition: LYNX Mobile App

As described previously [25], LYNX was developed through iterative formative research and pilot testing with young sexual minority men. The LYNX app is guided by the information-motivation-behavioral skills (IMB) model [34] and includes the following primary components: Sexual Health Promotion (Sex Pro) score that provides a personalized HIV risk score; sexual diary to help participants track sexual partners and encounters; reminders for HIV or STI testing; ability to order free home HIV and STI testing kits as well as condoms and lubricant; a geolocation map of HIV or STI testing sites and PrEP clinics; a PrEP information page including a roadmap on how to get on PrEP; a “trends and insights” page including infographics summarizing partner types and sexual activities; and a 2-way chat feature to answer questions and provide real-time assistance with PrEP navigation (Figure 1B).

Privacy

To enhance confidentiality, in order to access either app, users initially entered their email address and a registration key (supplied following randomization) and set up a user-generated password. On subsequent visits to the app, users were required to supply their username and password or create a PIN. The apps timed out every time users left the app or when the phone went into sleep mode.

Adjustments in Trial Design in Response to the COVID-19 Pandemic

Enrollment for the COMPARE study began in October 2019. At this point, all enrollment visits were conducted in person at the enrolling sites (n=83). All sites paused enrollment due to COVID-19 from March 2020 to June 2020. In June 2020, sites began to resume enrollment, with all enrollment visits occurring via videoconferencing. This continued until November 2021 when enrollment closed. Given the entirely remote procedures, in October 2021 and November 2021, enrollment was expanded to the entire United States. Participants were assigned to the study site that was closest to their zip code of residence. See Multimedia Appendix 2 for the enrollment timeline. In addition to changes in enrollment procedures, the COVID-19 pandemic significantly slowed down enrollment. In response, in collaboration with the ATN-wide Study Monitoring Committee, the decision was made to extend the enrollment period and shorten the full follow-up period for participants enrolled between June 2021 and November 2022 (follow-up ranging from 6 months to 11 months; n=40 participants were only eligible for a 6-month visit because they were enrolled in November 2022; the remaining 341 participants were eligible to complete an additional follow-up survey between 7 and 12 months post baseline). As such, our statistical analysis plan (also approved by the Study Monitoring Committee and IRB) was revised to shift the primary endpoint to be measured at the 6-month follow-up visit (rather than the 12-month follow-up visit). Importantly, based on eligibility criteria, all participants would be recommended to have an HIV test every 3 to 6 months according to CDC guidelines; as such, the outcome time frame remained clinically meaningful.

Assessments

Baseline assessments were conducted at the enrolling study site or via videoconferencing and self-administered using a secure web-based data entry system. All follow-up assessments were self-administered and conducted off-site through the same secure web-based system. Follow-up assessments occurred at 6-month postbaseline for all participants, and an additional follow-up assessment occurred between 7- and 12-month post baseline for all participants enrolled prior to November 2022.

Primary Efficacy Outcome Measures

HIV testing and PrEP initiation over study follow-up were assessed by self-report at each follow-up visit. Specifically, participants were asked “How many times have you had an HIV test in the past 3 months?” with response options: “I have not had an HIV test in the past 3 months, 1 time, 2 times, 3 times, and 4 or more times.” For PrEP, participants were asked “In the past 3 months, have you taken PrEP (pre-exposure prophylaxis) to reduce the chance of getting HIV?” with response options: “Yes, I am on PrEP right now,” “Yes, but I’m not on PrEP anymore,” and “No.”

Secondary Outcome Measures

Time-to-event outcomes. As secondary outcomes, we examined the time to first HIV test and time to PrEP initiation from the baseline assessment using all available follow-ups. At each assessment, participants were asked the month and year of all their HIV tests and PrEP Initiation (if applicable) since the previous assessment. Since the specific dates when HIV tests and PrEP initiation took place were not assessed, we considered these events to happen on the last day of the month in which they reported the events.

HIV Testing Frequency

Participants were asked the number of times they were tested for HIV in the past 3 months (at 3- and 6-month visits) and 6 months (at 12-month visits). Responses were summed to obtain the total number of times participants were tested for HIV through the 12 months of follow-up, which ranged from 0 to 4 or more times.

PrEP IMB Scale

We created a 5-item scale to assess the 3 domains of the IMB model related to PrEP uptake: information (1 item), motivation (1 item), and behavioral skills (3 items). Sample items included, “I understand whether PrEP would be a good fit for me,” “I feel motivated to get on PrEP,” and “I know how to get PrEP”. Responses on a 5-point scale were averaged to reach a total IMB score. The Cronbach α=0.83 indicated strong internal consistency.

PrEP Self-Efficacy

We adopted the PrEP Adherence Self-Efficacy Scale [35] to assess PrEP self-efficacy in the past month with 10 items asking how confident participants had been that they can or could “integrate PrEP into your daily routine” or “take a PrEP pill every day even when you aren’t feeling well.” Response options ranged from 0=“could not do at all” to 10=“completely certain could do.” We obtained a total PrEP self-efficacy score based on the average of all 10 responses. The Cronbach α=0.96 indicated strong internal consistency.

PrEP Stigma

We assessed PrEP stigma with 12 items drafted by the study team, including “I can tell my friends that I am using PrEP,” “I would not want a sexual partner to see my PrEP pills in my medicine cabinet,” and “People would think I am sexually risky if they find out that I take PrEP.” Responses on a 5-point scale were reverse-coded as needed and averaged across 12 items such that higher scores indicated higher PrEP stigma. The Cronbach α=0.76 indicated moderate internal consistency.

Discussing PrEP With Health Care Provider

In order to measure the first steps in the PrEP care continuum, we assessed whether participants talked to a health care provider about PrEP (yes vs no).

PrEP Interest

Participants who had not initiated PrEP were asked about their interest in taking PrEP with a single item with response options ranging from 1=“not at all interested” to 5=“extremely interested.”

App Use Measures

We assessed the overall number of log-ins in each of the apps, the number of calendar days with app logins, and the total amount of time spent in the apps (minutes).

Sociodemographic Measure

To characterize the sample and examine potential moderating effects, we collected participants’ age, race and Hispanic or Latinx ethnicity, current relationship status (ie, single vs not single), and zip code of residence (which was aggregated to census regions: South, Northeast, Midwest, and West). We also assessed recent history of condomless anal sex by asking participants a series of questions about past-3-month sex partners, type of sex with those partners, and condom use during sex [36,37].

Power and Sample Size

Using data from our pilot study, we powered the study to see a significant difference between each experimental condition compared to the standard of care assuming a baseline past-year HIV testing rate of 40%, HIV testing rate in the interventions arm at 77% (risk ratio [RR] 1.56), any app use in intervention arms of 60%, α=.05, and power=.80. With 70% retention at 12 months follow-up and equal allocation, we estimated a sample size of 450 (n=150 per arm, 33%). Notably, in June 2021, at the request of the ATN-wide Study Monitoring Committee, we reran our sample size calculation with all of the same assumptions but based on the higher-than-expected retention demonstrated at that point in the study (approximately 90% at each visit). The new targeted sample size was 351 (n=117 per arm, 33%).

Statistical Analysis

Means for continuous variables and frequencies for categorical variables were calculated to describe the participants at baseline, overall, and stratified by intervention condition to assess balance in baseline characteristics.

An intent-to-treat approach was used, analyzing participants in the study arm to which they were assigned. As missing data was minimal (less than 10%), we used complete case analysis for each study aim. The primary efficacy analyses compared HIV testing and PrEP Initiation between the study arms over the study follow-up (ie, 6 months) using chi-square tests. Risk ratios and their corresponding 95% CIs were also calculated using log-binomial regression. Moreover, we examined differences between the study arms in the hypothesized mediators, including information, motivation, and behavioral skills related to PrEP Initiation, PrEP self-efficacy, PrEP stigma, discussing PrEP with a health care provider, and PrEP interest.

Secondary outcomes included time-to-event analyses for HIV testing and PrEP Initiation through the 12 months of follow-up using Kaplan-Meier survival curves. We examined overall differences across study arms using log-rank tests, and Cox proportional hazard models to obtain hazard ratios by study arm. We also examined the frequency of HIV testing over 12 months using Poisson regression models with robust standard errors.

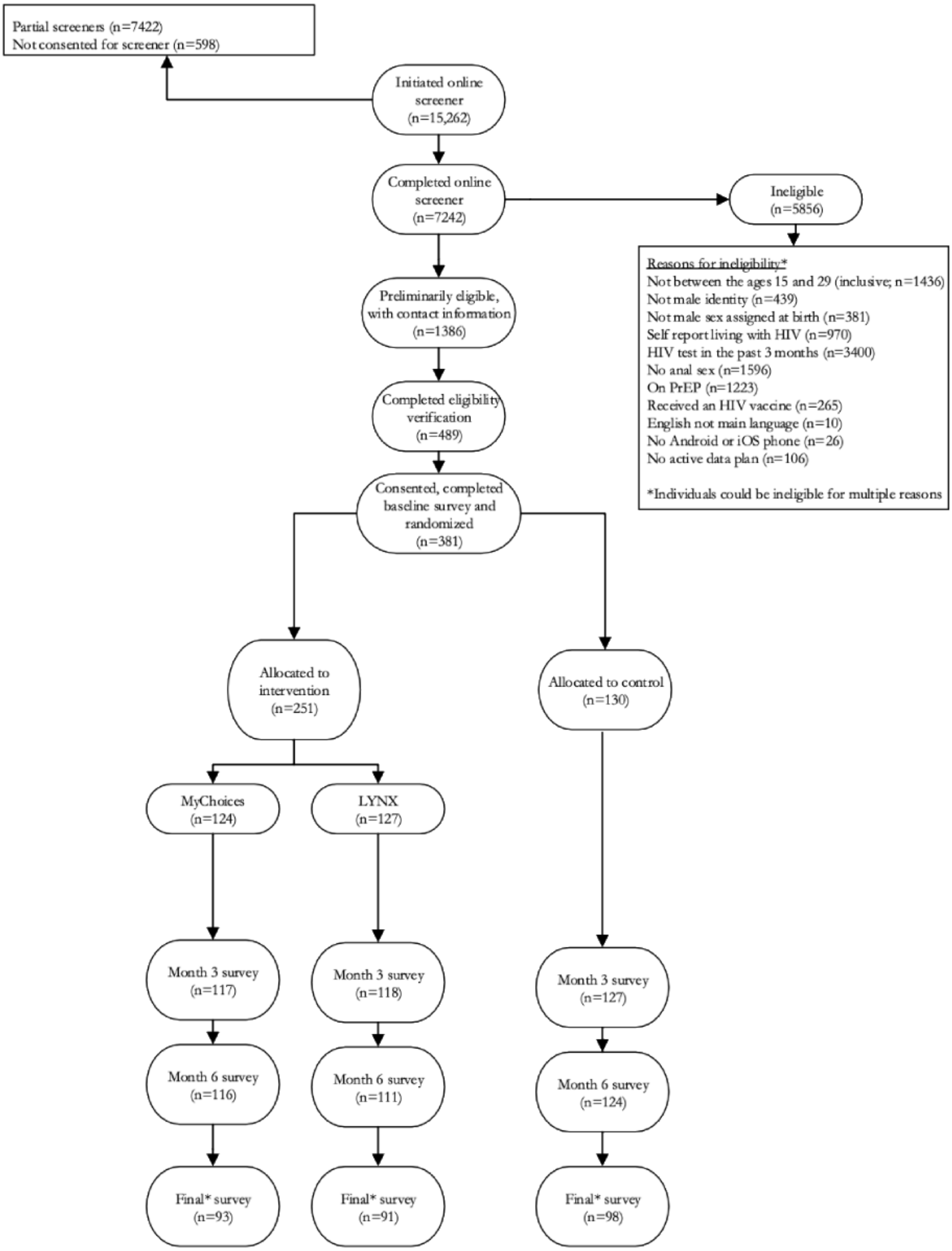
We also assessed whether the intervention effect on the outcomes differed by baseline characteristics (ie, age, race or ethnicity, region, condomless anal sex, relationship status, depression and anxiety, and prepandemic enrollment). These moderator analyses were chosen a priori (with the exception of the pandemic measure) and were evaluated with interaction terms; we probed potential moderation in post hoc stratified analyses if the *P* value of interaction terms was less than .20.

In secondary analyses, we explored dose-response relationships by examining the effect of measures of usage of each app on HIV testing and PrEP Initiation at 6 months using log-binomial regression. In all dose-response analyses, we only considered app usage until the date of the assessment when outcomes of interest (ie, HIV testing and PrEP Initiation) were first reported. Since app usage measures were right-skewed, we winsorized these variables at the 98th percentile in all regression analyses [38].

Results

Online eligibility screeners were completed by 7,242 individuals. Of those, 1,386 screened initially eligible and provided contact information. Common reasons for ineligibility included having a recent HIV test and not reporting recent anal sex. Four hundred and eighty-nine individuals were able to be contacted and attended the initial visit. Among these, 381 participants consented, completed a baseline assessment and were randomized to either the MyChoices app (n=124), the LYNX app (n=127) or the comparison condition (n=130). Among randomized participants, 95% completed a 3-month follow up assessment and 92% completed a 6-month follow up assessment. Figure 2 depicts participant flow throughout the study from screening to follow-up.

Figure 2. COMPARE (Comparison of Men's Prevention Apps to Research Efficacy) study CONSORT (Consolidated Standards of Reporting Trials) diagram, 2019-2022. PrEP: pre-exposure prophylaxis. *Final follow-up period ranged from 7-12 months postbaseline. Of the 381 enrolled, 341 were eligible to complete this final follow-up (n=111 in MyChoice; n=115 in LYNX; n=115 in standard of care).



Sample characteristics at baseline are described in Table 1. The mean age was 22.4 (SD 3.2); 51% were White, non-Hispanic (n=194), 18% were Black, non-Hispanic (n=67), and 18% were Latino (n=67). 60% (n=228) identified as gay. 53% (n=200) resided in the Southern United States. At baseline, participants self-reported the following: 29% (n=110) had never had an HIV test, and 67% (n=253) had not had an HIV test in the prior 3 months. 85% (n=324) had never used PrEP, and an additional 13% (n=51) had not used PrEP in the prior 3 months. Sociodemographic and behavioral characteristics did not differ significantly by study arm.

Table 1. Baseline characteristics for the COMPARE[a] study, overall and by study condition, 2019-2022.

	Total (N=381)	SOC[b] (n=130)	LYNX (n=127)	MyChoices (n=124)
Age (years), mean (SD)	22.4 (3.17)	22.8 (3.34)	22.3 (3.20)	22.2 (2.96)
Race and ethnicity, n (%)				
Asian, non-Hispanic	34 (9)	10 (7.8)	12 (9.5)	12 (9.8)
Black, non-Hispanic	67 (17.7)	25 (19.4)	26 (20.5)	16 (13)
Latino or Hispanic	67 (17.7)	29 (22.5)	23 (18.1)	15 (12.2)
White, non-Hispanic	194 (51.2)	58 (45)	59 (46.5)	77 (62.6)
Other or Mixed	17 (4.5)	7 (5.4)	7 (5.5)	3 (2.4)
Census region of current residence, n (%)				
Midwest	56 (14.7)	19 (14.6)	18 (14.2)	19 (15.3)
Northeast	112 (29.4)	36 (27.7)	39 (30.7)	37 (29.8)
South	200 (52.5)	68 (52.3)	67 (52.8)	65 (52.4)
West	13 (3.4)	7 (5.4)	3 (2.4)	3 (2.4)
Insurance status, n (%)				
Uninsured	43 (11.3)	12 (9.2)	12 (9.5)	19 (15.5)
Own insurance	136 (35.8)	46 (35.4)	47 (37)	43 (35)
Parent’s insurance	196 (51.6)	70 (53.9)	67 (52.8)	59 (48)
Do not know	5 (1.3)	2 (1.5)	1 (0.8)	2 (1.6)
Sexual orientation, n (%)				
Bisexual	69 (18.1)	23 (17.7)	23 (18.1)	23 (18.6)
Gay or homosexual	228 (59.8)	78 (60)	78 (61.4)	72 (58.1)
Queer or same sex or other	77 (20.2)	26 (20)	22 (17.3)	29 (23.4)
Heterosexual or straight	7 (1.8)	3 (2.3)	4 (3.2)	0 (0)
Relationship status, n (%)				
Single	182 (47.9)	62 (47.7)	61 (48.0)	59 (48)
Has boyfriend, girlfriend, partner, or lover	119 (31.3)	39 (30)	39 (30.7)	41 (33.3)
Casually dating	73 (19.2)	27 (20.8)	27 (21.3)	19 (15.5)
Married, civil union, or domestic partnership	6 (1.6)	2 (1.5)	0 (0)	4 (3.3)
Employment status, n (%)				
Unemployed	121 (31.9)	37 (28.7)	48 (37.8)	36 (29.3)
Employed part-time	126 (33.3)	41 (31.8)	40 (31.5)	45 (36.6)
Employed full-time	132 (34.8)	51 (39.5)	39 (30.7)	42 (34.2)
Student status, n (%)				
Currently enrolled	235 (61.8)	79 (60.8)	77 (60.6)	79 (64.2)
Not currently enrolled	145 (38.2)	51 (39.2)	50 (39.4)	44 (35.8)
Prior HIV test, n (%)				
Never	110 (29.1)	40 (30.8)	35 (27.9)	35 (28.7)
More than 3 months	253 (66.9)	87 (66.9)	85 (67.5)	81 (66.4)
Within past 3 months	15 (4.0)	3 (2.3)	6 (4.8)	6 (4.9)
Prior PrEP[c] use, n (%)				
Never	324 (85.3)	112 (86.2)	110 (86.6)	102 (82.9)
More than 3 months	51 (13.4)	15 (11.5)	17 (13.4)	19 (15.5)

	Total (N=381)	SOC[b] (n=130)	LYNX (n=127)	MyChoices (n=124)
Within past 3 months	5 (1.3)	3 (2.3)	0 (0)	2 (1.6)
Condomless anal or vaginal sex (past 3 months), n (%)				
Yes	150 (39.4)	56 (43.1)	52 (40.9)	42 (33.9)
No	231 (60.6)	74 (56.9)	75 (59.1)	82 (66.1)
Enrolled Pre–COVID-19 shutdown, n (%)				
Yes	82 (21.5)	27 (20.8)	30 (23.6)	25 (20.2)
No	299 (78.5)	103 (79.2)	96 (76.4)	98 (79.8)

[a] COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[b] SOC: standard of care.

[c] PrEP: pre-exposure prophylaxis.

HIV Testing

Compared to SOC (n=72, 59%), participants randomized to MyChoices were significantly more likely to have received at least 1 HIV test over 6 months of follow-up (n=87, 74%; *P*=.010; Table 2). Participants randomized to LYNX also had a higher proportion of testing but it did not meet the a priori threshold for statistical significance (n=80, 70%; *P*=.08). Individuals randomized to the MyChoices condition also reported more frequent HIV testing over follow-up compared to SOC (incidence rate ratio 1.25, 95% CI 1.01-1.54; *P*=.04), while this relationship did not reach significance for the LYNX condition compared to SOC (incidence rate ratio 1.13, 0.91-1.40; *P*=.26; Table 3). In time-to-event analyses (Table 4 and Multimedia Appendix 3), using all data (up to 12 months), those randomized to the MyChoices condition had a 45% higher likelihood (hazard ratio 1.45, 95% CI 1.07-1.95; *P*=.02) and those in the LYNX condition had a 34% higher likelihood (hazard ratio 1.34, 95% CI 1.00-1.81; *P*=.05) of receiving an HIV test compared to those in the SOC condition (log-rank *P*=.04).

Table 2. Estimated uptake of HIV testing and PrEP[a] in the COMPARE[b] study, 2019-2022.

	Any HIV test over 6 months			Any PrEP initiation over 6 months		
	n (%)	Risk ratio (95% CI)	*P* value	n (%)	Risk ratio (95% CI)	*P* value
SOC[c] (reference)	72 (58.5)	1.0	—[d]	19 (15.6)	1.0	—
MyChoices	87 (74.4)	1.27 (1.06-1.53)	.01	23 (21.3)	1.37 (0.79-2.37)	.26
LYNX	80 (70)	1.19 (0.98-1.44)	.08	21 (19.6)	1.26 (0.72-2.21)	.42

[a] PrEP: pre-exposure prophylaxis.

[b] COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[c] SOC: standard of care.

[d] Not applicable.

Table 3. Frequency of HIV testing in the COMPARE[a] study, 2019-2022.

	Incidence rate ratio (95% CI)	*P* value
SOC[b] (reference)	—[c]	—
MyChoices	1.25 (1.01-1.54)	.04
LYNX	1.13 (0.91-1.40)	.26

[a] COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[b] SOC: standard of care.

[c] Not applicable.

Table 4. Time to first HIV test and PrEP[a] initiation in the COMPARE[b] study, 2019-2022.

	Time to first HIV test over 12 months		Time to PrEP initiation over 12 months	
	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value
SOC[c] (reference)	—[d]	—	—	—
MyChoices	1.45 (1.07-1.95)	.015	0.90 (0.47-1.71)	.75
LYNX	1.34 (1.00-1.81)	.052	1.17 (0.64-2.13)	.61

[a] PrEP: pre-exposure prophylaxis.

[b] COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[c] SOC: standard of care.

[d] Not applicable.

According to the omnibus tests, the baseline characteristics did not significantly modify the effect of either app on HIV testing compared to SOC. However, in post hoc analyses (Table 5), both apps were significantly associated with HIV testing compared to SOC among White, non-Hispanic participants (MyChoices: RR 1.6, 95% CI 1.2-2.1; LYNX: RR 1.4, 95% CI 1.0-2.0); however, the conditions did not differ among non-White participants (MyChoices: RR 1.06, 95% CI 0.8-1.4; LYNX: RR 1.04, 95% CI 0.8-1.3). Among participants who reported engaging in condomless sex with a male or transgender female partner at baseline (Table 6), participants randomized to the MyChoices condition were nearly 6 times as likely to get an HIV test compared to SOC (RR 5.6, 95% CI 1.7-18.1; *P*=.004); the likelihood did not significantly differ among those not reporting recent condomless sex (RR 1.5, 95% CI 0.8-3.0; *P*=.21). Finally, participants who enrolled after the COVID-19 shutdown and were randomized to the MyChoices and LYNX conditions were significantly more likely to obtain an HIV test compared to the SOC condition (*P*=.005 and *P*=.02, respectively); the conditions did not differ for participants randomized prior to the COVID-19 shutdown (Table 7).

Table 5. Post hoc stratified analysis to assess moderation for HIV testing at 6 months by race or ethnicity in the COMPARE[a] study, 2019-2022.

	White, non-Hispanic			Non-White		
	n (%)	RR[b] (95% CI)	*P* value	n (%)	RR (95% CI)	*P* value
SOC[c] (reference)	28 (48.3)	1.0	—[d]	43 (67.2)	1.0	—
MyChoices	56 (75.7)	1.57 (1.17-2.11)	.003	30 (71.4)	1.06 (0.82-1.37)	.64
LYNX	35 (68.6)	1.42 (0.67-3.00)	.03	45 (70.3)	1.05 (0.83-1.32)	.70

[a] COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[b] RR: risk ratio.

[c] SOC: standard of care.

[d] Not applicable.

Table 6. Post hoc stratified analyses to assess moderation for HIV testing at 6 months by condomless anal or vaginal sex (past 3 months) in the COMPARE[a] study, 2019-2022.

	Yes			No		
	n (%)	RR[b] (95% CI)	*P* value	n (%)	RR (95% CI)	*P* value
SOC[c] (reference)	33 (62.3)	1.0	—[d]	39 (55.7)	1.0	—
MyChoices	37 (90.2)	5.61 (1.74-18.09)	.004	50 (65.8)	1.53 (0.78-2.98)	.21
LYNX	36 (76.6)	1.98 (0.83-4.75)	.13	44 (64.7)	1.46 (0.73-2.89)	.28

[a] COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[b] RR: risk ratio.

[c] SOC: standard of care.

[d] Not applicable.

Table 7. Post hoc stratified analyses to assess moderation for HIV testing at 6 months by pre–COVID-19 enrollment in the COMPARE[a] study, 2019-2022.

	Yes			No		
	n (%)	RR[b] (95% CI)	*P* value	n (%)	RR (95% CI)	*P* value
SOC[c] (reference)	13 (68.4)	1.0	—[d]	59 (56.7)	1.0	—
MyChoices	13 (59.1)	1.04 (0.68-1.59)	.86	74 (77.9)	1.37 (1.13-1.68)	.005
LYNX	12 (54.6)	0.90 (0.57-1.40)	.63	68 (73.1)	1.29 (1.04-1.59)	.02

[a] COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[b] RR: risk ratio.

[c] SOC: standard of care.

[d] Not applicable.

PrEP Initiation

While participants in both MyChoices (21%) and LYNX (20%) arms had higher rates of initiating PrEP compared to SOC (16%), these differences were not statistically significant (*P*=.52; Table 2). The time-to-event analysis also did not demonstrate a significant difference in PrEP Initiation between study conditions (log-rank *P*=.72; Table 4 and Multimedia Appendix 3). Omnibus tests did not demonstrate an effect modification of PrEP Initiation by baseline characteristics or time of enrollment (all *P*>.10).

Given the null findings for PrEP Initiation, we explored, in post hoc analyses, whether the apps were efficacious in improving the hypothesized PrEP-related mediators. Participants randomized to the MyChoices condition had significantly higher scores for the PrEP IMB scale compared to the SOC condition (β=.24; 95% CI 0.03-0.44; *P*=.02). Among participants who had not initiated PrEP, participants randomized to the LYNX arm had a lower interest in PrEP compared to SOC (β=–.43; 95% CI –0.78 to –0.08; *P*=.02); PrEP interest among those randomized to the MyChoices arm who did not initiate PrEP was not significantly different from SOC (β=–.17, 95% CI –0.53 to 0.18; *P*=.33). The arms did not differ for PrEP self-efficacy, PrEP stigma scores, nor having spoken to a provider about PrEP.

App Use

MyChoices and LYNX app use data are described in Table 4. In brief, for MyChoices, the mean number of log-ins was 8.55 (SD 17.0; range 1-186), and the mean total minutes on the app was 31.00 (SD 26.16; range 2.95-148.83). For LYNX, the mean number of log-ins was 12.43 (SD 13.10; range 1-74), and the mean total minutes on the app was 34.52 (SD 33.57; range 0.00-209.17). For either app, among those randomized to the experimental conditions, the number of app logins and time on the app were not significantly associated with HIV testing or PrEP Initiation (Table 8).

Table 8. Use and dose-response effects of MyChoices and LYNX apps on HIV testing and PrEP[a] uptake at 6 months in the COMPARE[b] study, 2019-2022.

	Range	Mean (SD)	HIV testing at 6 months		PrEP Initiation at 6 months	
			RR[c] (95% CI)	*P* value	RR (95% CI)	*P* value
LYNX (n=119)						
Number of log-ins	1-74	12.43 (13.10)	1.00 (0.98-1.02)	.97	0.97 (0.92-1.02)	.21
Number of days with log-in to app	1-49	10.0 (9.78)	1.00 (0.98-1.02)	.98	0.94 (0.87-1.01)	.09
Total minutes spent on app	<0.01-209.17	34.52 (33.57)	1.00 (0.99-1.01)	.72	1.00 (1.00-1.00)	.34
Minutes per login	<0.01-13.39	3.35 (2.38)	0.97 (0.92-1.03)	.39	0.99 (0.98-1.01)	.37
MyChoices (n=124)						
Number of log-ins	1-186	8.55 (16.96)	1.01 (0.99-1.04)	.29	0.94 (0.85-1.04)	.23
Number of days with app log-in	1-153	7.58 (14.03)	1.02 (0.99-1.06)	.24	0.95 (0.85-1.06)	.35
Total minutes spent on app	2.95-148.83	31.00 (26.16)	1.00 (1.00-1.00)	.44	0.99 (0.97-1.01)	.23
Minutes per log-in	0.51-16.58	4.86 (3.04)	1.01 (0.97-1.04)	.72	0.97 (0.85-1.12)	.68

[a] PrEP: pre-exposure prophylaxis.

[b] COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[c] RR: risk ratio

Discussion

HIV remains a significant public health problem in the United States, with particular subgroups at increased risk, including young sexual minority men. mHealth interventions, delivered through mobile apps, have the potential to reach this population given their patterns of use. Routine HIV testing and PrEP use are 2 strategies that have been shown to be efficient and cost-effective at reducing HIV transmission, particularly among groups at higher risk for HIV [39,40]. However, barriers to widespread acceptance of these interventions among young sexual minority men have reduced their population-level effects. Our 2 community-driven and theory-informed apps, MyChoices and LYNX, increased the likelihood that a large, diverse sample of at-risk young sexual minority men across the United States received at least 1 HIV test over 6 months of follow-up compared to SOC; however, the apps did not statistically significantly improve the PrEP initiation. While these results were mixed, they are illuminating.

Both MyChoices and LYNX had a meaningful impact on HIV testing compared to SOC; though the effect size for MyChoices was slightly higher and was statistically significant. An approximately 25% relative increase in HIV testing among young sexual minority men, as seen in this study, is similar to those seen in the evidence-based interventions promoted by the CDC for HIV testing [41]. It is estimated that people living with HIV who are unaware of their serostatus account for up to 40% of new sexually transmitted HIV cases, and knowing one’s status reduces HIV sexual transmission by 71% [42]. Given the potential reach and successful implementation, these apps could have a substantial impact on the epidemic among young sexual minority men in the United States.

Over the course of the study follow-up, the amount of time that youths engaged with the apps was relatively modest. Future work should explore the role of different types of reminders for repeat testing and assess whether additional information (eg, about having more fulfilling sex, signs and symptoms of STI, other sexual health topics of interest) might lead to more sustained use, and greater impact in establishing regular testing patterns and PrEP use [43]. That said, notably, among young sexual minority men randomized to MyChoices and LYNX, the number of log-ins and the time spent on the app were not associated with each outcome, suggesting that it may not necessarily be the dose but the content and functionalities that matter the most. Future studies should examine the role of app engagement to examine the optimal balance between quantity versus quality [44].

Despite omnibus tests for interactions not showing evidence for effect modification by baseline characteristics, stratified analyses demonstrated that both apps were significantly associated with HIV testing compared to SOC among White, non-Hispanic participants; however, the conditions did not differ among non-White participants. The null finding among non-White participants may have been due to the lack of statistical power due to small sample sizes. However, given that young sexual minority men of color are at disproportionate risk of acquiring HIV, it is essential to further explore how to support HIV testing uptake in this population [45], and future iterations of these apps should attempt to address any gaps. On the other hand, the MyChoices app was substantially more effective in improving HIV testing among participants who had engaged in recent condomless sex with a male or transgender female partner, suggesting that this app is appropriately routing young sexual minority men who are at higher risk of HIV acquisition toward HIV testing.

These mHealth interventions and the format of the efficacy trial facilitated the continuation of this study during the COVID-19 pandemic. Even more, post hoc analyses suggested that individuals enrolled after the COVID-19 shutdown received the greatest benefit of both MyChoices and LYNX. It is possible that the availability of at-home HIV tests, condoms, and lubes through the apps was even more essential to isolate young sexual minority men during this time frame. Given the prediction that additional pandemics are increasingly more likely due to climate change and ease of global travel, this is an important finding. Additionally, it suggests that young sexual minority men who may be isolated for reasons other than pandemics (eg, living in rural areas) might further benefit from these apps and the resources that they provide [19].

MyChoices and LYNX demonstrated a small and nonsignificant impact on PrEP Initiation among young sexual minority men in this study, including among subgroups. Modeling studies have suggested that even a modest increase in PrEP Initiation could have meaningful impacts on HIV incidence [46]. Reasons for the small effect sizes (particularly in comparison to the larger impact on HIV testing) are not fully understood; however, barriers to PrEP Initiation are greater than those for HIV testing—it is less widely available, requires clinical appointments (vs HIV testing within community-based organizations or even at home), has more cost implications and is a medicine that must be taken regularly. In this study, for those who did not initiate PrEP by 6 months, the most common reasons reported include not having a perceived need (47.3%), costs (31.2%), worry about side effects (32.2%), and concern about long-term effects (30.1%). Future app enhancements should further address these barriers, including highlighting the potential benefits of PrEP, linkage to a PrEP navigator and telehealth services for PrEP, more diverse content on side effects, and linkage to programs that help defray the costs. Additionally, since these apps were developed, several different PrEP modalities have been approved, including event-driven oral medication and injectable PrEP that can be administered every 2 months. Future iterations of these apps will need to accommodate the opportunities and challenges for youths now that a menu of PrEP choices is available.

This large, randomized efficacy trial has a number of strengths, including methodological rigor, community engagement throughout the app development and study planning, and the diversity of participants in terms of race or ethnicity, geography, employment, and insurance status. However, there are also limitations to consider. First, due to the timing of this study, the enrollment of participants was slowed due to the COVID-19 pandemic. As a result, we had to extend our enrollment dates and shorten our follow-up for a subset of participants to 6 months instead of the preplanned 12 months. However, given the enrollment criteria limited the sample to young sexual minority men at higher risk of HIV and who had not had a recent HIV test, according to the CDC, each participant would be recommended to an HIV test at least every 6 months, and thus this shorter follow-up time frame remains meaningful. Additionally, our time-to-event analysis, where all available data was used, demonstrated a similar pattern of results. This study also relied on self-reported HIV testing and PrEP Initiation, which increases the likelihood of social desirability bias and overestimation of the impact of the apps on these practices. This potential bias was reduced by the use of self-administered, web-based follow-up assessments. Additionally, the alternative would have required participants to show proof of HIV testing and PrEP Initiation, and we anticipated that the increased chances of missing data would have led to increased selection bias.

In spite of these limitations, the findings of this study are compelling and merit future exploration. In summary, future studies should enhance the functionalities related to PrEP to increase the apps' impact on this important outcome, as well as add content and functionalities related to emerging prevention technologies as they arise, including doxycycline prophylaxis for STI prevention. Additionally, given the 25% relative increase in HIV testing among young sexual minority men who used the MyChoices app and the potential for a large population-level effect, implementation studies that test and determine best practices for dissemination of this app to young sexual minority men across the United States are warranted [47].

Acknowledgments

The authors thank the study participants and community partners for their time and commitment. The authors would also like to acknowledge the work and contributions of Mr. Kenneth Coleman. This study was funded by the ATN for HIV Interventions, with primary funding by the Eunice Kennedy Shriver National Institute of Child Health and Human Development and supplemental funding from the National Institute of Mental Health, the National Institute on Drug Abuse, and the National Institute on Minority Health and Health Disparities (grant U01HD068040). This study also received support from the Emory Center for AIDS Research (P30AI050409), The Providence/Boston Center for AIDS Research (P30AI042853), and the MAC AIDS Fund. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or other funders.

Data Availability

The datasets generated during and/or analyzed during this study are available in the National Institute of Child Health and Human Development's Data and Specimen Hub data repository.

Authors' Contributions

KBB was responsible for conceptualization, methodology, formal analysis and investigation, original draft preparation, review and editing of the manuscript, funding acquisition, and supervision. KHM was responsible for conceptualization, methodology, review and editing of the manuscript, and funding acquisition. HS was responsible for conceptualization, methodology, review and editing of the manuscript, and funding acquisition. PKV was responsible for the formal analysis and review and editing of the manuscript. JHR was responsible for the review and editing of the manuscript and supervision. SB was responsible for the conceptualization, methodology, and review and editing of the manuscript. LAT was responsible for the formal analysis and review and editing of the manuscript. PS was responsible for conceptualization, methodology, review and editing of the manuscript, and funding acquisition. LHW was responsible for conceptualization, methodology, review and editing of the manuscript, and funding acquisition. AL was responsible for conceptualization, methodology, review and editing of the manuscript, and funding acquisition. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comparison of theoretical models, key components and user experiences for the LYNX and MyChoices apps.
[PNG File , 153 KB - publichealth_v11i1e63428_app1.png]

Multimedia Appendix 2
Enrollment by month for the COMPARE (Comparison of Men's Prevention Apps to Research Efficacy) study, 2019-2022. Key: Red solid line=COVID-19 Shutdown (March 15, 2020); Green dashed line=re-opening of initial sites for enrollment (July 2020); Green solid line=all sites re-opened for enrollment (November 2020); Gray dashed line=expanded to national enrollment (October 2021).
[PNG File , 26 KB - publichealth_v11i1e63428_app2.png]

Multimedia Appendix 3
Time to first HIV test (a) and PrEP Initiation (b) in the COMPARE study, 2019-2022.
[PNG File , 74 KB - publichealth_v11i1e63428_app3.png]

Multimedia Appendix 4
CONSORT eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1134 KB - publichealth_v11i1e63428_app4.pdf]

References

1. Agwu A. Sexuality, sexual health, and sexually transmitted infections in adolescents and young adults. Top Antivir Med 2020;28(2):459-462 [FREE Full text] [Medline: 32886466]
2. HIV Surveillance Supplemental Report: Estimated HIV incidence and prevalence in the United States, 2017?2021. Centers for Disease Control and Prevention. 2023. URL: https://stacks.cdc.gov/view/cdc/149080 [accessed 2024-09-12]
3. DiNenno EA, Prejean J, Irwin K, Delaney KP, Bowles K, Martin T, et al. Recommendations for HIV screening of gay, bisexual, and other men who have sex with men—United States, 2017. MMWR Morb Mortal Wkly Rep 2017;66(31):830-832 [FREE Full text] [doi: 10.15585/mmwr.mm6631a3] [Medline: 28796758]
4. US Preventive Services Task Force, Barry MJ, Nicholson WK, Silverstein M, Chelmow D, Coker TR, et al. Preexposure prophylaxis to prevent acquisition of HIV: US Preventive Services Task Force recommendation statement. JAMA 2023;330(8):736-745. [doi: 10.1001/jama.2023.14461] [Medline: 37606666]
5. Gleason N, Serrano PA, Muñoz A, French A, Hosek S. Limited interaction targeted epidemiology of HIV in sexual and gender minority American adolescents and adults: feasibility of the keeping it LITE study. JMIR Form Res 2021 Nov 26;5(11):e30761 [FREE Full text] [doi: 10.2196/30761] [Medline: 34346403]
6. Whitfield THF, Parsons JT, Rendina HJ. Rates of pre-exposure prophylaxis use and discontinuation among a large U.S. national sample of sexual minority men and adolescents. Arch Sex Behav 2020;49(1):103-112 [FREE Full text] [doi: 10.1007/s10508-019-01602-z] [Medline: 31845148]
7. Kamitani E, Johnson WD, Wichser ME, Adegbite AH, Mullins MM, Sipe TA. Growth in proportion and disparities of HIV PrEP use among key populations identified in the United States national goals: systematic review and meta-analysis of published surveys. J Acquir Immune Defic Syndr 2020;84(4):379-386 [FREE Full text] [doi: 10.1097/QAI.0000000000002345] [Medline: 32205721]
8. Wang Y, Mitchell JW, Zhang C, Liu Y. Evidence and implication of interventions across various socioecological levels to address pre-exposure prophylaxis uptake and adherence among men who have sex with men in the United States: a systematic review. AIDS Res Ther 2022;19(1):28 [FREE Full text] [doi: 10.1186/s12981-022-00456-1] [Medline: 35754038]
9. Estimated HIV incidence and prevalence in the United States, 2015-2019. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-supplemental-report-vol-26-1.pdf [accessed 2024-12-03]
10. Moskowitz D, Moran K, Matson M, Alvarado-Avila A, Mustanski B. The PrEP cascade in a national cohort of adolescent men who have sex with men. J Acquir Immune Defic Syndr 2021;86(5):536-543. [doi: 10.1097/QAI.0000000000002613] [Medline: 33399311]
11. Estem KS, Catania J, Klausner JD. HIV self-testing: a review of current implementation and fidelity. Curr HIV/AIDS Rep 2016;13(2):107-115. [doi: 10.1007/s11904-016-0307-y] [Medline: 26879653]
12. Dellucci TV, Jones SS, Adhemar C, Feldstein Ewing SW, Lovejoy TI, Starks TJ. Correlates of HIV Testing across the lifespan—adolescence through later adulthood—among sexual minority men in the US who are not on PrEP. J Behav Med 2022;45(6):975-982 [FREE Full text] [doi: 10.1007/s10865-022-00341-2] [Medline: 35916966]
13. Fisher CB, Fried AL, Macapagal K, Mustanski B. Patient-provider communication barriers and facilitators to HIV and STI preventive services for adolescent MSM. AIDS Behav 2018;22(10):3417-3428 [FREE Full text] [doi: 10.1007/s10461-018-2081-x] [Medline: 29546468]
14. Fuzzell L, Fedesco HN, Alexander SC, Fortenberry JD, Shields CG. "I just think that doctors need to ask more questions": Sexual minority and majority adolescents' experiences talking about sexuality with healthcare providers. Patient Educ Couns 2016;99(9):1467-1472. [doi: 10.1016/j.pec.2016.06.004] [Medline: 27345252]
15. Babel RA, Wang P, Alessi EJ, Raymond HF, Wei C. Stigma, HIV risk, and access to HIV prevention and treatment services among men who have sex with men (MSM) in the United States: a scoping review. AIDS Behav 2021;25(11):3574-3604 [FREE Full text] [doi: 10.1007/s10461-021-03262-4] [Medline: 33866444]
16. Newcomb ME, Moran K, Li DH, Mustanski B. Demographic, regional, and political influences on the sexual health care experiences of adolescent sexual minority men. LGBT Health 2020;7(1):28-36 [FREE Full text] [doi: 10.1089/lgbt.2019.0122] [Medline: 31750760]
17. LaBelle M, Strong C, Tseng YC. mHealth strategies to promote uptake and adherence to PrEP: a systematic review. 2020 Presented at: 12th International Conference on Cross-Cultural Design (CCD 2020); July 19–24, 2020; Copenhagen, Denmark. [doi: 10.1007/978-3-030-49913-6_9]

18. Mulawa MI, Rosengren AL, Amico KR, Hightow-Weidman LB, Muessig KE. mHealth to reduce HIV-related stigma among youth in the United States: a scoping review. Mhealth 2021;7:35 [FREE Full text] [doi: 10.21037/mhealth-20-68] [Medline: 33898604]
19. Sullivan PS, Hightow-Weidman L. Mobile apps for HIV prevention: How do they contribute to our epidemic response for adolescents and young adults? Mhealth 2021;7:36 [FREE Full text] [doi: 10.21037/mhealth-20-71] [Medline: 33898605]
20. Teens and internet, device access fact sheet. Pew Research Center. 2024. URL: https://www.pewresearch.org/internet/fact-sheet/teens-and-internet-device-access-fact-sheet/ [accessed 2024-09-12]
21. Reyes E, Silvis J, Gandhi M, Shi Y, Greene M. Telehealth access and experiences of older adults with HIV during the COVID-19 pandemic: lessons for the future. J Am Geriatr Soc 2024;72(9):2816-2824. [doi: 10.1111/jgs.18970] [Medline: 38819631]
22. Biello KB, Hill-Rorie J, Valente PK, Futterman D, Sullivan PS, Hightow-Weidman L, et al. Development and evaluation of a mobile app designed to increase HIV testing and pre-exposure prophylaxis use among young men who have sex with men in the United States: open pilot trial. J Med Internet Res 2021;23(3):e25107 [FREE Full text] [doi: 10.2196/25107] [Medline: 33759792]
23. Biello KB, Daddario SR, Hill-Rorie J, Futterman D, Sullivan PS, Hightow-Weidman L, et al. Uptake and acceptability of MyChoices: results of a pilot RCT of a mobile app designed to increase HIV testing and PrEP initiation among young American MSM. AIDS Behav 2022;26(12):3981-3990. [doi: 10.1007/s10461-022-03724-3] [Medline: 35829971]
24. Biello KB, Marrow E, Mimiaga MJ, Sullivan P, Hightow-Weidman L, Mayer KH. A mobile-based app (MyChoices) to increase uptake of HIV testing and pre-exposure prophylaxis by young men who have sex with men: protocol for a pilot randomized controlled trial. JMIR Res Protoc 2019;8(1):e10694 [FREE Full text] [doi: 10.2196/10694] [Medline: 30617042]
25. Liu A, Coleman K, Bojan K, Serrano PA, Oyedele T, Garcia A, et al. Developing a mobile app (LYNX) to support linkage to HIV/sexually transmitted infection testing and pre-exposure prophylaxis for young men who have sex with men: protocol for a randomized controlled trial. JMIR Res Protoc 2019;8(1):e10659 [FREE Full text] [doi: 10.2196/10659] [Medline: 30681964]
26. Lee S, Kapogiannis BG, Allison S. Improving the youth HIV prevention and care continuums: the Adolescent Medicine Trials Network for HIV/AIDS Interventions. JMIR Res Protoc 2019;8(3):e12050 [FREE Full text] [doi: 10.2196/12050] [Medline: 30912750]
27. Hightow-Weidman LB, Muessig K, Rosenberg E, Sanchez T, LeGrand S, Gravens L, et al. University of North Carolina/Emory Center for Innovative Technology (iTech) for addressing the HIV epidemic among adolescents and young adults in the United States: protocol and rationale for center development. JMIR Res Protoc 2018;7(8):e10365 [FREE Full text] [doi: 10.2196/10365] [Medline: 30076126]
28. PrEP brochure. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/hiv/media/pdfs/2024/04/cdc-hiv-stsh-prep-brochure-english.pdf [accessed 2023-12-11]
29. HIV testing 101. Centers for Disease Control and Prevention. 2018. URL: https://www.cdc.gov/hiv/resourcelibrary/infosheets/cdc-hiv-consumer-info-sheet-testing-2023-508.pdf [accessed 2024-09-18]
30. Sullivan PS, Stephenson R, Hirshfield S, Mehta CC, Zahn R, Bauermeister JA, et al. Behavioral efficacy of a sexual health mobile app for men who have sex with men: randomized controlled trial of mobile messaging for men. J Med Internet Res 2022;24(2):e34574 [FREE Full text] [doi: 10.2196/34574] [Medline: 35025755]
31. Biello KB, Horvitz C, Mullin S, Mayer KH, Scott H, Coleman K, et al. HIV self-testing and STI self-collection via mobile apps: experiences from two pilot randomized controlled trials of young men who have sex with men. Mhealth 2021;7:26 [FREE Full text] [doi: 10.21037/mhealth-20-70] [Medline: 33898595]
32. Sullivan PS, Driggers R, Stekler JD, Siegler A, Goldenberg T, McDougal SJ, et al. Usability and acceptability of a mobile comprehensive HIV prevention app for men who have sex with men: a pilot study. JMIR Mhealth Uhealth 2017;5(3):e26 [FREE Full text] [doi: 10.2196/mhealth.7199] [Medline: 28279949]
33. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
34. Fisher JD, Fisher WA, Shuper PA. The information-motivation-behavioral skills model of HIV preventive behavior. In: DiClemente RJ, Crosby RA, Kegler MC, editors. Emerging theories in health promotion practice and research (2nd ed.) 2009:21-63.
35. Golub SA, Starbuck L, Fikslin R, Gamarel KE. Psychometric evaluation and predictive validity of an adapted Adherence Self-Efficacy Scale for PrEP. AIDS Behav 2023 Jan;27(1):218-230. [doi: 10.1007/s10461-022-03758-7] [Medline: 35809144]
36. Donenberg GR, Emerson E, Bryant FB, Wilson H, Weber-Shifrin E. Understanding AIDS-risk behavior among adolescents in psychiatric care: links to psychopathology and peer relationships. J Am Acad Child Adolesc Psychiatry 2001;40(6):642-653 [FREE Full text] [doi: 10.1097/00004583-200106000-00008] [Medline: 11392341]
37. Kurth AE, Martin DP, Golden MR, Weiss NS, Heagerty PJ, Spielberg F, et al. A comparison between audio computer-assisted self-interviews and clinician interviews for obtaining the sexual history. Sex Transm Dis 2004;31(12):719-726. [doi: 10.1097/01.olq.0000145855.36181.13] [Medline: 15608586]
38. Barnett V, Lewis T. Outliers in Statistical Data. 3rd edition. Chichester, United Kingdom: John Wiley & Sons; 1994.

39. Juusola JL, Brandeau ML, Owens DK, Bendavid E. The cost-effectiveness of preexposure prophylaxis for HIV prevention in the United States in men who have sex with men. Ann Intern Med 2012;156(8):541-550 [FREE Full text] [doi: 10.7326/0003-4819-156-8-201204170-00001] [Medline: 22508731]
40. Sanders GD, Bayoumi AM, Sundaram V, Bilir SP, Neukermans CP, Rydzak CE, et al. Cost-effectiveness of screening for HIV in the era of highly active antiretroviral therapy. N Engl J Med 2005;352(6):570-585. [doi: 10.1056/NEJMsa042657] [Medline: 15703422]
41. HIV compendium of best practices. Centers for Disease Control and Prevention. 2023. URL: https://www.cdc.gov/hivpartners/php/compendium/index.html [accessed 2023-06-12]
42. Marks G, Crepaz N, Janssen RS. Estimating sexual transmission of HIV from persons aware and unaware that they are infected with the virus in the USA. AIDS 2006;20(10):1447-1450. [doi: 10.1097/01.aids.0000233579.79714.8d] [Medline: 16791020]
43. Hightow-Weidman LB, Bauermeister JA. Engagement in mHealth behavioral interventions for HIV prevention and care: making sense of the metrics. Mhealth 2020;6:7 [FREE Full text] [doi: 10.21037/mhealth.2019.10.01] [Medline: 32190618]
44. Iribarren SJ, Akande TO, Kamp KJ, Barry D, Kader YG, Suelzer E. Effectiveness of mobile apps to promote health and manage disease: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2021;9(1):e21563 [FREE Full text] [doi: 10.2196/21563] [Medline: 33427672]
45. Nittas V, Daniore P, Chavez SJ, Wray TB. Challenges in implementing cultural adaptations of digital health interventions. Commun Med 2024;4(1):7 [FREE Full text] [doi: 10.1038/s43856-023-00426-2] [Medline: 38182750]
46. Sewak A, Lodi S, Li X, Shu D, Wen L, Mayer KH, et al. Causal effects of stochastic PrEP interventions on HIV incidence among men who have sex with men. Am J Epidemiol 2024;193(1):6-16 [FREE Full text] [doi: 10.1093/aje/kwad097] [Medline: 37073419]
47. Li DH, Macapagal K, Mongrella M, Saber R, Mustanski B. "Your package could not be delivered": The state of digital HIV intervention implementation in the US. Curr HIV/AIDS Rep 2024;21(3):152-167. [doi: 10.1007/s11904-024-00693-1] [Medline: 38502421]

Abbreviations

ATN: Adolescent Medicine Trials Network
CDC: Centers for Disease Control and Prevention
COMPARE: Comparison of Men’s Prevention Apps to Research Efficacy
IMB: information-motivation-behavioral skills
IRB: institutional review board
mHealth: mobile health
PrEP: pre-exposure prophylaxis
RCT: randomized controlled trial
SOC: standard of care
STI: sexually transmitted infection

Edited by A Mavragani; submitted 19.06.24; peer-reviewed by S John, J Zucker, R Amico, AAS Sawitri; comments to author 13.08.24; revised version received 19.11.24; accepted 29.11.24; published 05.02.25.

Please cite as:
Biello KB, Mayer KH, Scott H, Valente PK, Hill-Rorie J, Buchbinder S, Ackah-Toffey L, Sullivan PS, Hightow-Weidman L, Liu AY
The Effects of MyChoices and LYNX Mobile Apps on HIV Testing and Pre-Exposure Prophylaxis Use by Young US Sexual Minority Men: Results From a National Randomized Controlled Trial
JMIR Public Health Surveill 2025;11:e63428
URL: https://publichealth.jmir.org/2025/1/e63428
doi:10.2196/63428
PMID:

Original Paper

Mapping Key Populations to Develop Improved HIV and AIDS Interventions: Multiphase Cross-Sectional Observational Mapping Study Using a District and City Approach

Pande Putu Januraga[1], MPH, MD, DrPH; Endang Lukitosari[2], MD; Lanny Luhukay[2], MD; Rizky Hasby[2], MSc; Aang Sutrisna[3], MSc

[1]Center for Public Health Innovation, Faculty of Medicine, Udayana University, Denpasar, Indonesia
[2]Ministry of Health of Indonesia, Jakarta, Indonesia
[3]Monitoring, Evaluation, and Learning Platform USAID, Jakarta, Indonesia

Corresponding Author:
Pande Putu Januraga, MPH, MD, DrPH
Center for Public Health Innovation
Faculty of Medicine
Udayana University
Jl PB Sudirman
Denpasar, 80232
Indonesia
Phone: 62 81246180389
Email: januraga@unud.ac.id

Abstract

Background: Indonesia’s vast archipelago and substantial population size present unique challenges in addressing its multifaceted HIV epidemic, with 90% of its 514 districts and cities reporting cases. Identifying key populations (KPs) is essential for effectively targeting interventions and allocating resources to address the changing dynamics of the epidemic.

Objective: We examine the 2022 mapping of Indonesia’s KPs to develop improved HIV and AIDS interventions.

Methods: In 2022, a district-based mapping of KPs was conducted across 201 districts and cities chosen for their HIV program intensity. This multiphase process included participatory workshops for hotspot identification, followed by direct hotspot observation, then followed by a second direct observation in selected hotspots for quality control. Data from 49,346 informants (KPs) were collected and analyzed. The results from individual hotspots were aggregated at the district or city level, and a formula was used to estimate the population size.

Results: The mapping initiative identified 18,339 hotspots across 201 districts and cities, revealing substantial disparities in hotspot distribution. Of the 18,339 hotspots, 16,964 (92.5%) were observed, of which 1822 (10.74%) underwent a second review to enhance data accuracy. The findings mostly aligned with local stakeholders’ estimates, but showed a lower median. Interviews indicated a shift in KP dynamics, with a median decline in hotspot attendance since the pandemic, and there was notable variation in mapping results across district categories. In “comprehensive” areas, the average results for men who have sex with men (MSM), people who inject drugs, transgender women, and female sex workers (FSWs) were 1008 (median 694, IQR 317-1367), 224 (median 114, IQR 59-202), 196 (median 167, IQR 81-265), and 775 (median 573, IQR 352-1131), respectively. “Medium” areas had lower averages: MSM at 381 (median 199, IQR 91-454), people who inject drugs at 51 (median 54, IQR 15-63), transgender women at 101 (median 55, IQR 29-127), and FSWs at 304 (median 231, IQR 118-425). “Basic” areas showed the lowest averages: MSM at 161 (median 73, IQR 49-285), people who inject drugs at 7 (median 7, IQR 7-7), transgender women at 59 (median 26, IQR 12-60), and FSWs at 161 (median 131, IQR 59-188). Comparisons with ongoing outreach programs revealed substantial differences: the mapped MSM population was >50% lower than program coverage; the estimates for people who inject drugs were twice as high as the program coverage.

Conclusions: The mapping results highlight significant variations in hotspots and KPs across districts and cities and underscore the necessity of adaptive HIV prevention strategies. The findings informed programmatic decisions, such as reallocating resources to underserved districts and recalibrating outreach strategies to better match KP dynamics. Developing strategies beyond identified hotspots, integrating mapping data into planning, and adopting a longitudinal approach to understand KP behavior over time are critical for effective HIV and AIDS prevention and control.

(JMIR Public Health Surveill 2025;11:e56820) doi:10.2196/56820

KEYWORDS

Indonesia; key population; mapping; pandemic; HIV; AIDS; hotspot

Introduction

Background

Indonesia faces a multifaceted HIV epidemic shaped by its extensive archipelago and significant population size: 90% of its 514 districts and cities, home to >260 million people, have reported HIV and AIDS cases [1]. This vast distribution presents unique challenges in the fight against HIV. It is projected that in 2023, there will be 543,509 people living with HIV aged ≥15 years, and the number of new HIV infections will reach 23,040 in that same year (Ministry of Health of Indonesia, unpublished data, 2023).

Although HIV transmission has been reported throughout the nation, the epidemic remains predominantly concentrated among key populations (KPs), including men who have sex with men (MSM), transgender women, female sex workers (FSWs), and people who inject drugs [2]. However, the dynamics of the HIV epidemic differ among these groups. The MSM group has seen a rise in HIV prevalence, in stark contrast to the other groups that have experienced a decline. The most significant reduction in HIV cases was observed in the FSW group: in 2007, the HIV prevalence was 10%, but this figure fell to 2.1% in 2018-2019. A similar decline was noted among the group of people who inject drugs: approximately 52.4% of HIV cases in 2017 decreased dramatically by almost 60% to 13.6% in 2018-2019. Transgender women experienced a 50% reduction in HIV prevalence—from 24.33% in 2007 to 12% in 2018. Conversely, the MSM group exhibited an upward trend: HIV prevalence surged from 5.33% in 2007 to 17.9% in 2018-2019 [3].

The observed decrease in HIV prevalence among KPs—except for the MSM group—is influenced by multiple factors and should not be interpreted as a significant reduction in risk or transmission at the community level. It is crucial to emphasize that the primary transmission mode in new cases reported in 2021 predominantly involved heterosexual groups, specifically transmission from husbands to wives or long-term partners. This includes long-term partners from among KPs. In addition, there remains a concerning deficiency in the HIV care cascade in Indonesia [4]: in 2022, only 42% of people living with HIV were on antiretroviral therapy (ART; Ministry of Health of Indonesia, unpublished data, 2023).

To effectively combat the HIV epidemic, it is paramount that HIV prevention and control initiatives prioritize KPs in Indonesia [4]. Given the country's unique characteristics, these programs must have precise data about their target populations. KPs were mapped to offer insights into these target groups' size and role in Indonesia's HIV dynamics. The most recent mapping was conducted in 2022. What distinguishes the mapping efforts of 2022 is their execution in the post–COVID-19 pandemic period, a time when the pandemic has profoundly altered the public health landscape [5]. Various Indonesian studies have documented shifts in health care behaviors and access among KPs due to the pandemic. Notably, the COVID-19 era has also driven significant changes in digital technology use, impacting both risky behaviors and access to health services among these populations [6,7]. The mapping conducted during this period considers these pandemic-induced changes, offering insights into the evolving dynamics of health service use and risk behaviors in a world altered by the COVID-19 pandemic. This makes the recent mapping efforts particularly relevant and timely, reflecting the new normal of health care engagement and the needs of KPs in Indonesia. Notably, the findings from this mapping exercise have not been reported in scientific publications.

Objectives

This study examines the outcomes of district- and city-level KP mapping in Indonesia. First, it aims to pinpoint hotspots for KPs, describing the types of places where these individuals gather, meet sexual partners, or obtain drugs for injection. In addition, the study seeks to estimate the size of KPs in selected districts and cities. Finally, it aims to compare the 2022 mapping data for 201 districts and cities with data from the pre–COVID-19 period, shedding light on the changes and continuities in KP distributions influenced by the pandemic's impact on public health and societal norms.

Methods

Mapping Frameworks

The 2022 mapping followed the basic methodology of the Priorities for Local AIDS Control Efforts (PLACE) strategy. PLACE is a cross-sectional data collection approach that recruited informants directly from hotspot locations. This method is a location based and aims to identify specific areas where targeted outreach services could effectively reach populations at high risk of HIV infection [8,9].

The core objective of the 2022 mapping effort using the PLACE methodology was to identify and map locations frequented by individuals engaged in high-risk behaviors, such as meeting sexual partners or obtaining drugs for injection. Once these locations were mapped, informants were selected and recruited from these areas, especially if the number of locations exceeded a predetermined maximum quota. Therefore, PLACE was a comprehensive strategy that involved not only identifying hotspot locations but also extensively mapping these areas. Subsequently, informants from the selected locations were systematically recruited and interviewed.

Definitions

For the purpose of mapping to identify KP sizes in Indonesia, standard definitions of the target groups were needed. FSWs were defined as women who sold sex for money or goods as their primary source of income. These women worked in brothels, streets, or public places where customers came to buy sex. They may or may not have worked for a manager or pimp.

This group also included women who worked in entertainment venues (such as karaoke, bars, massage parlors, etc) and sold sex to the customers encountered in these venues. Sex transactions occurred both within entertainment venues and outside, and the owner or manager of these venues may or may not have facilitated such transactions.

MSM, defined as men who have sex with their male partners, included those who identified as gay, bisexual, or straight, as well as individuals who sold and bought sex from other men (male sex workers). Transgender women included biologically male individuals who identified as female and behaved and dressed like women. People who inject drugs comprised men and women who had injected drugs in the past 12 months, excluding those prescribed by a medical professional. This group also included individuals receiving opiate substitution therapy or participating in abstinence-based programs.

The time variable of the past 12 months was included in the definition of people who inject drugs to ensure accurate and current data on drug injection behaviors, which could fluctuate more frequently and were subject to recent intervention impacts. By contrast, behaviors and identities within the populations of FSWs, MSM, and transgender women were generally more stable over time and did not typically require a recent time frame for accurate classification.

Mapping Procedures

The mapping procedure was carried out in multiple phases, beginning with the selection of districts and cities. For the 2022 KP mapping, 201 districts and cities were chosen proportionally based on a classification system established by the Ministry of Health, which categorized regions into 3 levels of HIV control programs: “basic,” “medium,” and “comprehensive.” Basic regions had minimal HIV testing and treatment services; medium regions provided additional services, including health promotion, HIV prevention, outreach, and condom distribution; and comprehensive regions offered a full range of HIV and AIDS services, including all levels of prevention and advocacy related to health laws in HIV control.

The selection criteria included various parameters, such as the estimated population numbers for MSM, transgender women, FSWs, and people who inject drugs as of 2020. Additional factors considered included the cumulative number of new HIV cases reported from 2011 to 2021, projected new cases for the period from 2022 to 2024, and the number of people living with HIV who were receiving ART in 2021.

The comprehensive category consisted of 100 districts and cities, accounting for 80.59% (42,681/52,955) of the total new HIV cases, individuals receiving ART, and most KPs. Of these 100 districts and cities, 97 (97%) were selected as mapping locations in 2022. The medium category included 138 districts and cities with a significant proportion of new cases and KPs, with 89 (64.5%) selected for mapping in 2022. The basic category comprised 276 districts and cities that had limited interventions in HIV testing services and ART, of which 15 (5.4%) were included in the 2022 mapping. The number of districts and cities selected for KP mapping in 2022 was proportional to the estimated number of KPs in 2020, as well as the district and city categories for HIV control programs. In addition, districts and cities that participated in the integrated biological and behavioral survey of KPs in 2023 were confirmed as locations for mapping.

Overall, the mapping activities covered 201 (39.1%) of the 514 districts and cities in Indonesia. Identifying hotspots within the 201 selected districts and cities was achieved through 1-day participatory workshops held in July and August 2022. These workshops included 4 to 13 stakeholders from each district or city, with 1809 participants across all locations. Stakeholders included representatives from the district health office, social welfare office, local police, outreach nongovernmental organizations (NGOs), and community representatives. During the workshops, working teams were formed in each district and city, consisting of a supervisor from the health office, field officers from health offices and health centers, local NGOs, and a data manager from the health office. The workshops produced a list of hotspots, incorporating insights from stakeholders about key individuals and peak activity times at each identified location.

After the workshop phase, field officers—1 individual each from the health office, local NGO, and health center—received training on the technical implementation of mapping before visiting the hotspots. The training was conducted in groups of districts, which was a cost-effective approach to maximizing the number of trained officers. Topics covered included how to perform mapping and observations at hotspots, conduct interviews, and fill out data using paper-based forms.

The mapping team carried out hotspot observations in the next phase of the project. The team members visited specific locations known as gathering points for KPs during peak hours from July to October 2022. For quality control purposes, 1822 (10.74%) of the 16,964 observed hotspots underwent a second observation. Trained enumerators conducted interviews at these hotspots using predetermined paper-based forms to collect data. The interviews addressed various aspects, such as the number of KPs present and how often they visited the hotspots within a single day. Before the interviews, all participants were asked for their informed consent to participate in the mapping initiative. Informant interviews were conducted at each hotspot with at least 2 individuals from KPs and 2 individuals from outside KPs, such as clients of sex workers or other relevant visitors.

On the basis of the information provided by the informants, the working teams from the districts and the mapping teams assessed the reliability or confidence level of the information obtained, categorizing it as low or high. Only information rated with a high confidence level was used for further analysis. The validation criteria included complete visit times and fully completed forms. These data were subsequently used to calculate the counts of KPs per hotspot and for aggregating data at the district level.

Community Participation

Involving community members in the field team was crucial for observing KPs in hotspots and gaining the trust of key individuals within these populations. To enhance the accuracy of the geomapping results, particularly for MSM and transgender women, district or city teams predominantly included members from MSM or transgender women communities in their field teams. A mixed approach was adopted for mapping FSWs, combining members from the FSW community with others who had experience with and knowledge about FSWs. When conducting mapping for people who inject drugs, the field team collaborated with the individuals responsible for accurately recording information on hotspot datasheets. This extensive engagement with individuals directly and indirectly involved with the KPs provided a comprehensive and nuanced understanding of the hotspot dynamics, which was crucial for developing effective HIV prevention and care strategies.

Data Management and Analysis

The mapping project used Microsoft Excel as a data management platform. This application allowed the mapping team to input, organize, and analyze information regarding observation times, hotspot characteristics, and informant details.

A careful mathematical and statistical approach was essential for accurately aggregating the number of KPs across districts, cities, and hotspots. By using a formula based on the average number of hotspots visited in the last 24 hours, we were able to estimate KP mapping results on a larger scale. This method ensured high accuracy and minimized the risk of duplication.

The aggregation process involved consolidating the mapping results from multiple hotspots into district and city totals. By focusing on the average hotspots visited, we ensured that the figures reflected the true distribution of KPs while avoiding excessive duplication. The following simple formula was used:

$$\sum〖\text{in district-city} = (\sum〖\text{hotspot mapping}〗) / (\text{average number of hotspots visited})〗$$

In this context, $\sum$(Hotspot mapping) represents the median number of individuals identified during peak times and Saturdays from all credible informants across all hotspots within a particular district or city. The “Average number of hotspots visited” corresponds to the mean number of distinct hotspots that a member of the KP reported visiting in the past 24 hours.

This simplified formula is the initial step in quantifying the size of KPs, although it does not fully encapsulate the complexities of KP movements and interactions [10-12]. Nonetheless, it provides a valuable baseline for refining the total KP figures at the district level, aiming to minimize the potential for duplication in the mapping process.

Ethical Considerations

The mapping initiative was approved after receiving ethics clearance from Atmajaya University (0010A/III/PPPE.PM.10.05/06/2022). Due to the legal implications associated with activities such as sex work and drug injection in Indonesia and the ongoing stigma and discrimination faced by other KPs, verbal consent was carefully obtained during the interview phase to ensure the protection of participants’ rights and confidentiality. No financial incentives were provided. The data obtained from the mapping initiative are securely stored in a password-protected folder at Udayana University and the Ministry of Health. The stored and analyzed data do not include any personally identifiable information.

Results

Mapping Results Based on Stakeholder-Provided Information During the District- or City-Level Workshops

Through collaborative, multisectoral hotspot identification at the district or city level workshops, 18,339 hotspots were identified nationwide, spanning 201 districts and cities. On the basis of stakeholder-provided information, it was estimated that, of these 18,339 hotspots, 8030 (43.79%) were frequented by FSWs, 6923 (37.75%) by MSM, 2718 (14.82%) by transgender women, and 668 (3.64%) by people who inject drugs.

The analysis of hotspot identification revealed a clear trend: districts and cities categorized as “comprehensive” have a significantly higher average number of hotspots for each KP group than those labeled “medium” and “basic.” Specifically, for MSM, on average, there were 48 hotspots per district or city in the comprehensive category compared to 24 in the medium category and 12 in the basic category; for transgender women, there were 18 hotspots on average in the comprehensive category, 11 in the medium category, and 7 in the basic category; FSWs reported an average of 54 hotspots in the comprehensive category, 30 in the medium category, and 16 in the basic category; while people who inject drugs had the fewest hotspots, averaging 12 in the comprehensive category, 6 in the medium category, and 1 in the basic category.

While the types of hotspots are often similar across districts with comprehensive, medium, and basic classifications for each KP, preferences for specific hotspot types varied among the different groups (Table 1). For MSM and people who inject drugs, open spaces were the predominant hotspot type at all priority levels of districts and cities, with the highest proportions in the comprehensive category. By contrast, salons were the primary hotspot type for transgender women, especially in the comprehensive and medium districts. Although slightly less prevalent in basic districts, salons remained a significant hotspot type. Entertainment venues consistently emerged as the most common hotspot for the FSW population across all categories, underscoring their role as central activity points for FSWs.

Overall, mapping the number of KPs at the district or city level, based on stakeholder-provided information, revealed broader disparities between the priority categories of districts and cities compared to the individual hotspot levels. These variations are depicted in Figure 1, illustrating the differences in the number of hotspots and KP numbers across the priority categories of districts and cities.

Table 1. Types of hotspots identified by the stakeholders (n=17,092).

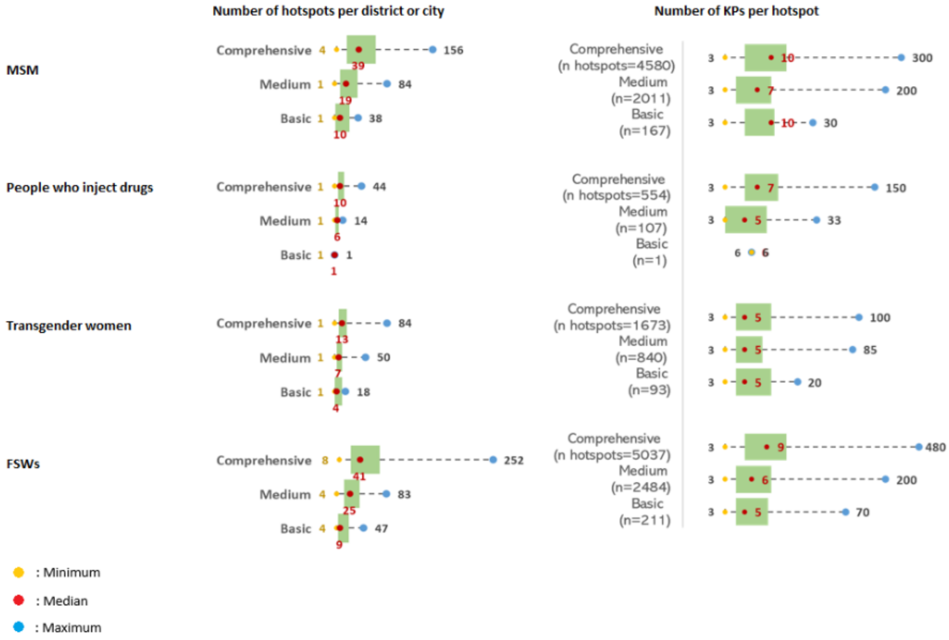
Types of hotspots	MSM[a] (n=6851), n (%)			People who inject drugs (n=651), n (%)			Transgender women (n=2687), n (%)			FSWs[b] (n=6903), n (%)		
	Comprehensive (n=97)	Medium (n=89)	Basic (n=15)	Comprehensive (n=97)	Medium (n=89)	Basic (n=15)	Comprehensive (n=97)	Medium (n=89)	Basic (n=15)	Comprehensive (n=97)	Medium (n=89)	Basic (n=15)
Open spaces	1178 (26)	562 (27)	49 (28)	336 (60)	62 (58)	1 (100)	353 (21)	169 (19)	20 (19)	562 (11)	276 (11)	43 (18)
Cafes	1127 (24)	529 (26)	53 (30)	42 (8)	6 (6)	—[c]	61 (4)	73 (8)	21 (20)	753 (15)	636 (25)	75 (31)
Recreation areas	1038 (23)	433 (21)	26 (15)	119 (21)	30 (28)	—	91 (5)	53 (6)	5 (5)	225 (4)	173 (7)	16 (7)
Salons	437 (9)	257 (12)	34 (19)	1 (0.2)	1 (1)	—	819 (48)	488 (55)	46 (45)	116 (2)	94 (4)	10 (4)
Entertainment venues	278 (6)	92 (4)	07 (4)	6 (1.1)		—	44 (3)	33 (4)	6 (6)	1,180 (23)	526 (21)	45 (19)
Residences	449 (10)	132 (6)	04 (2)	37 (7)	5 (5)	—	248 (15)	30 (3)	1 (1)	524 (10)	298 (12)	6 (2)
Fixed places	—	—	—	—	—	—	48 (3)	33 (4)	1 (1)	753 (15)	238 (9)	12 (5)
Massage parlors	23 (0.5)	8 (0.4)	1 (0.6)	—	—	—	2 (0.1)	2 (0.2)	—	707 (13.7)	159 (6.3)	14 (5.8)
Hotels	61 (1)	38 (2)	4 (2)	2 (0.4)	—	—	14 (0.8)	8 (0.9)	1 (1)	321 (6.2)	108 (4.3)	20 (8.3)
Malls or markets	6 (0.1)	3 (0.1)	—	—	—	—	—	0 (0)	—	—	7 (0.3)	—
Others	14 (0.3)	8 (0.4)	—	16 (3)	3 (3)	—	13 (0.8)	2 (0.2)	2 (2)	6 (0.1)	—	—

[a] MSM: men who have sex with men.

[b] FSW: female sex worker.

[c] Not applicable.

Figure 1. Differences in the number of key populations (KPs) identified by stakeholders across the priority categories of districts and cities. FSW: female sex worker; MSM: men who have sex with men.



Mapping Results Based on Observations and Interviews

Of the 18,339 hotspots identified across the 201 districts and cities, 16,964 (92.5%) were successfully observed. Despite these efforts, some of the hotspots remained unobserved: 6.4% (443/6923) of the MSM hotspots, 6.9% (46/668) of the hotspots frequented by people who inject drugs, 8.64% (235/2718) of the hotspots frequented by transgender women, and 8.1% (651/8032) of the FSW hotspots.

The primary reason for not observing some hotspots was changes in the field, as reported by the informants. Notably, several open spaces, previously gathering places for KPs, were no longer used. Other hotspot types, including eateries, entertainment venues, recreation areas, salons, fixed locations for selling sex, massage parlors, hotels, and malls or markets, were found to be closed or no longer operational. A few hotspots remained unobserved due to geographic challenges, where the distance from the district or city capital posed a barrier for the mapping officers.

The majority of the hotspot observations (13,832/16,964, 81.53%) for the 4 KPs were conducted from the afternoon into the evening, specifically between noon and 11:59 PM. The most active observation period was noted to be between noon and 5:59 PM, accounting for 46% (7803/16,964) of all observations across the KPs, indicating this as the prime time for data collection. Nighttime observations (6 PM- 11:59 PM) held particular importance for transgender women and FSWs, with nearly one-third (770/2483, 31.01%) to more than one-third (2730/7379, 37%) of their total observations commencing during these hours. Conversely, morning observations (6 AM-11:59 AM) were less frequent across all populations, with people who inject drugs mainly showing the lowest rate of observations during this time period.

The observed data revealed a pattern in the distribution of KP numbers, which generally corresponded with the estimates provided by stakeholders at the district or city level, albeit with a slightly lower median value. This consistency suggested that, despite the lower observed figures, the overall trends and relative differences among districts and cities of varying priority categories still mirrored the stakeholder estimates. The observations verified significant concentrations of KPs at certain hotspots. Nonetheless, the maximum values recorded indicate that while some hotspots were highly populated, their population densities were typically not as high as initially projected by the stakeholders. Figure 2 graphically displays these findings, illustrating the number of KPs per hotspot in different district or city categories.

Figure 2. Distribution of the number of key populations (KPs) observed per hotspot and per district or city. FSW: female sex worker; MSM: men who have sex with men.

Furthermore, the mapping results incorporate data from interviews conducted with the informants during the observations. During the hotspot observation period, the mapping team interviewed 49,346 informants, averaging 2 to 3 informants per hotspot. Of these 49,346 informants, 31,588 (64.01%) were members of KPs, while the remaining 17,758 (35.99%) were not classified as KPs. Specifically, on average, 2.2 informants from among people who inject drugs were interviewed per hotspot (1346 informants in total). In hotspots frequented by MSM and transgender women, on average, 1.9 informants were interviewed per hotspot (12,415 MSM and 4699 transgender women informants in total). In hotspots frequented by FSWs, on average, 1.8 informants were interviewed per hotspot (13,128 FSWs in total).

These data are categorized into 3 groups: hotspot conditions during peak times, hotspot conditions on Saturdays, and hotspot conditions before the COVID-19 pandemic. The aggregated mapping, displaying the distribution of KPs at hotspots during peak times across different districts and cities, revealed a widening gap between districts and cities categorized as comprehensive and those classified as medium or basic. This disparity is attributed to the higher number of hotspots in comprehensive districts and cities, which amplifies the difference in the mapped presence of KPs at these hotspots.

When the mapping of KP numbers at hotspots during peak times is compared with the stakeholders' estimates from Figure 1, it becomes evident that the mapped numbers during peak times are generally lower than the estimates provided by stakeholders. In addition, a comparison of the mapping distribution of KP numbers at peak times with previous data shows a typically lower Saturday attendance.

Using information from informants about the population density at hotspots before the pandemic compared to the current situation, it was found that 63.12% (10,708/16,964) of the hotspots had at least 1 informant who provided data on the number of KPs at peak times before the COVID-19 pandemic. The hotspots with the most prepandemic population data were FSW hotspots (2594/7379, 35.15%), followed by hotspots frequented by transgender women (677/2483, 27.27%) and hotspots frequented by people who inject drugs (129/622, 20.7%). Only 1 of the 6480 MSM hotspots lacked KP count information for peak times before the COVID-19 pandemic.

The analysis results regarding the median difference in the number of KPs at hotspots are detailed in Table 2. These results show a median decrease in the number of KPs present in almost all hotspot types at the time the study was conducted compared to the prepandemic period. For the MSM population, open spaces, cafes or stalls, and recreation or socialization areas exhibited a noticeable median decline: open spaces dropped from 15 (IQR 8-25) to 10 (IQR 7-20), cafes or stalls from 13 (IQR 8-20) to 10 (IQR 7-12), and recreation or socialization places from 15 (IQR 10-25) to 10 (IQR 7-20). The population of people who inject drugs also saw a decline, particularly in entertainment venues, with the prepandemic median of 15 (IQR 8.5-22.5) decreasing to 7.5 (IQR 6.25-8.5) at the time of the study was conducted.

Transgender women experienced a median decline in attendance across all hotspot types, with the most substantial decreases observed in massage parlors (from 15 to 8.5) and sex-selling locations (from 10 to 7.5). Similarly, FSW hotspots exhibited a median decline in attendance in most locations, with the most significant drops in entertainment venues (from 15 to 11) and massage parlors (from 8 to 6).

Finally, interviews with informants were conducted at these hotspot locations to account for KPs who did not frequent hotspots in our population estimates (Figure 3). For the MSM population in comprehensive districts and cities, the median number of individuals who had never visited a hotspot was 13, with a maximum of 270. In the context of the population of people who inject drugs within comprehensive districts and cities, the median was 4, escalating to a maximum of 400. This indicates that many people who inject drugs may not interact with existing hotspots. The median in comprehensive districts and cities for transgender women was 6, peaking at 78. Conversely, the FSW population in comprehensive districts displayed a median of 7, with the highest value recorded being 133.

Table 2. Distribution of key population numbers at hotspots before and during mapping (after the COVID-19 pandemic), categorized by hotspot type and key population.

Hotspot types	MSM[a], median (IQR)		People who inject drugs, median (IQR)		Transgender women, median (IQR)		FSWs[b], median (IQR)	
	Before the pandemic	During mapping	Before the pandemic	During mapping	Before the pandemic	During mapping	Before the pandemic	During mapping
Open spaces	15 (8-25)[c]	10 (7-20)[d]	8 (5-15)[c]	7 (5-13.75)[d]	10 (7-15)[c]	10 (6-13)[d]	12 (7-20)[c]	10 (6-15)[d]
Cafes	13 (8-20)[c]	10 (7-17)[d]	6 (4-10)[c]	5 (4.25-9.75)[d]	8.5 (5-15)[c]	6 (4-10)[d]	10 (6-17)[c]	8 (5-13)[d]
Salons	10 (7-15) [c]	8 (5-13)[d]	9 (9-9)[d]	12 (12-12)[c]	8 (5-12)[c]	6 (4-10)[d]	7 (5-11)[c]	5 (4-8)[d]
Entertainment venues	12 (8-20)[c]	10 (6-18.25)[d]	15 (8.5-22.5)[c]	7.5 (6.25-8.5)[d]	10 (6-13)[c]	6 (4-10.5)[d]	15 (7-28)[c]	11 (6-21)[d]
Recreation areas	15 (10-25)[c]	10 (7-20)[d]	9 (5-15)[c]	8 (5-15)[d]	10 (6-15)[c]	7.5 (5-12.75)[d]	9 (6-15)[c]	7 (5-10)[d]
Residences	12 (7-20)[c]	10 (6-15)[d]	10 (7-12.5)[c]	10 (5.5-12)[d]	10 (6-15)[c]	7 (5-10)[d]	12 (6-20)[c]	10 (6-15)[d]
Fixed places	—[e]	—	—	—	10 (7.75-15)[c]	7.5 (6-10)[d]	11 (6-25)[c]	9 (5-20)[d]
Massage parlors	14 (7.5-20)[c]	10 (6-15.5)[d]	—	—	15 (12.5-7.5)[c]	8.5 (7.75-9.25)[d]	8 (5-15)[c]	6 (4-12)[d]
Hotels	8 (5-20)[c]	9 (5-15)[d]	4 (4-4)[d]	5 (5-5)[c]	9.5 (5.75-15)[c]	8 (5.75-10)[d]	13 (8-20)[c]	10 (6-15)[d]
Malls or markets	7.5 (4.25-13.75)[c]	5.5 (4-13)[d]	—	—	—	—	5 (5-6)[c]	4 (3-6)[d]
Others	7 (6-8)[c]	6 (5-7)[d]	10 (4.75-25)[c]	5.5 (4-21.25)[d]	15 (9-22.5)[c]	12 (4.5-13.5)[d]	20 (10-20)[c]	10 (10-13)[d]

[a] MSM: men who have sex with men.

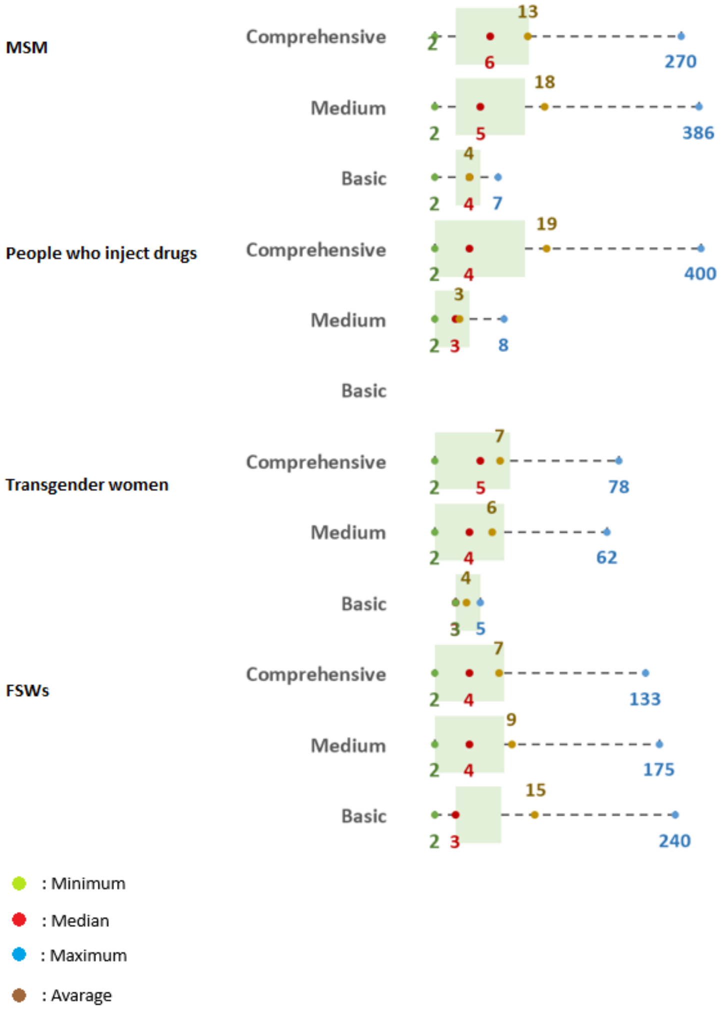
[b] FSW: female sex worker.

[c] Higher.

[d] Lower.

[e] Not applicable.

Figure 3. Distribution of key populations who never visited hotspots, based on mapping data. FSW: female sex worker; MSM: men who have sex with men.



Aggregation of Mapping Results in Districts and Cities

An aggregation of the mapping results was performed using the data collected from observations and interviews. The aggregations were carried out for each KP and district or city classification. Figure 4 illustrates the distribution of these aggregated mapping results across districts and cities, categorized by KPs and district or city classifications. The average number of mapping results per district or city for the MSM population in the comprehensive category was 1008 (median 694; range 381-6511). The wide range indicates a significant variation in mapping results across districts or cities in this category. By contrast, the medium category showed an average of 381 (median 199) mapping results. The basic category presented even lower figures, with an average of 161 (median 73), exhibiting more limited variation.

The population of people who inject drugs exhibits a similar trend but with overall lower numbers. The average aggregated mapping results in comprehensive districts and cities was 224 (median 114). The averages in the medium and basic categories were 51 (median 54) and 7, respectively, highlighting this population's disparity in aggregated mapping data.

For transgender women, the comprehensive category had an average of 196 (median 167) mapping results. In the medium category, the average decreased to 101 (median 55). The basic category showed an even lower average of 59 (median 26). Regarding the FSW population, comprehensive districts and cities averaged 775 (median 573) mapping results. The medium and basic categories saw a decrease in averages, dropping to 304 (median 231) and 161 (median 131), respectively.

Subsequently, the mapping results were compared with data from ongoing outreach programs. Table 3 provides the median number of populations mapped successfully and compares these figures with the outreach data. Generally, the mapping results yielded lower numbers than the outreach results collected throughout the program, and there was a notable discrepancy between the estimated numbers of MSM through mapping exercises and the coverage of HIV outreach and testing programs. These estimates indicated that the actual number of MSM was >50% lower than the program coverage. A similar disparity existed for FSWs, with estimates being 13% lower than outreach results. By contrast, for people who inject drugs and transgender women, the estimated numbers exceeded the extent of existing programs. Notably, the estimated figures for people who inject drugs were twice as high as the program coverage.

Figure 4. Distribution of aggregated mapping results across districts and cities, reported by key population type. FSW: female sex worker; MSM: men who have sex with men.

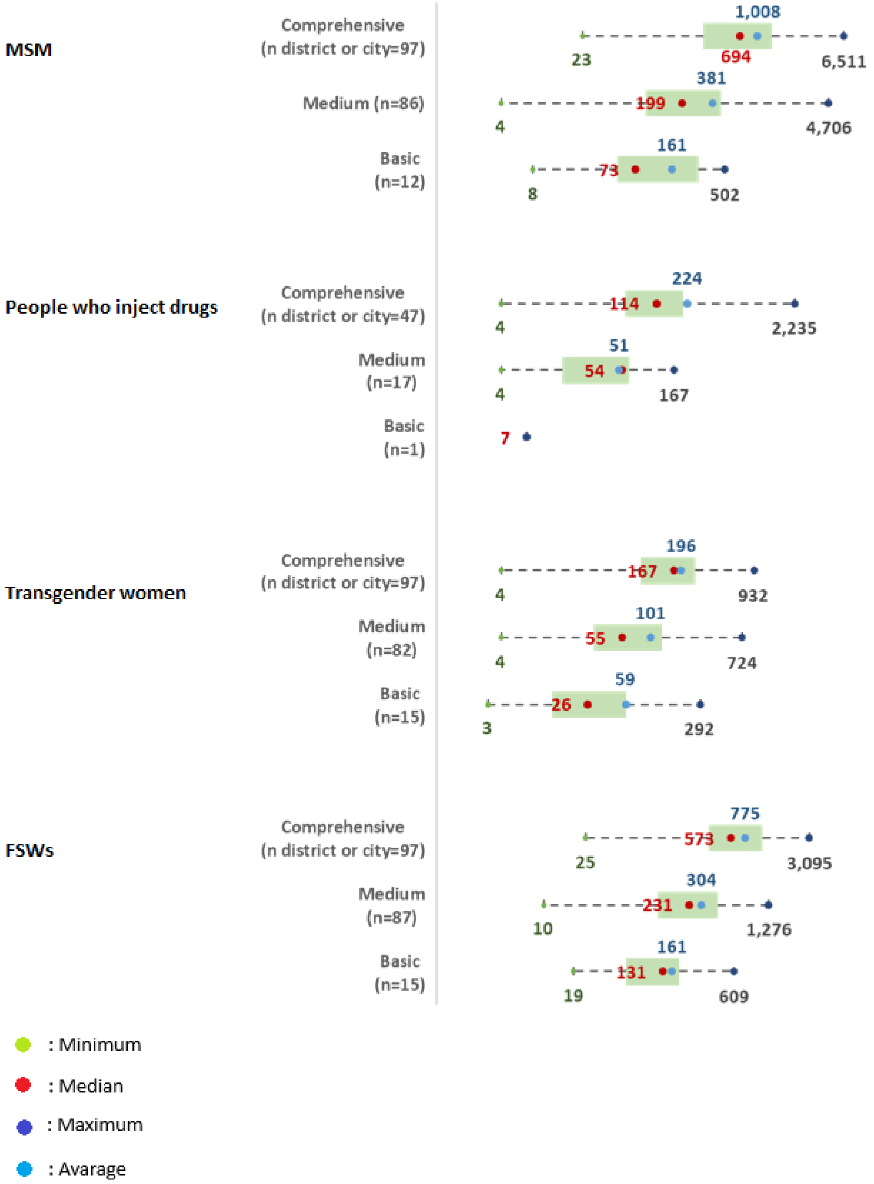


Table 3. Comparison of distribution of key population (KP) mapping data with program data by district or city.

KPs	Informed estimates	Observed	Saturdays	Peak times	Empirical estimates	Coverage by ORWs[a]	Had HIV test
MSM[b]							
Values, median (IQR; range)	238 (105.5-560; 2-4481)	189 (78-384; 2-2905)	275 (109-553; 4-4930)	323 (119-601; 4-5500)	376 (152-822; 4-6511)	2218 (1240-3412106-14,996)	934 (239-1899.5; 2-11,472)
Values, mean (SD)	425 (556)	299 (380)	470 (617)	499 (651)	697 (919)	2778 (2360)	1338 (1580)
People who inject drugs							
Values, median (IQR; range)	55 (21.5-119; 2-535)	36 (17-68; 2-278)	55 (19-97; 2-560)	64 (31-131; 2-555)	85 (36-179; 4-2235)	223.5 (80.25-441.5; 3-1473)	29.5 (24.75-167; 2-1046)
Values, mean (SD)	97 (120)	53 (59)	85 (103)	98 (108)	175 (364)	310 (323)	123 (191)
Transgender women							
Values, median (IQR; range)	63.5 (29-132.25; 1-717)	46 (22-871-369)	63 (31.75-143.5; 3-496)	73 (31-155; 3-507)	98.5 (38.75-208.5; 3-932)	158.5 (67-266.25; 14-814)	59 (16-167.5; 1-723)
Values, mean (SD)	95 (97)	63 (62)	100 (101)	106 (101)	148 (154)	194 (155)	102 (119)
FSWs[c]							
Values, median (IQR; range)	309 (143-538; 22-3603)	211 (119-371; 13-2718)	309 (150-536; 10-3087)	329 (175-551; 19-3099)	364 (180-660; 10-3095)	707 (309-1435; 69-3560)	291 (68.25-592.25; 1-2670)
Values, mean (SD)	465 (573)	317 (357)	436 (468)	465 (481)	528 (515)	956 (774)	431 (480)

[a] ORW: outreach worker.

[b] MSM: men who have sex with men.

[c] FSW: female sex worker.

Discussion

Principal Findings

The findings from district and city workshops showed that comprehensive districts and cities generally had significantly higher hotspot locations than medium and basic areas; for example, the average number of MSM hotspots per district or city in the comprehensive category was 48. This figure contrasts with 24 in the medium category and 12 in the basic category. It is important to target specific areas for HIV and AIDS interventions, such as open spaces for MSM, salons for transgender women, and entertainment venues for FSWs. Although there has been a decrease in hotspot density after the pandemic, the types of hotspots identified remain consistent with those commonly recognized. The preferences of KPs for these hotspots have also shown little change. Similar trends were observed in other mapping studies conducted in different geographic settings [13,14]. Despite varying proportions among district categories, these hotspot types are critical points for intervention across all levels of district priority.

Insights from observation data show that the concentration of KPs in hotspots varied, based on the district or city category; for instance, in the comprehensive category, the median number of MSM per district or city was 302, while for people who inject drugs, transgender women, and FSWs, the median numbers were 46, 65, and 360, respectively. The comprehensive category had a significantly higher concentration of KPs per hotspot than the medium and basic categories. However, these numbers were lower than the estimates from workshop results, where the comprehensive category had higher medians: 422 for MSM, 75 for people who inject drugs, 100 for transgender women, and 471 for FSWs. This indicates a greater focus on comprehensive areas for HIV and AIDS control in Indonesia, with some attention also given to medium category areas. It is worth noting that basic areas also receive attention, including through the Global Fund Program, although this does not imply the absence of HIV prevention programs in these areas. Standard HIV testing and treatment services are available in basic areas, following the concept of allocating interventions based on geographic epidemiological conditions [15]. Moreover, key findings from field observations indicated significant fluctuations in the presence of KPs in hotspots, particularly during peak periods. This highlights the necessity for careful interpretation of the data and underscores the complexity of addressing HIV and AIDS in varied district or city contexts [16].

The comparison between workshops' estimates of KP numbers and informant interview results revealed a consistent trend: stakeholder estimates were generally higher. This discrepancy likely stems from stakeholders basing their numbers on experience or expectations, whereas informant interviews provide a more realistic, albeit lower, representation of the actual situation. In addition, the variance might be attributed to the limited duration of observations, which often spanned only a few hours and may not have captured the full range of KP arrivals. This suggests that extending observation periods and broadening the time frame could yield a more comprehensive and accurate portrayal of KP activities [17].

The report highlights the significant impact of the COVID-19 pandemic on the mobility of KPs. There has been a decrease in the number of KPs in almost all types of hotspots compared to the prepandemic period. For the MSM population, there has been a noticeable median decline in attendance in open spaces (from 15 to 10), cafes or stalls (from 13 to 10), and recreation or socialization areas (from 15 to 10). The population of people who inject drugs also experienced a decline in attendance, particularly in entertainment venues, with the prepandemic median of 15 decreasing to 7.5 today. This aligns with academic research that documents the pandemic's effects on mobility and social interactions, especially among KPs of individuals living with HIV [6,7,18]. Studies worldwide indicate a postpandemic decline in public space social interactions [19-21]. Understanding these dynamics is crucial for developing effective postpandemic public health interventions, particularly restoring social interactions in hotspots and preventing disease spread. Strategies should evolve in response to the altered social and epidemiological landscapes, drawing on pandemic experiences to ensure that HIV prevention and response are effectively tailored to the new context.

A comparison of the mapping results for KP distribution with the coverage of HIV prevention and testing programs revealed notable discrepancies. The actual number of MSM was >50% lower than the program coverage. A similar disparity existed for FSWs, with estimates being 13% lower than outreach results. Mapping, which uses various methods such as stakeholder inputs and direct observations, often yields lower estimates compared to the broader scope of ongoing, sustainable programs; for instance, the median number of MSM per district identified through mapping was 238, significantly lower than the median recorded by outreach program coverage, which was 2218. Similar patterns were observed for people who inject drugs, transgender women, and FSWs, where the figures from HIV outreach and testing programs consistently exceeded those from mapping exercises. This suggests that mapping provides a temporal "snapshot" and might not fully capture the year-round dynamics of KPs [11]. Therefore, integrating mapping results with data from year-long programs is crucial to gaining a more comprehensive understanding of the needs and effectiveness of HIV intervention programs [14].

Finally, the mapping process conducted across 201 Indonesian districts and cities identified several limitations that are important to acknowledge for accurate analysis and future planning. These include incomplete hotspot identification and focusing only on certain segments of KPs, leading to incomplete representation. The mapping primarily accounts for visible KP members at hotspots during observation times, omitting those who might visit at other times. Another limitation is the heavy reliance on key informants, whose reports may be affected by memory bias and challenges in estimating precise numbers, potentially compromising data accuracy.

In addition, as highlighted previously, data collection was predominantly conducted in a single visit, which fails to consider temporal changes and affects the reliability of the findings. Focusing on high or peak periods could also lead to overestimations that do not represent typical daily conditions. Consequently, these factors create challenges in extrapolating the observed "snapshot" into a more extensive understanding of KP presence over extended periods. Recognizing these limitations also presents opportunities for future refinement of mapping methodologies and strategies. Further research is essential to validate and expand upon these findings. Longitudinal data collection would be particularly beneficial, providing a deeper understanding of the evolving behaviors of KPs over time and in response to various influences, including global pandemics. Despite its limitations, the mapping exercise has yielded valuable insights into the distribution of KPs based on district or city categories, which are instrumental in enhancing HIV control efforts.

Conclusions

In conclusion, the findings from this comprehensive mapping of KPs of people living with HIV across Indonesian districts and cities carry significant public health implications. First, identifying high-risk hotspots and their concentration in comprehensive districts highlights the urgent need for geographically targeted interventions, which can optimize resource allocation and enhance the effectiveness of HIV prevention efforts. Despite the postpandemic changes, the consistency in the hotspots frequented by KPs indicates stable behavioral patterns, which should inform the design of tailored intervention strategies. The discrepancies between stakeholder estimates and direct observations underscore the importance of robust, longitudinal data collection to understand KP dynamics better. In addition, the impact of the COVID-19 pandemic on mobility and social interactions among KPs necessitates the adaptation of intervention strategies to the new social and epidemiological realities. Addressing these challenges through improved mapping methodologies and sustained public health initiatives can significantly enhance HIV control efforts, reduce transmission rates, and ultimately improve health outcomes for KPs in Indonesia.

Acknowledgments

In the mapping effort, the vital involvement of multiple stakeholders was crucial. The authors' deepest thanks go to the Ministry of Health for the consistent guidance and foundational support. The provincial and district health offices were indispensable, contributing to data gathering and contextualizing the research. The authors are deeply grateful to the dedicated supervisors and outreach workers whose tireless work connected research with ground realities. Their firsthand knowledge and engagement deepened the authors' understanding of the key populations. Central to this study were the communities to whom the authors owe their utmost appreciation. Their active participation, insights, and trust formed the cornerstone of this project. Collectively, these contributors charted a clear course, ensuring that the authors’ interventions were both realistic and impactful. The authors would also like to acknowledge the primary recipient of the Global Fund Program in Indonesia, the Ministry of Health, for the support in funding the mapping initiative and the publication cost.

Data Availability

Data sharing is available upon request to the Ministry of Health of Indonesia. The authors do not possess the data; moreover, they are not authorized to share any data with other researchers.

Authors' Contributions

PPJ and AS conceptualized the study. RH, LL, EL, and AS curated the data. PPJ, LL, and AS were responsible for formal analysis. RH, LL, and EL were responsible for funding acquisition. RH, LL, EL, and AS were responsible for investigation. PPJ and AS were responsible for methodology. RH was responsible for project administration. RH, LL, EL, and AS supervised the project. PPJ was responsible for validation and wrote the original draft. All authors reviewed and edited the manuscript.

Conflicts of Interest

None declared.

References

1. Indonesia HIV AIDS quarterly report, fourth quarter 2022. Ministry of Health of Indonesia. 2022. URL: https://hivaids-pimsindonesia.or.id/download/file/Laporan_tahunan_HIV_AIDS_2022.pdf [accessed 2024-04-29]
2. Ministry of Health of Indonesia. HIV AIDS Estimates and Projections in Indonesia Update 2023. HIV AIDS Estimates and Projections in Indonesia Update 2023. Jakarta: Ministry of Health of Indonesia; 2023. URL: https:// [accessed 2024-04-29]
3. Integrated biological and behavioral survey (IBBS) 2018-2019. Ministry of Health of Indonesia. 2018. URL: https://www.unicef.org/indonesia/media/8496/file/IBBS%20Report%20YKAP%202019.pdf [accessed 2024-04-29]
4. Januraga PP, Reekie J, Mulyani T, Lestari BW, Iskandar S, Wisaksana R, et al. The cascade of HIV care among key populations in Indonesia: a prospective cohort study. Lancet HIV 2018 Oct;5(10):e560-e568. [doi: 10.1016/S2352-3018(18)30148-6] [Medline: 30143455]
5. Gedela K, Januraga PP, Luis H, Wignall FS, Irwanto I. COVID-19 lockdown in Indonesia: greater investment will be needed to mitigate the impact on people living with HIV. Asia Pac J Public Health 2020 Nov 19;32(8):461-462. [doi: 10.1177/1010539520962611] [Medline: 33073580]
6. Magnani RJ, Wirawan DN, Sawitri AA, Mahendra IG, Susanti D, Utami Ds NK, et al. The short-term effects of COVID-19 on HIV and AIDS control efforts among female sex workers in Indonesia. BMC Womens Health 2022 Jan 26;22(1):21 [FREE Full text] [doi: 10.1186/s12905-021-01583-z] [Medline: 35081950]
7. Wirawan GB, Wardhani BD, Pradnyani PE, Nurhalina A, Sulaiman N, Sukmaningrum E, et al. Behavioral changes, adaptation, and supports among Indonesian female sex workers facing dual risk of COVID-19 and HIV in a pandemic. Int J Environ Res Public Health 2022 Jan 26;19(3):1361 [FREE Full text] [doi: 10.3390/ijerph19031361] [Medline: 35162384]
8. Priorities for local AIDS control efforts (PLACE) tool kit. MEASURE Evaluation. URL: https://www.measureevaluation.org/place.html [accessed 2024-04-29]
9. Weir S, Ssengooba F, Ssekamatte-Sebuliba J. PLACE in Uganda: monitoring AIDS-prevention programs in Kampala, Uganda using the PLACE method priorities for local AIDS control. The United States Agency for International Development (USAID). URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=dc046dd8d0ae6c2c1bb794268dcf724f8ec0f1e9 [accessed 2024-04-29]
10. Wambura M, Nyato DJ, Makyao N, Drake M, Kuringe E, Casalini C, et al. Programmatic mapping and size estimation of key populations to inform HIV programming in Tanzania. PLoS One 2020 Jan 30;15(1):e0228618 [FREE Full text] [doi: 10.1371/journal.pone.0228618] [Medline: 31999810]
11. Odek WO, Githuka GN, Avery L, Njoroge PK, Kasonde L, Gorgens M, et al. Estimating the size of the female sex worker population in Kenya to inform HIV prevention programming. PLoS One 2014 Mar 3;9(3):e89180 [FREE Full text] [doi: 10.1371/journal.pone.0089180] [Medline: 24595029]
12. Bunjaku DG, Deva E, Gashi L, Kaçaniku-Gunga P, Comins CA, Emmanuel F. Programmatic mapping to estimate size, distribution, and dynamics of key populations in Kosovo. JMIR Public Health Surveill 2019 Mar 05;5(1):e11194 [FREE Full text] [doi: 10.2196/11194] [Medline: 30835241]
13. Lo J, Nwafor SU, Schwitters AM, Mitchell A, Sebastian V, Stafford KA, et al. Key population hotspots in Nigeria for targeted HIV program planning: mapping, validation, and reconciliation. JMIR Public Health Surveill 2021 Feb 22;7(2):e25623 [FREE Full text] [doi: 10.2196/25623] [Medline: 33616537]
14. Cheuk E, Isac S, Musyoki H, Pickles M, Bhattacharjee P, Gichangi P, et al. Informing HIV prevention programs for adolescent girls and young women: a modified approach to programmatic mapping and key population size estimation. JMIR Public Health Surveill 2019 Apr 01;5(2):e11196 [FREE Full text] [doi: 10.2196/11196] [Medline: 30932868]

15. Anderson S, Ghys PD, Ombam R, Hallett TB. HIV prevention where it is needed most: comparison of strategies for the geographical allocation of interventions. J Int AIDS Soc 2017 Dec 08;20(4):e25020 [FREE Full text] [doi: 10.1002/jia2.25020] [Medline: 29220115]
16. Singh SK, Malviya A, Panakadan S, Tshering N, Rehman A, Vishwakarma D. Methodological issues in estimating size of key affected population: illustrations with data from South Asia. Demogr India 2018;47(1):47-60 [FREE Full text]
17. Zhao X, Zhou Y, Chen W, Li X, Li X, Li D. Mapping hourly population dynamics using remotely sensed and geospatial data: a case study in Beijing, China. GISci Remote Sens 2021 Jun 17;58(5):717-732. [doi: 10.1080/15481603.2021.1935128]
18. Hung CC, Banerjee S, Gilada I, Green K, Inoue Y, Kamarulzaman A, et al. Impact of COVID-19 on the HIV care continuum in Asia: insights from people living with HIV, key populations, and HIV healthcare providers. PLoS One 2022 Jul 20;17(7):e0270831 [FREE Full text] [doi: 10.1371/journal.pone.0270831] [Medline: 35857755]
19. Horváth D, Ásványi K, Cosovan A, Csordás T, Faludi J, Galla D, et al. Online only: future outlooks of post-pandemic education based on student experiences of the virtual university. Soc Econ 2022 Mar 04;44(1):2-21. [doi: 10.1556/204.2021.00026]
20. Towner E, Tomova L, Ladensack D, Chu K, Callaghan B. Virtual social interaction and loneliness among emerging adults amid the COVID-19 pandemic. Curr Res Ecol Soc Psychol 2022;3:100058 [FREE Full text] [doi: 10.1016/j.cresp.2022.100058] [Medline: 35937279]
21. Long E, Patterson S, Maxwell K, Blake C, Bosó Pérez R, Lewis R, et al. COVID-19 pandemic and its impact on social relationships and health. J Epidemiol Community Health 2022 Feb;76(2):128-132 [FREE Full text] [doi: 10.1136/jech-2021-216690] [Medline: 34413184]

Abbreviations

ART: antiretroviral therapy
FSW: female sex worker
KP: key population
MSM: men who have sex with men
NGO: nongovernmental organization
PLACE: Priorities for Local AIDS Control Efforts

Edited by A Mavragani; submitted 27.01.24; peer-reviewed by S Jameela, K Iwata, I Ezirim; comments to author 08.05.24; revised version received 22.07.24; accepted 24.12.24; published 30.01.25.

Please cite as:
Januraga PP, Lukitosari E, Luhukay L, Hasby R, Sutrisna A
Mapping Key Populations to Develop Improved HIV and AIDS Interventions: Multiphase Cross-Sectional Observational Mapping Study Using a District and City Approach
JMIR Public Health Surveill 2025;11:e56820
URL: https://publichealth.jmir.org/2025/1/e56820
doi:10.2196/56820
PMID:39883483

Trends in Prescription of Stimulants and Narcoleptic Drugs in Switzerland: Longitudinal Health Insurance Claims Analysis for the Years 2014-2021

Tamara Scharf[1,2,3*], MD; Carola A Huber[4,5*], MPH, PhD; Markus Näpflin[4*]; Zhongxing Zhang[3,6*], PhD; Ramin Khatami[3,6,7*], Prof Dr Med

[1]Graduate School of Health Sciences, University of Bern, Bern, Switzerland
[2]Institute of Primary Health Care (BIHAM), University of Bern, Bern, Switzerland
[3]Centre of Sleep Medicine and Epileptology Barmelweid, Klinik Barmelweid AG, Aargau, Switzerland
[4]Department of Health Sciences, Helsana Group, Zürich, Switzerland
[5]Institute of Primary Care, University of Zürich, University Hospital Zürich, Zürich, Switzerland
[6]Barmelweid Academy, Klinik Barmelweid AG, Barmelweid, Switzerland
[7]Department of Neurology, Inselspital Bern University Hospital, University of Bern, Bern, Switzerland
[*]all authors contributed equally

Corresponding Author:
Ramin Khatami, Prof Dr Med
Centre of Sleep Medicine and Epileptology Barmelweid, Klinik Barmelweid AG, Aargau, Switzerland

Abstract

Background: Stimulants are potent treatments for central hypersomnolence disorders or attention-deficit/hyperactivity disorders/attention deficit disorders but concerns have been raised about their potential negative consequences and their increasing prescription rates.

Objective: We aimed to describe stimulant prescription trends in Switzerland from 2014 to 2021. Second, we aimed to analyze the characteristics of individuals who received stimulant prescriptions in 2021 and investigate the link between stimulant prescriptions and hospitalization rates in 2021, using hospitalization as a potential indicator of adverse health outcomes.

Methods: Longitudinal and cross-sectional data from a large Swiss health care insurance were analyzed from all insureds older than 6 years. The results were extrapolated to the Swiss general population. We identified prescriptions for methylphenidate, lisdexamfetamine, modafinil, and sodium oxybate and calculated prevalences of each drug prescription over the period from 2014 to 2021. For 2021 we provide detailed information on the prescribers and evaluate the association of stimulant prescription and the number and duration of hospitalization using logistic regression models.

Results: We observed increasing prescription rates of all stimulants in all age groups from 2014 to 2021 (0.55% to 0.81%, 43,848 to 66,113 insureds with a prescription). In 2021, 37.1% (28,057 prescriptions) of the medications were prescribed by psychiatrists, followed by 36.1% (n=27,323) prescribed by general practitioners and 1% (n=748) by neurologists. Only sodium oxybate, which is highly specific for narcolepsy treatment, was most frequently prescribed by neurologists (27.8%, 37 prescriptions). Comorbid psychiatric disorders were common in patients receiving stimulants. Patients hospitalized in a psychiatric institution were 5.3 times (odds ratio 5.3, 95% CI 4.63 - 6.08, *P*<.001) more likely to have a stimulant prescription than those without hospitalization. There were no significant associations between stimulant prescription and the total length of inpatient stay (odds ratio 1, 95% CI 1 - 1, *P*=.13).

Conclusions: The prescription of stimulant medication in Switzerland increased slightly but continuously over years, but at lower rates compared to the estimated prevalence of central hypersomnolence disorders and attention-deficit/hyperactivity disorders/attention deficit disorders. Most stimulants are prescribed by psychiatrists, closely followed by general practitioners. The increased odds for hospitalization to psychiatric institutions for stimulant receivers reflects the severity of disease and the higher psychiatric comorbidities in these patients.

(JMIR Public Health Surveill 2025;11:e53957) doi:10.2196/53957

KEYWORDS

prescription trends; claims data; cross-sectional data; narcolepsy; prescribers; prescribing practices; medical care; stimulants; stimulant medication

Introduction

Stimulants such as methylphenidate, lisdexamfetamine, and modafinil are highly potent pharmacologic treatment options for hypersomnolence disorders, including narcolepsy or attention-deficit/hyperactivity disorders/attention deficit disorders (ADHDs/ADDs).

ADHDs/ADDs are some of the most common diagnosed psychiatric disorders worldwide with a prevalence of 5.3% worldwide in the years of 1978 to 2010 among people aged 18 years or younger [1]. Prevalence can be different due to varying diagnostic methods per country. In Switzerland, a prevalence of 5.2% was found in children aged 7 to 17 years and in adult men a prevalence of 4% [2,3].

Prescriptions of stimulant drugs have been increasing at various rates over the past decades, with methylphenidate showing an 8.2 fold increase from 1996 to 2013 [4,5]. In the United States, amphetamine and methylphenidate increased from 5.6% to 6.1% in adults aged 20 years or older between 2014 and 2019 [4,5]. In New Zealand, 1.06% of adolescents received stimulants in 2016, an increase of 41.3% from 0.75% in 2011 [6]. Denmark observed a trend in stimulant prescriptions rising from 0.31 per 1000 person-years in 1996 to 7.29 per 1000 person-years in 2010 [7]. In Switzerland, data on stimulant prescription rates is scarce. A Swiss study from 2015 found a lifetime prevalence of stimulants and other substances enhancing cognitive abilities of about 1.4% in employees and students [8]. No sufficient data have, however, been collected in a nationwide study.

Stimulant medication is used for diseases often manifesting during childhood or adolescence and in many cases long-term pharmacological treatment throughout adulthood is needed. Prescription of stimulant agents in this age group is therefore of special interest to balance the need of medication and potential risk by over prescription or under prescription. Prescription of stimulants in young age groups increased from 0.02% to 0.26% over time in Asia, Australia, Europe, and North America in children and 0.003% to 1.48% in adults [9]. Among Swiss school children between 2002 and 2005, methylphenidate prescription increased from 0.74% to 1.02% for children aged 5 to 14 years [10].

This rapid prescription increase may indicate over prescription or even misuse. Misuse of stimulants is common, with up to 17% prevalence in US college students according to a meta-analysis. Misuse can lead to a range of negative consequences such as decreased appetite, insomnia, and increases in heart rate and blood pressure with increased long-term cardiovascular risk and possibly lead to increased hospitalization rates [11-13].

Properly identifying current stimulant prescription rates and discovering the prescription patterns or circumstances of their prescription (prescriber, package size, and comorbidities) may help to further identify alarming prescription increases and potential misuse in Switzerland, and the possible causes.

We therefore had 2 objectives: first, we aim to describe the rate of stimulant prescriptions in Switzerland from 2014 to 2021, focusing on both minors and adults. We hypothesize that the rate of stimulant prescriptions in Switzerland, similar to international trends, has increased over the last decade, with the highest prescription rates being recorded in 2021.

Second, we aimed at analyzing the characteristics of individuals who received stimulant prescriptions in 2021. This analysis will include factors such as comorbidities, age, and prescription details such as package size and health care providers most frequently issuing these prescriptions. Our hypotheses are that recipients likely mirror the characteristics of ADHDs/ADDs and central hypersomnolence disorders, most are minors with few comorbidities, and the likelihood of receiving a prescription decreases with age. Specialists, particularly psychiatrists and neurologists, are expected to be the primary prescribers.

Lastly, this study aims to investigate the link between stimulant prescriptions and hospitalization rates in 2021, using hospitalization as a potential indicator of adverse health outcomes. Given the expected increase in stimulant prescriptions until 2021, the hypothesis is that individuals prescribed stimulants may have a higher risk of hospitalization, particularly if prescription rates are on the rise.

Methods

Study Design

This study is a longitudinal and cross-sectional analysis of the Helsana health care insurance data of around 1.5 million people in Switzerland insured over the period of 2014 to 2021. Helsana belongs to a group of the biggest insurance companies in Switzerland and insures 14% of the Swiss population, with insureds in 26/26 cantons of Switzerland. Data describes general information on the insured persons and all their invoices for health services directed to the insurance. These invoices are representative of all health care costs of the insureds, except for the costs that were not sent as invoices to the insurance and paid by the insureds themselves (ie, over the counter drug costs and dental costs). We decided to provide more detailed descriptive statistics for data in 2021, since it was the most recent data available to us and presumably with the highest rate of stimulant prescription.

Identification of Drugs

We identified the drug invoices through their Anatomical Therapeutic Chemical code, which classifies chemical substances based on their therapeutical properties. The identification was performed for the following drugs: methylphenidate (N06BA04), lisdexamfetamine (N06BA12), modafinil (N06BA07), and sodium oxybate (N07XX04) which is a specific medication for narcolepsy treatment and used to estimate the treatment prevalence of narcolepsy patients in the dataset. Pitolisant (N07XX11) was identified too as a stimulant with specific use for narcolepsy but was excluded from further analysis, as it was only authorized for use in 2020 but had neglectable low prescription rates. We also identified the Swissmedic code of the medications, which is specific not only for the chemical substance but also the producer of the medication and package size. These drugs are only accessible through prescription by a medical professional and reimbursed by the insurance company. Overlapping prescriptions of the 4 drugs was defined as 1 prescription of one of the 4 drugs invoiced with at least one of the other drugs once or multiple times during the year of 2021.

Variables

The dataset consisted of all insureds aged ≥6 years with information on their age, sex, region of language, and region of residence. We categorized 5 age groups (6 - 17 y, 18 - 35 y, 36 - 65 y, 66 - 75 y, and 76+ y). We divided the insureds` residential regions into “rural,” “intermediate,” and “urban” subgroups according to the Swiss federal office for statistics.

The insureds had various health care plans including standard care and managed care models (eg family physician model). These health insurance plans were identified and categorized into standard care and managed care (ie, the combination of telemedical care and general practitioner [GP] care).

Chronic health condition status was identified by substance prescriptions related to chronic diseases. This was carried out according to approaches developed in previous research on the dataset [14]. We classified 22 different chronic conditions and categorized them into psychiatric, cardiologic, rheumatologic, respiratory comorbidities, or all other.

All invoices for hospitalization (ie, hospitals of all sizes providing acute care and psychiatric clinics) and the length of stay were included in our analysis.

Based on the medical prescriber who issued the invoices, several prescriber categories were defined: GP, psychiatrist, neurologist, other specialists (combining all other prescribers, such as nonspecific group practices, cardiologists, pulmonologists, rheumatologists, etc). Only health care personnel in Switzerland are allowed to prescribe medication. We further grouped them into “only prescriber” of the medication when a describer prescribed all medication exclusively for single individuals, “>50% prescriber” meaning more than 50% of the prescriptions for single individuals were invoiced by the prescriber, and “rest” with all other prescription proportions.

Within the prescriber categories (inpatient psychiatry, inpatient acute, and rehabilitation, nursing) we differentiated between inpatient (during a hospital stay) and outpatient invoices.

Statistical Analysis

For our first objective we provide descriptive statistics for prescription trends among different age groups across the years 2014 and 2021 by identifying individuals in the dataset with at least one prescription of the predefined drugs. To obtain representative data for Swiss population we extrapolated these data by current residency numbers and populations statistics of the Federal Statistical Office.

For our second objective we restricted the descriptive statistics to the year 2021 and provide detailed information on age, sex, region of residence, health insurance status, comorbidities, prescribers, and package size. Logistic regression models were performed to evaluate the association of at least one stimulant prescription versus no prescription (ie, the dependent variable is prescription yes or no), number of hospitalization, and length of stay, adjusted for age, sex, region of residence, health insurance status, and number of chronic diseases for the year 2021.

All data management, graphic generation, and analysis was performed with the statistics program R (version 4.2.1; R Foundation).

Data Availability

The authors were permitted access to the data by collaboration with the insurance companies research team. The datasets generated and analyzed during this study are available from the corresponding author on reasonable request. AI was not used in any way in data generation, analysis, and presentation of results.

Ethical Considerations

According to ethical and legal regulations in Switzerland no ethical approval or patient consent was needed for this study, as all data complied with privacy regulations and personal data protection, data was anonymized when presented to the research team. The Swiss Human Research Act (REQ-2017 - 00280) did not apply to this project. The exploratory statistical analyses of the feasibility test complied with the Swiss Federal Law on data protection. All data were anonymized and deidentified prior to the performed analysis to protect the privacy of patients, physicians, and hospitals. According to the national ethical and legal regulation, an ethical approval was not needed because the data were pre-existing and deidentified. Since data was anonymized, no consent of patients was required.

Results

Trend of Stimulant Prescription Per Year

As baseline we refer to the prescription period in 2014. Between baseline and 2021, on average 14% of Swiss people were at any time insured with the Helsana Group.

In the year 2014, 0.55% (42,848 insureds) of insured people of any age received at least 1 stimulant agent (or stimulant prescription). This number increased up to 0.8% (66,113 insureds) in the year of 2021. The largest growing percentage of stimulus prescriptions were in the youngest age group with an increase of prescription of 0.6 percentage points (17,972, 1.8% to 24,982, 2.4%) between 2014 and 2021 (Table 1).

Between 2014 and 2021, 32,2418 packages of methylphenidate were the most prescribed stimulant, followed by 46,074 packages of lisdexamfetamine, 8797 packages of modafinil, and 3115 packages of sodium oxybate. We found a prescription increase of all identified medications in younger and middle age groups over the years of 2014 to 2021, extrapolated by the Swiss population. The increase in prescription was steady in age groups aged 36 - 65 years, whereas in other age groups prescription stagnated from 2018 to 2020, with a steep increase from 2020 to 2021.

Methylphenidate prescription increased overall in all age groups with a steady increase in insureds aged 36 - 65 years. All other age groups experienced a steep increase after 2020. Only insureds aged 66 - 75 years and 76+ years ever experienced a smaller prescription rate than at baseline in 2014, with a drop to 89% (664/746) in 2017 (Figure 1).

Lisdexamfetamine prescription was low at baseline and increased steadily in all age groups with great increase from 0 prescriptions at baseline to 39 in insureds aged 60 ‐ 65 years. Smaller increase in prescription was seen in age groups aged 66 ‐ 75 years and 76+ years (Figure 1).

Prescription rates of modafinil—only prescribed to few—increased the most in the youngest age group (6 ‐ 17 y). The prescribing trend in this age group strongly fluctuated. From its highest peak in 2019 rates decreased from around 1100% (66/6) to slightly less than 300% (15/6) prescription compared to the baseline in 2014. In 2021, modafinil was most frequently prescribed in age groups aged 6 ‐ 17 years and 76+ years. All other age groups had only a moderate increase in modafinil prescriptions over the years 2014 to 2021.

Sodium oxybate overall increased the most in insureds aged 18 ‐ 35 years and 6 ‐ 17 years from 100% to 175% in 2021. Prescription rates for sodium oxybate only decreased overall in insureds aged 66 ‐ 75 years. Age groups aged 36 ‐ 65 years and 76+ experienced a similar prescription trend over the years with overall increase but declining prescription rates after 2020.

Table . Proportions of Swiss insureds with at least 1 stimulant agent or other narcolepsy treatment prescription within age groups for each year between 2014 and 2021.

Age (years)	2014, n (%)	2015, n (%)	2016, n (%)	2017, n (%)	2018, n (%)	2019, n (%)	2020, n (%)	2021, n (%)	*P* value (chi-square test for trend in proportions)
6 ‐ 17	17,972 (1.8)	17,624 (1.8)	17,702 (1.8)	18,711 (1.9)	20,209 (2)	21,296 (2.1)	21,714 (2.1)	24,982 (2.4)	<.001
18 ‐ 35	13,145 (0.68)	14,103 (0.72)	15,159 (0.77)	16,755 (0.85)	17,874 (0.91)	17,969 (0.92)	17,998 (0.92)	21,599 (1.10)	<.001
36 ‐ 65	10,708 (0.31)	11,661 (0.34)	12,355 (0.35)	13,584 (0.38)	14,537 (0.41)	15,277 (0.43)	16,394 (0.45)	18,251 (0.50)	<.001
66 ‐ 75	746 (0.1)	731 (0.1)	793 (0.1)	664 (0.1)	777 (0.1)	763 (0.1)	776 (0.1)	908 (0.1)	.10
76+	277 (0)	263 (0)	271 (0)	216 (0)	274 (0)	285 (0)	333 (0)	373 (0.10)	.05
All age groups	42,848 (0.55)	44,382 (0.57)	46,280 (0.59)	49,931 (0.63)	53,671 (0.67)	55,590 (0.69)	57,216 (0.70)	66,113 (0.81)	<.001

Figure 1. Trends of stimulant prescription from 2014 to 2021 per age group and active ingredient (indexed, base year=2014).

Factors Associated With Prescription in 2021

Characteristics of Stimulant Users

Most stimulant receivers in the year of 2021 were male representing 61% (42,803/70,396) of our population. Only modafinil was more often prescribed to women than to men with 55% (970/1776) female receivers. The highest proportion of prescriptions was provided to people living in urban areas with a proportion of 67% (46,968/70,396) of all stimulants compared to intermediate and rural area residents. Managed care was the preferred health care plan for patients receiving stimulants with 72% (50,513/70,396) receiving. A total of 45% (798/1776) of all modafinil receivers had 3+ chronic illnesses, 55% (775/1399) of those had psychiatric comorbidities, followed by cardiological and rheumatological diseases. A total of 42% (49/121) of sodium oxybate users had no comorbidities. Most common chronic illness in sodium oxybate users was psychiatric (43/72, 60%) or cardiologic (31/72, 43%). Additionally, half of the methylphenidate users had comorbidities (28,619/57,128, 50%, 28,509 had no chronic illness identified) with psychological (14,396/28,619, 50%) and other chronic conditions (7034/28,619, 25%) as the most common identified chronic diseases. Similar results were found for lisdexamfetamine users, that is, 43% (4908/11,371) users had no comorbidities and psychological (3556/6463, 55%) and other (1989/6463, 31%) comorbidities were the most common chronic diseases (Table 2).

Table . Characteristics of Swiss insureds receiving any stimulant prescription in the year 2021.

Characteristic	Modafinil (n=1776)	Sodium oxybate (n=121)	Methylphenidate (n=57,128)	Lisdexamfetamine (n=11,371)
Sex, n (%)				
Male	806 (45.4)	63 (52.1)	35,038 (61.3)	6896 (60.7)
Female	970 (54.6)	58 (48)	22,090 (38.7)	4475 (39.4)
Age (years)				
Median (IQR)	45 (33-6)	38 (29-5)	22 (14-4)	26 (16-4)
Mean (SD)	46 (16)	40 (18)	27 (16)	28 (14)
Age (years, in groups), n (%)				
6 - 17	16 (0.9)	10 (8.3)	22,948 (40.2)	3355 (29.5)
18 - 35	568 (32)	47 (38.8)	18,002 (31.5)	4401 (38.7)
36 - 65	976 (55)	48 (39.7)	15,157 (26.5)	3565 (31.4)
66 - 75	139 (7.8)	12 (9.9)	742 (1.3)	38 (0.3)
76+	77 (4.3)	4 (3.3)	279 (0.5)	12 (0.1)
Region of residence, n (%)				
Urban	1060 (59.7)	80 (66.1)	38,196 (66.9)	7632 (67.1)
Intermediate	449 (25.3)	29 (24)	11,554 (20.2)	2422 (21.3)
Rural	267 (15)	12 (9.9)	7378 (12.9)	1317 (11.6)
Health insurance status, n (%)				
Managed care	1064 (59.9)	82 (67.8)	41,178 (72.1)	8189 (72)
Standard care	712 (40.1)	39 (32.2)	15,950 (27.9)	3182 (28)
Comorbidities [14], **n (%)**				
0	377 (21.2)	49 (40.5)	28,509 (49.9)	4908 (43.2)
1	340 (19.1)	27 (22.3)	12,617 (22.1)	2601 (22.9)
2	261 (14.7)	25 (21.7)	6678 (11.7)	1619 (14.2)
3+	798 (44.9)	20 (16.5)	9324 (16.3)	2243 (19.7)
Most frequent, n (%)	psyd:[a] 775 (43.6)	psyd: 43 (36)	psyd: 14,396 (25.2)	psyd: 3556 (31.3)
2nd most frequent, n (%)	card:[b] 592 (33.3)	card: 31 (26)	ther:[c] 7034 (12.3)	ther: 1989 (17.5)
3rd most frequent, n (%)	rheu:[d] 430 (24.2)	rheu: 17 (14)	rheu: 6588 (11.5)	resp:[e] 1435 (12.6)

[a] psyd: psychiatric.

[b] card: cardiological.

[c] ther: other.

[d] rheu: rheumatological.

[e] resp: respiratory.

Package Size

All medication was predominantly prescribed more than once within a year, with ≥5 packages prescribed in 40.4% (715/1768) of all modafinil prescriptions, 95.8% (115/120) of all sodium oxybate prescriptions, 47.5% (27,124/57,093) of all methylphenidate prescriptions, and 59.1% (6727/11,392) of all lisdexamfetamine prescriptions (Table 3).

Table . Number of Swiss insureds and their stimulant prescriptions by number of packages and by prescriber (profession of the physician) in the year 2021.

	At least 1 stimulant agent (total)	At least 1 modafinil use	At least 1 sodium oxybate use	At least 1 methylphenidate use	At least 1 lisdexamfetamin use
Total patients (N)	66,113	1768	120	57,093	11,392
Number of packages, n (%)					
1	12,235 (17.4)	536 (30.3)	0 (0)	9984 (17.5)	1715 (15.1)
2	9301 (13.2)	228 (12.9)	5 (4.2)	7968 (14)	1099 (9.6)
3	7435 (10.6)	113 (6.4)	0 (0)	6300 (11)	1022 (9)
4	6721 (9.6)	176 (10)	0 (0)	5716 (10)	829 (7.3)
≥5	34,681 (49.3)	715 (40.4)	115 (95.8)	27,124 (47.5)	6727 (59.1)
Package sizes, median (IQR)[a]					
1	—[b]	30 (30-90)	—	30 (30-100)	30 (30-30)
2	—	90 (60-90)	—	50 (30-100)	30 (30-30)
3	—	90 (67.5 - 90)	—	50 (30-83)	30 (30-30)
4	—	90 (90-90)	—	50 (35-72)	30 (30-30)
≥5	—	90 (90-90)	—	45 (30-60)	30 (30-30)
Prescriber of the issued prescriptions, n (%)					
General practitioner (GP)	27,323 (41.3)	639 (36.1)	10 (8.3)	24,127 (42.3)	3529 (31)
Only	22,578 (82.6)	476 (74.5)	10 (100)	19,521 (80.9)	1762 (49.9)
>50	1911 (7)	63 (9.9)	0 (0)	1802 (7.5)	717 (20.3)
Rest	2834 (10.4)	100 (15.6)	0 (0)	2804 (11.6)	1050 (29.8)
Psychiatrist	28,057 (42.4)	336 (19)	5 (4.2)	23,553 (41.3)	6222 (54.6)
Only	23,267 (82.9)	220 (65.5)	0 (0)	18,167 (77.1)	3204 (51.5)
>50	2247 (8)	31 (9.2)	5 (100)	2260 (9.6)	1308 (21)
Rest	2543 (9.1)	85 (25.3)	0 (0)	3126 (13.3)	1710 (27.5)
Neurologist	748 (1.1)	181 (10.2)	37 (30.8)	545 (1)	63 (0.6)
Only	558 (74.6)	97 (53.6)	11 (29.7)	384 (70.5)	26 (41.3)
>50	99 (13.2)	31 (17.1)	20 (54.1)	48 (8.8)	25 (39.7)
Rest	91 (12.2)	53 (29.3)	6 (16.2)	113 (20.7)	12 (19)
Other specialists	7962 (12)	286 (16.2)	15 (12.5)	6680 (11.7)	1300 (11.4)
Only	4599 (57.8)	176 (61.5)	10 (66.7)	3782 (56.6)	437 (33.6)
>50	955 (12)	28 (9.8)	5 (33.3)	801 (12)	234 (18)
Rest	2408 (30.2)	82 (28.7)	0 (0)	2097 (31.4)	629 (48.4)
Inpatient psychiatry	6000 (9.1)	70 (4)	6 (5)	5027 (8.8)	1234 (10.8)
Only	3429 (57.2)	27 (38.6)	0 (0)	2857 (56.8)	362 (29.3)
>50	898 (15)	11 (15.7)	0 (0)	756 (15)	214 (17.3)
Rest	1673 (27.9)	32 (45.7)	6 (100)	1414 (28.1)	658 (53.3)
Inpatient acute, rehabilitation, and nursing	5549 (8.4)	496 (28.1)	60 (50)	4251 (7.4)	1110 (9.7)
Only	3312 (59.7)	284 (57.3)	11 (18.3)	2376 (55.9)	415 (37.4)
>50	756 (13.6)	47 (9.5)	32 (53.3)	699 (16.4)	136 (12.3)
Rest	1481 (26.7)	165 (33.3)	17 (28.3)	1176 (27.7)	559 (50.4)

[a] Median (IQR) package size (number of units) per number of prescribed packages. Not shown for sodium oxybate, since this substance is a liquid.

[b]Not applicable.

Prescribers

In 2021, psychiatrists were the most frequent prescribers with 42.4% (28,057/66,172) of all the prescribed medications. Lisdexamfetamine (6222/11,396, 54.6%), followed by methylphenidate with 41.3% (23,553/57,038) were mainly prescribed by them. By contrast modafinil and sodium oxybate were rarely prescribed by psychiatrists as 19% and 4.2% (336/1769 and 5/120), respectively. If chosen as a prescriber, psychiatrists are often the only source of prescription for 82.9% (23,267/28,057) insured individuals of all medication of 2021.

The second most frequent prescribers were GPs, with a similar high proportion of 41.3% (27,323/66,172) of all drug prescriptions. More specifically 42.3% (24,127/57,038) methylphenidate was prescribed by GPs followed by modafinil (639/1769, 36.1%), lisdexamfetamine (3529/11,396, 31%), and smaller proportions for sodium oxybate (10/120, 8.3%). GPs were most often the exclusive prescribers of the medications (22,578/27,323, 82.6%). Only 10.4% (2834/27,323) of prescriptions by GPs shared the prescribing job with other medical specialists.

Neurologists were rarely prescribers of stimulants or narcolepsy treatments, as only 1.1% (748/66,172) of all prescriptions were invoiced by them. Only sodium oxybate was the most frequent (37/120, 30.8%) medication prescribed by neurologists. If chosen as the prescriber, they are often the only source (558/748, 74.6%) from which individuals received the prescriptions in 2021 (Table 3).

Concerning invoices handed in by hospitals or psychiatric clinics or rehabilitation facilities, they only made up a small part of overall stimulant invoices. An exception is sodium oxybate, of which 45.1% (60/133) of invoices are issued by an acute clinic, rehabilitation clinic, or nursing home (Table 3).

Association of Stimulant Use and Outcomes

We found an association with patients receiving a stimulant or narcolepsy treatment prescription and increased hospitalizations in a psychiatric facility (odds ratio [OR] 5.30, 95% CI 4.63 - 6.08, *P*<.001). In contrast, there was a negative association between stimulant prescription and hospitalization in an acute medical care facility (OR 0.77, 95% CI 0.73 - 0.82, *P*<.001). There were no significant associations between stimulant prescription and the total length of inpatient stay (OR 1, 95% CI 1 - 1, *P*=.13; Table 4).

Table . Regression model of predicting the outcomes of hospitalization and length of stay in Swiss insureds, who received stimulant prescription in the year 2021.

Characteristic		Odds ratio (95% CI)	*P* value
Sex			
	Male	—[a] (—)	—
	Female	0.89 (0.86 - 0.92)	<.001
Age (years, in groups)			
	6-17	1.17 (1.13-1.21)	<.001
	18 - 35	— (—)	—
	36 - 65	0.62 (0.60 - 0.64)	<.001
	66 - 75	0.31 (0.27 - 0.34)	<.001
	76+	0.24 (0.20 - 0.28)	<.001
Region of residence			
	Urban	— (—)	—
	Intermediate	1.09 (1.05 - 1.13)	<.001
	Rural	0.97 (0.94 - 1.02)	.2
Health insurance status			
	Standard	— (—)	—
	Managed care	0.69 (0.67 - 0.71)	<.001
Number of comorbidities		1.56 (1.54 - 1.59)	<.001
Total inpatient length of stay		1 (1 - 1)	.13
Hospitalization acute (yes or no)		0.77 (0.73 - 0.82)	<.001
Hospitalization psychiatry (yes or no)		5.3 (4.63 - 6.08)	<.001

[a]Not applicable.

When only focusing on methylphenidate or lisdexamfetamine compared to a balanced sample of nonstimulant users, we found a positive association between their prescription and hospitalization in a psychiatric facility (OR 6.85, 95% CI 5.89 - 7.99, *P*<.001). No significant association was found between the prescription rate and the total inpatient length of stay (OR 1, 95% CI 1 - 1, *P*=.1). Hospitalization in an acute medical care facility was less likely (OR 0.76, 95% CI 0.71 - 0.8, *P*<.001) to prescribe methylphenidate or lisdexamfetamine (Table 5).

Table . Regression models of predicting the outcomes of hospitalization and length of stay in Swiss insureds who received stimulant prescriptions in the year 2021. Prescriptions of methylphenidate or lisdexamfetamine are shown in the left columns, prescriptions of modafinil or sodium oxybate are shown in the right columns.

Characteristic		Methylphenidate- or lisdexamfetamine-users versus nonusers		Modafinil- or sodium oxybate-users versus nonusers	
		OR[a] (95% CI)	*P* value	OR[a] (95% CI)	*P* value
Sex					
	Male	— (—)	—	— (—)	—
	Female	0.9 (0.87 - 0.93)	<.001	0.84 (0.72 - 0.97)	.02
Age (years, in groups)					
	6-17	1.17 (1.13-1.21)	<.001	1.17 (0.73-1.88)	.5
	18 - 35	— (—)	—	— (—)	—
	36 - 65	0.63 (0.6 - 0.65)	<.001	0.57 (0.48 - 0.68)	<.001
	66 - 75	0.31 (0.28 - 0.36)	<.001	0.25 (0.19 - 0.34)	<.001
	76+	0.26 (0.21 - 0.31)	<.001	0.15 (0.09 - 0.23)	<.001
Region of residence					
	Urban	— (—)	—	— (—)	—
	Intermediate	1.11 (1.07 - 1.15)	<.001	0.96 (0.8 - 1.15)	.7
	Rural	0.95 (0.91 - 0.99)	.019	1.11 (0.89 - 1.38)	.3
Health insurance status					
	Standard	— (—)	—	— (—)	—
	Managed care	0.69 (0.67 - 0.71)	<.001	0.58 (0.5 - 0.68)	<.001
Number of comorbidities		1.55 (1.53 - 1.57)	<.001	1.62 (1.54 - 1.71)	<.001
Total inpatient length of stay		1 (1-1)	.10	1 (1-1)	.8
Hospitalization acute (yes or no)		0.76 (0.71 - 0.8)	<.001	1.23 (0.96 - 1.56)	.10
Hospitalization psychiatry (yes or no)		6.85 (5.89 - 7.99)	<.001	2.15 (1.17 - 4.23)	.02

[a] Odds ratio

[b] Not applicable.

When only focusing on modafinil or sodium oxybate compared to a balanced sample of nonstimulant users, we found a positive association between their prescription and hospitalization in a psychiatric facility (OR of 2.15, 95% CI 1.17-4.23, *P*=.02). We found no significant association between total inpatient length of stay (OR 1, 95% CI 1 - 1, *P*=.80) and hospitalization in an acute medical care facility (OR 1.23, 95% CI 0.96 - 1.56, *P*=.10; Table 5).

Discussion

Summary

We found an increasing trend of stimulant and narcoleptic drug prescriptions in Switzerland from 0.6% to 0.8% over the years 2014 to 2021. Most stimulants are prescribed continuously with more than 5 packages in 1 year per insured to underaged individuals with no comorbidities. Psychiatrists and GPs are often the prescribers of stimulants, much more frequently than neurologists.

Context in Research

Compared with data on global stimulant prescription rates, our results are notably lower, showing a smaller increase in prescription rates than in the United States (5.6% - 6.1%), New Zealand (0.75% - 1.06%), and Denmark (0.03% - 0.73%) [5-7]. Given the ADHDs/ADDs prevalence rates of 4% in adult men and 5.3% in children in Switzerland [2,3], compared to a worldwide prevalence of between 5% and 11.4% [1,15,16], it is reasonable to interpret the lower Swiss prescription rates as either an indication of underprescription or as lower stimulant misuse rates in Switzerland [15]. Even when taking all stimulants together, current prescription rates do not reach the prevalence rate of ADHDs/ADDs in Switzerland. The fact that

ADHDs/ADDs are also treated nonpharmacologically is another argument for assuming that our result of low prescription prevalence is lower than the disease prevalence. Unfortunately, there are no data that quantifies the extent of drug or nonpharmacological treatment for ADHDs/ADDs and thus could help define the normal gap between prescription rate and disease prevalence. Our distribution of prescription rates corresponds to the disease distribution in different age groups, with prescriptions and prevalence of ADHDs/ADDs being higher in the younger age groups [17]. The distribution of prescriptions by gender reflects the current disease prevalence of ADHDs/ADDs, as we found a slightly lower proportion of females than males, in line with another summary by Thapar and Cooper [17]

Stimulants were predominantly prescribed in urban areas. A higher population compared to rural areas, a higher density of prescribing physicians, and a higher number of hospitals specialized in diagnosis and treatment of these diseases in urban settings may account for this predominant prescription pattern. Vice versa underprescription in rural areas could be due to reduced access to adequate health care services.

Methylphenidate was by far the most prescribed stimulant agent, followed by lisdexamfetamine, a new stimulant showing promising results for treatment of ADHDs/ADDs [18]. The prescription characteristics were very similar between the two stimulants, with lisdexamfetamine more often prescribed to older adults than methylphenidate. This shows that lisdexamfetamine is possibly used as a second medication after methylphenidate was prescribed in the young and did not give continuous results while the patients aged.

Prescriptions were most frequently issued by psychiatrists. Since methylphenidate and lisdexamfetamine are standard treatment for the psychiatric diseases ADHDs and ADDs, this prescription pattern makes sense from a health provider perspective and is confirmed by other findings about prescribers of stimulants [19].

Prescription characteristics of sodium oxybate should reflect its specificity for narcolepsy because it is not indicated for any other disease. Here we find most of the prescriptions in middle aged groups and fewer in children and teenagers, with a nearly even distribution between women and men. This grouping does not match with the expected narcolepsy features of young patients with a possible second peak in the late forties [20]. The lack of prescription in young patients is best explained by a missed or severely delayed diagnosis of narcolepsy, which is in line with the recently published delayed diagnosis for Switzerland and other European countries [21,22]. Surprisingly, neurologists who diagnose and treat narcolepsy, are rarely the patients' prescription source, even for highly specific and not easy to handle medication, such as sodium oxybate. Multiple reasons may account for this prescription practice, among them is the lower barrier for receiving an appointment with GPs before prescriptions expires, compared to neurologists.

We found no significant change in the odds of at least one stimulant prescription in patients hospitalized in an acute hospital but a significant increase in the odds of patients hospitalized in a psychiatric facility. This reflects the fact that ADHDs/ADDs are often overlapping or comorbid with other psychiatric diseases which can lead to hospitalization, such as addiction, disruptive disorders, anxiety disorders, or bipolar disorders [23-26], and by the severity of disease.

We assume that our data is representative for Switzerland, since we analyzed claims data extrapolated to the entire population from one of the biggest insurance companies in Switzerland, with a nearly equal distribution across the country. As stimulant agents are only accessed through prescription, we were able to register all invoices for the medication in question in real world; therefore, our data minimize sampling bias and recall bias that frequently influence the accuracy and reliability of retrospective studies.

Limitations

This study is an analysis of health insurance claims data, which does not contain any information on the clinical reason of why a medication was indicated. We therefore could not distinguish between prescriptions according to current treatment guidelines and prescriptions of pharmacological treatment for diseases without proper diagnosis.

We identified prescribers by Zahlsteller register (registered number for medical personel allowed to bill insurance companies) number and medication prescription pattern, which may lead in some cases to misclassification, as some physicians share Zahlsteller register numbers in group practices and GPs go sometimes through additional training as psychiatrists or neurologists.

Conclusion

The prescription of stimulants and sodium oxybate in Switzerland increased slightly but continuously over the past years, but at lower rates compared to the estimated prevalence of central hypersomnolence disorders and ADHDs/ADDs. Most stimulants are prescribed by psychiatrists, closely followed by GPs. The increased odds for hospitalization to psychiatric institutions for stimulant receivers reflects the severity of disease and the higher psychiatric comorbidities in these patients.

Conflicts of Interest

None declared.

References

1. Polanczyk GV, Willcutt EG, Salum GA, Kieling C, Rohde LA. ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. Int J Epidemiol 2014 Apr;43(2):434-442. [doi: 10.1093/ije/dyt261] [Medline: 24464188]

2. Steinhausen HC, Metzke CW, Meier M, Kannenberg R. Prevalence of child and adolescent psychiatric disorders: the Zürich epidemiological study. Acta Psychiatr Scand 1998 Oct;98(4):262-271. [doi: 10.1111/j.1600-0447.1998.tb10082.x] [Medline: 9821446]
3. Estévez N, Eich-Höchli D, Dey M, Gmel G, Studer J, Mohler-Kuo M. Prevalence of and associated factors for adult attention deficit hyperactivity disorder in young Swiss men. PLoS ONE 2014;9(2):e89298. [doi: 10.1371/journal.pone.0089298] [Medline: 24586672]
4. Safer DJ. Recent trends in stimulant usage. J Atten Disord 2016 Jun;20(6):471-477. [doi: 10.1177/1087054715605915] [Medline: 26486603]
5. Board AR, Guy G, Jones CM, Hoots B. Trends in stimulant dispensing by age, sex, state of residence, and prescriber specialty - United States, 2014-2019. Drug Alcohol Depend 2020 Dec 1;217:108297. [doi: 10.1016/j.drugalcdep.2020.108297] [Medline: 32961454]
6. Barczyk ZA, Rucklidge JJ, Eggleston M, Mulder RT. Psychotropic medication prescription rates and trends for New Zealand children and adolescents 2008-2016. J Child Adolesc Psychopharmacol 2020 Mar;30(2):87-96. [doi: 10.1089/cap.2019.0032] [Medline: 31633377]
7. Steinhausen HC, Bisgaard C. Nationwide time trends in dispensed prescriptions of psychotropic medication for children and adolescents in Denmark. Acta Psychiatr Scand 2014 Mar;129(3):221-231. [doi: 10.1111/acps.12155] [Medline: 23738593]
8. Maier LJ, Haug S, Schaub MP. Prevalence of and motives for pharmacological neuroenhancement in Switzerland--results from a national Internet panel. Addiction 2016 Feb;111(2):280-295. [doi: 10.1111/add.13059] [Medline: 26189457]
9. Raman SR, Man KKC, Bahmanyar S, et al. Trends in attention-deficit hyperactivity disorder medication use: a retrospective observational study using population-based databases. Lancet Psychiatry 2018 Oct;5(10):824-835. [doi: 10.1016/S2215-0366(18)30293-1] [Medline: 30220514]
10. Gumy C, Huissoud T, Dubois-Arber F. Prevalence of methylphenidate prescription among school-aged children in a Swiss population: increase in the number of prescriptions in the Swiss Canton of Vaud, from 2002 to 2005, and changes in patient demographics. J Atten Disord 2010 Nov;14(3):267-272. [doi: 10.1177/1087054709356386] [Medline: 20484708]
11. Benson K, Flory K, Humphreys KL, Lee SS. Misuse of stimulant medication among college students: a comprehensive review and meta-analysis. Clin Child Fam Psychol Rev 2015 Mar;18(1):50-76. [doi: 10.1007/s10567-014-0177-z] [Medline: 25575768]
12. Charach A, Ickowicz A, Schachar R. Stimulant treatment over five years: adherence, effectiveness, and adverse effects. J Am Acad Child Adolesc Psychiatry 2004 May;43(5):559-567. [doi: 10.1097/00004583-200405000-00009]
13. Lakhan SE, Kirchgessner A. Prescription stimulants in individuals with and without attention deficit hyperactivity disorder: misuse, cognitive impact, and adverse effects. Brain Behav 2012 Sep;2(5):661-677. [doi: 10.1002/brb3.78] [Medline: 23139911]
14. Huber CA, Szucs TD, Rapold R, Reich O. Identifying patients with chronic conditions using pharmacy data in Switzerland: an updated mapping approach to the classification of medications. BMC Public Health 2013 Oct 30;13(1):1-10. [doi: 10.1186/1471-2458-13-1030] [Medline: 24172142]
15. Danielson ML, Claussen AH, Bitsko RH, et al. ADHD prevalence among U.S. children and adolescents in 2022: diagnosis, severity, co-occurring disorders, and treatment. J Clin Child Adolesc Psychol 2024;53(3):343-360. [doi: 10.1080/15374416.2024.2335625] [Medline: 38778436]
16. Thomas R, Sanders S, Doust J, Beller E, Glasziou P. Prevalence of attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. Pediatrics 2015 Apr;135(4):e994-1001. [doi: 10.1542/peds.2014-3482] [Medline: 25733754]
17. Thapar A, Cooper M. Attention deficit hyperactivity disorder. The Lancet 2016 Mar;387(10024):1240-1250. [doi: 10.1016/S0140-6736(15)00238-X]
18. Roskell NS, Setyawan J, Zimovetz EA, Hodgkins P. Systematic evidence synthesis of treatments for ADHD in children and adolescents: indirect treatment comparisons of lisdexamfetamine with methylphenidate and atomoxetine. Curr Med Res Opin 2014 Aug;30(8):1673-1685. [doi: 10.1185/03007995.2014.904772] [Medline: 24627974]
19. Karanges EA, Stephenson CP, McGregor IS. Longitudinal trends in the dispensing of psychotropic medications in Australia from 2009–2012: focus on children, adolescents and prescriber specialty. Aust N Z J Psychiatry 2014 Oct;48(10):917-931. [doi: 10.1177/0004867414538675] [Medline: 24927734]
20. Dauvilliers Y, Montplaisir J, Molinari N, et al. Age at onset of narcolepsy in two large populations of patients in France and Quebec. Neurology (ECronicon) 2001 Dec 11;57(11):2029-2033. [doi: 10.1212/wnl.57.11.2029] [Medline: 11739821]
21. Oberle D, Drechsel-Bäuerle U, Schmidtmann I, Mayer G, Keller-Stanislawski B. Incidence of narcolepsy in Germany. Sleep 2015 Oct 1;38(10):1619-1628. [doi: 10.5665/sleep.5060] [Medline: 25902804]
22. Zhang Z, Dauvilliers Y, Plazzi G, et al. Idling for decades: a european study on risk factors associated with the delay before a narcolepsy diagnosis. Nat Sci Sleep 2022;14:1031-1047. [doi: 10.2147/NSS.S359980] [Medline: 35669411]
23. van de Glind G, Van Emmerik ‐ van Oortmerssen K, Carpentier PJ, et al. The International ADHD in Substance Use Disorders Prevalence (IASP) study: background, methods and study population. Int J Methods Psych Res 2013 Sep;22(3):232-244 [FREE Full text] [doi: 10.1002/mpr.1397]

24. Strålin P, Hetta J. First episode psychosis and comorbid ADHD, autism and intellectual disability. Eur Psychiatry 2019 Jan;55:18-22. [doi: 10.1016/j.eurpsy.2018.09.007] [Medline: 30384107]
25. Jensen PS, Hinshaw SP, Kraemer HC, et al. ADHD comorbidity findings from the MTA study: comparing comorbid subgroups. J Am Acad Child Adolesc Psychiatry 2001 Feb;40(2):147-158. [doi: 10.1097/00004583-200102000-00009] [Medline: 11211363]
26. Marangoni C, De Chiara L, Faedda GL. Bipolar disorder and ADHD: comorbidity and diagnostic distinctions. Curr Psychiatry Rep 2015 Aug;17(8):604. [doi: 10.1007/s11920-015-0604-y] [Medline: 26084666]

Abbreviations

ADHD/ADD: attention-deficit/hyperactivity disorder/attention deficit disorder
GP: general practitioner
OR: odds ratio

Edited by A Mavragani, T Sanchez; submitted 25.10.23; peer-reviewed by H Demir, K Karimova; revised version received 29.09.24; accepted 30.09.24; published 07.01.25.

Please cite as:
Scharf T, Huber CA, Näpflin M, Zhang Z, Khatami R
Trends in Prescription of Stimulants and Narcoleptic Drugs in Switzerland: Longitudinal Health Insurance Claims Analysis for the Years 2014-2021
JMIR Public Health Surveill 2025;11:e53957
URL: https://publichealth.jmir.org/2025/1/e53957
doi:10.2196/53957

Multidimensional Evaluation of the Process of Constructing Age-Friendly Communities Among Different Aged Community Residents in Beijing, China: Cross-Sectional Questionnaire Study

Yingchun Peng[1*], PhD; Zhiying Zhang[2*], MS; Ruyi Zhang[3], MS; Yiyao Zhang[1], BS; Runying Wang[1], BS; Jiaying Zhang[1], BS; Shaoqi Zhai[1], MS; Qilin Jin[4], MS; Jiaojiao Zhou[5], MS; Jingjing Chen[6], MS

[1]School of Medical Humanities, Capital Medical University, No.10, Xitoutiao, You An Men Wai, Beijing, China
[2]Outpatient Office, Beijing Luhe Hospital, Capital Medical University, Beijing, China
[3]Ethics Committee Office, Beijing Ditan Hospital, Capital Medical University, Beijing, China
[4]Cardiac Surgery Department, People’s Hospital of Beijing Daxing District, Beijing, China
[5]Medical Department, Fengtai District Xiluoyuan Community Health Service Center, Beijing, China
[6]Huairou District Liulimiao Community Health Service Center, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Yingchun Peng, PhD
School of Medical Humanities, Capital Medical University, No.10, Xitoutiao, You An Men Wai, Beijing, China

Abstract

Background: The World Health Organization (WHO) has made significant efforts to promote age-friendly community initiatives (AFCI) to address the challenges of population aging. Previous studies have discussed the construction of age-friendly communities (AFC) in urban cities, evaluating AFCs often rooted in the WHO’s Checklist and focused on a single group, namely older adults, overlooking the role of other age groups in community development.

Objective: This study aims to evaluate AFCs from multidimensional aspects, particularly the positive living experiences of older adults, summarize the deficiencies in both hardware and software aspects in the process of constructing AFCs in China, and provide some recommendations to promote AFCIs worldwide.

Methods: Using a multistage sampling strategy, 470 community residents from urban and suburban areas participated in this study. A self-designed questionnaire was designed to use a standardized method to evaluate older adults’ living experiences across five dimensions, including the degree of age-friendliness in the community, social support, sense of gain, sense of happiness, and sense of security. Respondents rated each dimension on a 10-point scale. This study defined community residents into 3 groups: residents younger than 45 years(Group 1: youth), those aged 45-59 years (Group 2: middle-aged), and those aged ≥60 years (Group 3: old-age).

Results: In this study, 382 (81.3%) community residents were unaware of the relevant concepts of AFCs. Most participants highlighted the importance of community support and health services, followed by respect and social inclusion, and outdoor spaces and buildings. The findings showed that the highest-rated dimension was the sense of security. The mean scores for the degree of the sense of security in urban and suburban areas were 7.88 (SD 1.776) and 7.73 (SD 1.853), respectively. For Group 2, the mean scores were 7.60 (SD 2.070) and 8.03 (SD 1.662), while Group 3 had mean scores of 7.34 (SD 2.004) and 7.91 (SD 1.940). The lowest-rated dimension was social support; the mean scores for Group 1 for the degree of social support in urban and suburban areas were 7.63 (SD 1.835) and 7.48 (SD 1.918), respectively. For Group 2, the mean scores were 6.94 (SD 2.087) and 7.36 (SD 2.228), while those for Group 3 were 6.37 (SD 2.299) and 6.84 (SD 2.062). Further, there were significant differences in the scores of residents among different age groups in urban areas regarding age-friendliness (*P*<.001), social support, (*P*<.001), and sense of gain (*P*=.01).

Conclusions: China is in the early stages of developing AFCs. We further highlight the importance of continued research on the collaboration and participation among multiple stakeholders. These outcomes have a direct and positive impact on the well - being of older adults.

(JMIR Public Health Surveill 2025;11:e66248) doi:10.2196/66248

KEYWORDS

age-friendly; positive living experience; active aging; healthy aging; community residents; older adults; age-friendly communities; multiple stakeholders; evaluating age-friendliness; urban and suburban areas

Introduction

The History and Imperative of Implementing Age-Friendly Communities

Population aging is poised to become one of the most significant problems in the 21st century, impeding the development of many countries globally. Data from *World Populations Prospects* [1] predicated that the number of persons aged ≥65 years in countries and areas with populations that have already peaked is projected to reach 409 million by 2027, surpassing the number of children under 18. The number of persons aged ≥80 years is increasing at an even faster rate and is likely to triple, from 85 million in 2024 to approximately 267 million by 2100. With the increasing trend of population aging, health issues in older adults have become a major concern for scholars and policymakers worldwide. As people age, many experience a myriad of changes impacting their health and functional ability to manage independent living in their homes and communities [2]. Meanwhile, age-related changes such as physiological and cognitive decline may lead to negative health outcomes such as frailty, mobility limitations, and disability. Moreover, many social and physical environmental factors also affect the health status and quality of life of older adults [3,4]. Therefore, finding accessible and affordable measures to help older adults live longer and healthier will be crucial.

To address the challenges of population aging and achieve the goal of healthy aging, the World Health Organization(WHO) has made significant efforts to promote age-friendly community initiatives (AFCI). The concept of age-friendly communities (AFCs) originated from the Ecology Theory of Aging in 1973 [5,6], which indicated that aging was a dynamic process at both the individual and community levels, revealing that a positive surrounding environment can promote well-being among older adults. In 2000, the American Association of Retired Persons first defined AFCs as communities with affordable housing, adequate infrastructure and services, and diverse options of transportation. Based on the report *A Guide: Global Age-Friendly Cities* [7], the AFCs incorporated three aspects and eight key domains. Three aspects included service support, space environment, social and humanistic environment [8], and the eight key domains consisted of outdoor spaces and buildings, transportation, housing, social participation, respect and social inclusion, civic participation and employment, communication and information, and community and health services [9,10]. In the current conceptualization of AFCs, the importance of social environment and physical environment are placed equally as the determinants of health and older adults are considered to be effective social resources involved in community construction [11]. The AFCs are designed to account for the wide diversity among older people and promote their autonomy, inclusion, and contributions in all areas of community life [12].

Meanwhile, the implementation of aging-in-place policies highlights the imperative for AFCIs [13]. Supporting older adults to continue living safely and comfortably within their communities is a shared vision of many governments and international organizations. Due to their continuous efforts, the WHO Global Network for AFCIs currently includes 1705 cities and communities across 60 countries, covering over 330 million people worldwide. However, the lack of measurability of the actual age-friendliness of AFCs has been a major weakness in moving AFCIs forward. As age-friendliness is a complex, dynamic, multidimensional, and highly context-dependent concept, it does not easily lend itself to standardization of measurement [14-16].

Evaluating the Age-Friendliness of Communities

While constructing AFCs, it is crucial to evaluate the age-friendliness of communities [17]. The previous studies have mainly focused on measuring the age-friendliness of a city or community in which older adults live, based on the WHO’s Checklist [18]. Jo-Ying Huang et al [19] developed indicators of age-friendliness for communities in Taiwan province that conform to international standards by referring to the WHO’s Checklist and Taiwan’s existing indicators. Yu, Wong, and Woo [20] examined the relationships between perceptions of neighborhood environment, sense of community, and self-rated health by relying on the WHO Checklist. Wood et al [21] conducted a systematic scoping review of 23 articles using citizen science or participatory approaches. An inductive and deductive thematic analysis was performed to identify local urban barriers and facilitators and map them against the WHO’s Checklist. Kim et al [22] used data from the American Association of Retired Persons AFC surveys to evaluate the reliability and concurrent validity of AFC assessment tools. They found that higher levels of perceived availability of community and health services were associated with worse health outcomes.

Additionally, the Age Friendly Cities and Communities Questionnaire (AFCCQ) is designed by Dikken for measuring age-friendliness. Compared with WHO’s Checklist, the Age Friendly Cities and Communities Questionnaire only has 23 items, covers eight domains of the WHO’s AFC model, and has an additional ninth domain of financial situation [23]. In addition to the above-mentioned measurement tools of age-friendliness, sets of priority indicators, frameworks, and conceptual surveys are proposed by other studies. Jiravanichkul et al [24] identified minimum standard indicators and explored the development of a well-being environment and AFC assessment criteria by using the analytic hierarchy process. Ide et al [17] developed a community-level AFC indicator based on the WHO AFC guidelines by incorporating dementia-friendly elements and tested its validity and reliability.

Beijing as a Case Example

Beijing is the capital of China, and a municipality directly governed by the Central Government. It is famous as China’s political, cultural, international communication, and science and technology innovation center. Beijing currently has entered a moderately aging society. By the end of 2023, the permanent resident population aged 60 years and above was 4.948 million, accounting for 22.6% of the total permanent resident population. The resident population aged 65 years and above was 3.469 million, accounting for 15.9% of the total resident population [25]. To enhance older adults’ sense of gain, happiness, and security in communities, the National Health Commission (National Office for the Elderly, 2020) launched an initiative to construct AFCs across the country. By the end of 2023, a total of 93 communities in Beijing had been named “ National Model AFCs,” which were at the forefront of AFC development nationwide.

Focus of the Current Study

Although many scholars have attempted to evaluate the age-friendliness of communities in various ways, there is limited research emphasizing the positive aspects of health, such as the sense of gain, happiness, and security that older adults experience in their communities. Additionally, most studies have focused on the construction of AFCs in urban cities [26,27], and evaluating the age-friendliness of communities was only based on assessments from older adults [22,28], which overlooked the perspectives of other community residents. Therefore, the current evaluation of constructing AFCs may lack objectivity and accuracy.

This study aims to explore the cognition of AFCs among the different aged communities’ residents in Beijing, evaluate AFCs from multidimensional aspects, especially the positive living experience of older adults, summarize the deficiencies in hardware and software infrastructure in the process of constructing AFCs in Beijing, and provide a reference for promoting AFCIs across the world.

Methods

Study Design

This is a cross-sectional study, designed to understand the awareness and perception of AFCs among different aged community residents, explore deficiencies in the process of constructing AFCs, and propose recommendations for improvement.

Study Setting and Sample

Considering the level of economic development and construction status of AFCs among 16 districts in Beijing, a cross-sectional quantitative survey was conducted on community residents between urban and suburban areas. This study adopted a multistage sampling strategy [29,30]. In the first stage, Xicheng District, Fengtai District, Daxing District, and Huairou District were selected based on the functional orientation of the 16 districts in Beijing and the current situation of constructing AFCs. To implement urban strategic positioning, promoting sustainable development, and considering the practical needs of the historical pattern of the ancient capital, Beijing issued the Beijing Urban Master Plan in 2017. According to the Beijing Urban Master Plan, there are six central urban areas (including Dongcheng District, Xicheng District, Chaoyang District, Haidian District, Fengtai District, and Shijingshan District), five new towns in plain areas (including Shunyi District, Daxing District, Yizhuang District, Fangshan District, and Changping District), and six ecological conservation areas (including Mentougou District, Pinggu District, Huairou District, Miyun District, and Yanqing District). For this study, plain new towns and ecological conservation areas were unified as suburban areas, while the central urban areas were unified as urban areas. Among the 16 districts, Xicheng District, Fengtai District, Daxing District, and Huairou Districts ranked second, third, fifth, and eighth, respectively. In the second stage, 3 - 5 communities were randomly picked from each district, amounting to 15 communities in the list. In the last stage, trained investigators visited each sampling community to randomly invite community residents to participate in our investigation. Potential participants were recruited via the invitation of the investigators, and we sought to balance the sample based on population characteristics such as gender, age, education, and region. To better understand the construction of AFCs, our goal was to recruit residents who lived in the communities and who could freely and voluntarily express their insights.

Measures

A self-designed questionnaire was developed to use a standardized method for assessing the residents’ living experience and their community’s age-friendliness in Beijing. It was designed in consultation with a group of AFC experts [31].

To explore the cognition of community residents regarding the concept of AFCs, the questionnaire included questions such as whether they had heard of or learned about AFCs. Based on our literature review and WHO guidelines, we asked residents to choose the most important domain among the eight domains (outdoor spaces and buildings, transportation, housing, social participation, respect and social inclusion, civic participation and employment, communication and information, and community and health services) in the process of constructing AFCs.

The positive living experience of older adults that was obtained within their community was measured across five dimensions: (1) the degree of age-friendliness in the community, (2) the level of social support for older adults, (3) the sense of gain in community, (4) the sense of happiness in the community, and (5) the sense of security in the community. Respondents rated each dimension on a 10-point scale (0 - 1: does not exist, 2 - 3: poor, 4 - 5: fair, 6 - 7: good, 8 - 9: very good, 10: excellent).

To understand the shortcomings in the current process of constructing AFCs, the deficiencies were categorized into hardware and software aspects. Hardware aspects included increasing the green spaces and beautiful buildings, building more nursing institutions and medical institutions for older adults, building more leisure and entertainment places, installing more age-friendly facilities for older adults, and expanding traffic roads, among others. Software aspects consisted of conducting more abundant recreational activities, creating an age-friendly atmosphere, providing personalized elderly care services and community and neighborhood support, providing community volunteer services, and more personalized medical services. Participants were asked to choose the items that required urgent improvement in their communities.

Data Analysis

Based on the latest WHO age definition criteria [32], this study defined community residents into three age groups: (1) Group 1 as youth (younger than 45 years) (2) Group 2 as middle-aged adults (45-59 years), and (3) Group 3 as old-aged adults (above 60 years).

All the gleaned data were recorded into EpiData software (version 3.1; EpiData Association) and analyzed by SPSS software (version 21.0; IBM Corp). The mean and standard deviation were used to describe continuous data, while the categorical data were presented by composition ratio, frequency distributions, and parity arrangement. Frequency and rank analyses were used to summarize the quantitative data of community residents including demographic characteristics such as gender, age, region, and education level. The rank-sum test was used to analyze the most important domains of AFCs and deficiencies in the construction of the AFCs. One-way ANOVA was used to explore the five dimensions of the community living experiences of older adults.

Ethical Considerations

The study was approved by the Medical Ethics Committee of Capital Medical University, Beijing, China (Reference number Z2023SY048). All participants provided informed consent for the collection, handling, and storage of their personal and health data. All procedures were performed in accordance with relevant guidelines and regulations. All participants joined the study voluntarily, and no compensation was provided to them. All data were kept confidential, deidentified, and anonymous.

Results

A total of 477 residents voluntarily agreed to participate in our survey. We collected 470 valid questionnaires were collected, resulting in a validity rate of 98.53%. Table 1 presents the demographic characteristics of community residents. Among the 470 community residents in our survey, the majority were female (n=339, 72.1%). Approximately 44.5% (n=209) of the participants were younger than 44 years, and the average age of the 470 community residents was 46.74 (SD 18.63) years. A total of 259 (55.1%) community residents were from urban areas. Additionally, 113 participants (24%) had lived in their current community for more than 30 years and 178 (37.9%) participants had obtained a bachelor’s degree.

Table . Demographic characteristics of community residents.

Items	Surveys (N=470), n (%)
Gender	
Male	131 (27.9）
Female	339 (72.1）
Age (years）	
≤44	209 (44.5）
45-59	132 (28.1）
≥60	129 (27.4）
Regions	
Urban areas	259 (55.1）
Suburban areas	211 (44.9）
Length of residence (years）	
≤10	157 (33.4）
11-30	200 (42.6）
31-50	53 (11.3）
51-70	42 (8.9）
≥71	18 (3.8）
Education	
Junior high school or below	129 (27.4）
High school or Junior college	99 (21.1）
Bachelor’s degree	178 (37.9）
Master’s degree or above	64 (13.6）

Table 2 shows community residents’ cognition of AFCs. Almost 81.3% of the community residents were not aware of the concept of AFCs; only 88 (18.7%) community residents had heard of AFCs.

Table . Community residents' cognition of the Age-Friendly Communities.

Items	Results
Community residents' cognition of the concept of the Age-Friendly Communities (N=470), Surveys, n (%)	
Know	88 (18.7)
Not know	382 (81.3)
The most important domain of the Age-Friendly Communities from the perspective of community residents, total points[a] (rank)	
Outdoor spaces and buildings	453 (3)
Transportation	260 (6)
Housing	293 (5)
Social participation	364 (4)
Respect and social inclusion	559 (2)
Job opportunities and civic participation	84 (8)
Communication and information	135 (7)
Community support and health services	666 (1)

[a] Note: total points: number of people selected for the first important domain ×3 + number of people selected for the second important domain ×2 +number of people selected for the third important domain ×1.

With regard to the eight important domains of creating an AFC, community support and health services were the most important domains from the perspective of residents when they lived in the community, followed by respect and social inclusion, and outdoor spaces and buildings. Job opportunities and civic participation were considered relatively less important for residents when they lived in the community.

Table 3 displays an evaluation of the living experiences of older adults, categorized by different aged community residents. Among the five dimensions of older adults' living experience, the scores of the sense of security in the community were relatively high, whereas the scores for the degree of social support for older adults were relatively low. There were significant differences in the scores of residents between different age groups in urban areas in terms of age-friendliness ($P<.001$), social support ($P<.001$), and sense of gain ($P=.01$), whereas no significant differences were found in terms of happiness and security. In the evaluation of age-friendliness, there were significant differences between young and middle-aged residents, and between young residents and elderly residents. In the evaluation of the degree of social support, there were significant differences between young residents and middle-aged residents, and between young residents and elderly residents. Additionally, the sense of gain differed significantly between young residents and elderly residents.

Table . Evaluation of living experience of older adults from the perspectives of different aged community residents.

Regions, items, and groups	Scores[a] , mean (SD）	*F* test (*df*)	*P* value
Urban areas			
The degree of age-friendliness		9.687（2, 256）	<.001
Group 1（age <45 years)	7.82 (1.807)[a]		
Group 2 (age 45-60 years)	7.34 (1.689)[a]		
Group 3 （age ≥60 years)	6.58 (2.208)[b]		
The degree of social support		8.914（2, 256）	<.001
Group 1（age <45 years)	7.63 (1.835)[a]		
Group 2 (age 45-60 years)	6.94 (2.087)[b]		
Group 3 (age ≥60 years)	6.37 (2.299)[b]		
The sense of gain in community		4.445（2, 256）	.01
Group 1 (age <45 years)	7.58 (1.693)[a]		
Group 2 (age 45-60 years)	7.10 (1.995)[ab]		
Group 3 (age ≥60 years)	6.79 (1.880)[b]		
The sense of happiness in community		2.741（2, 256）	.07
Group 1 (age <45 years)	7.72 (1.777)		
Group 2 (age 45-60 years)	7.27 (1.933)		
Group 3 (age ≥60 years)	7.09 (2.079)		
The sense of security in community		1.792（2, 256）	.17
Group 1 (age <45 years)	7.88 (1.776)		
Group 2 (age 45-60 years)	7.60 (2.070)		
Group 3 (age ≥60 years)	7.34 (2.004)		
Suburban areas			
The degree of age-friendliness		1.290（2, 208）	.28
Group 1 (age <45 years)	7.61 (1.914)		
Group 2 (age 45-60 years)	7.44 (2.232)		
Group 3 (age ≥60 years)	7.05 (1.910)		
The degree of social support		1.732（2, 208）	.18
Group 1 (age <45 years)	7.48 (1.918)		
Group 2 (age 45-60 years)	7.36 (2.228)		
Group 3 (age ≥60 years)	6.84 (2.062)		
The sense of gain in community		0.138（2, 208）	.87
Group 1 (age <45 years)	7.29 (1.923)		
Group 2 (age 45-60 years)	7.41 (2.045)		
Group 3 (age ≥60 years)	7.21 (2.145)		
The sense of happiness in community		0.519(2, 208）	.60
Group 1 (age <45 years)	7.53 (1.719)		
Group 2 (age 45-60 years)	7.51 (2.081)		
Group 3 (age ≥60 years)	7.82 (1.724)		
The sense of security in community		0.561（2, 208）	.57

Regions, items, and groups	Scores[a] , mean (SD）	*F* test (*df*)	*P* value
Group 1 (age <45 years)	7.73 (1.853)		
Group 2 (age 45-60 years)	8.03 (1.662)		
Group 3 (age ≥60 years)	7.91 (1.940)		

[a] Different letters within the same row indicate statistically significant differences between groups based on the Student-Newman-Keuls (SNK) multiple comparisons (*P*<.05). Groups sharing the same letter do not differ significantly.

Table 4 presents the shortcomings in the current process of constructing AFCs. Regarding the aspects of community hardware facilities, most residents expressed the need for building more nursing institutions for older adults, followed by building additional medical institutions and installing more suitable age-friendly facilities. Expanding traffic roads was an aspect that concerned fewer people. Regarding the community software facilities, creating an age-friendly atmosphere was most commonly reported by many community residents, followed by providing personalized elderly care services, more personalized medical services, and more recreational activities.

Table . The shortcomings in the current process of constructing age-friendly communities.

Items	Frequency	Rank
The aspects of community hardware facilities that should be improved		
Increase the green area and beautiful buildings	217	5
Build more nursing institutions for older adults	319	1
Build more medical institutions for older adults	257	2
Build more leisure and entertainment places	221	4
Install more suitable facilities for older adults	252	3
Expand traffic roads	78	6
Others	11	7
The aspects of community software facilities that should be improved		
Hold more abundant recreational activities	207	4
Create an age-friendly atmosphere	323	1
Provide personalized elderly care services	317	2
Provide community and neighborhood support	117	6
Provide community volunteer services	170	5
Provide more intimate medical services	241	3
Others	13	7

Discussion

Principal Findings

This study is the first to comprehensively evaluate AFCs from multidimensional aspects in China. It particularly examines the positive living experiences that residents derive from their communities, compares the cognition and perception of AFCs among different aged community residents in urban and suburban areas, and explores the deficiencies in the process of constructing AFCs from both hardware and software perspectives. The community residents' suggestions for improvement for constructing AFCs reflect the needs and interests of older adults and other age groups, providing a direction for constructing AFCs in the future.

Increasing Age-Friendly Awareness and Sense of Belonging Among Community Residents

China has only recently begun promoting AFCIs; compared to other developed countries, there remains significant progress to be made. Consequently, the public is not aware of the concepts of AFCs, which is consistent with our findings as only 88 (18.7%) residents were aware of AFCs. Recognizing the age-friendliness of communities is the premise to engage residents in the construction of AFCs. Relevant studies show that perceived age-friendliness of the community is positively associated with a sense of belonging to the community [33,34]. However, a gap exists between family belonging and community belonging. Given the changes in China's economic and social structures, the sense of community belonging has been weakened [35], which may have led to a lack of initiative and consciousness among residents in the process of creating AFCs. Therefore, it is vital to improve the consciousness of age-friendliness and strengthen the sense of belonging among community residents in the process of constructing AFCs.

Establishing an Age-Friendly Atmosphere and Improving Communities' Software Facilities

Previous studies have shown that constructing AFCs mainly focuses on eight domains, based on the WHO guideline [14,36]; however, little is known about which domains are considered most important for community residents. In our study, most participants highlighted the importance of community support and health services; they wish to build more medical institutions and provide more personalized medical services in their communities, which reflects the high demand for medical services by community residents in Beijing. However, a major challenge faced by both high-income and low-income countries today is inequitable access to health care resources [37-39]. For example, in suburban or medically underprivileged communities, residents may often be unable to access desired health care services. Telemedicine in China has been improving the dissemination of high-quality health care resources from urban cities to remote suburban areas, providing services such as teleconsultations and specialist diagnoses [37]. Meanwhile, respect and social inclusion are also key components for promoting AFCIs. Many community residents identified that creating an age-friendly atmosphere was the most important step in the process of AFCs to improve the aspects of community software facilities, which is consistent with other studies. Lui et al [13] noted that a supportive environment characterized by positive relations, engagement, and inclusion is a core prerequisite for aging well. Rémillard et al [40] have explored eleven case studies of AFCs worldwide and reported that the need to shift perception, change mindsets, and promote a more positive view of aging were identified as key priorities. Respect and promoting social participation of older adults is also seen as a way to challenge ageism. In addition, the range of recreational activities and social support for community residents within their communities should be increased. Some studies found a positive association between employment, social participation, and healthy aging [41]. These observations could be due to the potential of these activities to enhance individuals' social status, promote psychological well-being, and foster a sense of dignity.

Improving the Living Environment and Strengthening Communities' Hardware Facilities

Our results indicate that many community residents wish to improve the physical infrastructure of their communities; outdoor spaces and buildings were the third most important domain of AFCs. This finding is in line with the reality of communities' infrastructure in China. Most older communities in Beijing lack elevators, which causes major inconvenience for older adults when going out. In addition, there is limited public space for residents to participate in activities and exercise. A common phenomenon in many old communities is that public spaces meant for residents' activities are occupied by parked cars. As shown in Figure 1, limited public space, scarcity of age-appropriate fitness equipment, and the absence of elevators are key physical environment factors that may hinder older adults from participating in community activities.

Figure 1. Overview of hardware facilities in a surveyed community in the Daxing District, showing limited public space, scarcity of age-appropriate fitness equipment, and no elevators.

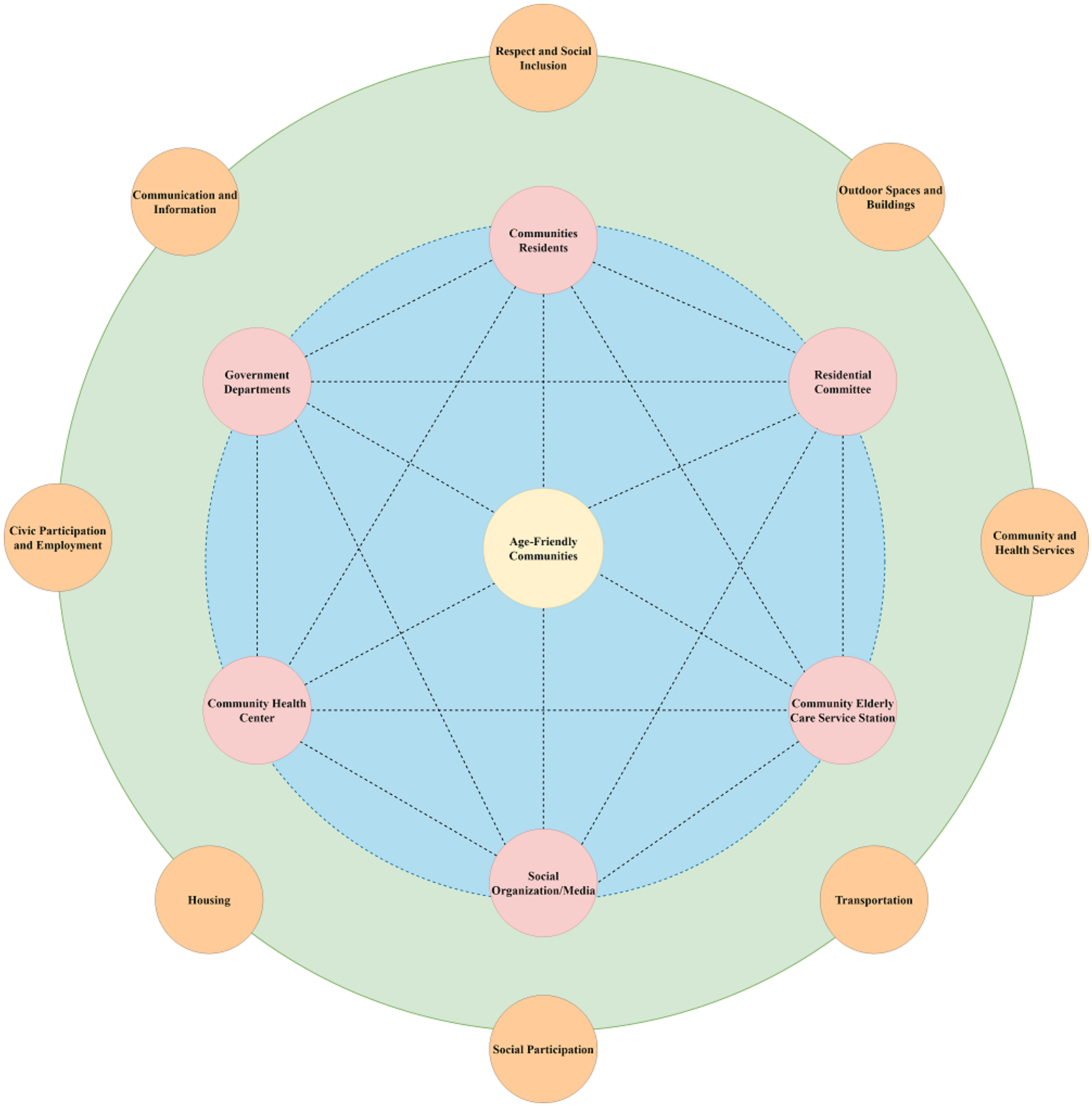
Actively Participating in Community Activities to Enhance Community Residents' Happiness

In the process of constructing AFCs, the notion of healthy aging should be considered the final goal. Older adults in good health can continue to be vital societal resources, experience greater job satisfaction and independence in their lives, and engage more actively as community residents. This study aimed to evaluate the positive living experiences that older adults obtained in their communities. Our findings show that among the five dimensions assessed, the sense of security scored the highest, while social support scored the lowest, particularly among urban older adults. Social support is a well-recognized social determinant of health, and it is obvious that obtaining social support from others is crucial throughout life, including old age [42]. Most older adults consistently prefer aging in place, which requires a high level of social support in the process of constructing AFCs [11]. Therefore, improving the degree of social support for older adults should be a key focus while constructing AFCs in the future. Further, there were significant differences in the evaluation of older adults' living experiences among different aged community residents in urban areas. Young residents rated the degree of social support, age-friendliness, sense of gain, and happiness significantly higher than older adults' self-evaluations. This disparity may stem from younger residents having greater physical capacity and healthier habits, which enables them to afford expenses related to transportation, health care, and retirement [41]. Consequently, they have more energy and capacity to actively engage in paid work, social participation, and environmental preservation, thereby fostering their sense of happiness and sense of gain. Besides, elderly residents in suburban areas have the highest scores for the sense of happiness, which suggests that older adults in suburban communities experience higher levels of happiness compared to their urban elderly residents. These findings may be attributed to both the physical environment (eg, accessible public facilities) and the social environment (eg, active engagement in volunteer activities), providing sufficient support to older adults in suburban areas [41]. Suburban seniors typically lived in the rural society for a long time. At the heart of rural society lies a deep-rooted connection to the land. Life in these communities is often dictated by the cycles of nature, with agricultural activities and environmental stewardship forming the backbone of their livelihoods. Therefore, compared with urban elderly residents, suburban seniors may engage in agricultural activities within their capacity, establish trust-based relationships with neighbors, and be more likely to participate in communities' activities.

Participation of Multiple Stakeholders in the Promotion of Constructing AFCs

The task of constructing AFCs is complex, dynamic, and multistage, which involves multiple stakeholders' cooperation. To address the challenges of an aging population and promote AFCIs, collaboration is essential between government departments, residential committees, community health centers, community elderly care service stations, social organizations, media, and community residents (Figure 2).

Figure 2. Constructing age-friendly communities as a central goal involves collaboration and participation from multiple stakeholders, including government departments, residential committees, community health centers, community elderly care service stations, social organizations, media, and community residents. Effective cooperation among these stakeholders is necessary to expand the reach of constructing age-friendly communities across all eight domains.



Government departments should focus on constructing AFCs, by formulating and improving relevant laws and regulations on AFCs, designing scientific and accurate tools for evaluating AFCs, and establishing long-term supervision and feedback mechanisms. Residential committees should first design medium- and long-term plans for promoting AFCIs based on three overarching areas [43]: (1) physical infrastructure (eg, outdoor spaces and buildings, transportation, housing); (2) social opportunities (eg, civic participation, social participation, respect, and social inclusion); and (3) supportive services (eg, communication and information, community and health services). Residential committees should conduct a comprehensive assessment of their respective communities based on AFC policies and tools developed by the government, and identify deficiencies within communities. Our findings show that community residents have urgent needs for old-age care and medical services. Therefore, promoting continuous and integrated community care that combines medical and old-age care is necessary. Additionally, residential committees should explore smart community care services with the help of information technology, treating the community as the basic unit. Besides, it is important to build a smart service platform that connects “Internet+community home care”, with community health centers and elderly care service stations, that can establish a connection between the elderly needs and community service resources. Finally, residential committees should create an atmosphere of age-friendliness within the communities, build more spaces where different aged residents can interact together, and encourage more older adults to participate in the construction of AFCs. Social organizations should adopt age-friendly practices or provide more activities catering to the needs of older adults. Social media should strengthen public awareness of AFCs, ensuring more residents understand the concept of AFCs and their benefits and actively participate in the construction of AFCs. Community residents should also actively cooperate with the work of the communities, improve their cognition of AFCs, and make suggestions for the construction of AFCs.

Limitations

The findings generated from this study should be interpreted considering several limitations. First, our survey included healthy and cognitively intact older adults, while those who were suffering from disabilities or dementia were excluded. However, understanding the living experience of this excluded group would be vital to help construct AFCs. Second, the representativeness of our study was limited, since we surveyed only a sample from 4 out of 16 districts in Beijing. Future studies should consider conducting similar studies in smaller cities across China and including more related subjects. Finally, our study followed a voluntary participation principle when conducting the questionnaire survey. Female respondents showed more willingness to participate in the survey, resulting in the gender ratio of female respondents (over 70%). Additionally, many surveyed residents never left their living communities, and they did not have a deep understanding of other communities. Future work should further explore the influence of gender, education level, and length of residence on community residents’ perceptions of AFCs.

Conclusions

China is in the early stages of becoming an age-friendly society. Findings from this study show that community residents in Beijing have a low level of cognition of AFCs, and face many barriers from the perspective of hardware facilities and software facilities. Therefore, it is imperative to take measures to (1) increase age-friendly awareness and sense of belonging among community residents; (2) establish an age-friendly atmosphere and improve communities’ software facilities; (3) improve the living environment and strengthen communities’ hardware facilities; (4) increase the active participation of older persons in the community to enhance their happiness; and (5) and promote AFCIs through joint efforts of multiple stakeholders, including government departments, residential committees, community health centers, community elderly care service stations, social organizations, media, and community residents. Our study provides suggestions for tangible central government policy and practice initiatives and resource allocation. These outcomes have a direct and positive impact on the well‐being of older adults.

Acknowledgments

The author would like to thank all the participants, experts, and researchers who participated in this study. This study was funded by the Beijing Social Science Foundation Project (Funding Number 22SRBO08). The funding organization had no further role in the study design, data collection and analysis, interpretation of the data, writing the paper, or the decision to submit the paper for publication.

Data Availability

Data are available upon reasonable request. Transcripts will not be shared for online access to protect the anonymity of the participants. Readers who wish to gain access to the data can write to the corresponding author.

Authors' Contributions

Conceptualization: YP, ZZ

Data curation: ZZ, RZ, SZ, JZ[1],YP

Formal analysis: YZ, RW, JZ[2], YP, ZZ
Methodology: YP, ZZ
Writing – original draft: YP, ZZ

Writing – review & editing: YP, ZZ, RZ, YZ, RW, JZ[1], SZ, JZ[2], QJ, JC

Conflicts of Interest

None declared.

References

1. World population prospects 2024: summary of results. United Nations. 2024. URL: https://desapublications.un.org/publications/world-population-prospects-2024-summary-results [accessed 2025-01-01]
2. Stenholm S, Westerlund H, Salo P, et al. Age-related trajectories of physical functioning in work and retirement: the role of sociodemographic factors, lifestyle and disease. J Epidemiol Community Health 2014 Jun;68(6):503-509. [doi: 10.1136/jech-2013-203555] [Medline: 24534071]
3. Hong A, Welch-Stockton J, Kim JY, Canham SL, Greer V, Sorweid M. Age-friendly community interventions for health and social outcomes: a scoping review. Int J Environ Res Public Health 2023 Jan 31;20(3):36767920. [doi: 10.3390/ijerph20032554] [Medline: 36767920]
4. Fang ML, Sixsmith J, Hamilton-Pryde A, et al. Co-creating inclusive spaces and places: towards an intergenerational and age-friendly living ecosystem. Front Public Health 2022;10(996520). [doi: 10.3389/fpubh.2022.996520] [Medline: 36685002]
5. Lawton MP, Nahemow L. Ecology and the aging process. In: The Psychology of Adult Development and Aging: American Psychological Association; 1973:619-674. [doi: 10.1037/10044-020]
6. Menec VH, Means R, Keating N, Parkhurst G, Eales J. Conceptualizing age-friendly communities. Can J Aging 2011 Sep;30(3):479-493. [doi: 10.1017/S0714980811000237] [Medline: 21745427]
7. Global age-friendly cities: a guide. World Health Organization. 2007. URL: https://www.who.int/publications/i/item/9789241547307 [accessed 2025-01-02]
8. Lai S, Zhou Y, Yuan Y. Associations between community cohesion and subjective wellbeing of the elderly in Guangzhou, China-a cross-sectional study based on the structural equation model. Int J Environ Res Public Health 2021 Jan 22;18(3):953. [doi: 10.3390/ijerph18030953] [Medline: 33499175]
9. Kim K, Buckley TD, Burnette D, Huang J, Kim S. Age-friendly communities and older adults' health in the United States. Int J Environ Res Public Health 2022 Jul 29;19(15):9292. [doi: 10.3390/ijerph19159292] [Medline: 35954648]
10. Hsu HC. Associations of city-level active aging and age friendliness with well-being among older adults aged 55 and over in Taiwan. Int J Environ Res Public Health 2020 Jun 16;17(12):4304. [doi: 10.3390/ijerph17124304] [Medline: 32560170]
11. Jeste DV, Blazer DG 2nd, Buckwalter KC, et al. Age-friendly communities initiative: public health approach to promoting successful aging. Am J Geriatr Psychiatry 2016 Dec;24(12):1158-1170. [doi: 10.1016/j.jagp.2016.07.021] [Medline: 27742528]
12. National programmes for age-friendly cities and communities: a guide. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/9789240068698 [accessed 2025-01-01]
13. Lui CW, Everingham JA, Warburton J, Cuthill M, Bartlett H. What makes a community age-friendly: a review of international literature. Australas J Ageing 2009 Sep;28(3):116-121. [doi: 10.1111/j.1741-6612.2009.00355.x] [Medline: 19845650]
14. Dikken J, van den Hoven RFM, van Staalduinen WH, Hulsebosch-Janssen LMT, van Hoof J. How older people experience the age-friendliness of their city: development of the age-friendly cities and communities questionnaire. Int J Environ Res Public Health 2020 Sep 20;17(18):6867. [doi: 10.3390/ijerph17186867] [Medline: 32962226]
15. Measuring the age-friendliness of cities: a guide to using core indicators. World Health Organization. 2015. URL: https://www.who.int/publications/i/item/9789241509695 [accessed 2025-01-01]
16. Meeks S. Age-friendly communities: introduction to the special issue. Gerontologist 2022 Jan 14;62(1):1-5. [doi: 10.1093/geront/gnab163] [Medline: 35029660]
17. Ide K, Jeong S, Tsuji T, et al. Suggesting indicators of age-friendly city: social participation and happiness, an ecological study from the JAGES. Int J Environ Res Public Health 2022 Apr 22;19(9):5096. [doi: 10.3390/ijerph19095096] [Medline: 35564490]
18. Checklist of essential features of age-friendly cities. World Health Organization. 2007. URL: https://extranet.who.int/agefriendlyworld/wp-content/uploads/2018/04/Age-Friendly-Checklist-WHOedit.pdf [accessed 2025-01-01]
19. Huang JY, Hsu HC, Hsiao YL, et al. Developing indicators of age-friendliness in Taiwanese communities through a modified Delphi method. Int J Environ Res Public Health 2022 Nov 4;19(21):14430. [doi: 10.3390/ijerph192114430] [Medline: 36361310]
20. Yu R, Wong M, Woo J. Perceptions of neighborhood environment, sense of community, and self-rated health: an age-friendly city project in Hong Kong. J Urban Health 2019 Apr;96(2):276-288. [doi: 10.1007/s11524-018-00331-3] [Medline: 30511137]
21. Wood GER, Pykett J, Daw P, et al. The role of urban environments in promoting active and healthy aging: a systematic scoping review of citizen science approaches. J Urban Health 2022 Jun;99(3):427-456. [doi: 10.1007/s11524-022-00622-w] [Medline: 35587850]
22. Kim K, Buckley T, Burnette D, Kim S, Cho S. Measurement indicators of age-friendly communities: findings from the AARP age-friendly community survey. Gerontologist 2022 Jan 14;62(1):e17-e27. [doi: 10.1093/geront/gnab055] [Medline: 33909074]
23. van Hoof J, Marston HR, Kazak JK, Buffel T. Ten questions concerning age-friendly cities and communities and the built environment. Build Environ 2021 Jul;199:107922. [doi: 10.1016/j.buildenv.2021.107922]
24. Jiravanichkul S, Pinich S, Sreshthaputra A, Jarutach T. The development of a well-being environment and age-friendly communities assessment criteria using the analytic hierarchy process: a case of Thailand. NJEDP 2024 Dec;23(3):416. [doi: 10.54028/NJ202423416]

25. Beijing municipal bureau of statistics [Chinese]. Beijing Statistical YearBook. 2024. URL: https://nj.tjj.beijing.gov.cn/nj/main/2024-tjnj/zk/indexch.htm [accessed 2025-01-01]
26. Buckner S, Mattocks C, Rimmer M, Lafortune L. An evaluation tool for age-friendly and dementia friendly communities. Work Older People 2018;22(1):48-58. [doi: 10.1108/WWOP-11-2017-0032] [Medline: 30166937]
27. Loos E, Sourbati M, Behrendt F. The role of mobility digital ecosystems for age-friendly urban public transport: a narrative literature review. Int J Environ Res Public Health 2020 Oct 14;17(20):33066528. [doi: 10.3390/ijerph17207465] [Medline: 33066528]
28. Molina-Martínez M, Marsillas S, Sánchez-Román M, Del Barrio E. Friendly residential environments and subjective well-being in older people with and without help needs. Int J Environ Res Public Health 2022 Nov 28;19(23):15832. [doi: 10.3390/ijerph192315832] [Medline: 36497900]
29. Wu D, Wang Y, Lam KF, Hesketh T. Health system reforms, violence against doctors and job satisfaction in the medical profession: a cross-sectional survey in Zhejiang Province, Eastern China. BMJ Open 2014 Dec 31;4(12):e006431. [doi: 10.1136/bmjopen-2014-006431] [Medline: 25552614]
30. Zhang T, Wang X. Association of continuity of general practitioner care with utilisation of general practitioner and specialist services in China: a mixed-method study. Healthcare (Basel) 2021 Sep 13;9(9):1206. [doi: 10.3390/healthcare9091206] [Medline: 34574980]
31. Plouffe L, Kalache A. Towards global age-friendly cities: determining urban features that promote active aging. J Urban Health 2010 Sep;87(5):733-739. [doi: 10.1007/s11524-010-9466-0] [Medline: 20549569]
32. Nemoto Y, Nonaka K, Kuraoka M, et al. Effects of intergenerational contact on social capital in community-dwelling adults aged 25-84 years: a non-randomized community-based intervention. BMC Public Health 2022 Sep 24;22(1):1815. [doi: 10.1186/s12889-022-14205-6] [Medline: 36153514]
33. Choi YJ. Understanding aging in place: home and community features, perceived age-friendliness of community, and intention toward aging in place. Gerontologist 2022 Jan 14;62(1):46-55. [doi: 10.1093/geront/gnab070] [Medline: 34043782]
34. Chaudhury H, Oswald F. Advancing understanding of person-environment interaction in later life: one step further. J Aging Stud 2019 Dec;51(100821):100821. [doi: 10.1016/j.jaging.2019.100821] [Medline: 31761094]
35. Lai S, Lu L, Zhou Z, et al. The effects of family physician-contracted service on health-related quality of life and equity in health in China. Int J Equity Health 2021 Jan 6;20(1):15. [doi: 10.1186/s12939-020-01348-4] [Medline: 33407523]
36. Menec VH, Nowicki S. Examining the relationship between communities' "age-friendliness" and life satisfaction and self-perceived health in rural Manitoba, Canada. Rural Remote Health 2014;14(2594):24437338 [FREE Full text] [Medline: 24437338]
37. Gao J, Fan C, Chen B, et al. Telemedicine Is becoming an increasingly popular way to resolve the unequal distribution of healthcare resources: evidence from China. Front Public Health 2022;10(916303):916303. [doi: 10.3389/fpubh.2022.916303] [Medline: 35874991]
38. Jiang J, Wang P. Health status in a transitional society: urban-rural disparities from a dynamic perspective in China. Popul Health Metr 2018 Dec 27;16(1):22. [doi: 10.1186/s12963-018-0179-z] [Medline: 30591053]
39. Ismail M. Regional disparities in the distribution of Sudan's health resources. East Mediterr Health J 2020 Sep 24;26(9):1105-1114. [doi: 10.26719/emhj.20.056] [Medline: 33047802]
40. Rémillard-Boilard S, Buffel T, Phillipson C. Developing age-friendly cities and communities: eleven case studies from around the world. Int J Environ Res Public Health 2020 Dec 27;18(1):133. [doi: 10.3390/ijerph18010133] [Medline: 33375466]
41. Cheng Y, Chen ZL, Wei Y, Gu N, Tang SL. Examining dynamic developmental trends: the interrelationship between age-friendly environments and healthy aging in the Chinese population-evidence from China Health and Retirement Longitudinal Study, 2011-2018. BMC Geriatr 2024 May 15;24(1):429. [doi: 10.1186/s12877-024-05053-7] [Medline: 38750429]
42. Gudlavalleti MVS, John N, Allagh K, et al. Access to health care and employment status of people with disabilities in South India, the SIDE (South India Disability Evidence) study. BMC Public Health 2014 Nov 1;14(1125):1125. [doi: 10.1186/1471-2458-14-1125] [Medline: 25361570]
43. Scharlach AE. Aging in context: individual and environmental pathways to aging-friendly communities-the 2015 Matthew A. Pollack award lecture. Gerontologist 2017 Aug 1;57(4):606-618. [doi: 10.1093/geront/gnx017] [Medline: 28460019]

Abbreviations

AFC: Age-Friendly Communities
AFCI: Age-Friendly Communities Initiatives
WHO: World Health Organization

Edited by A Mavragani; submitted 07.09.24; peer-reviewed by JK Kazak, Z Chen; revised version received 27.01.25; accepted 28.01.25; published 21.03.25.

Please cite as:
Peng Y, Zhang Z, Zhang R, Zhang Y, Wang R, Zhang J, Zhai S, Jin Q, Zhou J, Chen J
Multidimensional Evaluation of the Process of Constructing Age-Friendly Communities Among Different Aged Community Residents in Beijing, China: Cross-Sectional Questionnaire Study
JMIR Public Health Surveill 2025;11:e66248
URL: https://publichealth.jmir.org/2025/1/e66248
doi:10.2196/66248

Migrant-Local Differences in the Relationship Between Oral Health, Social Support, and Loneliness Among Older Adults in Weifang, China: Cross-Sectional Study

Hui Liu[1,2,3*], MPH; Jieru Wang[1,2,3,4*], MPH; Rui Chen[1,2,3], MPH; Xixing Xu[1,2,3], MPH; Mingli Pang[1,2,3], MPH; Kaiyuan Feng[1,2,3], MPH; Bingsong Li[1,2,3], MPH; Qinling Li[1,2,3], MPH; Ziwei Qin[1,2,3], MPH; Shuyi Yan[1,2,3], MPH; Nabila Ibn Ziyat[1,2,3], MPH; Fanlei Kong[1,2,3], PhD

[1] Department of Social Medicine and Health Management, School of Public Health, Cheeloo College of Medicine, Shandong University, 44 Wenhuaxi Road, Jinan, China

[2] NHC Key Lab of Health Economics and Policy Research, Shandong University, Jinan, China

[3] Center for Health Management and Policy Research, Shandong University (Shandong Provincial Key New Think Tank), Jinan, China

[4] Editorial Office of Chinese Rural Health Service Administration of Publishing Center, Anhui Medical University, Hefei, China

[*] these authors contributed equally

Corresponding Author:
Fanlei Kong, PhD
Department of Social Medicine and Health Management, School of Public Health, Cheeloo College of Medicine, Shandong University, 44 Wenhuaxi Road, Jinan, China

Abstract

Background: Increased aging and accelerated urbanization have led to the migration of older adults within China. Migrant older adults (MOAs) may experience physical and psychological discomfort in influx cities, and they are a vulnerable group that has emerged in the course of fast urbanization. Previous studies have confirmed the association between oral health and loneliness as well as the relationship between social support and loneliness; however, no research has been done to clarify the underlying mechanisms and the migrant-local difference between oral health, social support, and loneliness.

Objective: This study aimed to test the association between oral health, social support, and loneliness among Chinese older adults, as well as the migrant-local difference on the above relationship.

Methods: Multistage cluster random sampling was used to enroll a total of 1205 participants, including 613 MOAs and 592 local older adults (LOAs). Loneliness was assessed by the 6-item short-form UCLA Loneliness Scale; oral health was measured via the Chinese version Geriatric Oral Health Assessment Index (GOHAI); social support was evaluated by the Social Support Rating Scale (SSRS). Descriptive analysis, χ^2 tests, and *t* tests were conducted. Multigroup structural equation modeling (SEM) was employed to clarify the migrant-local difference on the association between oral health, social support, and loneliness among MOAs and LOAs.

Results: The mean score of loneliness was 8.58 (SD 3.032) for MOAs and 8.00 (SD 2.790) for LOAs. Oral health and social support were found to be negatively related to loneliness among MOAs and LOAs; the standardized direct effects for MOAs were −0.168 and −0.444 (*P*<.001), and they were −0.243 and −0.392 (*P*<.001) for LOAs, respectively. Oral health generated a direct positive effect on social support, and the direct effect was 0.186 for MOAs (*P*<.001) and 0.247 for LOAs (*P*<.001).

Conclusions: Loneliness was fairly low among older adults in Weifang, China, while MOAs showed higher loneliness than LOAs. Oral health had both direct and indirect negative effects on loneliness among MOAs and LOAs, with no significant path differences between MOAs and LOAs. Social support was found to be negatively associated with loneliness for both MOA and LOA, while the association was stronger among MOAs than LOAs. Oral health exerted a significantly positive effect on social support for both MOAs and LOAs, while no significant difference existed between them. Measures should be taken by the government, society, and families to increase social support, improve oral health, and further reduce loneliness among MOAs and LOAs.

(JMIR Public Health Surveill 2025;11:e66061) doi:10.2196/66061

KEYWORDS

loneliness; oral health; social support; migrant older adults; local older adults

Introduction

China has one of the fastest growing aging populations in the world [1]. According to the data of China's Seventh National Census, the number of people aged 60 years or above was 264.02 million, accounting for 18.70% (the number of people aged 65 years or above was 190.64 million, accounting for 13.50%). Compared with 2010, the proportion of people aged 60 years or above increased by 5.44% [2]. As the population ages, concern regarding older adults' health status has also increased.

In the past few decades, China has experienced a rapid increase in economic level and an acceleration of urbanization, which also caused an economic gap across different cities and may further lead to population migration. Due to the traditional Chinese culture, which highlights the family union, more and more older adults move to live with their adult children. Existing studies refer to the older adults who leave their hometown to migrate with their children as migrant older adults (MOAs) [3]. Previous studies found that MOAs had difficulties in developing social networks and social integration, which would further affect their mental health [4,5], life satisfaction [6], and quality of life [7]. Based on "adaptive theory" [8] and "social support buffering model" [9], MOAs face increased adaptation stress and social isolation, which may lead to higher reliance on social support networks to cope with the challenges posed by migration. Thus, it is also important to pay more attention to the health status of MOAs, in addition to that of local older adults (LOAs).

Loneliness is defined by Weiss [10] as a subjective feeling formed when individuals perceive that they lack satisfying interpersonal relationships and there is a gap between their desire to connect and the actual situation of connection. Donovan et al [11] found that 17.6% of US older adults felt lonely much of the time during the preceding week. Guo et al's [12] nationwide study found that 24.3% of older people reported they were lonely in China. Another study found that nearly one-fourth of Chinese older adult participants felt lonely [13]. In terms of MOAs, one longitudinal study showed that migrants from non–English-speaking countries reported higher levels of loneliness, as compared with the native-born, non-Indigenous Australians [14]. Some studies indicated that loneliness was associated with some negative health outcomes [15,16], such as poor physical health [17] and poor mental health [18,19]. Moreover, another national study in China clarified that a high degree of loneliness could further reduce the life satisfaction of older adults [20]. A cohort study from northern and southern Europe showed that loneliness was associated with a decreased quality of life among older adults [21]. Therefore, it is important to pay more attention to older adults' loneliness and find possible ways to reduce their loneliness.

Oral health has been defined by the World Health Organization (WHO) as a state of being free from chronic mouth and facial pain, oral and throat cancer, oral infection and sores, periodontal (gum) disease, tooth decay, tooth loss, and other diseases and disorders that limit an individual's capacity in biting, chewing, smiling, speaking, and psychosocial well-being [22]. The Fourth Chinese Oral Health Epidemiological Survey Report demonstrated that the prevalence of dental caries in older adults' permanent teeth was fairly high, implying serious oral health conditions [23]. Oral health, as a part of general health, is an important determinant of loneliness. One study examined the relationship between oral health and mental health issues and found that oral health was negatively associated with depression [24]. However, few studies have examined the association between oral health and loneliness. A study among English older adults demonstrated that oral health–related quality of life was identified as an independent risk factor for loneliness both cross-sectionally and longitudinally, and older adults with oral impacts had a significantly higher risk of being lonely than their counterparts without any oral impacts [25]. A previous study also found that the fewer number of teeth, the more likely both Japanese and British older adults were to report social isolation or loneliness [26]. Studies also clarified that the degenerative changes of oral function could lead to eating difficulties and imbalanced nutritional intake, which in turn affects physical and mental health and may increase the loneliness of older adults [27,28]. Thus, it is important to test the empirical association between oral health and loneliness among older adults.

Social support refers to the material and moral help provided by various parties in society, including family, relatives, friends, colleagues, organizations, and trade unions [29]. A meta-analysis has pointed out that increased social support would be supportive of successful aging among older adults [30]. Existing research also found that social support could not only provide the necessary resources (such as financial help and emotional help) for older adults to cope with challenges, but also had a huge positive impact on their physical and mental health [30,31]. Conversely, previous studies clarified that a lack of social support would cause adverse outcomes, such as limitations on activities of daily living [32] and poor quality of life [33]. The association between social support and loneliness has been widely explored. A previous study pointed out that social support was negatively associated with loneliness among older adults [34]. Chung and Kim [35] found that social support was critical in lowering loneliness in middle-aged and older adults. Studies from different countries indicated that the changes in older people's social networks and social support resulted in a significant onset of loneliness during the COVID-19 pandemic [36,37].

A previous study indicated that social support was related to older adults' oral health [38]. A nationwide study in Britain demonstrated social support was associated with the oral health status and oral health behavior of older people [39]. A study conducted in Germany clarified that impaired oral health–related quality of life was positively correlated with lower social support among older seniors [40]. A study had shown that oral condition was closely related to facial appearance, which can affect social image and confidence and negatively affect relationships [41]. A cross-national comparative study found that oral health was strongly associated with social isolation, with poorer oral health accelerating social isolation and thus reducing individuals' perceived social support [26]. Therefore, oral health has a critical impact on social support in older adults, and it is essential to study the effects of both in MOAs and LOAs.

Based on the above literature review, no study was found that determined the association between oral health, social support, and loneliness, and no study has ever compared the difference between MOAs and LOAs regarding the above relationship. Thus, this study aimed to (1) clarify the association between oral health, social support, and loneliness and (2) test whether a statistically significant difference exists for the above relationship.

Methods

Sample Collection

The data were collected in Weifang City, Shandong Province, China in August 2021. The gross domestic product of Weifang City was 701.06 billion Chinese yuan in 2021 (US $96.9 billion at 2025 conversion rate) [42]. Weifang governed 10 districts and 2 counties (59 subdistricts and 59 towns) until July 2020 [43]. The total population of Weifang City was 9.3 million by the end of 2020 according to the Seventh National Census. In 2020, nearly 2.38 million of its whole population constituted migrants from other counties, cities, or provinces, with a variety of sociodemographic and cultural backgrounds [44]. Thus, two groups of older adults were recruited in this study. For MOAs, the inclusion criteria were: (1) aged ≥60 years; (2) their Hukou is not in the present place (Hukou is one of China’s oldest tools for population control; it is essentially a household registration permit, similar to an internal passport, which defines where people are registered, not where they currently reside [45]); and (3) an ability and willingness to communicate with surveyors. For LOAs, the inclusion criteria were (1) and (3) from the previous list.

Multistage cluster random sampling was conducted to select the samples of MOAs and LOAs. In the first stage, 4 of 12 districts were selected as the primary sampling units (PSUs), considering the economic development and geographic location. In the second stage, 1 subdistrict was selected from each district (PSU), and a total of 4 subdistricts were taken as the secondary sampling units (SSUs). In the last stage, 4 communities were chosen from the SSUs as the tertiary sampling units (TSUs); that is, 1 community was selected from each of the subdistricts chosen previously. All the MOAs as well as the LOAs who met the above criteria constituted the total study sample. Initially, 616 MOAs and 592 LOAs were selected for interviews. However, 3 MOAs were excluded for answering the questionnaire incorrectly or incompletely. Ultimately, 613 MOAs and 592 LOAs were included in the database. The detailed sample selection process is shown in Multimedia Appendix 1.

Measurement

Sociodemographic Characteristics

The section on sociodemographic characteristic traits included the following: gender (man, woman); age group (60 - 64, 65 - 69, 70 - 74, 75 - 79, ≥80 years); Hukou (rural, urban); marital status (married, single); education level (illiterate, primary school, junior high school, high school and above); pension (Yes, No).

Oral Health

The Chinese version of the Geriatric Oral Assessment Index (GOHAI) was used to measure participants’ oral health status. This tool is primarily used to assess the self-reported oral health of older adults and is widely used in China [46]. The Chinese version of the GOHAI is divided into 12 items and 3 subdimensions designed to assess different aspects of oral health: (1) physical functioning (four items), (2) psychosocial functioning (five items), and (3) pain or discomfort (three items). GOHAI scores could be divided into the following 3 categories: (1) 50 and below are defined as low oral health, (2) 51 - 56 as fair-to-moderate oral health, and (3) 57 - 60 as high oral health. In the previous study, the GOHAI scores also had good reliability and validity [47]. In this study, the Cronbach α coefficient was 0.853, indicating that the scale had good reliability.

Social Support

The Social Support Rating Scale (SSRS) was used to measure the social support of MOAs and LOAs, including 10 types of support: friends, residents, neighbors, colleagues, family members, financial, comfort, conversation, help, and activities [48]. The social support scale has 10 items, including 3 dimensions of objective support (3 items), subjective support (4 items), and utilization of social support (3 items). The range of the total score of the scale is 12 - 66. A higher total social support score means the subject received more social support. The SSRS has been proven to have good reliability and validity in practice and has been widely used in China [49,50]. The Cronbach α coefficient was 0.822, implying this scale had good reliability.

Loneliness

The 6-item version of the UCLA Loneliness Scale (ULS-6) was used to assess the loneliness of participants, which excluded 2 reverse-scored items from the 8-item UCLA Loneliness Scale. The scale measures loneliness caused by the discrepancy between the level of desired social engagement and that which is actually experienced. The options are on a Likert scale: 1=never, 2=rarely, 3=sometimes, and 4=always, with a total score ranging from 6 - 24. Previous studies have demonstrated good reliability and validity in a Chinese population [51,52]. In this study, the Cronbach α coefficient of the ULS-6 was 0.82, indicating the scale had good reliability.

Statistical Analysis

This study used SPSS (version 24.0; IBM Corp) and AMOS (version 24.0; IBM Corp) to perform the data analysis.

The samples’ sociodemographic characteristics in this study were characterized by using descriptive statistics, including frequency (%) and mean (SD). The difference between sociodemographic characteristics of MOAs and LOAs was determined using the χ^2 test, while the *t* test was used to explore the difference in oral health (3 dimensions), life satisfaction (5 items), and loneliness (6 questions) of MOAs and LOAs. *P* values less than .05 were considered statistically significant. The above analyses were performed using SPSS (version 24.0; IBM Corp).

A hypothesized structural equation model was set to analyze the relationship between oral health, social support, and loneliness among MOAs and LOAs in Shandong Province. The maximum likelihood estimation method was used to evaluate the hypothesized model's fit. The structural equation modeling (SEM) process model consisted of 2 categories of variables: latent variables and observed variables. The latent variables were oral health, life satisfaction, and loneliness. The 3 observed variables of oral health included physical functioning, psychosocial functioning, and pain and discomfort. The 3 observed variables of social support were objective support, subjective support, and utilization of social support. The 6 observed variables of loneliness were the 6 items of the ULS-6. All SEM analyses were performed using AMOS (version 24.0; IBM Corp).

The following model fitness indexes were used to judge the fit of the hypothesized model: CMIN (Chi-square value, χ^2), degrees of freedom (*df*), *P* value of the χ^2 test, root mean square error of approximation (RMSEA), the goodness-of-fit index (GFI), the adjusted goodness-of-fit index (AGFI), and the comparative fit index (CFI). In this study, the models would be regarded to be well-fitted when $P>.05$, GFI>0.90, AGFI>0.90, and RMSEA<0.05 [53]. *P* value is easily influenced by sample size under many conditions, so it was only demonstrated in this study and not used as a criterion for judgment [54].

In the multigroup analysis, various parameters are restricted to find the most suitable path model. Five models were displayed in this study, namely M1 (MOA model), M2 (LOA model), M3 (unconstrained model), M4 (measurement weights model), and M5 (structural weights model). M1 and M2 were models fitted based on the sample data of two groups, and M3-M5 were obtained by adding conditions gradually restricted from the initial unconstrained model [55]. The multigroup model invariance was determined before the discussion of the MOA and LOA difference in the structural model of SEM. The change in CFI (ΔCFI) and the change in RMSEA (ΔRMSEA) were used to assess the measurement invariance between unconstrained and constrained multigroup analyses [55]. ΔCFI is independent of both model complexity and sample size, as well as the overall fit measurements. ΔCFI<0.010 indicates that we obtained measurement invariance across groups [56]. For the ΔRMSEA, with more than 300 samples, ΔRMSEA less than 0.015 implies that measurement invariance has been successfully established [57]. After the multigroup model invariance test was passed, we determined whether there were path differences between the different groups based on the model outputs, and paths with an absolute value of the critical ratio greater than 1.96 indicate a significant difference in the coefficients between the two groups ($P<.05$) [58].

Ethical Considerations

The research program was reviewed and approved by the ethical committee of Shandong University (number 20180225). For the original data collected, all participants had given informed consent to our study and were well aware of their right to withdraw from the study at any time. Our data have been completely anonymized and there is no information to identify the participants.

Results

Sample Characteristics

Table 1 shows the demographic characteristics of the participants, with a total of 1205 older adults included in the data analysis, of which 613 were MOAs and 592 were LOAs. Overall, 885 (73.4%) of the total sample were women while 320 (26.6%) were men; more than half (64.7%) of participants belonged to the 60 - 69 year old group; nearly half of older adults had a rural Hukou; 971 of 1205 participants (80.6%) were married; approximately four-fifths of older adults (n=975, 80.6%) were educated; and 785 (65.1%) older adults have a pension.

Table . Characteristics of participants and disparity between MOAs[a] and LOAs[b].

Variables		Total (n=1205), n (%)	MOA (n=613), n (%)	LOA (n=592), n (%)	Chi-square (*df*)	*P* value
Gender					0.083 (1)	.79
	Men	320 (26.6)	165 (26.9)	155 (26.1)		
	Women	885 (73.4)	448 (73.1)	437 (73.8)		
Age group (years)					139.631 (4)	<.001
	60 - 64	436 (36.2)	271 (44.2)	165 (27.9)		
	65 - 69	344 (28.5)	215 (35.1)	129 (21.8)		
	70 - 74	175 (14.5)	76 (12.4)	99 (16.7)		
	75 - 79	105 (8.7)	27 (4.4)	78 (13.2)		
	≥80	145 (12)	24 (3.9)	121 (20.4)		
Hukou					507.268 (1)	<.001
	Rural	649 (53.9)	525 (85.6)	124 (20.9)		
	Urban	556 (46.1)	88 (14.4)	468 (79.1)		
Marital status					43.045 (1)	<.001
	Married	971 (80.6)	539 (87.9)	432 (73)		
	Single	234 (19.4)	74 (12.1)	160 (27)		
Education level					41.675 (3)	<.001
	Illiterate	230 (19.1)	161 (26.2)	69 (11.7)		
	Primary school	403 (33.4)	185 (30.2)	218 (36.8)		
	Junior high school	338 (28)	158 (25.8)	180 (30.4)		
	High school and above	234 (19.4)	109 (17.8)	125 (21.1)		
Pension					106.504 (1)	<.001
	Yes	785 (65.1)	314 (51.2)	471 (79.6)		
	No	420 (34.9)	299 (48.8)	121 (20.4)		

[a] MOAs: migrant older adults.

[b] LOAs: local older adults.

The disparity between MOAs and LOAs was statically significant for age (*P*<.001), Hukou (*P*<.001), marital status (*P*<.001), education level (*P*<.001), and pension (*P*<.001). Specifically, nearly four-fifths of MOA participants (486/613, 79.3%) were aged 60 - 69 years, while less than 50% (294/592, 48.9%) of LOAs were in that age group. In total, 525 (85.6%) of MOAs were rural Hukou while 468 (79.1%) of LOAs were urban Hukou; in addition, there were more than twice as many single LOAs than single MOAs (n=160, 27% vs n=74, 12.1%); over one-fourth of MOAs (n=161, 26.2%) were illiterate, while approximately one-tenth of LOAs (n=69, 11.7%) were illiterate. Finally, 471 (79.6%) LOAs had a pension while only 121 (51.2%) MOAs had one.

Oral Health, Social Support, and Loneliness of the Participants

Table 2 illustrates the general characteristics of participants' oral health, life satisfaction, and loneliness, and the difference between MOAs and LOAs for the above variables. The total scores of GOHAI, SSRS, and ULS-6 for MOAs were 54.95 (SD 6.469), 38.89 (SD 6.629), 8.58 (SD 3.032), and 54.40 (SD 7.024), 39.51 (SD 6.856), and 8.00 (SD 2.790) for LOAs. Statistical differences between MOAs and LOAs were found in total ULS-6 score (t_{1203}=3.442, *P*=.001), SSRS score (objective support*:* t_{1203}=4.545, *P*<.001; subjective support*:* t_{1203}=−3.608, *P*<.001), and GOHAI score (psychosocial function*:* t_{1203}=2.028, *P*=.04). It is noted that there were no statistically significant differences in Q4 (t_{1203}=1.760, *P*=.08) and Q6 (t_{1203}=1.265, *P*=.21) of the total score of loneliness between MOAs and LOAs.

Table . General characteristics of the loneliness, social support, and oral health of MOA[a] and LOA[b] participants.

Variables	Total (n=1205), mean (SD)	MOA (n=613), mean (SD)	LOA (n=592), mean (SD)	*t* test (*df*=1203)	*P* value
Loneliness (ULS-6 [c])					
Total	8.29 (2.929)	8.58 (3.032)	8.00 (2.790)	3.442	.001
Often feel a lack of friends	1.45 (0.757)	1.53 (0.789)	1.36 (0.714)	3.704	<.001
Feel no one can be trusted	1.42 (0.722)	1.47 (0.728)	1.37 (0.713)	2.326	.02
Often feel left out	1.32 (0.587)	1.37 (0.622)	1.28 (0.547)	2.672	.008
Feel separated from others	1.34 (0.654)	1.37 (0.670)	1.30 (0.636)	1.760	.08
Often feel shy	1.31 (0.583)	1.36 (0.636)	1.25 (0.518)	3.365	.001
Surrounded but no one cares	1.46 (0.714)	1.49 (0.720)	1.43 (0.707)	1.265	.21
Social support (SSRS [d])					
Total	39.20 (6.746)	38.89 (6.629)	39.51 (6.856)	−1.612	.11
Objective support	8.22 (2.005)	8.47 (1.636)	7.95 (2.297)	4.545	<.001
Subjective support	23.94 (4.627)	23.47 (4.789)	24.43 (4.404)	−3.608	<.001
Utilization of support	7.04 (2.347)	6.94 (2.257)	7.14 (2.435)	−1.432	.15
Oral health (GOHAI [e])					
Total	54.68 (6.750)	54.95 (6.469)	54.40 (7.024)	1.408	.16
Physical function	17.14 (3.567)	17.35 (3.442)	16.93 (3.683)	2.028	.04
Psychological function	24.05 (2.217)	24.10 (2.061)	23.99 (2.369)	0.832	.41
Pain and discomfort	13.49 (2.162)	13.50 (3.115)	13.48 (2.211)	0.196	.84

[a] MOA: migrant older adults.

[b] LOA: local older adults.

[c] ULS-6: 6-item short-form UCLA Loneliness Scale.

[d] SSRS: Social Support Rating Scale.

[e] GOHAI: Geriatric Oral Health Assessment Index.

Structural Equation Model

Measurement Invariance Across Migration Status

Table 3 shows 5 selected models, which revealed related fit statistics of the measurement invariance across migration status and the fitness indexes. The fitness indexes of MOAs and LOAs should be compared to check whether the variable “migration state” was suitable for the group comparison.

In this study, χ^2, *df*, *P* value, GFI, AGFI, CFI, and RMSEA were the fitness indexes. As shown in Table 3, the fitness indexes of the MOAs were GFI=0.969, AGFI=0.952, CFI=0.971, and RMSEA=0.045 (M1), while for the LOAs, they were GFI=0.964, AGFI=0.944, CFI=0.940, and RMSEA=0.051 (M2). All fitness indexes showed values over 0.90 and very slight differences between the MOA and LOA groups, implying that we could further compare the differences between the MOA and LOA groups with the other models. Although the RMSEA value of M2 was more than 0.05, these variables were mainly used to calculate the change of RMSEA, not to assess the model’s fitness. Then, the ΔCFI and ΔRMSEA between M3 (unconstrained model), M4 (measurement weights model), and M5 (structural weights model) were used to evaluate the measurement invariance. The M3 did not restrict any coefficient in the model, the M4 assumed the indicator loadings for the corresponding construct of each group are equal, and the M5 constrained the indicator loadings of the corresponding construct and the structural coefficients between the groups.

As seen in Table 3, the ΔCFI between M4 and M3 was 0.002; between M5 and M4, it was 0. All of the ΔCFI values were less than 0.010, indicating that measurement invariance was established between the models of M1, M2, M3, M4, and M5 between the MOA and LOA groups. The ΔRMSEA between M4 and M3 was 0.001, and it was 0 between M5 and M4. All of the ΔRMSEA values were less than 0.015, also indicating that measurement invariance was established between the models of M1, M2, M3, M4, and M5 across the MOA and LOA groups.

Table . Multigroup model invariance. The variables in 5 models were oral health, social support, and loneliness among MOAs[a] and LOAs[b] (n=1205).

Model	Chi-square (*df*)	*P* value	Chi-square/*df*	GFI[c]	AGFI[d]	CFI[e]	RMSEA[f]	ΔCFI[g]	ΔRMSEA[h]
M1[i]	112.152 (50)	<.001	2.243	0.969	0.952	0.971	0.045	—[j]	—
M2[k]	127.660 (50)	<.001	2.553	0.964	0.944	0.940	0.051	—	—
M3[l]	239.812 (100)	<.001	2.398	0.967	0.948	0.967	0.034	—	—
M4[m]	255.289 (109)	<.001	2.682	0.965	0.950	0.965	0.033	0.002	0.001
M5[n]	258.639 (112)	<.001	2.309	0.964	0.950	0.965	0.033	0.000	0.000

[a] MOAs: migrant older adults.

[b] LOAs: local older adults.

[c] GFI: goodness of fit index.

[d] AGFI: adjusted goodness of fit index.

[e] CFI: comparative fitness index.

[f] RMSEA: root mean square error of approximation.

[g] ΔCFI: change of CFI.

[h] ΔRMSEA: change of RMSEA.

[i] M1: MOA model.

[j] Not applicable.

[k] M2: LOA model.

[l] M3: unconstrained model.

[m] M4: measurement weights model.

[n] M5: structural weights model.

Model Fitness Indexes

Figures 1 and 2 show the proposed models for MOAs and LOAs, respectively, which contained 3 variables: oral health, social support, and loneliness. Table 3 demonstrates the model fitness indexes for variable models (M1=MOA, M2=LOA). The MOA and LOA groups both had the same estimated value for model fitness: GFI=0.967, AGFI=0.948, CFI=0.967, and RMSEA=0.034. All fitness indexes implied that the theoretical model perfectly matched the empirical data for both the MOA and LOA groups.

Figure 1. Structural equation modeling analysis of the association between oral health, social support, and loneliness of MOAs (n=613). AGFI: adjusted goodness of fit index; CFI: comparative fitness index; CMIN: chi-square value; e: residual variables; GFI: goodness of fit index; MOA: migrant older adults; RMSEA: root mean square error of approximation. All parameter estimates were statistically significant (*P*<.05).

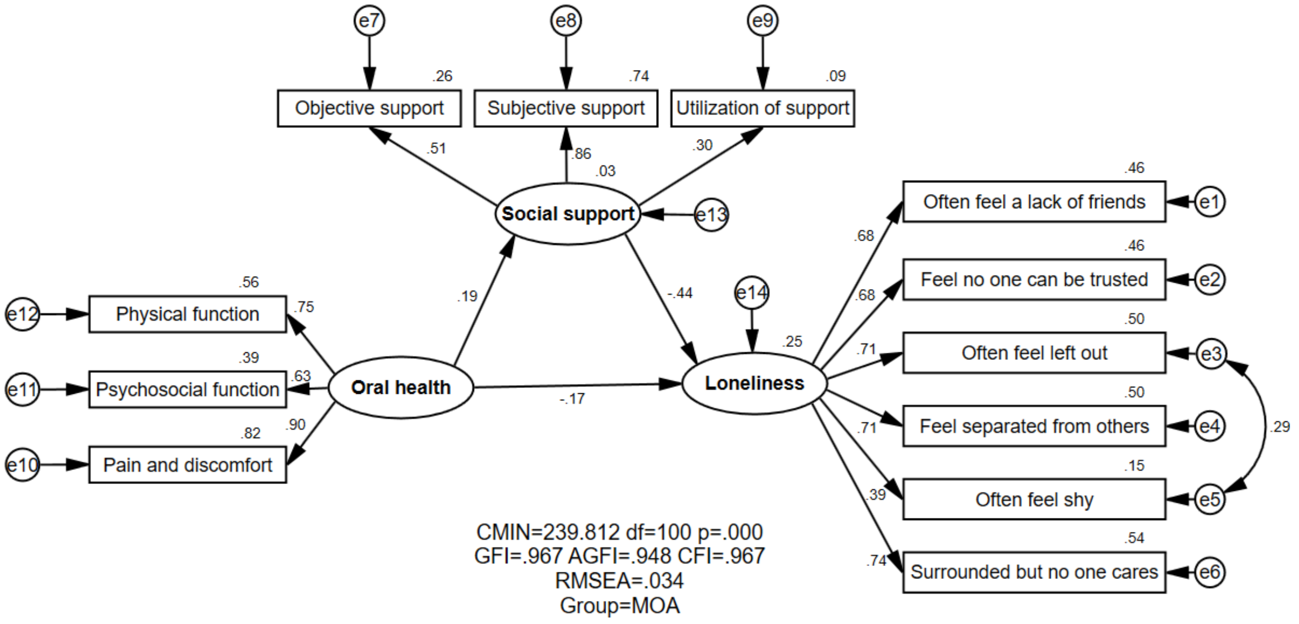
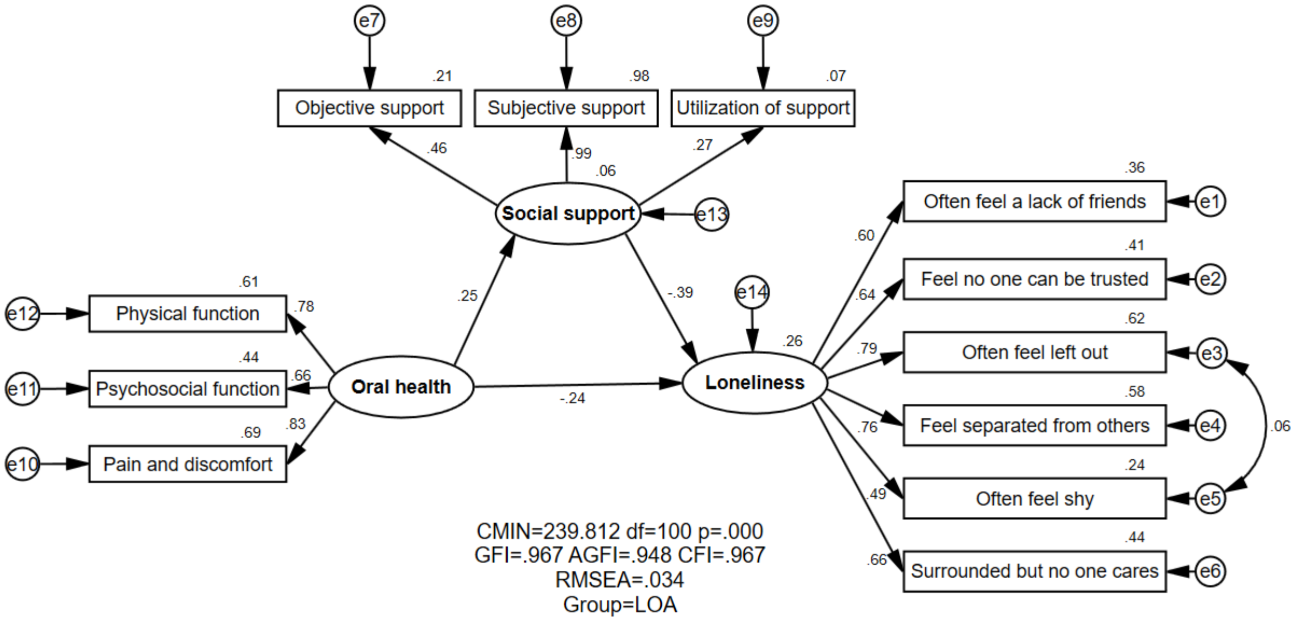


Figure 2. Structural equation modeling analysis of the association between oral health, social support, and loneliness of LOAs (n=592). AGFI: adjusted goodness of fit index; CFI: comparative fitness index; CMIN: chi-square value; e: residual variables; GFI: goodness of fit index; LOA: local older adult; RMSEA: root mean square error of approximation. All parameter estimates were statistically significant (*P*<.05).



Relationship Between Oral Health, Social Support, and Loneliness Assessed by SEM

Association Between Oral Health and Loneliness of Participants

The association between oral health, social support, and loneliness was shown in Figures 1 and 2 and Table 4. A negative and direct association was observed between oral health and loneliness among the MOAs (standardized direct effect=−0.168) and LOAs (standardized direct effect=−0.243). Moreover, oral health could exert a negative effect on loneliness indirectly via social support (standardized indirect effect=−0.083 for MOAs; standardized indirect effect=−0.097 for LOAs). It is noted that oral health was negatively associated with loneliness, which meant that MOAs and LOAs with higher oral health would generally have lower loneliness. A statistically significant relationship between oral health and loneliness was found in both the MOA and LOA groups.

Table . Standardized effects between oral health, social support, and loneliness among MOAs[a] and LOAs[b].

Path	Direct effect		Indirect effect		Total effect		Difference (critical ratio)
	MOA	LOA	MOA	LOA	MOA	LOA	
Oral health → Loneliness	−0.168[c]	−0.243[c]	−0.083[c]	−0.097[c]	−0.251[c]	−0.340[c]	−0.538
Oral health → Social support	0.186[c]	0.247[c]	—[d]	—	0.186[c]	0.247[c]	1.145
Social support → Loneliness	−0.444[c]	−0.392[c]	—	—	−0.444[c]	−0.392[c]	2.741[e]

[a] MOAs: migrant older adults.

[b] LOAs: local older adults.

[c] *P*<.001.

[d] Not applicable.

[e] *P*<.01.

Association Between Social Support and Loneliness of Participants

As for the relationship between social support and loneliness, a negative and direct effect was demonstrated among both MOAs (standardized direct effect=−0.444) and LOAs (standardized direct effect=−0.392), which meant less social support among both MOAs and LOAs would generally indicate higher loneliness. Concerning the group difference, a significantly negative correlation was slightly stronger in the MOA group than in the LOA group (critical ratio=2.741, *P*<.01).

Association Between Oral Health and Social Support of Participants

Oral health had a positive and direct effect on social support for both MOAs and LOAs (standardized direct effect=0.186 for MOAs; standardized direct effect=0.247 for LOAs), indicating that the higher the oral health of MOAs and LOAs, the higher their social support. It was found that a statistically significant relationship existed between oral health and social support among MOAs and LOAs.

Discussion

Principal Findings

This study examined the severity of loneliness as well as the association between social support, oral health, and loneliness among older adults in Weifang. The results further showed a statistical difference in loneliness between MOAs and LOAs; the empirical associations between oral health, social support, and loneliness (including the local-migrant difference) were also clarified.

Loneliness Among MOAs and LOAs

The mean score of loneliness among MOAs (8.58) and LOAs (8.00) was lower than in a previous study conducted among rural empty-nest older adults in China (16.19) [51], indicating a lower level of loneliness among the MOAs and LOAs in this study. Moreover, this study found loneliness in MOAs was higher than in LOAs, which was similar to a study that showed that immigrant groups were lonelier than older adults born in Canada [59]. This may be due to the fact that, because of MOAs' migration, they need more time to adapt to their new environment while LOA have been in a familiar environment for a long time.

Association Between Oral Health and Loneliness

A negative association between oral health and loneliness was found among both MOAs and LOAs, which was similar to one existing study, which reported that older people who had a poor oral health status had higher odds of experiencing loneliness [60]. Ma and Chen [27] found that masticatory function, swallowing function, tooth loss, tooth function, and toothache were the influencing factors of loneliness among Chinese older people in the community. Another cross-sectional study in Indonesia showed similar findings, where older adults who had a poor oral status had a higher chance of feeling lonely [61]. Some studies pointed out that poor oral health increases psychological stress in communication among older adults, consequently limiting social interaction with others and causing loneliness [62,63]. However, the results showed that the effect of oral health on loneliness was not statistically different between MOAs and LOAs. This may be due to the fact that the relationship between oral health and loneliness is generalizable across older populations [25]. Another possible reason is that when facing oral health problems, both MOAs and LOAs will adopt different coping styles to decrease their loneliness. Some studies suggested that people with higher socioeconomic status and health literacy tend to proactively utilize medical resources to mitigate the negative impacts of oral problems [64]. In contrast, MOAs had poorer oral health services in their hometowns [65] and have experienced and adapted to more oral problems, which were more common and socially acceptable in their lives [66]. These further result in fewer psychological changes and lower loneliness even though oral problems occurred in the inflow cities.

Association Between Social Support and Loneliness

A negative relationship between social support and loneliness was found among both MOAs and LOAs, indicating that older adults with higher social support could reduce their loneliness. This was consistent with a previous nationwide cohort study, which also showed that social support decreased the odds of loneliness incidence among older adults in China [67]. The effect of social support on loneliness could be explained through the interaction theory, which views loneliness as a response to a lack of satisfying social networks and attachment partners [68]. Moreover, this study showed that the negative association between social support and loneliness was stronger among the MOAs than the LOAs. The economic development gap between urban and rural areas led to higher medical and education services in urban areas, which further resulted in population mobility mainly from rural to urban areas in China [69,70]. As for the MOAs, their migration may lead to a decline in the quantity and quality of social connections [71], which may further result in higher levels of loneliness among them than in the LOAs.

Association Between Oral Health and Social Support

The SEM results illustrated that oral health and social support were positively correlated, implying the better the oral health status, the higher the social support. This was consistent with previous findings, which suggested the experience of oral pain was associated with physical discomfort [72] and social barriers [73]. In addition, people with poor oral health may face additional mental health challenges due to problems such as halitosis, creating social anxiety and exacerbating barriers to interaction [74]. The results also noted that no significant difference was found between LOAs and MOAs when examining the impact of oral health on loneliness. This meant that the gradual loss of teeth and changes in facial morphology in older adults as they age, whether migrating or not, could lead to similar social limitations and increased psychological stress, which can affect their social participation and social support.

Implications

In order to reduce older adults' loneliness, the following measures should be taken. First, this study found that MOAs had higher loneliness than LOAs. Hence, strengthening social integration and promoting equality may be beneficial for reducing the loneliness of MOAs. Second, the results indicated that loneliness was negatively correlated with oral health; thus, it is important to aid older adults in maintaining their oral health. The government could speed up the process of including oral health services in health insurance reimbursements for older patients and enhance oral health education for them. Third, the findings showed that high social support would also reduce the loneliness of older adults. In particular, for MOAs, their social networks will be affected and they may face challenges such as language barriers and cultural differences after migration, making it more urgent for them to implement effective measures to alleviate their loneliness. Therefore, there is a need to enhance social support through the efforts of family members, communities, and the government. For instance, the government could implement policies to support young people whose families include older adults since older adults may receive financial and emotional support from their children. It is necessary to promote the equalization of social security and welfare benefits for older adults who migrate with their families and further enhance their belonging in the inflow city.

Limitations

Some limitations existed in this study and should be addressed in future research. First, this study used cross-sectional data, therefore causal relationships could not be made and a longitudinal design is needed for follow-up research. Second, oral health status was assessed by a self-report scale with a lack of clinical evidence. In the future, clinical oral health examinations could be used to evaluate the oral health of older adults. Third, in addition to social support, there may be other variables that exert an indirect effect on loneliness among older adults and the variables used in this study may also be influenced by other confounding factors; therefore, more research is needed to verify their association.

Conclusion

Loneliness levels were fairly low among older adults in Weifang, China, while MOAs showed higher loneliness than LOAs. Oral health had both direct and indirect negative effects on loneliness among MOAs and LOAs, with no significant path differences between MOAs and LOAs. Social support was found to be negatively associated with loneliness for both MOAs and LOAs, while the association was stronger among MOAs than LOAs. Oral health exerted a significantly positive effect on social support for both MOAs and LOAs, while no significant difference existed between LOAs and MOAs.

Acknowledgments

The research team greatly appreciates the funding support obtained and thanks the research participants for their cooperation and support. This study was supported and funded by the National Natural Science Foundation of China (number 71804094), China Postdoctoral Science Foundation (number 2016M592161), Natural Science Foundation of Shandong Province (number ZR2016GB02), Postdoctoral Science Foundation of Shandong Province (number 201603021), and Fundamental Research Funds of Shandong University (number 2015HW002, number 2018JC055).

Data Availability

Under reasonable requirements, the data and material of this study can be obtained from the corresponding author. The data are not publicly available due to privacy restrictions.

Authors' Contributions

HL and JW conceptualized and designed the study; analyzed the data; and wrote the first draft of the manuscript. RC, XX, and MP contributed to surveys, data curation, and software and methodology usage. KF, BL, QL, ZQ, and SY contributed to data curation and visualization and validation of results. NIZ provided several valuable comments on the draft and corrected some issues related to grammar, vocabulary, and other aspects. FK applied for the funding used to support this study; designed the study; completed the questionnaire design; supervised and combined the data collected; instructed the writing, statistical analysis, and data processing; and provided comments on the modification of the manuscript; FK also polished the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
A flow diagram of the sample selection process.
[PNG File, 199 KB - publichealth_v11i1e66061_app1.png]

References

1. Ageing and health in China. World Health Organization. URL: https://www.who.int/china/health-topics/ageing [accessed 2025-02-28]
2. The seventh national population census [Website in Chinese]. National Bureau of Statistics. 2021. URL: https://www.stats.gov.cn/zt_18555/zdtjgz/zgrkpc/dqcrkpc [accessed 2025-02-28]
3. Pang M, Wang J, Zhao M, et al. The migrant-local difference in the relationship between social support, sleep disturbance, and loneliness among older adults in China: cross-sectional study. JMIR Public Health Surveill 2024 Jan 9;10:e49253. [doi: 10.2196/49253] [Medline: 38194253]
4. Xu J, Liu G, Li H, et al. Association between social integration and loneliness among the female migrant older adults with children: the mediating effect of social support. BMC Geriatr 2024 Jan 3;24(1):4. [doi: 10.1186/s12877-023-04569-8] [Medline: 38172722]

5. Schoenmakers D, Lamkaddem M, Suurmond J. The role of the social network in access to psychosocial services for migrant elderly-a qualitative study. Int J Environ Res Public Health 2017 Oct 11;14(10):1215. [doi: 10.3390/ijerph14101215] [Medline: 29019961]
6. Wang J, Pang M, Kong F. Association between self-reported oral health and life satisfaction among China's migrant elderly following children: the mediating effect of social support. Front Public Health 2023;11:950942. [doi: 10.3389/fpubh.2023.950942] [Medline: 36866087]
7. Wang J, Zhao J, Tian T, et al. The effects of oral health and social support on health-related quality of life of migrant older with children in Weifang, China. BMC Public Health 2022 Aug 6;22(1):1505. [doi: 10.1186/s12889-022-13843-0] [Medline: 35933414]
8. Deutsch S, House JS. Work stress and social support. Contemp Sociol 1983 May;12(3):329. [doi: 10.2307/2069001]
9. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychol Bull 1985 Sep;98(2):310-357. [doi: 10.1037/0033-2909.98.2.310] [Medline: 3901065]
10. Weiss RS. Loneliness: The Experience of Emotional and Social Isolation: MIT Press; 1973. [doi: 10.2307/2062224]
11. Donovan NJ, Wu Q, Rentz DM, Sperling RA, Marshall GA, Glymour MM. Loneliness, depression and cognitive function in older U.S. adults. Int J Geriatr Psychiatry 2017 May;32(5):564-573. [doi: 10.1002/gps.4495] [Medline: 27162047]
12. Guo L, An L, Luo F, Yu B. Social isolation, loneliness and functional disability in Chinese older women and men: a longitudinal study. Age Ageing 2021 Jun 28;50(4):1222-1228. [doi: 10.1093/ageing/afaa271] [Medline: 33352582]
13. Dong T. Association between social support and loneliness in the elderly in China: based on the 2014 Chinese Longitudinal Elderly Social Survey. The World of Survey and Research 2017(8):3-9. [doi: 10.13778/j.cnki.11-3705/c.2017.08.001]
14. Lam J. Neighborhood characteristics, neighborhood satisfaction, and loneliness differences across ethnic-migrant groups in Australia. J Gerontol B Psychol Sci Soc Sci 2022 Nov 23;77(11):2113-2125. [doi: 10.1093/geronb/gbab219] [Medline: 34888639]
15. Wang J, Mann F, Lloyd-Evans B, Ma R, Johnson S. Associations between loneliness and perceived social support and outcomes of mental health problems: a systematic review. BMC Psychiatry 2018 May 29;18(1):156. [doi: 10.1186/s12888-018-1736-5] [Medline: 29843662]
16. Guo L, Wang W, Shi J, Zheng X, Hua Y, Lu C. Evaluation of social isolation trajectories and incident cardiovascular disease among middle-aged and older adults in China: national cohort study. JMIR Public Health Surveill 2023 Jun 30;9:e45677. [doi: 10.2196/45677] [Medline: 37389914]
17. Yu K, Wu S, Jang Y, et al. Longitudinal assessment of the relationships between geriatric conditions and loneliness. J Am Med Dir Assoc 2021 May;22(5):1107-1113. [doi: 10.1016/j.jamda.2020.09.002] [Medline: 33071157]
18. Paquet C, Whitehead J, Shah R, et al. Social prescription interventions addressing social isolation and loneliness in older adults: meta-review integrating on-the-ground resources. J Med Internet Res 2023 May 17;25:e40213. [doi: 10.2196/40213] [Medline: 37195738]
19. Margolis R, Chai X, Verdery AM, Newmyer L. The physical, mental, and social health of middle-aged and older adults without close kin in Canada. J Gerontol B Psychol Sci Soc Sci 2022 Jul 5;77(7):1350-1360. [doi: 10.1093/geronb/gbab222] [Medline: 34864999]
20. Jiang C, Jiang S. Elder mistreatment and life satisfaction of older adults: mediating roles of emotional closeness with children and loneliness. J Elder Abuse Negl 2021 Oct 20;33(5):351-367. [doi: 10.1080/08946566.2021.1980171]
21. Beridze G, Ayala A, Ribeiro O, et al. Are loneliness and social isolation associated with quality of life in older adults? Insights from Northern and Southern Europe. Int J Environ Res Public Health 2020 Nov 20;17(22):8637. [doi: 10.3390/ijerph17228637] [Medline: 33233793]
22. World Health Organization. Oral health. 2024. URL: https://www.afro.who.int/health-topics/oral-health [accessed 2025-02-28]
23. Feng X. Oral health status of chinese residents: report of the fourth oral health epidemiological survey in china. Presented at: 18th Academic Annual Meeting of the Preventive Dentistry of Chinese Stomatological Association; 2018; Xi'an, Shaanxi Province, China p. 14-17.
24. Kunrath I, Silva AER. Oral health and depressive symptoms among older adults: longitudinal study. Aging Ment Health 2021 Dec;25(12):2265-2271. [doi: 10.1080/13607863.2020.1855104] [Medline: 33291966]
25. Rouxel P, Heilmann A, Demakakos P, Aida J, Tsakos G, Watt RG. Oral health-related quality of life and loneliness among older adults. Eur J Ageing 2017 Jun;14(2):101-109. [doi: 10.1007/s10433-016-0392-1] [Medline: 28579932]
26. Koyama S, Saito M, Cable N, et al. Examining the associations between oral health and social isolation: a cross-national comparative study between Japan and England. Soc Sci Med 2021 May;277:113895. [doi: 10.1016/j.socscimed.2021.113895] [Medline: 33882441]
27. Ma X, Chen C. The relationship between the degenerative changes of visual, auditory, and oral functions and loneliness in the elderly in the community. Chin J Gerontol 2021;41(11):2403-2405. [doi: 10.13778/j.cnki.11-3705/c.2017.08.001]
28. Song J, Wu H, Lan H, Xu D, Wang W. The influence of disease status on loneliness of the elderly: evidence from rural China. Int J Environ Res Public Health 2022 Mar 4;19(5):3023. [doi: 10.3390/ijerph19053023] [Medline: 35270713]
29. Langford CPH, Bowsher J, Maloney JP, Lillis PP. Social support: a conceptual analysis. J Adv Nurs 1997 Jan;25(1):95-100. [doi: 10.1046/j.1365-2648.1997.1997025095.x]

30. Qian LFW, Xiaoling S, ling L, Jingyi H, Lunwei L. Correlation between successful aging and social support for the elderly in China: a meta-analysis. Chin J Geriat Care 2022;20(5):46-50.
31. Bai Y, Bian F, Zhang L, Cao Y. The impact of social support on the health of the rural elderly in China. Int J Environ Res Public Health 2020 Mar 18;17(6):2004. [doi: 10.3390/ijerph17062004] [Medline: 32197415]
32. Mahmud NA, Shahein NA, Yoep N, et al. Influence of social support on limitation in daily living among older persons in Malaysia. Geriatrics Gerontology Int 2020 Dec;20(S2):26-32 [FREE Full text] [doi: 10.1111/ggi.14029]
33. Wang J, Xue J, Jiang Y, Zhu T, Chen S. Mediating effects of depressive symptoms on social support and quality of life among rural older Chinese. Health Qual Life Outcomes 2020 Jul 20;18(1):242. [doi: 10.1186/s12955-020-01490-1] [Medline: 32690060]
34. Zhao X, Zhang D, Wu M, et al. Loneliness and depression symptoms among the elderly in nursing homes: a moderated mediation model of resilience and social support. Psychiatry Res 2018 Oct;268:143-151. [doi: 10.1016/j.psychres.2018.07.011] [Medline: 30025285]
35. Chung S, Kim M. Age-friendly environment, social support, sense of community, and loneliness among middle-aged and older adults in Korea. Aging Ment Health 2023;27(7):1352-1359. [doi: 10.1080/13607863.2022.2116395] [Medline: 36036282]
36. Vlachantoni A, Evandrou M, Falkingham J, Qin M. The impact of changing social support on older persons’ onset of loneliness during the COVID-19 pandemic in the United Kingdom. Gerontologist 2022 Sep 7;62(8):1147-1159. [doi: 10.1093/geront/gnac033] [Medline: 35235949]
37. Sugaya N, Yamamoto T, Suzuki N, Uchiumi C. Loneliness and social isolation factors under the prolonged COVID-19 pandemic in Japan: 2-year longitudinal study. JMIR Public Health Surveill 2024 Sep 9;10:e51653. [doi: 10.2196/51653] [Medline: 39250195]
38. Wen C, Zhang Q. The role of intergenerational support in shaping oral healthcare-seeking behavior among older adults in China. Front Public Health 2023;11:1234539. [doi: 10.3389/fpubh.2023.1234539] [Medline: 37744500]
39. McGrath C, Bedi R. Influences of social support on the oral health of older people in Britain. J Oral Rehabil 2002 Oct;29(10):918-922. [doi: 10.1046/j.1365-2842.2002.00931.x] [Medline: 12421323]
40. Weber S, Hahnel S, Nitschke I, Schierz O, Rauch A. Older seniors during the COVID-19 pandemic-social support and oral health-related quality of life. Healthcare (Basel) 2021 Sep 7;9(9):1177. [doi: 10.3390/healthcare9091177] [Medline: 34574950]
41. Larsson P, Bondemark L, Häggman-Henrikson B. The impact of oro-facial appearance on oral health-related quality of life: a systematic review. J Oral Rehabil 2021 Mar;48(3):271-281. [doi: 10.1111/joor.12965] [Medline: 32196720]
42. Statistical bulletin of Weifang national economic and social development 2021 [Website in Chinese]. People’s Goverment of Weifang City. 2022. URL: http://www.weifang.gov.cn/162/55356/1506907571745853440.html [accessed 2025-02-28]
43. Weifang overview [Website in Chinese]. People’s Goverment of Weifang City. 2023. URL: http://www.weifang.gov.cn/rwyd/wfgl [accessed 2025-02-28]
44. Main data of the seventh national population census in Weifang [Website in Chinese]. People’s Government of Weifang City. 2022. URL: http://www.wfkjcs.com/ZiXun/XwDetail/dd4132ba55f743329c60b575b29d452e [accessed 2025-02-28]
45. Chan KW. Post‐Mao China: a two‐class urban society in the making. Int J Urban Regional Res 1996 Mar;20(1):134-150 [FREE Full text] [doi: 10.1111/j.1468-2427.1996.tb00305.x]
46. A-Dan W, Jun-Qi L. Factors associated with the oral health-related quality of life in elderly persons in dental clinic: validation of a Mandarin Chinese version of GOHAI. Gerodontology 2011 Sep;28(3):184-191. [doi: 10.1111/j.1741-2358.2009.00360.x] [Medline: 21692833]
47. Shao R, Hu T, Zhong YS, et al. Socio-demographic factors, dental status and health-related behaviors associated with geriatric oral health-related quality of life in Southwestern China. Health Qual Life Outcomes 2018 May 21;16(1):98. [doi: 10.1186/s12955-018-0925-8] [Medline: 29784008]
48. Xiao S. Theoretical foundations and research applications of the Social Support Rating Scale. J Clin Psychiatry 1994;02:98-100 [FREE Full text]
49. Chu H, Yang Y, Zhou J, et al. Social support and suicide risk among Chinese university students: a mental health perspective. Front Public Health 2021;9:566993. [doi: 10.3389/fpubh.2021.566993] [Medline: 33681117]
50. Zhao S, Long M, Diao Y, et al. Culture-related grief beliefs and social support influence depressive symptoms of Shidu parents in rural China. Eur J Psychotraumatol 2021;12(1):1945748. [doi: 10.1080/20008198.2021.1945748] [Medline: 34367525]
51. Wang G, Hu M, Xiao SY, Zhou L. Loneliness and depression among rural empty-nest elderly adults in Liuyang, China: a cross-sectional study. BMJ Open 2017 Oct 6;7(10):e016091. [doi: 10.1136/bmjopen-2017-016091] [Medline: 28988166]
52. Niu L, Jia C, Ma Z, Wang G, Yu Z, Zhou L. The validity of proxy-based data on loneliness in suicide research: a case-control psychological autopsy study in rural China. BMC Psychiatry 2018 May 1;18(1):116. [doi: 10.1186/s12888-018-1687-x] [Medline: 29716552]
53. Wu M. Structural Equation Modeling: Operation and Application of AMOS: Chongqing University Press; 2010.
54. Fan X, Thompson B, Wang L. Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. Struct Equ Modeling 1999 Jan;6(1):56-83. [doi: 10.1080/10705519909540119]

55. Kline RB. Principles and Practice of Structural Equation Modeling, 4th edition: The Guilford Press; 2016. URL: https://api.semanticscholar.org/CorpusID:164967627 [accessed 2025-03-09]
56. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Modeling 2002 Apr;9(2):233-255. [doi: 10.1207/S15328007SEM0902_5]
57. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. Struct Equ Modeling 2007 Jul 31;14(3):464-504. [doi: 10.1080/10705510701301834]
58. Rong T. Amos and Research Methodology: Chongqing University Press; 2009.
59. De Jong Gierveld J, Van der Pas S, Keating N. Loneliness of older immigrant groups in Canada: effects of ethnic-cultural background. J Cross Cult Gerontol 2015 Sep;30(3):251-268. [doi: 10.1007/s10823-015-9265-x] [Medline: 25982532]
60. Kisely S. No mental health without oral health. Can J Psychiatry 2016 May;61(5):277-282. [doi: 10.1177/0706743716632523] [Medline: 27254802]
61. Susanty S, Chung MH, Chiu HY, et al. Prevalence of loneliness and associated factors among community-dwelling older adults in Indonesia: a cross-sectional study. Int J Environ Res Public Health 2022 Apr 18;19(8):4911. [doi: 10.3390/ijerph19084911] [Medline: 35457781]
62. Bernabé E, de Oliveira C, de Oliveira Duarte YA, Bof de Andrade F, Sabbah W. Social participation and tooth loss, vision, and hearing impairments among older Brazilian adults. J Am Geriatr Soc 2023 Oct;71(10):3152-3162. [doi: 10.1111/jgs.18423] [Medline: 37227109]
63. Igarashi A, Aida J, Yamamoto T, Hiratsuka Y, Kondo K, Osaka K. Associations between vision, hearing and tooth loss and social interactions: the JAGES cross-sectional study. J Epidemiol Community Health 2021 Feb;75(2):171-176. [doi: 10.1136/jech-2020-214545] [Medline: 32972921]
64. Yuan L, Yuan Y, Ren H, et al. Decomposition analysis of the prevalence of denture use between rural and urban older individuals with edentulism in China: cross-sectional study. Interact J Med Res 2024 Dec 13;13:e48778. [doi: 10.2196/48778] [Medline: 39671593]
65. Xu J, Zheng J, Xu L, Wu H. Equity of health services utilisation and expenditure among urban and rural residents under universal health coverage. Int J Environ Res Public Health 2021 Jan 12;18(2):593. [doi: 10.3390/ijerph18020593] [Medline: 33445637]
66. Olofsson H, Ulander EL, Gustafson Y, Hörnsten C. Association between socioeconomic and health factors and edentulism in people aged 65 and older - a population-based survey. Scand J Public Health 2018 Nov;46(7):690-698. [doi: 10.1177/1403494817717406] [Medline: 28768459]
67. Yang F, Gu D. Predictors of loneliness incidence in Chinese older adults from a life course perspective: a national longitudinal study. Aging Ment Health 2020 Jun;24(6):879-888. [doi: 10.1080/13607863.2018.1558174] [Medline: 30621448]
68. Beal C. Loneliness in older women: a review of the literature. Issues Ment Health Nurs 2006;27(7):795-813. [doi: 10.1080/01612840600781196] [Medline: 16849264]
69. Zhang K, Burr JA, Mutchler JE, Lu J. Internet use and loneliness among urban and non-urban Chinese older adults: the roles of family support, friend support, and social participation. The Journals of Gerontology, Series B 2024 Jul 1;79(7):gbae081. [doi: 10.1093/geronb/gbae081]
70. Zhang X, Dong S. The relationships between social support and loneliness: a meta-analysis and review. Acta Psychol (Amst) 2022 Jul;227:103616. [doi: 10.1016/j.actpsy.2022.103616] [Medline: 35576818]
71. Jia W, Liu L, Peng G. The impact of social media on users' self-efficacy and loneliness: an analysis of the mediating mechanism of social support. Psychol Res Behav Manag 2024;17:593-612. [doi: 10.2147/PRBM.S449079] [Medline: 38379636]
72. Jönsson H, Michaud M, Neuman N. What is commensality? A critical discussion of an expanding research field. Int J Environ Res Public Health 2021 Jun 9;18(12):6235. [doi: 10.3390/ijerph18126235] [Medline: 34207626]
73. Masood M, Newton T, Bakri NN, Khalid T, Masood Y. The relationship between oral health and oral health related quality of life among elderly people in United Kingdom. J Dent 2017 Jan;56:78-83. [doi: 10.1016/j.jdent.2016.11.002] [Medline: 27825838]
74. Goergen J, Costa RSA, Rios FS, et al. Oral conditions associated with oral health related quality of life: a population-based cross-sectional study in Brazil. J Dent 2023 Feb;129:104390. [doi: 10.1016/j.jdent.2022.104390] [Medline: 36526085]

Abbreviations

AGFI: adjusted goodness of fit Index
CFI: comparative fitness index
CMIN: chi-square value
GFI: goodness of fit index
GOHAI: Geriatric Oral Health Assessment Index
LOA: local older adults
MOA: migrant older adults
RMSEA: root mean square error of approximation

SSRS: Social Support Rating Scale
ULS-6: 6-item short-form UCLA Loneliness Scale

Edited by A Mavragani; submitted 03.09.24; peer-reviewed by AK Chun Chau, J Zhao; revised version received 17.02.25; accepted 18.02.25; published 13.03.25.

Please cite as:
Liu H, Wang J, Chen R, Xu X, Pang M, Feng K, Li B, Li Q, Qin Z, Yan S, Ibn Ziyat N, Kong F
Migrant-Local Differences in the Relationship Between Oral Health, Social Support, and Loneliness Among Older Adults in Weifang, China: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e66061
URL: https://publichealth.jmir.org/2025/1/e66061
doi:10.2196/66061

Long-Term Exposure to Fine Particulate Matter (PM2.5) Components and Precocious Puberty Among School-Aged Children: Cross-Sectional Study

Xuelian Zhou[1*], MD; Xiaochi Zhang[2,3*], BS; Guannan Bai[4], PhD; Guanping Dong[1], MD; Xinyi Li[2,3], BS; Ruimin Chen[5], MM; Shaoke Chen[6], MM; Rongxiu Zheng[7], MD; Chunlin Wang[8], MD; Haiyan Wei[9], MM; Bingyan Cao[10], MD; Yan Liang[11], MD; Hui Yao[12], MD; Zhe Su[13], MD; Mireguli Maimaiti[14], MD; Feihong Luo[15], MD; Pin Li[16], MD; Min Zhu[17], MM; Hongwei Du[18], MD; Yu Yang[19], MD; Lanwei Cui[20], MD; Jinling Wang[1], MD; Jinna Yuan[1], MD; Zhuang Liu[21], MD; Wei Wu[1], MD; Qi Zhao[2,3,22], PhD; Junfen Fu[1], PhD

[1]Department of Endocrinology, Children’s Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, National Children’s Regional Medical Center, 3333 Binsheng Road, Hangzhou, China

[10]Department of Endocrinology, Beijing Children’s Hospital, Capital Medical University, National Medical Center for Children’s Health, Beijing, China

[11]Department of Pediatrics, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[12]Department of Pediatrics, Wuhan Children’s Hospital, Tongji Medical College, Huazhong University of Science & Technology, Wuhan, China

[13]Department of Endocrinology, Shenzhen Children’s Hospital, Shenzhen, China

[14]Department of Pediatrics, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, China

[15]Department of Pediatric Endocrinology and Inherited Metabolic Diseases, Children’s Hospital of Fudan University, Shanghai, China

[16]Department of Endocrinology, Children’s Hospital of Shanghai Jiaotong University, Shanghai, China

[17]Department of Endocrinology, Children’s Hospital of Chongqing Medical University, Chongqing, China

[18]Department of Pediatric Endocrinology, The First Bethune Hospital of Jilin University, Changchun, China

[19]Department of Endocrinology, Jiangxi Provincial Children’s Hospital, Nanchang, China

[2]Department of Epidemiology, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China

[20]Department of Pediatrics, The First Affiliated Hospital of Harbin Medical University, Harbin, China

[21]Department of Reproductive Medicine, Hospital of Jining Medical University, Jining, China

[22]Faculty of Health, Deakin University, Melbourne, Australia

[3]Shandong University Climate Change and Health Center, Shandong University, Jinan, China

[4]Department of Child Health Care, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, China

[5]Department of Endocrinology, Children’s Hospital of Fuzhou, Fujian Province, Fuzhou, China

[6]Department of Pediatrics, Maternal and Child Health Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

[7]Department of Pediatrics, Tianjin Medical University Central Hospital, Tianjin, China

[8]Department of Pediatrics, The First Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China

[9]Department of Endocrinology, Zhengzhou Children’s Hospital, Zhengzhou, China

[*]these authors contributed equally

Corresponding Author:
Junfen Fu, PhD
Department of Endocrinology, Children’s Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, National Children’s Regional Medical Center, 3333 Binsheng Road, Hangzhou, China

Abstract

Background: The increasing incidence of precocious puberty is a major health challenge for Chinese children, while related risk factors remain less well explored. Exposure to ambient fine particulate matter ($PM_{2.5}$) is a leading environmental hazard in China. Although certain components of $PM_{2.5}$ have been reported to be endocrine disruptors for sex hormones, population-based evidence is still lacking on the association between $PM_{2.5}$ exposure and precocious puberty in China.

Objective: Based on a cross-sectional survey covering 30 cities in 2017 to 2019, this study was designed to explore the association between long-term exposure to $PM_{2.5}$ and its 5 major components with precocious puberty in China and to check the potential modifying effects of family-related and personal factors.

Methods: We included 34,105 children aged 6 to 9 years. We collected the 5-year average concentrations of $PM_{2.5}$ and its 5 major components (sulfate, nitrate, ammonium, organic matter, and black carbon) in the area (at a spatial resolution of 0.1° × 0.1°) where each school was located. We used mixed effect logistic regression to estimate the effect sizes of the total mass of $PM_{2.5}$ and each of its components on precocious puberty, and we examined the modifying effects of family-related and personal factors using an additional interactive term. A weighted quantile sum (WQS) regression model was applied to identify the weights of each component in explaining the effect size of the total mass of $PM_{2.5}$.

Results: We found that the odds ratio (OR) for precocious puberty per IQR increase in the concentration of total $PM_{2.5}$ mass was 1.27 (95% CI 0.92-1.75) for the whole population, 2.12 (95% CI 1.27-3.55) for girls, and 0.90 (95% CI 0.62-1.30) for boys. Similarly, the effect sizes of the 5 major components were all substantial for girls but minimal for boys. Results of the WQS analysis showed that organic matter could explain the highest proportion of the effect of $PM_{2.5}$, with the weight of its contribution being 0.71. Modification effects of family income and dietary habits were only observed in certain population subgroups.

Conclusions: Long-term exposure to total $PM_{2.5}$ mass was significantly associated with precocious puberty in girls, with organic matter identified as the major effect contributor. The results add evidence on the detrimental effects of $PM_{2.5}$ on children's development and growth.

(JMIR Public Health Surveill 2025;11:e62861) doi:10.2196/62861

KEYWORDS

fine particulate matter; PM2.5; PM2.5 components; air pollution; precocious puberty; children; long-term exposure

Introduction

Precocious puberty is an endocrine and metabolic disorder that may promote the development of pathophysiological or psychosocial outcomes in childhood and later life, including stunted growth, obesity, hypertension, type 2 diabetes, ischemic heart disease, stroke, and breast cancer [1]. Over the past decades, there has been a notable increase in the incidence of precocious puberty across countries. For instance, in Denmark, the annual incidence of precocious puberty among girls younger than 8 years rose from 2.6 to 14.6 cases per 10,000 individuals between 1998 and 2017 [2]. Similar increasing trends have also been observed in a few other countries, such as South Korea, where the incidence among girls younger than 9 years rose 4.7 times between 2008 and 2014 [3], and Spain, where the incidence among girls younger than 8 years increased 16.7 times between 1997 and 2006 [4]. Among these countries, China stands out because the incidence is increasing even more rapidly and there is a larger population base [5]. This raises the urgent necessity of clarifying risk factors to facilitate efficient health promotion strategies. Certain factors or preexisting health conditions are speculated to explain the rising incidence of precocious puberty, including genetic predisposition, obesity, environmental exposures, and socioeconomic conditions [1,6-8]. However, specific risk factors involved in the pathogenesis of precocious puberty and their potential interactions still remain unclear and underexplored.

Air pollution is the second leading cause of global disease burden, and exposure to fine particulate matter ($PM_{2.5}$) could underlie most cases of air pollution–related disease, considering the complex toxicity of its components and the fact that over 90% of the world's population lives in highly polluted locations [9]. Some $PM_{2.5}$ components, such as phthalate, bisphenol analogs, and polycyclic aromatic hydrocarbons (PAHs), have been recognized as endocrine-disrupting chemicals (EDCs), affecting the normal production and signaling of sex hormones [10-12]. Some previous studies explored the association between long-term exposure to $PM_{2.5}$ and the risk of precocious puberty or related symptoms in children, but the conclusions remained inconsistent [7,13-15]. Explanations may be related to differing susceptibility across populations and the varied concentration of the essential components of $PM_{2.5}$ across countries and regions.

Ambient $PM_{2.5}$ pollution is a significant environmental hazard in China. Although air pollution has decreased dramatically in the past decade due to a series of new laws and advancements in environmental protection, the annual mean $PM_{2.5}$ concentration in 2023 was still 5 times higher than the recommended threshold of the World Health Organization (WHO) Air Quality Guidelines [9,16]. Considering the substantial number of children and the increasingly early age at onset of puberty, the adverse impact of $PM_{2.5}$ on precocious puberty, if it indeed exists, could stand out as a significant public health concern in China. However, to date, limited studies have clarified this research question.

Using data from a survey covering 30 cities of 11 provinces, our study aims to estimate the association between long-term exposure to $PM_{2.5}$ components and precocious puberty among school-aged children and to identify the components with the highest effect size in China. This will be indispensable for understanding the harmful effects of $PM_{2.5}$ and for health care promotion and environmental protection in this country.

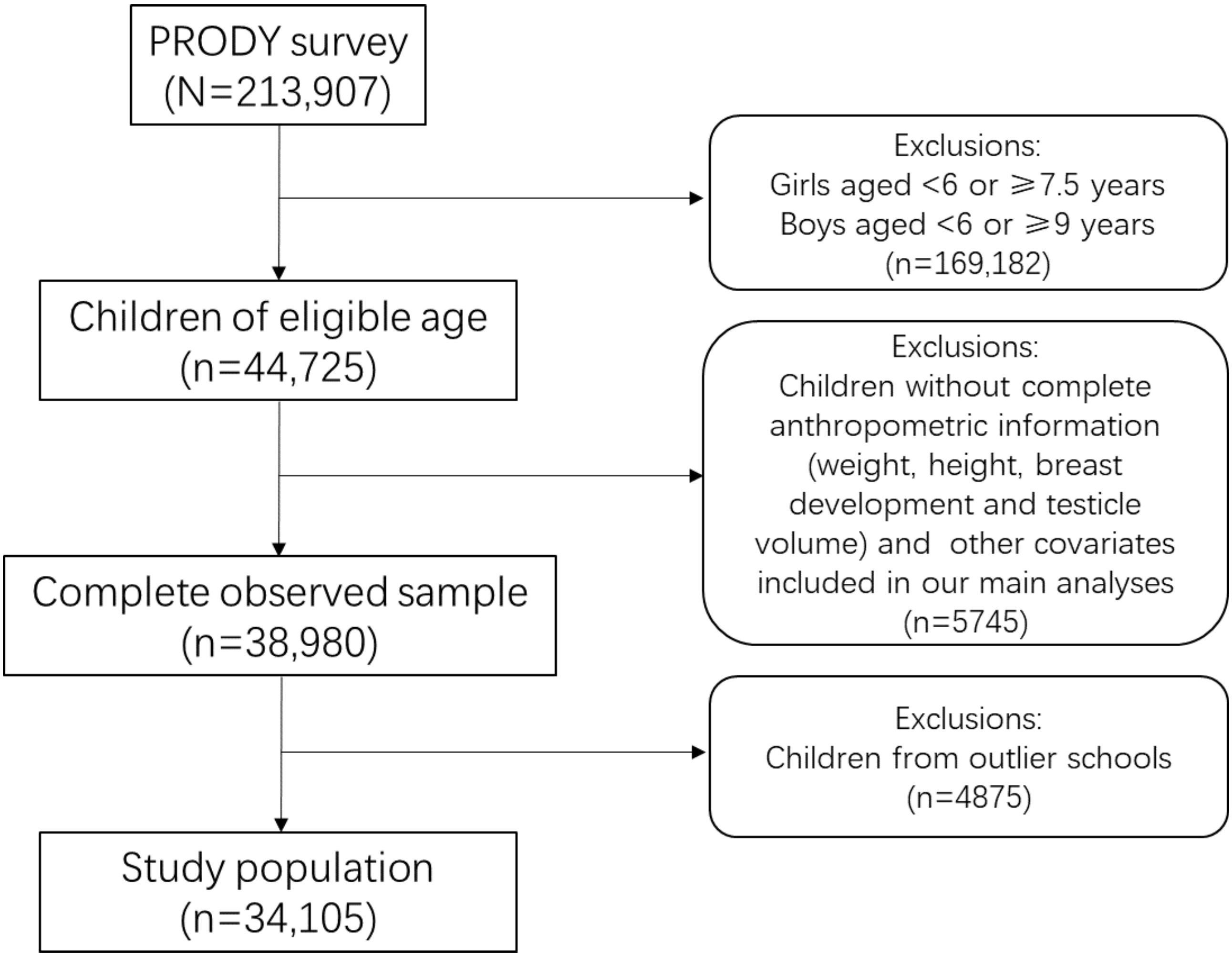
Methods

Study Population

We analyzed data from eligible-aged children in the Prevalence and Risk Factors for Obesity and Diabetes in Youth (PRODY) study. A detailed profile of the PRODY data has been described previously [17]. Briefly, PRODY is a multicenter cross-sectional study conducted among children aged 3 to 18 years in China from 2017 to 2019. Participants were recruited using a stratified cluster random sampling strategy by considering their geographic region, urbanization level, and economic level. Questionnaires and physical examinations were used to collect demographic characteristics, family information, dietary and behavioral habits, growth and development status, and other information, such as weight, height, parental educational level (low: high school or below; high: bachelor or higher), annual family income (first quartile [Q1]: <50,000 yuan [1 yuan=US $0.14]; Q2: 50,000 - 100,000 yuan; Q3: 100,000 - 200,000 yuan; Q4: >200,000 yuan), physical activity time (Q1:<90 min/week; Q2: 90 - 150 min/week; Q3:>150 min/week), frequency of consumption of dairy products, meat, fried foods, soy products, and junk foods (Q1:<1 times/month; Q2: 1 - 2 times/month; Q3: 3 - 4 times/month; Q4:>4 times/month), whole grains (Q1:<1 times/week; Q2: 1 - 2 times/week; Q3: 3 - 4 times/week; Q4:>4 times/week), and the consumption of sugary drinks, sweet fruit, and acidic fruit (yes or no). BMI (in kg/m^2) was calculated according to the growth reference of the WHO [18] and then categorized into 2 groups: low (normal weight or below) and high (overweight or obese).

This study initially included 213,907 participants across 11 provinces, autonomous regions, or municipalities in China (Xinjiang Uyghur Autonomous Region, Henan, Hubei, Guangdong, Guangxi Zhuang Autonomous Region, Zhejiang, Shanghai, Fujian, Jilin, Beijing, and Tianjin). Among the participants, 44,725 were at the eligible age for studying precocious puberty, and 38,980 of these (87.2%) accepted the related physical examination. Inclusion and exclusion criteria for the participants are shown in Figure 1. Briefly, based on previous studies and the diagnostic criteria for precocious puberty in China [5,19], we selected girls aged 6 to 7.5 years and boys aged 6 to 9 years as the study population. Children without complete anthropometric information or necessary covariates were excluded. To ensure the robustness of the analysis, schools were excluded as outliers if their prevalence rate of precocious puberty was 1.5 times greater or less than the IQR of the prevalence rates across all schools. The sample size for final analysis was 34,105. Differences in the main demographic characteristics of the study population and the original population were not statistically significant (Table S1 in Multimedia Appendix 1).

Figure 1. Flowchart of the inclusion and exclusion of study population. PRODY: Prevalence and Risk Factors for Obesity and Diabetes in Youth.



Ethical Considerations

This study protocol was approved by the Institutional Review Board of the Children's Hospital of Zhejiang University School of Medicine (approval number 2016-JRB-018). Written informed consent to participate in this study was provided by the participants' legal guardians or next of kin. We intended to use survey weights, but this possibility was denied by the data custodian due to data confidentiality concerns. All data used in this study were nonidentifiable.

Outcome Measurements

The children’s breast development, testicle volume, and pubic hair were measured by a well-trained endocrinologist. Pubertal stages were assessed following the 5-stage scale described by Marshall and Tanner [20]. Precocious puberty was defined as the initiation of secondary sexual characteristics, designated as Tanner stage 2, before the age of 7.5 years for girls and 9 years for boys [19]. Previous studies have demonstrated that over 98% of Chinese children younger than 6 years have not initiated pubertal development [5]. In line with previous studies, we confined the study population to girls aged 6 to 7.5 years and boys aged 6 to 9 years to improve modeling stability.

Exposure Measurements

Data on annual average concentrations of $PM_{2.5}$ and its 5 major components, sulfate (SO_4^{2-}), nitrate (NO_3^-), ammonium (NH_4^+), organic matter (OM), and black carbon (BC), were collected from the Tracking Air Pollution in China database [21] at a spatial resolution of 0.1° × 0.1° for the entire area of China for the period 2012 to 2019. The concentration at each school location was then calculated. Tracking Air Pollution developed a 2-stage machine learning model to predict daily $PM_{2.5}$ concentration using data from $PM_{2.5}$ measurements, satellite aerosol optical depth retrievals, online community multiscale air quality simulations, meteorological reanalysis data, land use information, and population distribution. Methodological details have been described elsewhere, with the data quality also warranted as high [22,23]. The 5-year average concentrations of $PM_{2.5}$ and its 5 major components, calculated as the average over the 0 to 4 years preceding the children’s physical examinations, were assigned to each school address as a surrogate for personal exposure.

Statistical Methods

Single Exposure Model

Mixed-effects logistic regression was applied to quantify the association between $PM_{2.5}$ concentration and each of its 5 components with the binary status of precocious puberty, with the “province” indicator entered as the random intercept. Considering that certain dietary variables were correlated (Figure S1 in Multimedia Appendix 1), we used a least absolute shrinkage and selection operator regression model with the minimum mean squared error to select important dietary confounders related to precocious puberty (Table S2 in Multimedia Appendix 1). The main models were adjusted for gender, age, parental educational level, annual family income, physical activity time, BMI, and dietary habits, including frequency of consumption of dairy products, meat, fried food, and whole grains, as well as the intake of sugary drinks as a snack, fruit as a snack, sweet fruit, and acidic fruit. The results were described as odds ratios (ORs) for precocious puberty with the 95% CI per IQR increase in the concentration of each air pollutant. In this study, $PM_{2.5}$ and its 5 major components were not entered into a single model at the same time to avoid potential multicollinearity problems. The potential interaction effects were examined individually by adding an additional interaction term between $PM_{2.5}$ and each of these variables, including parental educational level, annual family income, physical activity time, BMI, and dietary habits.

Joint Exposure Model

Given the high correlations and uncertain relationships among the various components of $PM_{2.5}$, we used a weighted quantile sum (WQS) regression model to assess the joint exposure effects. The WQS model was used to construct a cumulative linear index by grouping $PM_{2.5}$ components into quartiles, with the corresponding weight of each component indicating its contribution to the overall index. WQS regression assumes that all exposures included in the index are associated with the outcome in the same direction. Therefore, we first evaluated the direction of the association between each pollutant and precocious puberty through the single exposure model. As suggested previously, the training and validation datasets were split 40% and 60%, and the bootstrap was set to 1000 times [24].

Sensitivity Analyses

A series of sensitivity analyses were conducted to examine the robustness of our main findings. First, the main models were expanded by additionally adjusting for other dietary variables selected from the literature that failed the selection procedure. Second, we modified the personal exposure from 0 - 4 years to 1 - 4 years and 0 - 2 years to explore whether the results might differ with exposure duration. Third, we further adjusted for gaseous pollutants, such as SO_2, CO_2, and O_3. The quantile-based g-computation (QGC) model was used for sensitivity analysis of the joint exposure model, as it allows for the evaluation of combined effects of exposure in different directions [25]. Meta regression was used to test the significance of differences in the results of the main analyses and sensitivity analyses.

R (version 4.3.0; R Foundation for Statistical Computing) was used for analyses. *P* values <.05 were considered statistically significant.

Results

Table 1 presents the demographic characteristics of the study participants. A total of 34,105 children (n=23,683, 69.44% boys) with an average age of 7.25 (SD 0.82) years were enrolled in this study. Among them, 62.39% (21,279/34,105) had an annual family income exceeding 100,000 yuan (US $14,327), 63.25% (21,570/34,105) had parents with a college degree or higher, and 26.01% (8872/34,105) were categorized as overweight or obese. A total of 401 of the 34,105 children (1.18%) were diagnosed with precocious puberty, of which 214 were boys (0.9%) and 187 were girls (1.79%). The average individual exposure over a 5-year period was 47.93 (SD 16.52) μg/m^3 for $PM_{2.5}$, 2.31 (SD 0.59) μg/m³ for BC, 6.34 (SD 2.25) μg/m³ for NH_4^+, 11.74 (SD 3.75) μg/m³ for OM, 9.52 (SD 3.95) μg/m³ for NO_3^- and 8.74 (SD 2.11) μg/m³ for SO_4^{2-}, respectively (Table S3 in Multimedia Appendix 1). A map of $PM_{2.5}$ concentration and the study population during the study period is shown in Figure S2 in Multimedia Appendix 1.

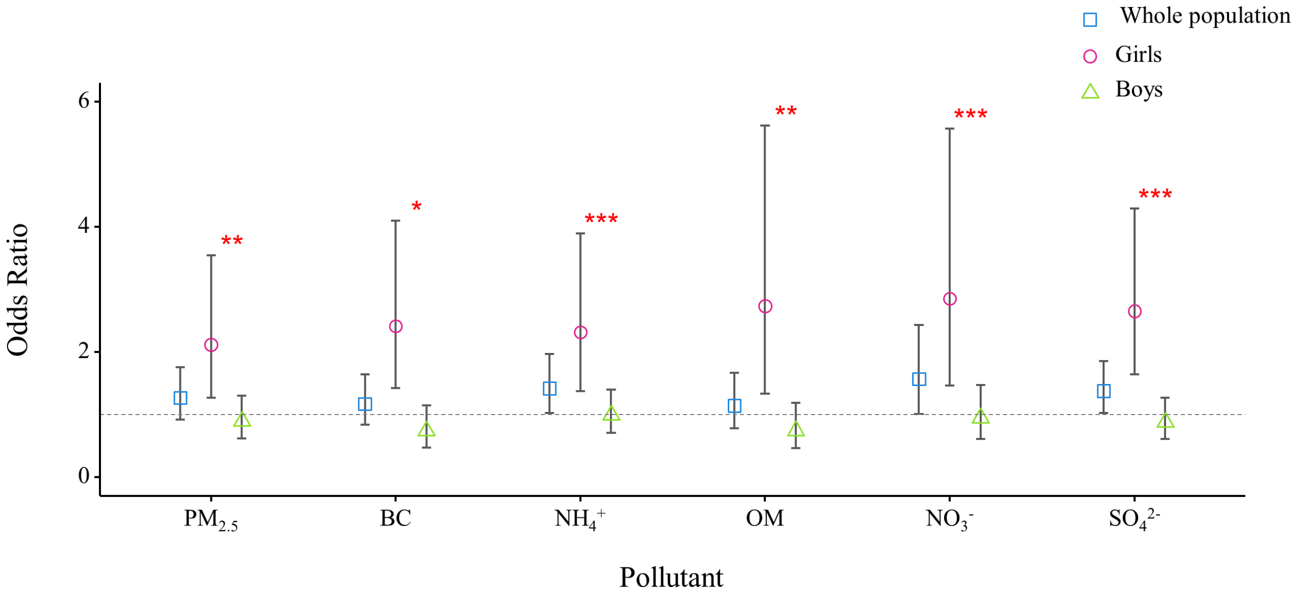
Table . Basic characteristics of participants enrolled during 2017 and 2019 in 30 cities in China.

Variables		Precocious puberty (n=401)	Nonprecocious puberty (n=33,704)
$PM_{2.5}$ concentration (μg/m^3), mean (SD)		47.92 (16.79)	47.93 (16.52)
Age (years), mean (SD)		7.56 (0.85)	7.24 (0.82)
Gender, n (%)			
	Female	187 (46.63)	10,235 (30.37)
	Male	214 (53.37)	23,469 (69.63)
Parental educational level, n (%)			
	High school or below	180 (44.89)	12,355 (36.66)
	Bachelor or higher	221 (55.11)	21,349 (63.34)
Annual family income (yuan; 1 yuan=US $0.14), n (%)			
	<50,000	67 (16.71)	5530 (16.41)
	50,000 - 10,0000	98 (24.44)	7131 (21.16)
	100,000 - 200,000	120 (29.93)	10,902 (32.35)
	>200,000	116 (28.93)	10,141 (30.09)
Physical activity (minutes/week) , n (%)			
	<90	154 (38.4)	13,700 (40.65)
	90 - 150	168 (41.9)	12,634 (37.49)
	>150	79 (19.7)	7370 (21.87)
BMI, n (%)			
	Underweight or normal	211 (52.62)	25,022 (74.24)
	Overweight or obese	190 (47.38)	8682 (25.76)
Consumed sugary drinks, n (%)			
	No	283 (70.57)	25,630 (76.04)
	Yes	118 (29.43)	8074 (23.96)
Consumed sweet fruit, n (%)			
	No	71 (17.71)	4709 (13.97)
	Yes	330 (82.29)	28,995 (86.03)
Consumed acidic fruit, n (%)			
	No	197 (49.13)	14,884 (44.16)
	Yes	204 (50.87)	18,820 (55.84)
Junk food frequency (times/month), n (%)			
	<1	354 (88.28)	30,226 (89.68)
	1 - 2	27 (6.73)	2462 (7.30)
	3 - 4	15 (3.74)	750 (2.23)
	>4	5 (1.25)	266 (0.79)
Whole grains frequency (times/week), n (%)			
	<1	63 (15.71)	5196 (15.42)
	1 - 2	248 (61.85)	20,788 (61.68)
	3 - 4	71 (17.71)	5636 (16.72)
	>4	19 (4.74)	2084 (6.18)
Dairy product frequency (times/day), mean (SD)		0.66 (0.54)	0.72 (0.57)
Meat frequency (times/day), mean (SD)		2.60 (1.46)	2.73 (1.46)
Fried food frequency (times/day), mean (SD)		0.18 (0.55)	0.14 (0.46)

Variables	Precocious puberty (n=401)	Nonprecocious puberty (n=33,704)
Soy product frequency (times/day), mean (SD)	0.93 (0.88)	0.98 (0.89)

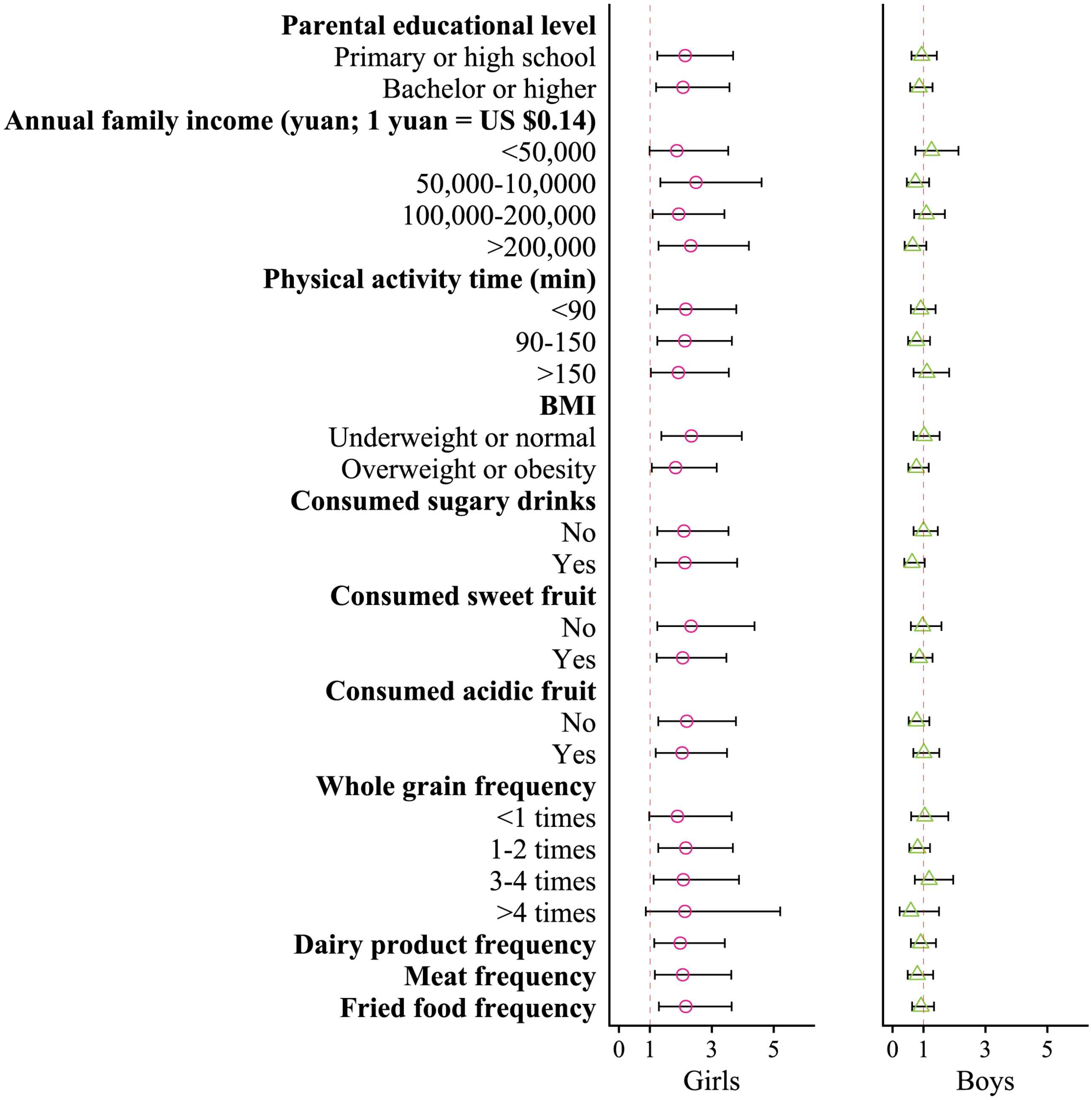
The analysis of the single exposure model indicated that the ORs for precocious puberty for the whole study population were 1.27 (95% CI 0.92-1.75) for total $PM_{2.5}$ mass, 1.42 (95% CI 1.03 - 1.97) for NH_4^+, 1.57 (95% CI 1.01 - 2.43) for NO_3^-, and 1.38 (95% CI 1.03 - 1.86) for SO_4^{2-} per IQR increase in their 5-year average concentrations (Figure 2). A clear gender variation in the effect size of $PM_{2.5}$ exposure was observed: per IQR increase in $PM_{2.5}$ mass, the ORs were 2.12 (95% CI 1.27 - 3.55) for girls and 0.90 (95% CI 0.62 - 1.30) for boys (*P*=.01). In addition, the effect sizes of all 5 $PM_{2.5}$ components were also only significant for girls.

Figure 2. Odds ratio for precocious puberty per IQR increase in $PM_{2.5}$ mass and its 5 major components. $PM_{2.5}$:fine particulate matter; BC: black carbon; OM: organic matter; NO_3^-: nitrate; NH_4^+: ammonium; SO_4^{2}–: sulfate. Asterisks indicate the significance of differences between boys and girls: $^*P<.05$, $^{**}P<.01$, $^{***}P<.001$.



The effect size of $PM_{2.5}$ was nonsignificantly higher in groups with certain demographic characteristics, family information, and behavioral habits (Figure 3). For example, the OR for precocious puberty was 2.49 (95% CI 1.34 - 4.61) for girls with a family income at the Q2 level and 1.87 (95% CI 0.99 - 3.53) for those with family income at the Q1 level. Similarly, the ORs were 2.34 (95% CI 1.37 - 3.97) and 1.83 (95% CI 1.06 - 3.16) for groups with low BMI and high BMI, respectively. Certain modifying effects were observed in boys, such as family income (<50,000 yuan group vs >200,000 yuan group; *P*=.03) and sugary drink consumption (no vs yes; *P*=.03).

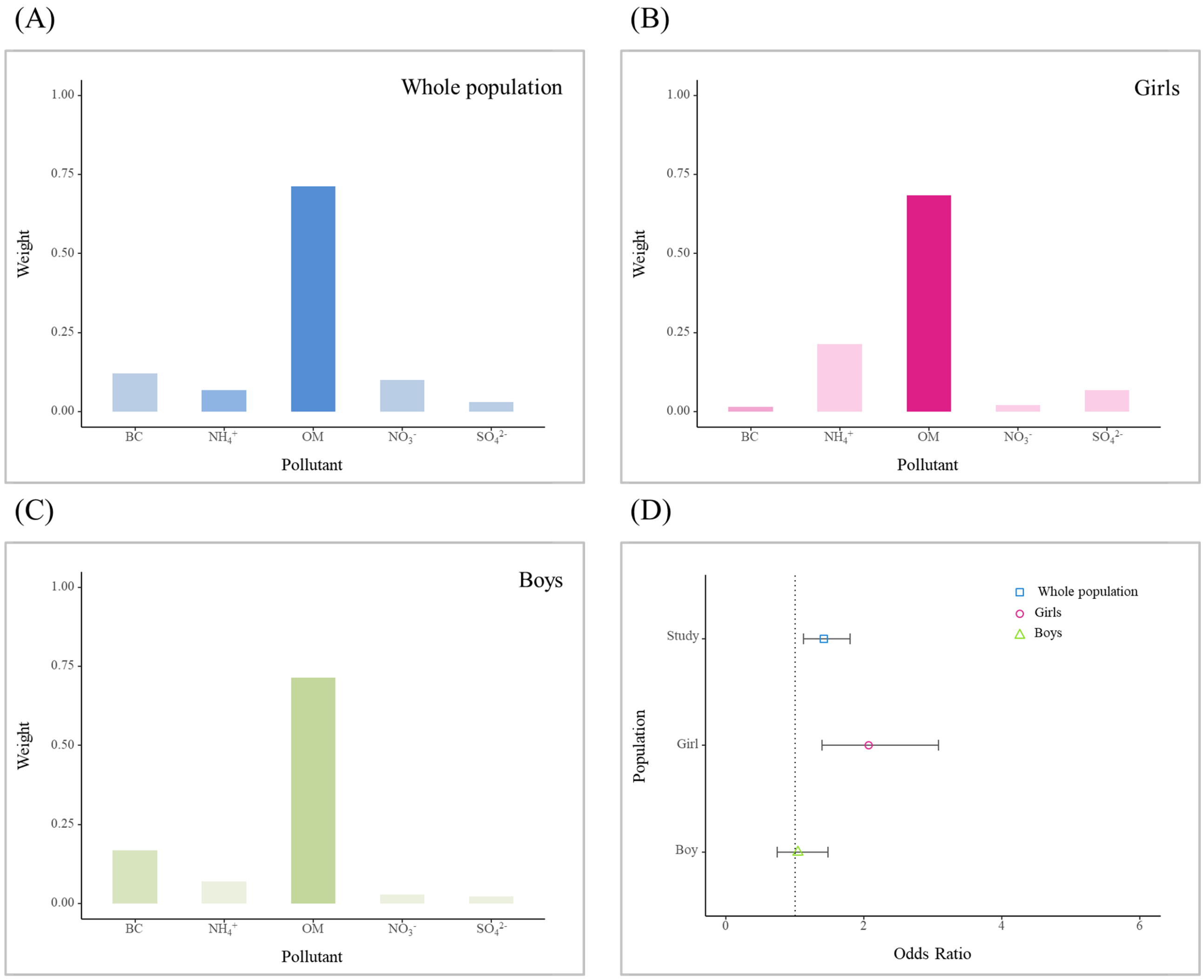
Figure 3. The modification effect of individual characteristics and dietary habits on the association between $PM_{2.5}$ exposure and precocious puberty.



As shown in Figure 4, the WQS index was significantly associated with precocious puberty (OR 1.42, 95% CI 1.12 - 1.80). The major contributors were OM, BC, and NO_3^-, with their weights in the overall effect size being 0.71, 0.12, and 0.10, respectively. After stratification by gender, a significant association was observed only in girls.

The strength of the association was stable after adjusting for additional covariables and gaseous pollutants or changing the exposure period in both the single-air-pollutant models and WQS models. The association between $PM_{2.5}$ components and precocious puberty remained unchanged in the QGC model, with OM continuing to be the major positive contributor to the overall effect of total $PM_{2.5}$ mass (Tables S4-5 in Multimedia Appendix 1).

Figure 4. Weights of the contributions of the 5 major components of fine particulate matter ($PM_{2.5}$) to its total effect size. Mean weights for (**A**) the whole study population; (**B**) girls; (**C**) boys. (**D**) Odds ratios for precocious puberty per IQR increase in the weighted quantile sum regression (WQS) index. WQS models were adjusted for gender; age; maximum level of parental education; annual family income; weekly physical activity time; BMI; consumption frequency of dairy products, meat, fried foods, and coarse food grain; intake of sugary drinks, sweet fruit, and acidic fruit; and province. BC: black carbon; OM: organic matter; NO_3^-: nitrate; NH_4^+: ammonium; SO_4^{2-}: sulfate.



Discussion

Principal Results

This study explored the association between exposure to $PM_{2.5}$ and its major components with precocious puberty based on a large survey of school-aged children in China. We observed that $PM_{2.5}$ mass was significantly associated with precocious puberty only in girls. OM was detected as the major effective component when the 5 major components of $PM_{2.5}$ were jointly exposed. Modification effects of certain factors at the individual or family levels were observed (eg, family income and sugary drink consumption), especially for boys.

Comparison With Prior Work

Recent evidence suggests that there is significant variation in the prevalence of precocious puberty across regions and study periods; the prevalence ranges from <0.05% to 2.53% in Denmark, Korea, and China [15,26,27]. To our knowledge, our PRODY study is to date the only large-scale survey on precocious puberty in China. We found that the prevalence in the study areas was 0.9% in boys and 1.79% in girls. Previously, only one study (the predecessor of PRODY) investigated precocious puberty in 5 Chinese provinces and reported a prevalence of 0.43% [26]. The differing prevalence across studies may be partly explained by differences in diagnostic criteria, population characteristics, the years of investigation, and certain limitations in sampling.

Previously, some studies have explored the effect of long-term $PM_{2.5}$ exposure on pubertal development or related sexual hormones and obtained inconsistent results. A cohort study in Poland revealed that children exposed to an average annual concentration of $PM_{2.5}$ exceeding 25 μg/m^3 were at high risk of menarche occurring before the age of 11 years [14]. Similarly, Jung et al [28] found that each 1 μg/m^3 increase in the 3-year average of annual mean PM_{10} concentration was associated with age acceleration at menarche by 0.031 years in 639 girls aged 13 to 17 years in South Korea. Additionally, a recent large-scale (n=1,205,784) nationwide retrospective cohort study reported that 4-year exposures to $PM_{2.5}$ and PM_{10} were associated with precocious puberty in girls in South Korea [15]. By contrast, in Hong Kong's cohort of children born in 1997, Huang et al [7] found a nonsignificant association between childhood PM_{10} exposure and precocious puberty. A German study reported a weak association between $PM_{2.5}$ exposure and decreased estradiol level in girls in Munich [13].

There may be several explanations for the inconsistent findings across studies. For example, in comparison to the Hong Kong study, which used the age of 11 years as the diagnostic threshold for pubertal stage, our study applied a younger diagnostic age (<7.5 years for girls and <9 years for boys), following the latest official recommendation in mainland China. This lower threshold may more effectively capture the early onset of pubertal development in children and the related effect size of $PM_{2.5}$. In addition, the inconsistent results of these studies may also be attributed to the differences in study area, population characteristics, outcome definitions, and other potential confounders like socioeconomic status, as well as exposure period and levels. This study provides supporting evidence for the adverse effect of childhood $PM_{2.5}$ exposure on precocious puberty among a nationwide sample in China.

The potential mechanisms underpinning the adverse effect of $PM_{2.5}$ on precocious puberty have been rarely discussed. One suggested hypothesis is the endocrine disruption caused by $PM_{2.5}$. Certain EDCs, such as phthalates, bisphenol analogs, and PAHs, have been detected as $PM_{2.5}$ components in several major Chinese cities [29-31]. These chemicals may target the estrogen receptor or kisspeptin in the hypothalamic nuclei and exhibit estrogenic or antiandrogenic activity, thereby promoting hypothalamic maturation and consequently causing precocious puberty [8,32-34]. In addition, some heavy metals found in $PM_{2.5}$, especially lead and cadmium, have also been proven to disrupt estrogen levels in female individuals and influence pubertal development [35,36].

The health effects of $PM_{2.5}$ depend on its chemical components. An increasing number of epidemiological studies have paid attention to the association between $PM_{2.5}$ chemical components and a range of health outcomes, such as pregnancy disorders, lung function, mortality, and body weight [37-40]. To the best of our knowledge, our study is the first to report associations of $PM_{2.5}$ components with precocious puberty. We found that OM is the most important component of $PM_{2.5}$ associated with precocious puberty in China. OM is produced by the incomplete combustion of solid fuels, which have higher toxicity than noncombustion aerosols [41]. In China, OM is the primary component of $PM_{2.5}$ in several major cities [42], posing a significant health risk to the population. Although the differential impact of components of $PM_{2.5}$ on precocious puberty has not yet been adequately examined, some underlying mechanisms may explain the observed effects. As previously mentioned, the chemical components in $PM_{2.5}$ that exert endocrine-disrupting effects are predominantly OM, such as PAHs. In addition, OM has the potential to induce mutagenic, inflammatory, and lipid metabolic disturbances in vivo [43-45]. However, without data at the cellular and molecular levels, it is beyond the scope of this study to speculate further about the potential pathways mediating the impact of OM on pubertal development.

Previous studies have suggested that the effects of $PM_{2.5}$ on children's pubertal development may differ by gender [13,15]. Consistent with this, our study observed a more significant association between $PM_{2.5}$ and precocious puberty in girls than in boys. It has been speculated that gender differences are due to the varying effects of PM exposure on different hormone levels in vivo [46], as well as the fact that prepubertal girls are particularly sensitive to low levels of exogenous estrogens [47]. In addition, gender differences may be related to the criteria for measuring secondary sexual characteristics, which depend on physical examinations of breast development and pubic hair in girls and testicle volume and pubic hair in boys, rather than on sex hormones and gonad B-ultrasounds.

Significant modifying effects of family income and sugary drink consumption were observed. Previously, several studies explored potential modifiers of the adverse effect of long-term $PM_{2.5}$ exposure, although few have focused on pubertal development. For example, a large cohort study in Canada found that the impact of $PM_{2.5}$ on mortality was greater for low-income earners than for high-income earners [48]. A study in the United States indicated that the cardiovascular risks associated with $PM_{2.5}$ exposure were higher among low-income groups compared to high-income groups [49]. The inequalities in $PM_{2.5}$ exposure between high-income populations and low-income populations may partly explain the modifying effect of family income.

Pubertal timing is sensitive to nutritional regulation; for example, girls with the highest levels of dietary isoflavone intake may experience early onset of breast development [50]. Some animal studies have found that fructose diet–fed rats had lower plasma testosterone levels, which can lead to a high sensitivity of the pituitary-gonadal axis in Leydig cells [51]. It can be speculated that dietary factors may potentially involve pathways that overlap with $PM_{2.5}$ and promote or hinder the effects of $PM_{2.5}$ on precocious puberty. However, there is still a lack of evidence to explain why significant modifying effects of family income and sugary drink consumption were observed more prominently in boys than in girls. Due to the absence of detailed information, further exploration is not feasible. Future studies may be able to elucidate the underlying mechanisms.

Strengths and Limitations

This study has several strengths that can be mentioned. First, our findings came from a survey covering 30 cities in 11 provinces, improving the generalizability when compared to studies with a smaller geographical scale, such as a single city. Second, the PRODY survey incorporated measurements of breast development and testicle volume conducted by medical professionals, resulting in a more accurate diagnosis of precocious puberty compared to studies using self-reported or parent-reported data. Third, we analyzed the effects of joint exposure to $PM_{2.5}$ components on precocious puberty using WQS regression, which can identify major contributors and therefore provide targeted recommendations for air quality–related health improvement.

We acknowledge some limitations of our analysis. In this study, air pollutant concentrations at the school level may not fully reflect all sources of air pollution, such as indoor air pollution within households, potentially leading to exposure misclassification. Because of the large sample size in this study and the limitations imposed by survey costs, we were unable to measure children's air pollution exposure at a finer scale. Future studies are strongly encouraged to consider a more holistic approach by incorporating direct measurements of indoor air quality, possibly using portable air monitoring devices. Due to ethical considerations, only children willing to undergo physical examinations were included in the study, which might have introduced a certain sampling bias. Without access to data on appropriate survey weights, we were unable to evaluate the accurate generalization of our findings to the whole target population of the country, which warrants further exploration when data are available in the future. Due to the cross-sectional study design, the inference of causality was limited. Furthermore, owing to constraints in data availability, our study did not account for the adverse effects of some other toxic components (ie, heavy metals) or confounding variables such as parental history of precocious puberty, sleep patterns, and psychological status. We suggest that future research should include these additional variables to better understand their potential impacts on the results.

Conclusion

Our study suggests that OM may be a major contributor to the association between long-term exposure to $PM_{2.5}$ and precocious puberty. Our findings further endorse the call for reducing fossil fuel emissions, concurrently advocating the development of specific and effective strategies to protect children's developmental health.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2021YFC2701900 and 2016YFC1305300), Key R&D Program of Zhejiang (2023C03047), National Natural Science Foundation of China (82370863, 81570759, and 81270938), Natural Science Foundation of Shandong Province in China (ZR2021QH318), Shandong Excellent Young Scientists Fund Program (Overseas) (2022HWYQ-055), Taishan Young Scholar Program of Shandong Province (tsqn202306063), and Key R&D Projects of Jining Science and Technology Bureau (2022YXNS086).

Data Availability

The data sets generated and analyzed during this study are available from the corresponding authors upon reasonable request.

Authors' Contributions

X Zhou contributed to data curation, conceptualization, and validation. X Zhang contributed to methodology, formal analysis, visualization, and writing (original draft). GB and WW contributed to conceptualization. KH, GD, and XL contributed to data curation. RC, SC, RZ, CW, HW, BC, YL, HY, ZS, MM, FL, PL, MZ, HD, YY, LC, JW, and JY contributed to investigation. ZL contributed to funding acquisition and conceptualization. QZ and JF contributed to conceptualization, supervision, and writing (review and editing). All authors have read and approved the final version of the manuscript. All authors have final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary materials showing the distribution of air pollutant concentrations and the study population during the study period, a comparison of characteristics between the original recruited population and the study population, and the results of the sensitivity analysis.

[DOCX File, 615 KB - publichealth_v11i1e62861_app1.docx]

References

1. Latronico AC, Brito VN, Carel JC. Causes, diagnosis, and treatment of central precocious puberty. Lancet Diabetes Endocrinol 2016 Mar;4(3):265-274. [doi: 10.1016/S2213-8587(15)00380-0] [Medline: 26852255]
2. Bräuner EV, Busch AS, Eckert-Lind C, Koch T, Hickey M, Juul A. Trends in the incidence of central precocious puberty and normal variant puberty among children in Denmark, 1998 to 2017. JAMA Netw Open 2020 Oct 1;3(10):e2015665. [doi: 10.1001/jamanetworkopen.2020.15665] [Medline: 33044548]
3. Kim YJ, Kwon A, Jung MK, et al. Incidence and prevalence of central precocious puberty in Korea: an epidemiologic study based on a national database. J Pediatr 2019 May;208:221-228. [doi: 10.1016/j.jpeds.2018.12.022] [Medline: 30857777]
4. Soriano-Guillén L, Corripio R, Labarta JI, et al. Central precocious puberty in children living in Spain: incidence, prevalence, and influence of adoption and immigration. J Clin Endocrinol Metab 2010 Sep;95(9):4305-4313. [doi: 10.1210/jc.2010-1025] [Medline: 20554707]

5. Liang X, Huang K, Dong G, et al. Current pubertal development in Chinese children and the impact of overnutrition, lifestyle, and perinatal factors. J Clin Endocrinol Metab 2023 Aug 18;108(9):2282-2289. [doi: 10.1210/clinem/dgad102] [Medline: 36881937]
6. Shi L, Jiang Z, Zhang L. Childhood obesity and central precocious puberty. Front Endocrinol (Lausanne) 2022;13:1056871. [doi: 10.3389/fendo.2022.1056871] [Medline: 36465655]
7. Huang JV, Leung GM, Schooling CM. The association of air pollution with pubertal development: evidence from Hong Kong's "children of 1997" birth cohort. Am J Epidemiol 2017 May 15;185(10):914-923. [doi: 10.1093/aje/kww200] [Medline: 28444106]
8. Brito VN, Canton APM, Seraphim CE, et al. The congenital and acquired mechanisms implicated in the etiology of central precocious puberty. Endocr Rev 2023 Mar 4;44(2):193-221. [doi: 10.1210/endrev/bnac020] [Medline: 35930274]
9. WHO global air quality guidelines: particulate matter (PM25 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240034228 [accessed 2025-01-14]
10. Fang B, Bravo MA, Wang H, et al. Polycyclic aromatic hydrocarbons are associated with later puberty in girls: a longitudinal study. Sci Total Environ 2022 Nov 10;846:157497. [doi: 10.1016/j.scitotenv.2022.157497] [Medline: 35868395]
11. Zhou Q, Chen J, Zhang J, et al. Toxicity and endocrine-disrupting potential of PM2.5: association with particulate polycyclic aromatic hydrocarbons, phthalate esters, and heavy metals. Environ Pollut 2022 Jan 1;292(Pt A):118349. [doi: 10.1016/j.envpol.2021.118349] [Medline: 34653588]
12. Gore AC, Chappell VA, Fenton SE, et al. EDC-2: the Endocrine Society's second scientific statement on endocrine-disrupting chemicals. Endocr Rev 2015 Dec;36(6):E1-E150. [doi: 10.1210/er.2015-1010] [Medline: 26544531]
13. Zhao T, Triebner K, Markevych I, et al. Outdoor air pollution and hormone-assessed pubertal development in children: results from the GINIplus and LISA birth cohorts. Environ Int 2021 Jul;152:106476. [doi: 10.1016/j.envint.2021.106476] [Medline: 33714142]
14. Wronka I, Kliś K. Effect of air pollution on age at menarche in Polish females, born 1993-1998. Sci Rep 2022 Mar 21;12(1):4820. [doi: 10.1038/s41598-022-08577-3] [Medline: 35315430]
15. Oh J, Choi JE, Lee R, et al. Long-term exposure to air pollution and precocious puberty in South Korea. Environ Res 2024 Jul 1;252(Pt 2):118916. [doi: 10.1016/j.envres.2024.118916] [Medline: 38614201]
16. Xiao Q, Geng G, Liu S, Liu J, Meng X, Zhang Q. Spatiotemporal continuous estimates of daily 1 km PM 2.5 from 2000 to present under the Tracking Air Pollution in China (TAP) framework. Atmos Chem Phys 2022;22(19):13229-13242. [doi: 10.5194/acp-22-13229-2022]
17. Huang K, Si S, Chen R, et al. Preterm birth and birth weight and the risk of type 1 diabetes in Chinese children. Front Endocrinol (Lausanne) 2021;12:603277. [doi: 10.3389/fendo.2021.603277] [Medline: 33935963]
18. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. Bull World Health Organ 2007 Sep;85(9):660-667. [doi: 10.2471/blt.07.043497] [Medline: 18026621]
19. Subspecialty Group of Endocrinologic, Hereditary and Metabolic Diseases, the Society of Pediatrics, Chinese Medical Association, Editorial Board, Chinese Journal of Pediatrics. Expert consensus on the diagnosis and treatment of central precocious puberty(2022). Zhonghua Er Ke Za Zhi 2023 Jan 2;61(1):16-22. [doi: 10.3760/cma.j.cn112140-20220802-00693] [Medline: 36594116]
20. Emmanuel M, Bokor BR. Tanner stages. In: StatPearls: StatPearls Publishing; 2024. [Medline: 29262142]
21. TAP: tracking air pollution in China. TAPData. URL: http://tapdata.org.cn [accessed 2025-01-08]
22. Liu S, Geng G, Xiao Q, et al. Tracking daily concentrations of PM 2.5 chemical composition in China since 2000. Environ Sci Technol 2022 Nov 15;56(22):16517-16527. [doi: 10.1021/acs.est.2c06510] [Medline: 36318737]
23. Geng G, Xiao Q, Liu S, et al. Tracking air pollution in China: near real-time PM2.5 retrievals from multisource data fusion. Environ Sci Technol 2021 Sep 7;55(17):12106-12115. [doi: 10.1021/acs.est.1c01863] [Medline: 34407614]
24. Carrico C, Gennings C, Wheeler DC, Factor-Litvak P. Characterization of weighted quantile sum regression for highly correlated data in a risk analysis setting. J Agric Biol Environ Stat 2015 Mar;20(1):100-120. [doi: 10.1007/s13253-014-0180-3] [Medline: 30505142]
25. Keil AP, Buckley JP, O'Brien KM, Ferguson KK, Zhao S, White AJ. A quantile-based g-computation approach to addressing the effects of exposure mixtures. Environ Health Perspect 2020 Apr;128(4):47004. [doi: 10.1289/EHP5838] [Medline: 32255670]
26. Zhu M, Fu J, Liang L, et al. Epidemiologic study on current pubertal development in Chinese school-aged children. Zhejiang Da Xue Xue Bao Yi Xue Ban 2013 Jul;42(4):396-402. [doi: 10.3785/j.issn.1008-9292.2013.04.005] [Medline: 24022927]
27. Teilmann G, Pedersen CB, Jensen TK, Skakkebaek NE, Juul A. Prevalence and incidence of precocious pubertal development in Denmark: an epidemiologic study based on national registries. Pediatrics 2005 Dec;116(6):1323-1328. [doi: 10.1542/peds.2005-0012] [Medline: 16322154]
28. Jung EM, Kim HS, Park H, Ye S, Lee D, Ha EH. Does exposure to PM10 decrease age at menarche? Environ Int 2018 Aug;117:16-21. [doi: 10.1016/j.envint.2018.04.020] [Medline: 29704753]

29. Duan L, Yu H, Wang Q, et al. PM2.5-bound polycyclic aromatic hydrocarbons of a megacity in eastern China: Source apportionment and cancer risk assessment. Sci Total Environ 2023 Apr 15;869:161792. [doi: 10.1016/j.scitotenv.2023.161792] [Medline: 36702280]
30. Zhang X, Wang Q, Qiu T, et al. PM2.5 bound phthalates in four metropolitan cities of China: concentration, seasonal pattern and health risk via inhalation. Sci Total Environ 2019 Dec 15;696:133982. [doi: 10.1016/j.scitotenv.2019.133982] [Medline: 31470327]
31. Munyaneza J, Qaraah FA, Jia Q, Cheng H, Zhen H, Xiu G. Seasonal trends, profiles, and exposure risk of pm2.5-bound bisphenol analogs in ambient outdoor air: a study in Shanghai, China. Aerosol Air Qual Res 2022;22(9):210324. [doi: 10.4209/aaqr.210324]
32. Lopez-Rodriguez D, Franssen D, Heger S, Parent AS. Endocrine-disrupting chemicals and their effects on puberty. Best Pract Res Clin Endocrinol Metab 2021 Sep;35(5):101579. [doi: 10.1016/j.beem.2021.101579] [Medline: 34563408]
33. Zhang Y, Dong S, Wang H, Tao S, Kiyama R. Biological impact of environmental polycyclic aromatic hydrocarbons (ePAHs) as endocrine disruptors. Environ Pollut 2016 Jun;213:809-824. [doi: 10.1016/j.envpol.2016.03.050] [Medline: 27038213]
34. Lopez-Rodriguez D, Franssen D, Bakker J, Lomniczi A, Parent AS. Cellular and molecular features of EDC exposure: consequences for the GnRH network. Nat Rev Endocrinol 2021 Feb;17(2):83-96. [doi: 10.1038/s41574-020-00436-3] [Medline: 33288917]
35. Gerald AC, Ganapathy S, Zhu J, Wei Y. Exposure to endocrine-disrupting metals and serum estrogen levels among US women. Reprod Toxicol 2023 Jun;118:108392. [doi: 10.1016/j.reprotox.2023.108392] [Medline: 37150308]
36. Li F, Yan J, Wei Y, et al. PM2.5-bound heavy metals from the major cities in China: spatiotemporal distribution, fuzzy exposure assessment and health risk management. J Clean Prod 2021 Mar;286:124967. [doi: 10.1016/j.jclepro.2020.124967]
37. Shen Y, Yu G, Liu C, et al. Prenatal exposure to PM2.5 and its specific components and risk of hypertensive disorders in pregnancy: a nationwide cohort study in China. Environ Sci Technol 2022 Aug 16;56(16):11473-11481. [doi: 10.1021/acs.est.2c01103] [Medline: 35914180]
38. Zhang Y, Yin Z, Zhou P, et al. Early-life exposure to PM2.5 constituents and childhood asthma and wheezing: findings from China, Children, Homes, Health study. Environ Int 2022 Jul;165:107297. [doi: 10.1016/j.envint.2022.107297] [Medline: 35709580]
39. Guo Q, Zhang K, Wang B, et al. Chemical constituents of ambient fine particulate matter and obesity among school-aged children: a representative national study in China. Sci Total Environ 2022 Nov;849:157742. [doi: 10.1016/j.scitotenv.2022.157742] [Medline: 35917963]
40. Raaschou-Nielsen O, Antonsen S, Agerbo E, et al. PM2.5 air pollution components and mortality in Denmark. Environ Int 2023 Jan;171:107685. [doi: 10.1016/j.envint.2022.107685] [Medline: 36502699]
41. Park M, Joo HS, Lee K, et al. Differential toxicities of fine particulate matters from various sources. Sci Rep 2018 Nov 19;8(1):17007. [doi: 10.1038/s41598-018-35398-0] [Medline: 30451941]
42. Weagle CL, Snider G, Li C, et al. Global sources of fine particulate matter: interpretation of PM2.5 chemical composition observed by SPARTAN using a global chemical transport model. Environ Sci Technol 2018 Oct 16;52(20):11670-11681. [doi: 10.1021/acs.est.8b01658] [Medline: 30215246]
43. DeMarini DM, Warren SH, Brooks LR. Mutagenicity of the organic fraction of World Trade Center dust. Environ Mol Mutagen 2023 Jan;64(1):16-25. [doi: 10.1002/em.22519] [Medline: 36433931]
44. Jiang X, Han Y, Qiu X, et al. Organic components of personal PM2.5 exposure associated with inflammation: evidence from an untargeted exposomic approach. Environ Sci Technol 2021 Aug 3;55(15):10589-10596. [doi: 10.1021/acs.est.1c02023] [Medline: 34297563]
45. Zhang SY, Shao D, Liu H, et al. Metabolomics analysis reveals that benzo[a]pyrene, a component of PM2.5, promotes pulmonary injury by modifying lipid metabolism in a phospholipase A2-dependent manner in vivo and in vitro. Redox Biol 2017 Oct;13:459-469. [doi: 10.1016/j.redox.2017.07.001] [Medline: 28715731]
46. Wang L, Luo D, Liu X, et al. Effects of PM2.5 exposure on reproductive system and its mechanisms. Chemosphere 2021 Feb;264(Pt 1):128436. [doi: 10.1016/j.chemosphere.2020.128436] [Medline: 33032215]
47. Aksglaede L, Juul A, Leffers H, Skakkebaek NE, Andersson AM. The sensitivity of the child to sex steroids: possible impact of exogenous estrogens. Hum Reprod Update 2006;12(4):341-349. [doi: 10.1093/humupd/dml018] [Medline: 16672247]
48. Chen H, Quick M, Kaufman JS, et al. Impact of lowering fine particulate matter from major emission sources on mortality in Canada: a nationwide causal analysis. Proc Natl Acad Sci U S A 2022 Dec 6;119(49):e2209490119. [doi: 10.1073/pnas.2209490119] [Medline: 36442082]
49. Luo J, Craver A, Jin Z, et al. Contextual deprivation, race and ethnicity, and income in air pollution and cardiovascular disease. JAMA Netw Open 2024 Aug 1;7(8):e2429137. [doi: 10.1001/jamanetworkopen.2024.29137] [Medline: 39158908]
50. Cheng G, Buyken AE, Shi L, et al. Beyond overweight: nutrition as an important lifestyle factor influencing timing of puberty. Nutr Rev 2012 Mar;70(3):133-152. [doi: 10.1111/j.1753-4887.2011.00461.x] [Medline: 22364156]

51. Hsia SM, Chiang YF, Chen HY, Ali M, Wang PS, Wang KL. Effect of high-fructose diet-induced metabolic syndrome on the pituitary-gonadal axis in male rats. Biomedicines 2022 Nov 23;10(12):3009. [doi: 10.3390/biomedicines10123009] [Medline: 36551765]

Abbreviations

BC: black carbon
EDC: endocrine-disruptor chemical
NH_4^+: ammonium
NO_3^-: nitrate
OM: organic matter
OR: odds ratio
PAH: polycyclic aromatic hydrocarbon
$PM_{2.5}$: fine particulate matter
PRODY: Prevalence and Risk Factors for Obesity and Diabetes in Youth
Q: quartile
QGC: quantile-based g-computation
SO_4^{2-}: sulfate
WHO: World Health Organization
WQS: weighted quantile sum

Edited by A Mavragani, T Sanchez; submitted 03.06.24; peer-reviewed by J Ma, P Tang; revised version received 12.12.24; accepted 12.12.24; published 07.02.25.

Please cite as:
Zhou X, Zhang X, Bai G, Dong G, Li X, Chen R, Chen S, Zheng R, Wang C, Wei H, Cao B, Liang Y, Yao H, Su Z, Maimaiti M, Luo F, Li P, Zhu M, Du H, Yang Y, Cui L, Wang J, Yuan J, Liu Z, Wu W, Zhao Q, Fu J
Long-Term Exposure to Fine Particulate Matter (PM2.5) Components and Precocious Puberty Among School-Aged Children: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e62861
URL: https://publichealth.jmir.org/2025/1/e62861
doi:10.2196/62861

Interpreting the Influence of Using Blood Donor Residual Samples for SARS-CoV-2 Seroprevalence Studies in Japan: Cross-Sectional Survey Study

Ryo Kinoshita[1], PhD; Sho Miyamoto[2], DVM, PhD; Tadaki Suzuki[2], MD, PhD; Motoi Suzuki[1], MD, PhD; Daisuke Yoneoka[1], PhD

[1]Center for Surveillance, Immunization, and Epidemiologic Research, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-Ku, Tokyo 162-0052 Japan, Tokyo, Japan

[2]Department of Pathology, National Institute of Infectious Diseases, Tokyo, Japan

Corresponding Author:
Daisuke Yoneoka, PhD
Center for Surveillance, Immunization, and Epidemiologic Research, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-Ku, Tokyo 162-0052 Japan, Tokyo, Japan

Abstract

Background: Residual blood donor samples are commonly used in SARS-CoV-2 seroepidemiological studies; however their use may introduce bias due to the healthy donor effect, wherein blood donors are generally healthier than the general population. This potential bias is critical for accurately interpreting seroepidemiological data, as blood donors might not fully represent broader population-level infection rates. **Objective:** This study aims to assess the potential bias in SARS-CoV-2 seroprevalence estimates derived from blood donor samples in Japan by examining the association between blood donation history and COVID-19 diagnosis. By quantifying the healthy donor effect, we seek to refine the interpretation of SARS-CoV-2 seroepidemiological studies using residual blood donor samples. **Methods:** We conducted a web-based survey from December 14 to 28, 2023, recruiting 10,781 Japanese residents aged 16 - 69, stratified by demographic factors to match national representation. Participants provided information on demographics, socioeconomic status, COVID-19 vaccination history, comorbidities, and blood donation experience. A logistic regression model adjusting for confounders such as age, sex, education, occupation, comorbidities, and vaccination status, was used to estimate the odds of COVID-19 diagnosis among blood donors compared to nondonors. **Results:** Of the 10,781 participants, 3583 (33.2%) reported a history of COVID-19 diagnosis, and 5015 (46.5%) indicated they had donated blood at least once in their lifetime, and 1128 (10.5%) donated within the last year. Blood donors had mean of 13.5 (SD 43.6) donations and were older, with a mean age of 46.4 (SD 13.9) years, compared to 38.5 (SD 14.1) years for nondonors. Among blood donors, 39.9% had comorbidities (95% CI 38.5 - 41.2) compared to 27.9% (95% CI 26.7 - 29.0) of nondonors. Blood donors had 1.62 (95% CI: 1.48 - 1.78) times higher odds of COVID-19 diagnosis compared to nondonors. The higher diagnosis rate among blood donors likely reflects increased social interactions and health-seeking behaviors, a phenomenon we refer to as the inverse healthy donor effect. This suggests that blood donor samples could overestimate SARS-CoV-2 seroprevalence when generalized to the broader Japanese population. **Conclusions:** Higher COVID-19 diagnosis rates among blood donors may reflect increased community involvement and health-seeking behaviors, suggesting an inverse healthy donor effect. This pattern indicates that in terms of SARS-CoV-2 infection, blood donors might not represent the healthiest segment of the population. Consequently, seroprevalence studies using blood donor samples could overestimate SARS-CoV-2 infection rates in the general Japanese population. For more accurate public health surveillance, the development of statistical methods to adjust for this bias is recommended.

(JMIR Public Health Surveill 2025;11:e60467) doi:10.2196/60467

KEYWORDS

SARS-CoV-2; COVID-19; seroprevalence; blood donor; selection bias; healthy donor effect; coronavirus; pandemic; Japan; cross-sectional study; residual blood; epidemiology; blood donation; web-based; logistic regression; social economic; comorbidity; COVID-19 vaccination; public health

Introduction

The COVID-19 pandemic, triggered by the rapid dissemination of SARS-CoV-2, has highlighted the critical need for robust epidemiological surveillance to decipher the virus's transmission dynamics and shape effective public health strategies. As the pandemic intensified, the exponential increase in cases coupled with the presence of mild and asymptomatic infections strained traditional case-based surveillance systems. These systems primarily capture symptomatic cases seeking healthcare, leaving a substantial proportion of the infected population—particularly those with mild or asymptomatic infections—undetected. This complication led to the adoption of seroprevalence studies, which estimate the extent of infection by measuring the presence of nucleocapsid or spike antibodies in populations to provide a broader estimation of infection spread, inclusive of undiagnosed cases.

While seroprevalence studies are invaluable for assessing SARS-CoV-2 exposure at the population level, they are subject to inherent biases [1,2]. For example, antibody levels can wane over time, and variations in assay sensitivity or specificity may underestimate or overestimate the seroprevalence. When residual blood samples from blood donors are used, an additional potential bias known as the healthy donor effect is introduced [3]. This effect arises because blood donors are typically healthier than the general population [4-7], potentially resulting in an underestimation of true seroprevalence. Although blood donor–based SARS-CoV-2 seroepidemiological studies have been widely conducted across various countries [8-11], these studies have not fully accounted for the extent of this effect. Recognizing and addressing such bias is critical to improving the reliability of seroprevalence estimates and their use in public health planning.

Despite recognition of the healthy donor effect, its quantitative impact on SARS-CoV-2 seroprevalence estimates remains largely unexplored, introducing uncertainty in interpreting findings derived from blood donor samples [12]. Our study addresses this gap by focusing on the quantification of the healthy donor effect within Japan's blood donor population, with the goal of improving the interpretation of SARS-CoV-2 seroprevalence estimates and providing a more accurate assessment of the population's exposure to SARS-CoV-2.

Methods

Study Participants and Questionnaire

The study surveyed Japanese residents aged 16 - 69 years—the age group eligible for blood donation in Japan—from all 47 prefectures who could respond in Japanese. The sample size was set at 10,829, based on an approximately 6.1% national blood donation rate reported in a 2022 Japanese Red Cross survey [13], with a 5% alpha error and 80% power. To ensure national representation, a quota-controlled sampling approach was applied, stratified by age, sex, and region, with quotas informed by the 2020 national census. Participants were recruited from Cross Marketing Inc.'s (Japan) panel of 5 million active and diverse members as of March 2024, who volunteered and were incentivized with point-based rewards redeemable for goods and services from affiliated businesses upon completion of the questionnaire. Panel members were accepted on a first-come-first-served basis until quotas for each age and prefecture category were met, ensuring coverage across all 47 prefectures, including both urban and rural populations. The survey was conducted from November 14 to 28, 2023, in alignment with a national seroepidemiological survey among blood donors in Japan. To ensure data completeness, respondents were requested to answer all questions.

The online questionnaire was developed based on a comprehensive literature review on blood donation and epidemiological surveys. For questions related to blood donation, we referenced validated items from governmental surveys [14]. The survey items were divided into three modules: (1) sociodemographic characteristics such as sex, age at the survey, residential prefecture, occupation, education level, and household income; (2) blood donation–related characteristics such as blood donation experience (yes or no) and the total number of blood donations; and (3) COVID-19-related and clinical questions such as the number and timing of prior COVID-19 infections, the number of COVID-19 vaccinations, and medical history. The definitions for COVID-19 vaccination, blood donation, and comorbidity are provided in Multimedia Appendix 1. A binary variable was created from medical history records to identify comorbidities that rendered participants ineligible for blood donation and medically unable to receive the COVID-19 vaccine (see Multimedia Appendix 2). These comorbidities included current treatment for periodontal disease or dental caries, a diagnosis of high blood pressure, diabetes, asthma, or chronic obstructive pulmonary disease, and a history of angina, myocardial infarction, stroke, or cancer.

Statistical Analysis

First, sociodemographic data were tabulated separately for participants with and without blood donation experience to provide basic information on the study population. Continuous variables were analyzed using Mann-Whitney U test and categorical variables were analyzed using Fisher exact test. To estimate the association between COVID-19 infection and blood donation experience, a logistic regression model was used and the associated odds ratio was estimated. Based on a directed acyclic graph, these confounders were adjusted: age groups (16-29, 30-39, 40-49, 50-59, 60-69 years), sex, residential region, education level, occupation, number of blood donations, number of vaccinations, and comorbidities (Multimedia Appendix 3). As behavioral changes induced by vaccination are expected to differ by age group, an interaction term between age groups, and the number of vaccinations was also adjusted in the model. Further, to explore the model robustness, we conducted subgroup analyses using Firth logistic regression to account for potential bias due to small sample sizes or sparse data. Model performance was evaluated by calculating c-statistics, performing the Hosmer-Lemeshow test, and assessing multicollinearity using the adjusted generalized variance inflation factor (GVIF), where GVIF $^{1/(2 \times df)}$, with df representing the degrees of freedom. All data analyses were conducted using R software (version 4.3.2; R Foundation for Statistical Computing).

Ethical Considerations

This study was reviewed and approved by the Institutional Review Board at the National Institute of Infectious Diseases (authorization no. 1579). Informed consent was obtained from all participants within the web survey, with only those providing consent allowed to proceed. All data used in the analysis were fully anonymized by Cross Marketing Inc. before being provided to the researchers. Participants received point-based rewards as compensation for survey completion, in line with institutional review board–approved ethical guidelines.

Results

The main characteristics of the study participants are presented in Multimedia Appendix 4. A total of 10,781 participants were enrolled in Japan during the study period from December 14 to 28, 2023. Of these, 3583 (33.2%) had a history of COVID-19 diagnosis, and 5015 (46.5%) reported having donated blood at least once in their lifetime. Furthermore, 1128 (10.5%) had donated blood within the past year (Multimedia Appendix 5). Blood donors (defined as those with a history of blood donation) reported mean 13.5 (SD 43.6) donations. . Multimedia Appendix 6 shows the detailed distribution of blood donation frequency in our web survey. The mean age of blood donors was 46.4 (SD 13.9) years, whereas nondonors were younger, with a mean age of 38.5 (SD 14.1) years. Among blood donors, 39.9% (95%CI 38.5 - 41.2) had comorbidities, compared to 27.9% (95% CI 26.7 - 29.0) of nonblood donors (see Multimedia Appendix 7).

Figure 1A depicts the proportion of blood donors categorized by age group and geographic region in Japan. The corresponding geographic regions are illustrated in Figure 1B. In general, individuals who have been diagnosed with COVID-19 had a higher prevalence of prior blood donation compared to those without a history of diagnosis. The estimated results of the logistic regression model shown in Table 1. After excluding individuals with comorbidities that render them ineligible to donate blood or receive the COVID-19 vaccine, and after adjusting for potential confounders, we found that the odds of infection among blood donors were 1.62 times greater (95%CI 1.48 - 1.78) than nondonors. Overall, the positive odds ratio was similar across most subgroup analysis with different stratifications, with most results being statistically significant (see Multimedia Appendix 8). Model validation resulted in a c-statistic of 0.66, a Hosmer-Lemeshow test value of χ^2=6.78 (*P*=.56), and adjusted GVIF values below 5, with a mean of 1.34.

Figure 1. Proportion of blood donors and COVID-19 diagnosis status among Japanese residents aged 16 - 69 years, based on a web-based survey conducted across all 47 prefectures in Japan from December 14 to 28, 2023. (A) Overall proportion of participants with blood donation experience. (B) Map showing geographic regions represented in the study. (C) Proportion of blood donors among individuals diagnosed with COVID-19. (D) Proportion of blood donors among uninfected (undiagnosed) individuals.

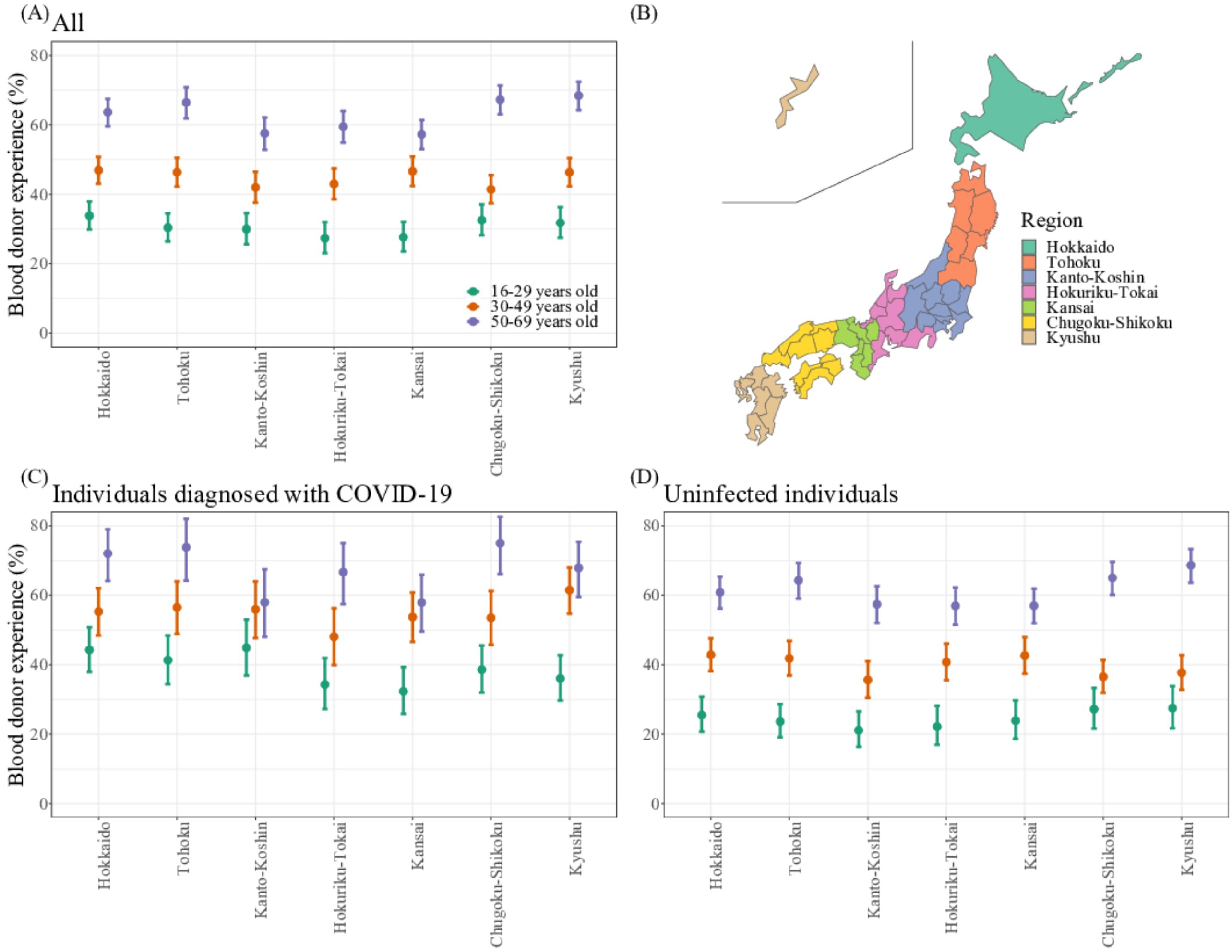


Table . Logistic regression results examining the association between COVID-19 diagnosis (outcome) and blood donor experience (main variable of interest) among Japanese residents. Odds ratios with 95% CI and *P* values are provided. Additional variables, including demographic, socioeconomic, and health-related factors, were included as adjustment variables based on a directed acyclic graph (DAG) to control for potential confounding. Vaccination status, household income, and blood donor frequency were included as continuous variables.

Variables	Odds Ratio (OR)	(95% CI)	*P* value
Age group (years)			
16 - 29 (Ref)[a]	1.00	_[b]	–
30 - 39	0.88	(0.67 - 1.14)	.37
40 - 49	0.59	(0.48 - 0.74)	<.001
50 - 59	0.59	(0.46 - 0.77)	<.001
60 - 69	0.55	(0.40 - 0.77)	<.001
Vaccination status	1.05	(1.01-1.10)	.04
Sex			
Female (Ref)[a]	1.00	–	–
Male	0.90	(0.82 - 0.99)	.05
Region			
Hokkaido	1.13	(0.97 - 1.32)	.14
Tohoku	0.97	(0.82 - 1.14)	.73
Kanto-Koshin (Ref)[a]	1.00	–	–
Hokuriku-Tokai	1.11	(0.94 - 1.31)	.27
Kansai	1.28	(1.09 - 1.50)	.004
Chugoku-Shikoku	1.07	(0.91 - 1.26)	.47
Kyushu	1.39	(1.19 - 1.62)	<.001
Highest level of education			
Middle school / High school	1.01	(0.92 - 1.11)	.85
Junior. college / Vocational school / University (Ref)	1.00	–	–
Graduate school (Master / PhD)	0.72	(0.58 - 0.89)	.004
Occupation			
Commerce	1.72	(1.44 - 2.06)	<.001
Construction / Manufacturing / Transportation	1.75	(1.46 - 2.09)	<.001
Education / Student	2.39	(1.96 - 2.92)	<.001
Food / Beverage / Accommodation	1.51	(1.10 - 2.06)	.01
Homemaker	1.88	(1.49 - 2.37)	<.001
Information / Communication	1.92	(1.52 - 2.44)	<.001
Medical / Social welfare	2.53	(2.05 - 3.13)	<.001
Primary industries	1.75	(1.15 - 2.65)	.01
Public servant	2.22	(1.77 - 2.80)	<.001
Other	1.38	(1.10 - 1.75)	.01
Unemployed (Ref)	1.00	–	–
Household Income	1.06	(1.04 - 1.07)	<.001
Blood donor experience			
No (Ref)[a]	1.00	–	–
Yes	1.62	(1.48 - 1.78)	<.001

Variables	Odds Ratio (OR)	(95% CI)	*P* value
Blood donor frequency	1.00	(1.00 - 1.00)	.85
Comorbidity			
No (Ref)[a]	1.00	–	–
Yes	1.40	(1.28 - 1.53)	<.001
Age group (years) * Vaccination Status			
Age group 16 - 29 * Vaccination status (Ref)[a]	1.00	–	–
Age group 30 - 39 * Vaccination status	0.93	(0.86 - 1.01)	.12
Age group 40 - 49 * Vaccination status	0.95	(0.89 - 1.02)	.16
Age group 50 - 59 * Vaccination status	0.87	(0.81 - 0.94)	<.001
Age group 60 - 69 * Vaccination status	0.86	(0.79 - 0.93)	<.001

[a]Ref: Reference variable.

[b]–: Not available.

Discussion

Our study showed a significant association between blood donation experience and COVID-19 diagnosis. However, the diagnosis rates did not increase with the frequency of donations, suggesting that the act of donating blood itself did not contribute to the transmission risk. Instead the higher incidence of COVID-19 among donors may be attributed to increased social interactions and greater health awareness, which could be characterized as an ‘inverse’ healthy donor effect [15]. Donors’ community involvement could raise exposure risks, while their health vigilance could lead to more frequent SARS-CoV-2 testing, leading to higher COVID-19 diagnosis. Further studies are needed to understand the potential underlying mechanism behind this association.

We found that using blood donor samples could potentially bias seroprevalence estimates in Japan. The observed positive association suggests that these samples may overestimate the actual rates compared to the general population. Previous efforts to adjust for demographic variances, such as age, gender, and prefecture, have not completely resolved the bias inherent in using blood donor samples [11]. This highlights the need to measure the healthy donor effect for accurate SARS-CoV-2 seroprevalence interpretations and emphasizes the challenge of correctly determining incidence from donor data, underscoring the need for improved research methodologies. Development of a statistical method to adjust for this bias of blood donor samples in estimating the seroprevalence of the general population of Japan using blood donor samples remains an ongoing focus of our research group.

There are a few limitations to note. First, the identification of COVID-19 cases was based on clinical diagnoses rather than antibody-confirmed infection rates, suggesting that the study may reflect differences in health-seeking behavior rather than actual SARS-CoV-2 infection rates. Second, although we applied quota sampling based on demographic distributions from the national census to improve representation, the sample population may not fully represent the general population in Japan, as the study was conducted via a web-based survey. This approach may introduce potential selection bias, particularly excluding individuals without internet access or those less comfortable with technology. Third, the study relied on self-reported data, which may introduce recall bias or social desirability bias. Fourth, we may not have fully adjusted for residual confounding from variables that were not included in the survey. Fifth, the cross-sectional design has inherent limitations, including difficulty determining temporal relationships and tracking changes over time. Finally, our findings may be specific to the Japanese context and might not necessarily apply to other countries due to differences in demographics, healthcare systems, and social behavior. Further evaluation is needed to assess the external validity across different countries or regions.

In conclusion, our study has identified a potential bias in SARS-CoV-2 seroprevalence studies that use residual samples from blood donors. Quantifying the direction and magnitude of this bias is essential for accurately interpreting seroprevalence surveys, especially when these surveys are used to guide infectious disease surveillance.

Acknowledgments

This study was supported by the Health Labour Sciences Research Grant from the Ministry of Health, Labour and Welfare (23HA2017, 24HA2016). RK and DY received funding from the Japan Society for the Promotion of Science (JSPS) KAKENHI (21K17307 and 21K17292).

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: RK, DY
Data curation: RK, DY
Formal analysis: RK, DY
Writing – original draft: RK, DY
Writing – review and editing: RK, SM, TS, MS, DY

Conflicts of Interest

None declared.

Multimedia Appendix 1
Description of COVID-19 vaccination and blood donation-related variables included in the model.
[DOCX File, 14 KB - publichealth_v11i1e60467_app1.docx]

Multimedia Appendix 2
Description of comorbidities that affect eligibility for blood donation in Japan, used to create the binary variable Comorbidity.
[DOCX File, 15 KB - publichealth_v11i1e60467_app2.docx]

Multimedia Appendix 3
Directed acyclic graph (DAG) illustrating the relationship between blood donor experience (exposure) and COVID-19 infection (outcome), with other variables included as potential confounders.
[DOCX File, 324 KB - publichealth_v11i1e60467_app3.docx]

Multimedia Appendix 4
Demographic, socioeconomic, and health characteristics of study participants in the web survey, categorized by COVID-19 diagnosis status (uninfected or infected) and blood donation status (nonblood donor or blood donor).
[DOCX File, 28 KB - publichealth_v11i1e60467_app4.docx]

Multimedia Appendix 5
Comparison of blood donation rates within 1 year among the sampled population (circle points) and Japanese Red Cross Society data for 2022 (star points), categorized by age group and region. The figure displays blood donation proportions across different age groups (16–29, 30–49, 50–69 years) and regions, with whiskers representing 95% confidence intervals.
[DOCX File, 58 KB - publichealth_v11i1e60467_app5.docx]

Multimedia Appendix 6
Frequency distribution of blood donation experience among web survey participants. (A) Number of blood donations over participants' entire lifetime. (B) Number of blood donations within the past year. The X-axis is presented on a logarithmic scale to better illustrate the range and variability in donation frequencies across the study population.
[DOCX File, 43 KB - publichealth_v11i1e60467_app6.docx]

Multimedia Appendix 7
Proportion of comorbidities among participants by blood donor status (0 = nondonor, 1 = donor). Whiskers represent 95% confidence intervals.
[DOCX File, 35 KB - publichealth_v11i1e60467_app7.docx]

Multimedia Appendix 8

Supplementary Table 4.
[DOCX File, 24 KB - publichealth_v11i1e60467_app8.docx]

References

1. Accorsi EK, Qiu X, Rumpler E, et al. How to detect and reduce potential sources of biases in studies of SARS-CoV-2 and COVID-19. Eur J Epidemiol 2021 Feb;36(2):179-196. [doi: 10.1007/s10654-021-00727-7] [Medline: 33634345]
2. Takahashi S, Greenhouse B, Rodríguez-Barraquer I. Are seroprevalence estimates for severe acute respiratory syndrome coronavirus 2 biased? J Infect Dis 2020 Nov 9;222(11):1772-1775. [doi: 10.1093/infdis/jiaa523] [Medline: 32856712]
3. Atsma F, Veldhuizen I, Verbeek A, de Kort W, de Vegt F. Healthy donor effect: its magnitude in health research among blood donors. Transfusion 2011 Aug;51(8):1820-1828. [doi: 10.1111/j.1537-2995.2010.03055.x] [Medline: 21342203]
4. Brodersen T, Rostgaard K, Lau CJ, et al. The healthy donor effect and survey participation, becoming a donor and donor career. Transfusion 2023 Jan;63(1):143-155. [doi: 10.1111/trf.17190] [Medline: 36479702]
5. O'Brien SF, Drews SJ, Lewin A, et al. How do we decide how representative our donors are for public health surveillance? Transfusion 2022 Dec;62(12):2431-2437. [doi: 10.1111/trf.17140] [Medline: 36193865]
6. Saeed S, Uzicanin S, Lewin A, et al. Current challenges of severe acute respiratory syndrome coronavirus 2 seroprevalence studies among blood donors: a scoping review. Vox Sang 2022 Apr;117(4):476-487. [doi: 10.1111/vox.13221] [Medline: 34862614]
7. Su S, Sun Y, Gu X, et al. Exploration of the healthy donor effect among 0.6 million blood donors in China: longitudinal study. JMIR Public Health Surveill 2024 Feb 22;10:e48617. [doi: 10.2196/48617] [Medline: 38386403]
8. Jones JM, Stone M, Sulaeman H, et al. Estimated US Infection- and Vaccine-Induced SARS-CoV-2 Seroprevalence Based on Blood Donations, July 2020-May 2021. JAMA 2021 Oct 12;326(14):1400-1409. [doi: 10.1001/jama.2021.15161] [Medline: 34473201]
9. Chang L, Hou W, Zhao L, et al. The prevalence of antibodies to SARS-CoV-2 among blood donors in China. Nat Commun 2021 Mar 2;12(1):1383. [doi: 10.1038/s41467-021-21503-x] [Medline: 33654063]
10. Erikstrup C, Laksafoss AD, Gladov J, et al. Seroprevalence and infection fatality rate of the SARS-CoV-2 Omicron variant in Denmark: A nationwide serosurveillance study. Lancet Reg Health Eur 2022 Oct;21:100479. [doi: 10.1016/j.lanepe.2022.100479] [Medline: 35959415]
11. Kinoshita R, Arashiro T, Kitamura N, et al. Infection-Induced SARS-CoV-2 Seroprevalence among Blood Donors, Japan, 2022. Emerg Infect Dis 2023 Sep;29(9):1868-1871. [doi: 10.3201/eid2909.230365] [Medline: 37506681]
12. Patel EU, Bloch EM, Tobian AAR. Severe acute respiratory syndrome coronavirus 2 serosurveillance in blood donor populations. J Infect Dis 2022 Jan 5;225(1):1-4. [doi: 10.1093/infdis/jiab517] [Medline: 34626464]
13. Blood donation statistics. Japanese Red Cross Society. 2022. URL: https://www.jrc.or.jp/donation/blood/data/ [accessed 2024-06-01]
14. Survey report on blood donation awareness among young people: questionnaire (in japanese). Ministry of Health Labour and Welfare. 2006. URL: https://www.mhlw.go.jp/new-info/kobetu/iyaku/kenketsugo/7n-04.html [accessed 2024-10-15]
15. Studte S, Clement M, Soliman M, Boenigk S. Blood donors and their changing engagement in other prosocial behaviors. Transfusion 2019 Mar;59(3):1002-1015. [doi: 10.1111/trf.15085] [Medline: 30549293]

Abbreviations

GVIF: generalized variance inflation factor

Edited by A Mavragani; submitted 15.05.24; peer-reviewed by L Aljerf, MJ Miller; revised version received 11.11.24; accepted 12.11.24; published 10.02.25.

Please cite as:
Kinoshita R, Miyamoto S, Suzuki T, Suzuki M, Yoneoka D
Interpreting the Influence of Using Blood Donor Residual Samples for SARS-CoV-2 Seroprevalence Studies in Japan: Cross-Sectional Survey Study
JMIR Public Health Surveill 2025;11:e60467
URL: https://publichealth.jmir.org/2025/1/e60467
doi:10.2196/60467

Association Among BMI, Self-Esteem, and Nonsuicidal Self-Injury in Young Adults to Understand the Influence of Socioenvironmental Factors: Longitudinal Study

Yi Zhang[1,2*], MD; Ruixue Ying[1*], MD; Wan Lu[1*], MD; Xuemeng Liu[1*], MD; Keyan Hu[3], MD; Qing Feng[1], MD; Zixiang Yu[4], MD; Zhen Wang[1], MD; Fangting Lu[1], MD; Yahu Miao[1], MD; Nanzhen Ma[5], MD; Fangbiao Tao[1], MD; Tian Jiang[1], MD; Qiu Zhang[1], PhD

[1]Department of Endocrinology, First Affiliated Hospital of Anhui Medical University, No 81 Meishan Road, Anhui, Hefei, China
[2]Department of Maternal, Child, and Adolescent Health, School of Public Health, Anhui Medical University, Hefei, China
[3]The First Affiliated Hospital and College of Clinical Medicine of Henan, University of Science and Technology, Luoyang, China
[4]Department of Pediatrics, The First Affiliated Hospital of Anhui Medical University, Hefei, China
[5]Hospital of Anhui Medical University, Hefei, China
[*]these authors contributed equally

Corresponding Author:
Qiu Zhang, PhD
Department of Endocrinology, First Affiliated Hospital of Anhui Medical University, No 81 Meishan Road, Anhui, Hefei, China

Abstract

Background: Nonsuicidal self-injury (NSSI) is a major public health problem leading to psychological problems in adolescents and young adults, similar to disorders such as depression and anxiety.

Objective: The aims of this study were to investigate (1) the interaction between BMI and socioenvironmental factors (including chronotype and mental health) that contribute to NSSI, and (2) whether self-esteem plays a mediating role in this association.

Methods: From May to June 2022, the multistage cluster sampling method was used to sample college students in four grades, including freshmen and seniors. The baseline participants were followed up 6 months later, excluding those who did not qualify, and the participants included 1772 college students. Socioenvironmental factors (chronotype/mental health), self-esteem, and NSSI were measured using a questionnaire. Multivariate linear regression models and chi-square analysis were used to evaluate the linear relationship between BMI, socioenvironmental factors, and self-esteem and the NSSI status. We use a process approach (mediation-moderation analysis) to explore the complex relationships between these variables.

Results: The mean age of the participants was 20.53 (SD 1.65) years at baseline. A significant association was revealed, suggesting that a high BMI (β=.056, 95% CI 0.008 - 0.086, *P*=.018) was associated with a higher NSSI. There was also an interaction among BMI, socioenvironmental factors, and NSSI. Socioenvironmental factors played both moderating and mediating roles in the relationship between BMI and NSSI, whereas self-esteem only played a mediating role.

Conclusions: Paying attention to factors such as overweight and obesity is important for early BMI control to identify other potential risk factors for NSSI and to evaluate how self-esteem can be improved considering multiple perspectives to improve the effect of BMI on NSSI in adolescents.

(JMIR Public Health Surveill 2025;11:e52928) doi:10.2196/52928

KEYWORDS

nonsuicidal self-injury; chronotype; BMI; self-esteem; body mass index; adolescent; young adult; teenager; social environmental factor; self-injury; sampling method; undergraduate; college student; linear regression; regression; regression model

Introduction

Nonsuicidal self-injury (NSSI) refers to physical harm, not the intention to commit suicide, and is not a socially recognized behavior [1]. NSSI, while being distinct from suicidal behavior, is a strong predictor of suicide and is also a significant mental health problem among adolescents worldwide [2]. NSSI is gradually increasing in the world; it seriously damages people's physical and mental health and is a psychopathological behavior, which has aroused people's attention and concern. The *DSM-5* identifies NSSI as a separate behavioral category and lists it as "an entry for further research" [3]. Adolescence and college years are important periods of psychological development and changes in life. In a meta-study, the reporting rate of NSSI was 22.5% (95% CI 17.2% - 28.9%) among adolescents [2]. The prevalence of NSSI ranged from 3.2% to 44.8% among nonclinical populations [4]. For clinical populations, the incidence of NSSI in the past 12 months ranged from 5% to 16.4% [4]. The harm of NSSI goes far beyond that, and after the emergence of NSSI in adolescence, it has maintained a certain “trajectory” into adulthood, including a continuous increase or fluctuation between the increase and decrease of NSSI, as well as an increase in psychobehavioral problems that largely affect people’s lives [5,6]. A systematic review and Bayesian meta-analysis concluded that in the entire developmental stage of adolescents, the NSSI frequency (but not the frequency of occurrence) in younger adolescents increased, that in middle adolescents remained stable, and that in older adolescents decreased [3]. On proceeding into the aftermath of the COVID-19 outbreak and trying to recover from its consequences, preventing NSSI and suicidal behaviors has continued to be a priority. In general, the current NSSI epidemic in China is serious [7], and it has become a major public health problem that severely endangers adolescents’ physical and mental health. Being an ongoing social health problem, NSSI is a major concern for mental health and it deserves our attention.

With such a high incidence rate of NSSI that has not been effectively controlled thus far, it is important to explore the potential factors of and prevention measures for NSSI, including what factors affect NSSI, how they interact with each other, and what needs to be done to protect those with high NSSI frequency. The specific factors that affect NSSI are as follows: (1) individual psychological factors, including self-abasement, impulsivity, early traumatic experience, self-esteem, and other characteristics; (2) external environmental factors, including negative life events (events that individuals feel unpleasant); (3) neurobiological factors, including the influence of the metabolism of substances such as 5-hydroxytryptamine, glutamate, and dopamine, which may also be related to genetics; (4) family factors, including family environment, family atmosphere, and family economic conditions; and (5) school factors, including school environment, teacher-student relationship, and classmate relationship, which may also have an impact on children and adolescents’ NSSI. The occurrence of NSSI behavior in adolescents is not the result of a single factor, but the result of the interaction of multiple factors. In addition, obesity has attracted considerable attention as a cause of many diseases, including physical diseases as well as psychological and behavioral problems [8]. Changes in BMI are correlated with suicide [9], and there is a significant association with suicide attempts in underweight and perceived overweight groups [10]. There is an association between psychological problems and NSSI [10,11]. A previous study demonstrated that while posttraumatic stress disorder leads to suicidal ideation, BMI plays a significant moderating role, indicating that the association between posttraumatic stress disorder and suicidal ideation is the strongest in individuals with high BMI values [11]. Another study has reported a correlation between obesity and NSSI and verified the correlation between BMI and major depressive disorder [12], but no precise research has explored the correlation between BMI and NSSI. NSSI is quite different from suicidal behavior in many aspects such as motivation, lethality, and manifestation; however, it remains a crucial risk factor for suicidal ideation and suicide attempts. This finding also suggests that psychological problems play an essential role in determining how BMI leads to NSSI.

Although some studies identified BMI as a risk factor for self-injury or suicide, other studies have not found positive results, which means that not everyone with an increased BMI necessarily shows NSSI behavior [13], as other factors may also be at play. To identify other possible psychological mechanisms that support this relationship, it is necessary to understand complex associations and mechanisms beyond the current one. In addition to the association between BMI and NSSI, it is significant to understand the mediators and moderators of BMI and NSSI for public health prevention and clinical practice. Research on borderline personality disorder has revealed that sleep deprivation, disrupted sleep, and prolonged sleep duration are associated with suicidal behavior or NSSI [14]. Sleep occupies a very important position in the circadian rhythm, especially the lack of sleep; the circadian rhythm refers to the 24-hour physiological and behavioral rhythm of individuals, which is divided into endogenous rhythm and exogenous rhythm. When the endogenous rhythm and exogenous rhythm are “mismatched,” circadian rhythms can be disrupted, which can lead to sleep disturbances [15]. Chronotype is a self-description of an individual’s circadian preference, reflecting the differences in their daily activity pattern and sleep-wake cycle; it can be generally divided into three types: morning, intermediate, and evening [16]. Moreover, the evening type has been reported to be strongly associated with suicidal behavior [17].

These studies provide evidence that NSSI moderates the relationship between BMI and socioenvironmental factors. A higher BMI has been associated with a lower risk of suicide in a large prospective study [18]; however, the mechanisms underlying this link require elucidation. Some studies have verified the correlation between self-esteem and NSSI [17,19]. However, little is known about how self-esteem moderates the relationship between BMI and NSSI. Socioenvironmental factors underlying these associations also remain unclear. Moreover, only limited research has been conducted on the relationship between BMI and socioenvironmental factors, and the outcome of NSSI during adolescence. A systematic review and meta-analysis reported that treatment of childhood obesity resulted in increased self-esteem after intervention [19]. The results of studies evaluating the self-esteem and NSSI suggest that although low self-esteem and self-injury are associated with both clinical and nonclinical populations, there are many factors influencing this relationship [20]. Therefore, based on a previous study [19], we hypothesize that there was an association among SE, BMI, and NSSI.

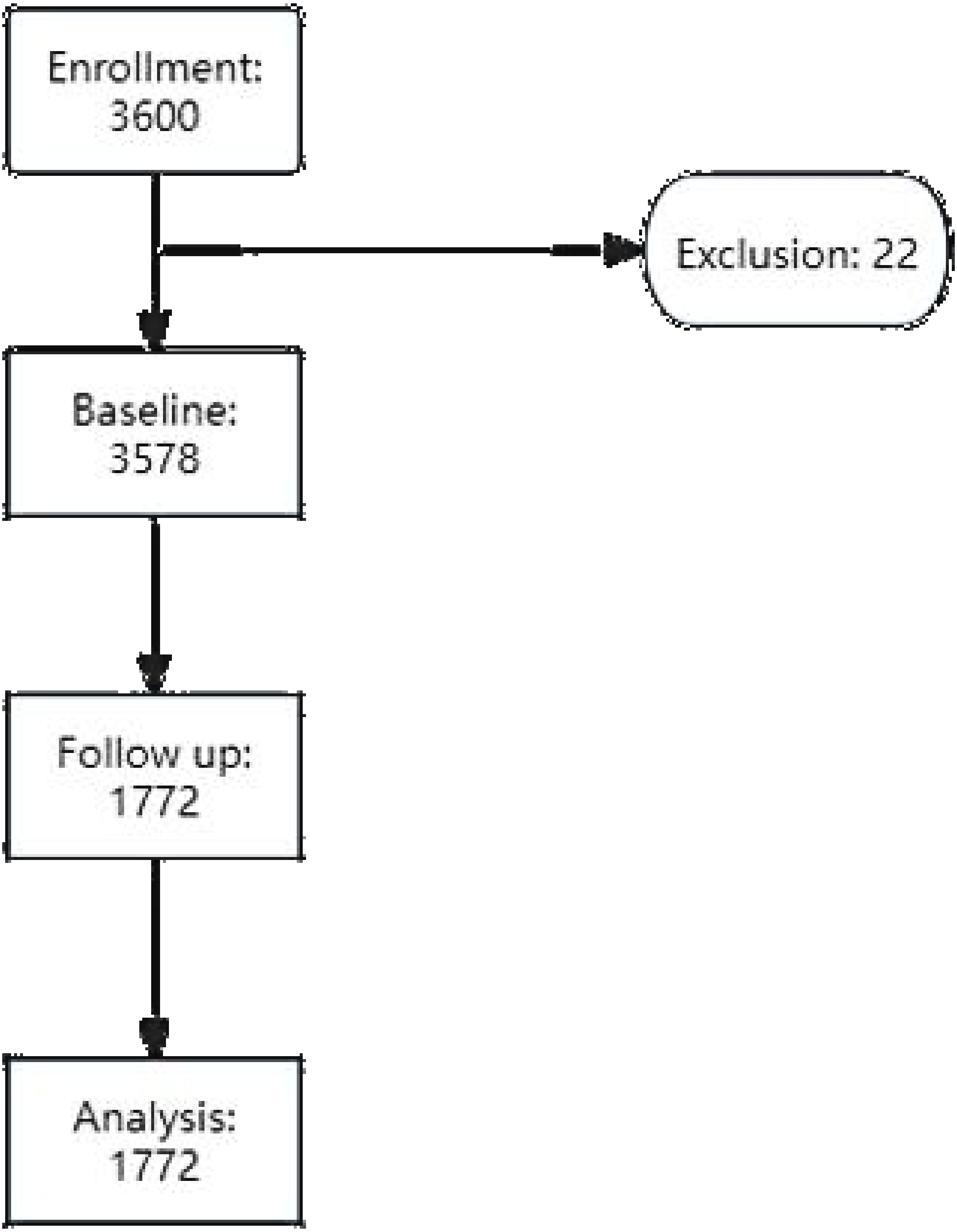
From a management and prevention perspective, these modifiable risk factors are crucial, and how they interact with each other requires further exploration. This study aimed to investigate the influencing factors and possible mechanisms of NSSI in college students and to observe the influence of some possible factors on BMI that could result in NSSI. Therefore, this study aimed to (1) examine the correlation among BMI, socioenvironmental factors, self-esteem, and NSSI and (2) determine whether socioenvironmental factors and self-esteem would play direct and indirect roles in the relationships between BMI and NSSI among Chinese college students.

Methods

Study Design and Study Setting for Recruitment

This was a longitudinal study, and the research methods have been described in previous studies [21,22]. The study setting for recruitment was as follows: we first selected a university in a city in the Anhui Province, and then selected college students in four consecutive grades according to the cluster sampling method. A total of 3600 college students were initially recruited using an anonymous electronic questionnaire survey about their health at baseline. The cluster sampling method was used for the schools surveyed, and most of the students from each grade were included in the study. The survey was conducted from May to June 2022. Some participants were excluded from the study owing to unwillingness to answer the questionnaire, absence from class, high levels of missing data (questionnaires with missing values greater than 5% were eliminated), or false responses [23-25] at baseline. The baseline group was followed up with a questionnaire after 6 months. Finally, a total of 1772 participants were included in the follow-up survey (Figure 1).

Figure 1. A flowchart for the enrollment and follow-up of participants.



Inclusion Criteria

Participants were included if (1) informed consent from them and their guardians were obtained, (2) they were college students aged 15 - 26 years, (3) they had no history of mental illness, and (4) they were attending the chosen school.

Exclusion Criteria

Participants were excluded if (1) their informed consent was not obtained, (2) they were junior and high school students, (3) they failed to submit the questionnaire, or (4) they had congenital or acquired immunodeficiency.

Exposure

Chronotype

The morningness-eveningness questionnaire is used to measure college students' sleep habits. According to the previous analysis results, personal sleep habits is classified into three types: morning, intermediate, and evening [26]. The suggested demarcations were further categorized as follows: sleeping from 4 AM to 7 AM is definitely eveningness, sleeping from 8 AM to 11 AM is medium eveningness, sleeping from 12 PM to 5 PM is intermediate type, sleeping from 6 PM to 9 PM is medium morningness, and sleeping from 10 PM to 12 AM is definitely morningness.

Self-Esteem

The self-esteem scale was first developed by Ronsenberg to evaluate an individual's overall feelings of self-worth and self-acceptance [27]. It was later translated and revised into Chinese by Ji Fuyi and Yu Xin [28]. There were 10 items on the scale: 5 items on self-denial and 5 items on self-affirmation. A 4-point Likert scale was used, ranging from 1 indicating "very inconsistent" to 4 indicating "very consistent." The 5 self-denial items were scored inversely. The higher the score, the higher the self-esteem.

Outcome

NSSI

One of the most important question in the questionnaire was "Have you intentionally hurt yourself in the past 12 months, but not for suicide?" The questionnaire also listed several NSSI methods: hit yourself with a fist or palm, pulled your hair, hit a hard object with your head or fist, pinched or scratched yourself, bit yourself, cut yourself, and stabbed yourself. Those who had an NSSI were asked about their NSSI frequency. The participants owning up to 5 or more methods were considered to be prone to NSSI [29].

Statistics Analysis

Conventional means and SDs were used for continuous variables. Univariate ANOVA was used to describe the association between independent and outcome variables. Frequency and percentage tests were used for categorical variables, and chi-square tests were used to explore the association between independent and outcome variables. The PROCESS method was used to conduct a mediation-moderation analysis to explore the relationships among BMI, chronotype, self-esteem, and NSSI [30]. The bootstrap method was used to resample 1000 samples, and the 95% CI was calculated. All data were analyzed using SPSS software (Windows version 23.0; IBM Corp).

Ethical Considerations

The current study is designed and reported according to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist. The design and data collection procedures were approved by the Ethics Committee of the First Affiliated Hospital of Anhui Medical University (review number PJ 2024-06-06). Written informed consent was obtained from the parents or guardians of all the students.

Results

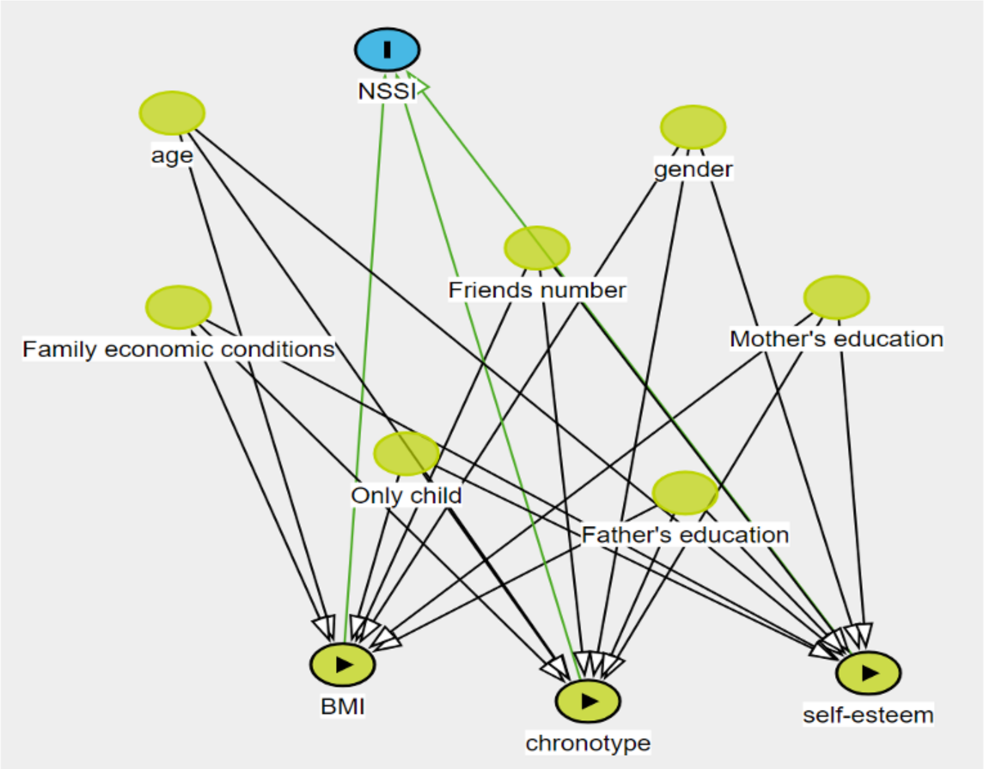
General Demographic Characteristics

Table 1 presents the general distribution of the NSSI variables examined in this study considering the demographic characteristics, while Figure 2 is a directed acyclic graph showing the covariates. Among the 1772 participants included in the questionnaire analysis, 908 (51.1%) were male students; the mean age was 20.53 (SD 1.65) years at baseline; and 853 (48.1%) dwelled in rural areas, 548 (30.9%) in towns, and 371 (20.9%) in urban areas. The prevalence of NSSI was 5.6% (N=1772).

Table . The distribution of the nonsuicidal self-injury (NSSI) variables considering demographic characteristics.

Variables	Total, n (%)	Absence of NSSI, n (%)	Presence of NSSI, n (%)	*t* test/chi-square (*df*)	*P* value
Sex				0.46 (1)	.54
Male	908 (51.1)	854 (51.0)	54 (54.5)		
Female	864 (48.9)	819 (49.0)	45 (45.5)		
Residential areas				0.53 (2)	.76
Rural	788 (44.5)	744 (44.5)	44 (44.4)		
Town	401 (22.6)	376 (22.5)	25 (25.3)		
Urban	583 (32.9)	553 (33.1)	30 (30.3)		
Only child				0.29 (1)	.66
Yes	581 (32.8)	551 (32.9)	30 (30.3)		
No	1191 (67.2)	1122 (67.1)	69 (69.7)		
Father’s education				2.31 (5)	.81
Not known	55 (3.1)	4 (4.0)	51 (3.0)		
Below primary	153 (8.6)	7 (7.1)	146 (8.7)		
Primary	205 (11.6)	11 (11.1)	194 (11.6)		
Junior high	661 (37.3)	32 (32.3)	629 (37.6)		
High school or technical secondary school	379 (21.4)	24 (24.2)	355 (21.2)		
Junior college or above	319 (18.0)	21 (21.2)	298 (17.8)		
Mother’s education				3.14 (5)	.68
Not known	67 (3.8)	63 (3.8)	4 (4.0)		
Below primary	345 (19.5)	323 (19.3)	22 (22.2)		
Primary	316 (17.8)	299 (17.9)	17 (17.2)		
Junior high	535 (30.2)	504 (30.1)	31 (31.3)		
High school or technical secondary school	301 (17.0)	290 (17.3)	11 (11.1)		
Junior college or above	208 (11.7)	194 (11.6)	14 (14.1)		
Family economic conditions				3.15 (4)	.53
Bad	93 (5.2)	90 (5.4)	3 (3.0)		
Worse	428 (24.2)	402 (24.0)	26 (26.3)		
Medium	1142 (64.4)	1075 (64.3)	67 (67.7)		
Better	100 (5.6)	97 (5.8)	3 (3.0)		
Good	9 (0.5)	9 (0.5)	0 (0)		
Number of friends				4.92 (3)	.18
0	36 (2.0)	32 (1.9)	4 (4.0)		
1 - 2	587 (33.1)	548 (32.8)	39 (39.4)		
3 - 5	854 (48.2)	810 (48.4)	44 (44.4)		
6 or more	295 (16.6)	283 (16.9)	12 (12.1)		

Figure 2. A directed acyclic graph showing the covariates. NSSI: nonsuicidal self-injury.



Association Between Independent Variables and Adolescent Psychological and Behavioral Problems

Table 2 presents the correlation between independent variables and depression symptoms through multilevel linear regression analyses. There was a significant relationship between different BMI and NSSI (β=.056); after controlling for covariates, these relationships were also significant. Furthermore, there was a significant relationship between self-esteem and NSSI (β=−.077); after controlling for covariates, this relationship was also significant. There was no significant correlation between chronotype and NSSI.

Table . Multilevel linear regression between independent variables and nonsuicidal self-injury (NSSI).

	NSSI					
	R^2	β	*t* test (*df*)	*P* value	*F* test (*df*)	95% CI
BMI						
Model 1[a]	0.003	.056	2.359 (1)	.02	5.566 (1)	0.008 to 0.086
Model 2[b]	0.008	.054	2.191 (9)	.03	1.551 (9)	0.005 to 0.085
Chronotype						
Model 1	0.00	.022	0.938 (1)	.99	0.88 (1)	−0.033 to 0.094
Model 2	0.006	.025	1.055 (9)	.29	1.139 (9)	−0.03 to 0.099
Self-esteem						
Model 1	0.006	−.077	−3.234 (1)	<.001	10.461 (1)	−0.136 to −0.033
Model 2	0.01	−.078	−2.896 (9)	.004	1.951 (9)	−0.130 to −0.025

[a] Model 1: crude model.

[b] Model 2: controlled for parent educational level, gender, economic level, whether only child, number of friends, residential area, and age.

Impact of the Mediation Analysis Among BMI, Chronotype, and Self-Esteem on Adolescent NSSI

Mediation-moderation analyses were performed for BMI, self-esteem, chronotype, and NSSI scores. The results are listed in Tables 3 and 4. Among NSSI, findings have suggested the following: a positive association between BMI exposure and NSSI, chronotype was associated with NSSI, and chronotype played a moderating role in the increased risk of NSSI induced by BMI mediated by self-esteem.

Table . Testing the mediation-moderation effects of BMI[a] and chronotype on the nonsuicidal self-injury (NSSI)[b] among college students.

Variables	Self-esteem			NSSI		
	ß	*t* test (*df*)	*P* value	ß	*t* test (*df*)	*P* value
BMI	.0775	−0.4799 (3)	.54	.2262	2.5466 (5)	.01
Chronotype	−.0387	0.6206 (3)	.63	.8102	2.9399 (5)	.003
BMI* chronotype[c]	−.0005	−0.0869 (3)	.93	−.0127	−2.1149 (5)	.04
Self-esteem[d]	N/A[e]	N/A	N/A	.1613	1.3426 (5)	.18
Self-esteem* chronotype	N/A	N/A	N/A	−.0167	−2.0839 (5)	.04
R^2	0.0065	N/A	N/A	.0140	N/A	N/A
F test (*df*)	3.8342	N/A	N/A	5.0212	N/A	N/A

[a] BMI: independent variables.

[b] NSSI: dependent variables.

[c] chronotype: moderated variables.

[d] Self-esteem: mediate variables.

[e] N/A: not applicable.

Table . Bootstrap method showing the conditional direct and indirect effects between self-esteem, chronotype, and nonsuicidal self-injury (NSSI).

	Nonsuicidal self-injury		
	Effect of chronotype	Self-esteem	Lower limit to upper limit
Direct effect	0.0818	0.0270	0.0288 to 0.1347
Predictor	0.0409	0.0199	0.0020 to 0.0799
Moderator	0.0001	0.024	−0.0556 to 0.0558
Indirect effect	11.3542	0.0013	−0.0011 to 0.0053
Predictor	14.5643	0.0038	0.0006 to 0.008
Mediator	17.7744	0.004	−0.0002 to 0.0179

Impact of Mediation Analysis Among BMI, Mental Health, and Self-Esteem on Adolescent NSSI

Mediation-moderation analyses were performed for BMI, self-esteem, mental health, and NSSI. The results are presented in Tables 5 and 6. In NSSI, results have suggested a positive association between BMI exposure and NSSI, where depression was associated with NSSI, and chronotype played a moderating role in the increased risk of NSSI induced by BMI and mediated by self-esteem (Tables 5 and 7). Similar results were obtained for stress and NSSI (Tables 6 and 8).

Table . Testing the mediation-moderation effects of BMI[a] and anxiety on the nonsuicidal self-injury (NSSI)[b] among college students.

Variables	Self-esteem			NSSI		
	ß	*t* test (*df*)	*P* value	ß	*t* test (*df*)	*P* value
BMI	−.0099	−0.1054 (3)	.91	−.2475	−2.2641 (5)	.02
Anxiety	−.1302	−2.5744 (3)	.01	.0233	0.222 (5)	.82
BMI* anxiety[c]	−.0007	−0.2931 (3)	.77	.0070	2.6944 (5)	.007
Self-esteem[d]	N/A[e]	N/A	N/A	.0233	1.2468 (5)	.21
Self-esteem* anxiety	N/A	N/A	N/A	−.0051	−1.8281 (5)	.07
R^2	.1014	N/A	N/A	.0180	N/A	N/A
F test (*df*)	66.5210	N/A	N/A	6.4595	N/A	N/A

[a] Independent variables: BMI.

[b] Dependent variables: NSSI.

[c] Moderated variables: depression.

[d] Mediate variables: self-esteem.

[e] N/A: not applicable.

Table . Bootstrap method showing the conditional direct and indirect effects between self-esteem, anxiety, and nonsuicidal self-injury (NSSI).

	Nonsuicidal self-injury		
	Effect of anxiety	Self-esteem	Lower limit to upper limit
Direct effect	−0.0242	0.0343	−0.0914 to 0.0430
Predictor	0.02390	0.0212	−0.0177 to 0.0655
Moderator	0.0719	0.0238	0.0252 to 0.1186
Indirect effect	0.0001	0.0018	−0.0026 to 0.0051
Predictor	0.0016	0.0017	−0.0006 to 0.0063
Mediator	0.0044	0.0028	0.0001 to 0.0121

Table . Bootstrap method showing the conditional direct and indirect effects self-esteem, depression, and nonsuicidal self-injury (NSSI).

	Nonsuicidal self-injury		
	Effect of depression	Self-esteem	Lower limit to upper limit
Direct effect	−0.0273	0.0325	−0.0911 to 0.0364
Predictor	0.0328	0.0201	−0.0066 to 0.0722
Moderator	0.0929	0.0273	0.0393 to 0.1465
Indirect effect	0.0005	0.0020	−0.0016 to 0.0081
Predictor	0.0022	0.0017	0.00 to 0.0078
Mediator	0.0044	0.0035	−0.000 to 0.0148

Table . Testing the mediation-moderation effects of BMI[a] and depression on the nonsuicidal self-injury (NSSI)[b] among college students.

Variables	Self-esteem			NSSI		
	ß	*t* test (*df*)	*P* value	ß	*t* test (*df*)	*P* value
BMI	.0549	0.7267	.47	−.1616	−1.8212	.07
Depression[c]	−.0766	−2.3582	.02	−.0008	−0.0106	.99
BMI* depression	−.0016	−1.1464	.25	.0040	2.3457	.02
Self-esteem[d]	N/A[e]	N/A	N/A	.0516	0.5076	.61
Self-esteem* depression	N/A	N/A	N/A	−.0026	−1.2582	.21
R^2	.1236	N/A	N/A	.0130	N/A	N/A
F test (*df*)	83.1489	N/A	N/A	4.6602	N/A	N/A

[a] Independent variables: BMI.

[b] Dependent variables: NSSI.

[c] Moderated variables: depression.

[d] Mediate variables: self-esteem.

[e] N/A: not applicable.

Discussion

Principal Findings

In our study, the prevalence of NSSI in a sample of college students was 5.6%. Here, we examined the combination of chronotype and mental health as well as the interaction effect among BMI, self-esteem, and NSSI. BMI was positively correlated with the development of NSSI during adolescence, that is, the higher the BMI, the higher was the incidence of NSSI. Consistent with our hypothesis, our study provided evidence that self-esteem mediates the relationship between BMI and NSSI; this relationship was also moderated by chronotype. The relationship among BMI, self-esteem, and NSSI is complex and may involve a moderating effect of chronotype. Therefore, this study developed and evaluated a mediation-moderation model to elucidate the role of multilevel factors, including BMI, self-esteem, and chronotype, in relation to NSSI [31].

Comparison With Other Studies

With the change in social living environment, adolescents are facing increasing pressure and challenges, and adolescent NSSI has become a problem that cannot be ignored. NSSI is a common mental health problem among adolescents [32]. In a college student–related survey, approximately 1 in 5 (20.4%) students reported lifelong NSSI, and the frequency of persistent NSSI was estimated at 56.4%, with 15.6% of students reporting high-frequency repeat patterns (≥5 times per year) [33]. This high prevalence is worth considering, and NSSI is not only associated with suicidal behavior but is also a specific risk factor for suicide. In addition to the burden on individuals and families, including NSSI and self-harm with suicidal intent, the increased incidence of NSSI requires significant health care and economic costs and can lead to death in severe cases, further emphasizing the need to study why people should be concerned about NSSI [34]. Therefore, it is necessary to pay attention to the factors that affect NSSI; some factors have been explored [35-38], such as adverse life events, stressful life events, negative coping styles, problematic internet use, sleep disorders, traumatic experiences, problematic parent-child relationships, and mental health issues. In the same way that obesity causes physical diseases (eg, cardiovascular disease and diabetes), obesity has a significant triggering effect on the number of times of suicide attempts [39]. However, a population-based telephone survey of US adults conducted by the Centers for Disease Control and Prevention found that traditional suicide risk factors do not show any correlation with BMI; therefore, there is unlikely to be any relationship between BMI and a lower risk of suicide [18]. That said, in some cases, it is not entirely the case that a higher BMI is associated with a greater number of risk factors. Further research into the relationship between BMI and suicide may lead to new modifiable risk factors that can result in or prevent this key cause of death [18]. This also encourages us to look beyond the known causes of suicidal behavior, as there are other factors at play in psychological research.

In the aforementioned research, individual factors are very important, and self-esteem is an important individual factor [20]. Such studies might shed light on how obesity interacts with chronotype factors and can affect self-esteem in people with NSSI [40]. Regarding the association between obesity and suicide-related variables, 3 out of 5 studies investigating these suicide attempts showed a positive association between these attempts and obesity, and this relationship may be related to the reduction in self-worth and self-esteem caused by obesity-related stigma, suggesting that self-esteem plays a key role in BMI-induced suicide attempts [20, 41]; the results suggest that although low self-esteem and NSSI are correlated in both clinical and nonclinical populations, many factors influencing this relationship still exist [20], which needs to be confirmed by further research. According to the developmental psychopathological model, individuals' negative understanding and evaluation of themselves will lead to the emergence of NSSI, so low self-esteem is an important variable that triggers NSSI. Therefore, another study demonstrated the mediating role of self-esteem in the relationship between BMI and depression or suicidality, as well as the moderating role of sex in the mediated pathways [20], which also provides an effective theoretical basis for our research. It is noteworthy that self-esteem is a crucial factor given its significance in depression and NSSI. This study, thus, converges with the existing literature and extends it. Moreover, we noticed that chronotype patterns could moderate these relationships, and sleep deprivation itself could influence an increase in psychobehavioral problems [42]. Another review concluded that people with bipolar disorder are characterized by extreme mood swings, high rates of suicide, sleep problems, and dysfunction in psychological characteristics, such as self-esteem (a feeling ranked below depression and above mania), which further provides a theoretical basis for understanding the relationship between sleep, self-esteem, and metabolism [43]. The current findings highlight the importance of assessing self-esteem in mediating BMI and NSSI, perhaps targeting self-esteem to mitigate self-esteem–related mental health problems (eg, NSSI and suicide). We also considered the effect of the moderator mental health on BMI and NSSI. The stigma associated with being overweight has increased in recent decades, which may have exacerbated the harmful effects of being overweight on mental health [44], thereby further elevating the occurrence of NSSI.

Another study explored the relationship between adolescents' BMI and mental health issues, including social phobia, depression, suicidal tendencies, and low self-esteem, and the mediating role of sex in participation in bullying [45]. Therefore, this provides a theoretical basis for our study that BMI is closely related to self-esteem, mental health, and suicidal behavior, and the correlation among the variables is also affected by other variables. According to the US Health and Retirement Study, baseline BMI, BMI transition, and BMI trajectory play significant moderating roles in the positive association between adverse childhood experience with new-onset all-cause dementia and Alzheimer disease [46]. The possible mechanisms include increased insulin and glucagon levels in the plasma and cerebrospinal fluid [47]. Additionally, our findings highlight the need for a multidisciplinary, collaborative, and integrated approach to optimize patient care. Different mechanisms have been identified, including gut-brain mechanisms, inflammatory responses (changes in the neutrophil-lymphocyte ratio, monocyte-lymphocyte ratio, and platelet-lymphocyte ratio), and changes in cytokine levels (interferon-γ, interleukin-1β, interleukin-6, interleukin-8, monocyte chemoattractant protein-1, tumor necrosis factor-α, and transforming growth factor-β1) [48].

Clinical Practices

Overweight and obese adolescents often experience prejudice due to the surrounding environment, which causes them to have low self-esteem, depression, and other negative emotions; engage in bad behaviors; cause self-injury; and harbor suicidal thoughts. As sociality is an essential attribute of adolescents, overweight and obese adolescents should communicate with others and obtain support, thereby effectively reducing the occurrence of NSSI. By further exploring the two variables of chronotype and self-esteem, this study aimed to explore the sociological pathway of BMI leading to NSSI.

Strengths and Limitations

The strengths of this study are its longitudinal design and large number of participants. The purpose of mediating is to strengthen the quasi-causal inference about the mechanism by which independent variables influence outcomes, whereas the purpose of moderating is to examine variables that influence the strength and direction of the predictor-outcome relationship. The use of moderated mediation could help identify the interdependence between the BMI and chronotypes. We also explored the role of the processes between early BMI and NSSI later in life.

However, this study has some limitations. First, although this was a follow-up study, only one follow-up survey was conducted, and changes in BMI were not evaluated during this process. These data are self-reported survey data, which can inevitably produce potential recall bias; therefore, we will further correct for the bias caused by recall bias in the future. Second, the study examined results in only one city; it is not clear how representative this sample is. Therefore, follow-up surveys are needed with samples from different regions and cultures across the country. Third, although this study mainly focuses on the relationship between BMI and NSSI under the influence of different factors, biological indicators will be included for further discussion in the future. Finally, we must consider critical attitudes toward the potential mechanism. As only some variables were evaluated in this study, additional variables with psychosocial factors should be considered in the future [48].

Conclusion

In this study, there was a combination effect of BMI, self-esteem, and NSSI, as well as an interaction effect among BMI, self-esteem, and NSSI. NSSI prevention programs should focus on evidence-based approaches relevant to public health and address them in conjunction with relevant influencing factors, including a wide range of interventions such as self-help programs, education, policy change, and legislation. Stronger inferences about the role of BMI in the etiology of adolescent NSSI can be made by exploring sex and ameliorating unfavorable mental health in prospective cohort studies. Future studies should expand upon our small subset of measured variables to paint a broader picture of adolescent NSSI in response to early-life social, psychological, and environmental stimuli.

Acknowledgments

We would like to acknowledge all hospital action teams, the staff and participants from the participating hospitals, and our cooperators for assistance in data collection. This research was funded by the China Postdoctoral Science Foundation (grant 2023M740022), the Research Fund of the Anhui Institute of Translational Medicine (grant 2023 hyx-C33), the National Natural

Science Foundation of China (grant 82370836), the Anhui Medical University Research Fund Project Youth Science Fund (grant 2022xkj307), the National Natural Science Foundation of China Youth Science 2018 Fund Training Program (grant 2022xkj307), The First Affiliated Hospital of Anhui Medical University (grant 2018kj20), and the Hunan Sanuo Diabetes Charity Foundation "Sweet Doctor Cultivation" project (grant 2021SD08).

Data Availability

The datasets generated for this study are available on request to the corresponding author.

Authors' Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work

Conflicts of Interest

None declared.

References

1. Nock MK. Self-injury. Annu Rev Clin Psychol 2010;6:339-363. [doi: 10.1146/annurev.clinpsy.121208.131258] [Medline: 20192787]
2. Knipe D, Padmanathan P, Newton-Howes G, Chan LF, Kapur N. Suicide and self-harm. Lancet 2022 May 14;399(10338):1903-1916. [doi: 10.1016/S0140-6736(22)00173-8] [Medline: 35512727]
3. De Luca L, Pastore M, Palladino BE, Reime B, Warth P, Menesini E. The development of non-suicidal self-injury (NSSI) during adolescence: a systematic review and Bayesian meta-analysis. J Affect Disord 2023 Oct 15;339:648-659. [doi: 10.1016/j.jad.2023.07.091] [Medline: 37479039]
4. Haregu T, Chen Q, Arafat SMY, Cherian A, Armstrong G. Prevalence, correlates and common methods of non-suicidal self-injury in South Asia: a systematic review. BMJ Open 2023 Nov 22;13(11):e074776. [doi: 10.1136/bmjopen-2023-074776] [Medline: 37993150]
5. Tilton-Weaver L, Latina D, Marshall SK. Trajectories of nonsuicidal self-injury during adolescence. J Adolesc 2023 Apr;95(3):437-453. [doi: 10.1002/jad.12126] [Medline: 36437557]
6. Adrian M, Zeman J, Erdley C, Whitlock K, Sim L. Trajectories of non-suicidal self-injury in adolescent girls following inpatient hospitalization. Clin Child Psychol Psychiatry 2019 Oct;24(4):831-846. [doi: 10.1177/1359104519839732] [Medline: 30947520]
7. Ribeiro JD, Franklin JC, Fox KR, et al. Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: a meta-analysis of longitudinal studies. Psychol Med 2016 Jan;46(2):225-236. [doi: 10.1017/S0033291715001804] [Medline: 26370729]
8. Stefanovics EA, Grilo CM, Potenza MN, Pietrzak RH. Obesity in Latinx and White U.S. military veterans: mental health, psychosocial burden, non-suicidal self-injury and suicidal behavior. Psychiatry Res 2024 May;335:115844. [doi: 10.1016/j.psychres.2024.115844] [Medline: 38484606]
9. Hecht LM, Yeh HH, Braciszewski JM, et al. Weighing the association between BMI change and suicide mortality. Psychiatr Serv 2021 Aug 1;72(8):920-925. [doi: 10.1176/appi.ps.202000475] [Medline: 33882679]
10. Kim B, Kim HS, Park S, Kwon JA. BMI and perceived weight on suicide attempts in Korean adolescents: findings from the Korea youth risk behavior survey (KYRBS) 2020 to 2021. BMC Public Health 2023;23(1):1107. [doi: 10.1186/s12889-023-16058-z]
11. Geoffroy MC, Bouchard S, Per M, et al. Prevalence of suicidal ideation and self-harm behaviours in children aged 12 years and younger: a systematic review and meta-analysis. Lancet Psychiatry 2022 Sep;9(9):703-714. [doi: 10.1016/S2215-0366(22)00193-6] [Medline: 35907406]
12. Freitas-Rosa M, Gonçalves S, Antunes H. Is being overweight associated with engagement in self-injurious behaviours in adolescence, or do psychological factors have more "weight"? Eat Weight Disord 2016 Sep;21(3):493-500. [doi: 10.1007/s40519-015-0251-7] [Medline: 26753879]
13. BMI and suicide. Harv Ment Health Lett 2006;22(12).
14. Fitzpatrick S, Varma S, Ip J. The impact of homeostatic and circadian sleep processes on non-suicidal self-injury and suicide urges in borderline personality disorder. Arch Suicide Res 2022;26(3):1556-1571. [doi: 10.1080/13811118.2021.1932647] [Medline: 34348588]
15. Jagannath A, Taylor L, Wakaf Z, Vasudevan SR, Foster RG. The genetics of circadian rhythms, sleep and health. Hum Mol Genet 2017 Oct 1;26(R2):R128-R138. [doi: 10.1093/hmg/ddx240] [Medline: 28977444]

16. Walsh NA, Repa LM, Garland SN. Mindful larks and lonely owls: the relationship between chronotype, mental health, sleep quality, and social support in young adults. J Sleep Res 2022 Feb;31(1):e13442. [doi: 10.1111/jsr.13442] [Medline: 34272788]
17. Park C, Rosenblat JD, Brietzke E, et al. Stress, epigenetics and depression: a systematic review. Neurosci Biobehav Rev 2019 Jul;102:139-152. [doi: 10.1016/j.neubiorev.2019.04.010] [Medline: 31005627]
18. Mukamal KJ, Miller M. BMI and risk factors for suicide: why is BMI inversely related to suicide? Obesity (Silver Spring) 2009 Mar;17(3):532-538. [doi: 10.1038/oby.2008.538] [Medline: 19219065]
19. Wang S, Xu H, Zhang S, Wan Y, Tao F. Mediating effects of self-esteem in the relationship between childhood maltreatment and non-suicidal self-injury among adolescents: the roles of sex and only-child status. Soc Sci Med 2020 Feb 14;249(3):112847. [doi: 10.1016/j.socscimed.2020.112847] [Medline: 32086153]
20. Forrester RL, Slater H, Jomar K, Mitzman S, Taylor PJ. Self-esteem and non-suicidal self-injury in adulthood: a systematic review. J Affect Disord 2017 Oct 15;221:172-183. [doi: 10.1016/j.jad.2017.06.027] [Medline: 28647667]
21. Li S, Wang R, Thomas E, et al. Patterns of adverse childhood experiences and depressive symptom trajectories in young adults: a longitudinal study of college students in China. Front Psychiatry 2022;13:918092. [doi: 10.3389/fpsyt.2022.918092] [Medline: 35958653]
22. Song X, Wang S, Wang R, et al. Mediating effects of specific types of coping styles on the relationship between childhood maltreatment and depressive symptoms among Chinese undergraduates: the role of sex. Int J Environ Res Public Health 2020 Apr 30;17(9):3120. [doi: 10.3390/ijerph17093120] [Medline: 32365812]
23. Wang S, Xu H, Li S, Jiang Z, Wan Y. Sex differences in the determinants of suicide attempt among adolescents in China. Asian J Psychiatr 2020 Mar;49:101961. [doi: 10.1016/j.ajp.2020.101961] [Medline: 32078954]
24. Li S, Wang S, Gao X, et al. Patterns of adverse childhood experiences and suicidal behaviors in adolescents: a four-province study in China. J Affect Disord 2021 Apr;285:69-76. [doi: 10.1016/j.jad.2021.02.045]
25. Jiang Z, Xu H, Wang S, et al. Parent-child relationship quality, childhood maltreatment, and psychological symptoms in Chinese adolescent. J Interpers Violence 2022 Oct;37(19-20):18130-NP18151. [doi: 10.1177/08862605211035869] [Medline: 34330168]
26. Patterson F, Malone SK, Grandner MA, Lozano A, Perkett M, Hanlon A. Interactive effects of sleep duration and morning/evening preference on cardiovascular risk factors. Eur J Public Health 2018 Feb 1;28(1):155-161. [doi: 10.1093/eurpub/ckx029] [Medline: 28371850]
27. Rosenberg M. Rosenberg self-esteem scale (RSE). Acceptance and commitment therapy. Meas Pack 1965;61:52.
28. Yifu J, Xin Y. The self-esteem scale. Chin J Ment Health 1999;Supplement:318-320.
29. Wan Y, Chen R, Ma S, et al. Associations of adverse childhood experiences and social support with self-injurious behaviour and suicidality in adolescents. Br J Psychiatry 2019 Mar;214(3):146-152. [doi: 10.1192/bjp.2018.263] [Medline: 30477603]
30. Jiang T, Zhang Y, Dai F, Liu C, Hu H, Zhang Q. Advanced glycation end products and diabetes and other metabolic indicators. Diabetol Metab Syndr 2022 Jul 25;14(1):104. [doi: 10.1186/s13098-022-00873-2] [Medline: 35879776]
31. Zhang Y, Suo X, Zhang Y, et al. The relationship between negative life events and nonsuicidal self-injury among Chinese adolescents: a moderated-mediation model. Neuropsychiatr Dis Treat 2022;18:2881-2890. [doi: 10.2147/NDT.S386179] [Medline: 36540672]
32. Brown RC, Plener PL. Non-suicidal self-injury in adolescence. Curr Psychiatry Rep 2017 Mar;19(3):20. [doi: 10.1007/s11920-017-0767-9] [Medline: 28315191]
33. Kiekens G, Claes L, Hasking P, et al. A longitudinal investigation of non-suicidal self-injury persistence patterns, risk factors, and clinical outcomes during the college period. Psychol Med 2023 Oct;53(13):6011-6026. [doi: 10.1017/S0033291722003178] [Medline: 36325723]
34. Hepp J, Carpenter RW, Störkel LM, Schmitz SE, Schmahl C, Niedtfeld I. A systematic review of daily life studies on non-suicidal self-injury based on the four-function model. Clin Psychol Rev 2020 Dec;82:101888. [doi: 10.1016/j.cpr.2020.101888] [Medline: 32949907]
35. Herbison CE, Allen K, Robinson M, Newnham J, Pennell C. The impact of life stress on adult depression and anxiety is dependent on gender and timing of exposure. Dev Psychopathol 2017 Oct;29(4):1443-1454. [doi: 10.1017/S0954579417000372] [Medline: 28397629]
36. Fan YY, Liu J, Zeng YY, Conrad R, Tang YL. Factors associated with non-suicidal self-injury in Chinese adolescents: a meta-analysis. Front Psychiatry 2021;12:747031. [doi: 10.3389/fpsyt.2021.747031] [Medline: 34916971]
37. Baetens I, Greene D, Van Hove L, et al. Predictors and consequences of non-suicidal self-injury in relation to life, peer, and school factors. J Adolesc 2021 Jul;90:100-108. [doi: 10.1016/j.adolescence.2021.06.005] [Medline: 34182197]
38. Wang YJ, Li X, Ng CH, Xu DW, Hu S, Yuan TF. Risk factors for non-suicidal self-injury (NSSI) in adolescents: a meta-analysis. EClinMed 2022 Apr;46:101350. [doi: 10.1016/j.eclinm.2022.101350] [Medline: 35330803]
39. Powell-Wiley TM, Poirier P, Burke LE, et al. Obesity and cardiovascular disease: a scientific statement from the American Heart Association. Circulation 2021 May 25;143(21):e984-e1010. [doi: 10.1161/CIR.0000000000000973] [Medline: 33882682]

40. Chauhan S, Pandey R, Vakani K, Norbury R, Ettinger U, Kumari V. Sleep quality mediates the association between chronotype and mental health in young Indian adults. Npj Ment Health Res 2024 Jun 24;3(1):31. [doi: 10.1038/s44184-024-00076-9] [Medline: 38914742]
41. Ye Z, Xiong F, Li W. A meta-analysis of co-occurrence of non-suicidal self-injury and suicide attempt: Implications for clinical intervention and future diagnosis. Front Psychiatry 2022;13:976217. [doi: 10.3389/fpsyt.2022.976217] [Medline: 36032240]
42. Yeo SC, Jos AM, Erwin C, et al. Associations of sleep duration on school nights with self-rated health, overweight, and depression symptoms in adolescents: problems and possible solutions. Sleep Med 2019 Aug;60:96-108. [doi: 10.1016/j.sleep.2018.10.041] [Medline: 30611714]
43. Rantala MJ, Luoto S, Borráz-León JI, Krams I. Bipolar disorder: an evolutionary psychoneuroimmunological approach. Neurosci Biobehav Rev 2021 Mar;122:28-37. [doi: 10.1016/j.neubiorev.2020.12.031] [Medline: 33421542]
44. Daly M, Robinson E, Sutin AR. Perceived overweight and suicidality among US adolescents from 1999 to 2017. Int J Obes (Lond) 2020 Oct;44(10):2075-2079. [doi: 10.1038/s41366-020-0620-9] [Medline: 32546863]
45. Yen CF, Liu TL, Ko CH, Wu YY, Cheng CP. Mediating effects of bullying involvement on the relationship of body mass index with social phobia, depression, suicidality, and self-esteem and sex differences in adolescents in Taiwan. Child Abuse Negl 2014 Mar;38(3):517-526. [doi: 10.1016/j.chiabu.2013.07.015] [Medline: 24011767]
46. Ren Z, Su B, Du Y, Zhou T, Zheng X, Liu J. Effect modifications of BMI transition and trajectory in the associations of adverse childhood experiences with new-onset dementia and its subtypes in older US adults. Gen Psychiatr 2023;36(4):e101092. [doi: 10.1136/gpsych-2023-101092] [Medline: 37622031]
47. Bendix M, Uvnäs-Moberg K, Petersson M, Kaldo V, Åsberg M, Jokinen J. Insulin and glucagon in plasma and cerebrospinal fluid in suicide attempters and healthy controls. Psychoneuroendocrinology 2017 Jul;81:1-7. [doi: 10.1016/j.psyneuen.2017.03.019]
48. Ninla-Aesong P, Puangsri P, Kietdumrongwong P, Jongkrijak H, Noipha K. Being overweight and obese increases suicide risk, the severity of depression, and the inflammatory response in adolescents with major depressive disorders. Front Immunol 2023;14:1197775. [doi: 10.3389/fimmu.2023.1197775] [Medline: 38022570]

Abbreviations

NSSI: Nonsuicidal self-injury

Edited by A Mavragani; submitted 19.09.23; peer-reviewed by D Li, D Di, T Zhang, Y Liu; revised version received 28.08.24; accepted 30.08.24; published 21.02.25.

Please cite as:
Zhang Y, Ying R, Lu W, Liu X, Hu K, Feng Q, Yu Z, Wang Z, Lu F, Miao Y, Ma N, Tao F, Jiang T, Zhang Q
Association Among BMI, Self-Esteem, and Nonsuicidal Self-Injury in Young Adults to Understand the Influence of Socioenvironmental Factors: Longitudinal Study
JMIR Public Health Surveill 2025;11:e52928
URL: https://publichealth.jmir.org/2025/1/e52928
doi:10.2196/52928

Mpox Prevention Self-Efficacy and Associated Factors Among Men Who Have Sex With Men in China: Large Cross-Sectional Study

Quyige Gao[1,2,3*]; Shangbin Liu[2*], MPH; Muzaibaier Tuerxunjiang[3], BPH; Huifang Xu[3], BPH; Jiechen Zhang[1], PhD; Gang Xu[3], PhD; Jianyu Chen[4], PhD; Yong Cai[5], PhD; Fan Hu[6], PhD; Ying Wang[2,3], PhD

[1]Dermatology Department, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]Public Health Research Center, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, 1111 Xianxia Road, Changning District, Shanghai, China

[3]School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[4]College of Public Health, Shanghai University of Medicine and Health Sciences, Shanghai, China

[5]Institute of Community Medical Research, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[6]Hongqiao International Institute of Medicine, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Ying Wang, PhD
Public Health Research Center, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, 1111 Xianxia Road, Changning District, Shanghai, China

Abstract

Background: Self-efficacy in mpox (formerly known as monkeypox) prevention plays a pivotal role in promoting preventive behaviors by fostering a sense of control and motivation, especially among men who have sex with men (MSM), the population most affected by mpox in many countries.

Objective: This study aims to assess the mpox prevention self-efficacy among MSM in China and identify factors influencing it, using a validated mpox prevention self-efficacy scale.

Methods: From October 2023 to March 2024, a nationwide cross-sectional study was conducted among MSM (aged ≥18 years) across 6 geographic regions in China using a snowball sampling method. The recruited participants (effective response rate=2403/2481, 96.9%) were asked to complete an anonymous questionnaire designed based on prior knowledge of mpox and social cognitive theory. The mpox prevention self-efficacy scale was evaluated for construct validity using exploratory factor analysis and confirmatory factor analysis, and its reliability was assessed using the Cronbach α coefficient. Univariate and multivariable logistic regression analyses were used to examine the factors associated with mpox prevention self-efficacy among MSM.

Results: A total of 2403 MSM participants were included, with a mean age of 29 (IQR 19‐39) years. Of these, 1228 (51.1%) were aged 25‐34 years, 1888 (78.6%) held a college degree or higher, and 2035 (84.7%) were unmarried. The median mpox prevention self-efficacy score was 23 (IQR 18‐28). Exploratory factor analysis retained 6 items of the mpox prevention self-efficacy scale. Confirmatory factor analysis confirmed a strong model fit (χ^2=32.1, n=1225; P<.001; comparative fit index=0.991; root mean square error of approximation=0.067; standardized root mean square residual=0.02; goodness-of-fit index=0.992; normed fit index=0.990; incremental fit index=0.991; Tucker-Lewis index=0.974), with all indices within acceptable ranges. The scale demonstrated good internal consistency, with a Cronbach α of 0.859. The positive factors associated with mpox prevention self-efficacy were mpox-related knowledge (OR 1.107, 95% CI 1.070‐1.146), perceived risk awareness (OR 1.338, 95% CI 1.132‐1.583), and mpox risk perception (OR 1.154, 95% CI 1.066‐1.250), while the negative factor was age, with individuals aged 25 years and older exhibiting lower self-efficacy in mpox prevention (25‐34 years: OR 0.789, 95% CI 0.642‐0.970; 35‐44 years: OR 0.572, 95% CI 0.444‐0.736; 45 years and older: OR 0.569, 95% CI 0.394‐0.823).

Conclusions: These findings highlight the critical role of targeted interventions to enhance mpox prevention self-efficacy, particularly through increasing knowledge, perceived risk awareness, and risk perception. Such interventions are especially important for middle-aged and older MSM, who may experience a decline in self-efficacy. Strengthening self-efficacy in these areas is essential for promoting sustained preventive behaviors, improving mental well-being, and contributing to more effective mpox prevention and control within the MSM community.

(JMIR Public Health Surveill 2025;11:e68400) doi:10.2196/68400

KEYWORDS

mpox; self-efficacy; men who have sex with men; MSM; monkeypox; cross-sectional study; mpox prevention self-efficacy scale

Introduction

Mpox (formerly known as monkeypox) is a viral illness caused by the *Monkeypox virus*, typically characterized by a skin rash or mucosal lesions lasting 2 ‐ 4 weeks. Other symptoms include headache, fever, back pain, muscle aches, fatigue, and swollen lymph nodes [1]. The World Health Organization declared the mpox outbreak a Public Health Emergency of International Concern twice: in July 2022 and again in August 2024 [2,3]. As of June 30, 2024, there were 99,176 confirmed cases and 208 deaths reported across 116 countries, including 2460 cases in China [4]. The outbreak predominantly affects men who have sex with men (MSM), who account for 85.8% of cases with known sexual behavior data [4,5]. China, with the world's largest MSM population—estimated at 8.3 million, nearly 1.8 times the size of the United States' MSM population—faces unique public health challenges in controlling mpox transmission [6]. In response, the National Disease Control Bureau of China, in collaboration with the National Health Commission, implemented a prevention and control strategy for mpox [7]. This includes educating high-risk populations, promoting self-protection among MSM, and encouraging timely medical consultations. Although the outbreak in China has been effectively controlled since September 2023, the high mobility and risk behaviors of the MSM population require continued vigilance and proactive measures [8].

According to social cognitive theory, self-efficacy, defined as the belief in one's ability to perform actions necessary to achieve desired outcomes [9], is a key factor in initiating and maintaining health-related behaviors [10]. Self-efficacy has been found to mediate the relationship between HIV self-testing provision and testing frequency among MSM [11], while it also directly influences the uptake of pre-exposure prophylaxis in this population [12], suggesting that enhancing self-efficacy may serve as an effective strategy to improve preventive health behaviors in at-risk populations. In the context of mpox prevention, self-efficacy is particularly critical for MSM, as it influences their confidence in adopting preventive measures, such as consistent condom use, vaccination, testing, and seeking timely medical care [13-16]. High self-efficacy has been associated with greater motivation to engage in mpox prevention behaviors and better adherence to preventive strategies, even among MSM who may continue to engage in high-risk behaviors despite awareness of mpox [17,18]. Given these findings, understanding mpox prevention self-efficacy in MSM is essential for informing targeted interventions and enhancing prevention strategies by identifying priority groups and potential barriers to mpox prevention and control.

A significant challenge in studying self-efficacy in mpox prevention is the lack of standardized measurement tools. Previous studies have primarily relied on single-item measures to assess the association between self-efficacy and mpox vaccination intentions [16,19] or used generalized instruments such as the General Self-Efficacy Scale or isolated questions [19,20]. Although these tools provide a general assessment, they fail to capture the nuanced dimensions of self-efficacy specific to mpox prevention. This limitation is particularly pronounced in the MSM population, where tailored assessments are essential to address unique behavioral and psychosocial factors. The lack of specificity and comprehensiveness in these existing measures not only undermines the validity of research findings, but also hampers the development of targeted interventions. Therefore, there is a clear need for tools specifically tailored to mpox prevention, integrating established scales with the unique characteristics of mpox.

Existing research indicates that the self-efficacy of MSM in preventing mpox may be influenced by multiple factors at various levels, including sociodemographic characteristics (eg, age, education level) [21], behavioral factors (eg, engagement in unsafe sexual practices) [22], and disease-related factors (eg, disease-related knowledge, risk perception) [23]. Research indicates that disease-related knowledge forms the basis for accurate health beliefs and enhances confidence in adopting safe sexual behaviors. A positive correlation has been observed between the level of disease-related knowledge and self-efficacy [24,25]. Additionally, mpox perceived risk awareness refers to individuals' recognition of the potential harm or likelihood of mpox in their surroundings [26,27], while risk perception pertains to the subjective judgment of one's personal susceptibility to mpox [28]. Evidence suggests that risk perception is a key trigger for both preventive awareness and action [29]. Along with disease-specific knowledge and awareness of environmental health risks, risk perception significantly influences mpox prevention awareness and behaviors [30]. However, the relationship between mpox perceived risk awareness, mpox risk perception, and self-efficacy, particularly among MSM, remains an underexplored area.

Given the critical role of self-efficacy in sustaining preventive motivation and goal-directed behavior among MSM, it is essential to assess it within the context of mpox prevention. While existing research on mpox prevention in MSM has primarily focused on vaccination willingness and testing behavior [15,31-33], there is a notable gap in understanding the self-efficacy of MSM in mpox prevention and the factors that influence it. Identifying these factors can help guide the development of targeted interventions. This study aims to evaluate mpox prevention self-efficacy among MSM in China, examining the impacts of sociodemographic characteristics, mpox-related knowledge, perceived risk awareness, and risk perception. By identifying these factors, the study seeks to provide a theoretical basis for designing more effective mpox prevention strategies tailored to this vulnerable population, while simultaneously contributing to improved behavioral outcomes and mental well-being within the MSM community.

Methods

Study Design

From October 2023 to March 2024, a nationwide cross-sectional study was conducted among MSM (age ≥18 years) across 6 geographic regions in China. Survey sites were located in the Northwest (Xinjiang Uyghur Autonomous Region), Northeast (Liaoning Province), Central (Shaanxi Province), Southwest (Yunnan Province), Southeast (Guangdong Province), and Eastern coastal region (Shanghai Municipality). Details of the program have been described previously [34].

Participant Recruitment and Data Collection

Eligibility criteria for participants included: (1) born male; (2) aged ≥18 years old; (3) have ever engaged in sex with men within the last six months; (4) primarily reside in one of the selected locations during the same period; and (5) agree to participate in the survey. Simultaneously, participants were excluded if they: (1) completed the questionnaire within 300 seconds; (2) failed the quality control questions (eg, Is Guangzhou, Shanghai, Beijing, or Shenzhen the capital city in China?); and (3) had an IP address indicating a location outside the target regions.

Participants were recruited from 6 cities in China with the assistance of local governmental centers for disease control and non-governmental organizations (NGOs), using a snowball sampling method. The NGOs routinely engaged with the MSM population in the region and were well-acquainted with local disease prevention and control initiatives. Initially, NGO staff, serving as primary investigators in each region, underwent training on recruitment procedures. They were tasked with recruiting 5 - 10 “seed” participants (MSM individuals with good outreach potential) who met the inclusion criteria from each region. These seed participants were then asked to recruit additional participants from their social networks until the target sample size was achieved. Data collection was conducted entirely anonymously through an online survey. Participants accessed the electronic questionnaire via QR code scanning at designated survey locations. NGO staff assisted participants in completing the online survey, providing technical support and non-directive guidance as necessary, in accordance with the study protocol outlined previously [34].

After thorough verification, a total of 2481 questionnaires were collected, of which 2403 were deemed valid, yielding an effective response rate of 96.9%. The collected data underwent thorough cleaning and processing, including checks for missing data, outliers, and logical inconsistencies, with necessary adjustments and corrections made accordingly.

Ethical Considerations

The study was approved by Shanghai University of Medicine and Health Sciences (approval number 2023-MSMMPOX-22 - 310222197604080237) on October 8, 2023. All participants signed electronic informed consent forms. Data collection was conducted anonymously using an electronic questionnaire administered through a secure, privacy-compliant platform (Wenjuanxing) via QR code scanning. To encourage participation, each individual who completed the survey and passed quality checks received Chinese yuan (CNY) 80 (about US $10.96). To ensure confidentiality, deidentified data were used for subsequent analysis.

Variables

Sociodemographic Characteristics

A range of sociodemographic characteristics (including age, education level, marital status, and monthly income) and disease diagnoses (including hypertension, diabetes, and hyperlipidemia) were assessed via self-report.

Mpox Prevention Self-Efficacy

Based on Bandura’s self-efficacy theory and established scales [19], such as the HIV Prevention Self-Efficacy Scale [35] and the Self-Efficacy for HIV Prevention Behaviors Scale [36], we developed a 6-item mpox prevention self-efficacy scale tailored to the characteristics of MSM and mpox prevention. The scale aimed to assess individuals’ confidence in their ability to effectively prevent mpox. Each item was rated on a 5-point Likert scale, ranging from “1 strongly disagree” to “5 strongly agree,” as described in Table S1 in Multimedia Appendix 1. The composite variable was calculated by summing the scores of these items (range: 6 - 30, Cronbach α=0.859), with a higher score indicating higher mpox prevention self-efficacy. A total score was categorized into 2 groups based on the median (median=23, skewness=−0.94): no/low self-efficacy (<23) and self-efficacy (≥23). When the independent variables are uncorrelated and exhibit a skewed distribution, the median is considered an effective tool for segmentation, which is the case for the variables in our study [37,38].

Mpox-Related Knowledge

Mpox-related knowledge included 12 yes/no questions designed to evaluate participants’ understanding of mpox, covering various aspects such as pathogenesis (eg, “Mpox is a viral infectious disease”), epidemiological characteristics (eg, “Mpox can be transmitted through mucous membranes and broken skin”), clinical manifestations (eg, “The only symptom of mpox is a rash”), and preventive measures (eg, “Smallpox vaccination can be used to prevent mpox”). The overall score was calculated by counting the number of correct answers, with each correct response scoring 1 point. Higher scores indicated greater knowledge of mpox, with a possible range of 0 - 12 (Cronbach α=0.80).

Mpox Perceived Risk Awareness

Perceived risk awareness, primarily utilized in public environmental health research, refers to the level of participants’ attentiveness to potential risks in their environment [26,27]. We developed a composite variable for mpox perceived risk awareness based on 4 statements [28], each rated on a 5-point Likert scale (from “1=strongly agree” to “5=strongly disagree”). A composite scale score was calculated by summing the item scores, with higher scores indicating greater awareness of risks associated with mpox (range: 0 - 20; Cronbach α=0.906). A total score was categorized into 2 groups based on the median: low perceived risk awareness (<17.00) and high perceived risk awareness (≥17.00).

Mpox Risk Perception

Risk perception refers to how people perceive the likelihood of experiencing harm or loss due to a disease and represents a subjective judgment of disease susceptibility [28]. It is a core construct in many health behavior theories, such as the Protection Motivation Theory and the Health Belief Model [39]. In the Health Belief Model, risk perception encompasses perceived susceptibility (belief in the likelihood of contracting a disease) and perceived severity (the perceived seriousness of the disease) [40]. In this study, considering the substantial variability in the clinical manifestations of mpox among individuals, mpox risk perception specifically focuses on participants' beliefs regarding the likelihood of contracting mpox. The statement "I believe I am someone who is likely to contract mpox" was used to assess personal mpox risk perception on a 5-point Likert scale, ranging from "strongly disagree" to "strongly agree" [28].

Statistical Analysis

First, descriptive statistics were used to characterize the distribution of the variables. Continuous variables, which were non-normally distributed, were presented as medians and IQRs, while categorical variables were described using frequencies and percentages. Second, the content validity of the questionnaire was evaluated using the critical ratio method and Spearman correlation coefficient (N=2403). The total scores of the mpox prevention self-efficacy scale were ranked from low to high, with the bottom 27% classified as the low-score group and the top 27% as the high-score group. The discrimination of each item was tested using an independent samples *t* test between these 2 groups. Additionally, the total sample was randomly split into 2 subsets using a random number method: one subset (n=1178) was used for exploratory factor analysis (EFA), and the other (n=1225) for confirmatory factor analysis (CFA), to cross-validate the scale. For EFA, the feasibility of factor analysis was assessed using the Kaiser-Meyer-Olkin test and Bartlett test of sphericity. For CFA, model fit was assessed using several fit indices: the chi-square goodness-of-fit (χ^2/df), goodness-of-fit index (GFI), root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), normed fit index (NFI), comparative fit index (CFI), Tucker-Lewis index (TLI), and incremental fit index (IFI). The average variance extracted and composite reliability were also computed to assess the scale's convergent and discriminant validity. Third, univariate logistic regression was employed to examine the crude associations between the outcome variable and the variables of interest (ie, mpox-related and demographic variables). Fourth, a multivariable logistic regression analysis was performed to identify factors associated with mpox prevention self-efficacy, using the Enter method. Odds ratios (OR) and 95% CIs were estimated. Variance inflation factors were calculated to ensure the independence of each variable in the model, as described in Table S2 in Multimedia Appendix 2. The Hosmer-Lemeshow test indicated a good model fit (χ^2_8=6.047, *P*=.642). Descriptive, univariate, and multivariable analyses were conducted using SPSS software (version 22.0; IBM Corp). The reliability and validity of the questionnaire were assessed using AMOS software (version 28.0; IBM Corp). Forest plots were generated using GraphPad Prism software (version 9.5; GraphPad Software Inc).

Results

Demographic Characteristics

The median age of participants was 29 (IQR 19 - 39) years old. In the 30-point self-efficacy score, the median mpox prevention self-efficacy score was 23 (IQR 18 - 28). The majority held a college degree or higher (n=1888, 78.6%), were unmarried (n=2035, 84.7%), and did not have diabetes (n=2334, 97.1%), hypertension (n=2248, 93.6%), or hyperlipidemia (n=2181, 90.8%). Among the participants, those with a college education or higher (1017/1888, 53.9%) and those who were divorced or widowed (62/113, 54.9%) demonstrated higher levels of mpox prevention self-efficacy (Table 1).

Table . Background characteristics for all included MSM[a] in 6 Chinese cities according to mpox prevention self-efficacy (N=2403), from a survey conducted from October 2023 to March 2024.

Characteristic	Overall	No/low self-efficacy	Self-efficacy	*P* value
Age (years)				
Median (IQR)	29 (19 - 39)	30 (19 - 41)	28 (19 - 37)	<.001
18 - 24, n (%)	564 (23.5)	225 (20)	339 (26.5)	<.001
25 - 34, n (%)	1228 (51.1)	563 (50.1)	665 (52)	
35 - 44, n (%)	459 (19.1)	249 (22.2)	210 (16.4)	
≥45, n (%)	152 (6.3)	86 (7.7)	66 (5.2)	
Education level, n (%)				
Junior high school and below	157 (6.5)	85 (54.1)	72 (45.9)	.16
Senior high school	358 (14.9)	167 (46.7)	191 (53.4)	
College or above	1888 (78.6)	871 (46.1)	1017 (53.9)	
Marital status, n (%)				
Unmarried	2035 (84.7)	937 (46)	1098 (54)	.11
Married	255 (10.6)	135 (52.9)	120 (47.1)	
Divorced or widowed	113 (7.4)	51 (45.1)	62 (54.9)	
Income (CNY[b]), n (%)				
≤3000 (US $414)	502 (20.9)	254 (50.6)	248 (49.4)	.01
3001 - 6000 (US $414-$828)	851 (35.4)	403 (47.4)	448 (52.6)	
6001 - 12,000 (US $828-$1656)	802 (33.4)	340 (42.4)	462 (57.6)	
>12,000 (US $1656)	248 (10.3)	126 (50.8)	122 (49.2)	
Hypertension, n (%)				
No	2248 (93.6)	1049 (46.7)	1199 (53.3)	.80
Yes	155 (6.5)	74 (47.7)	81 (52.3)	
Diabetes, n (%)				
No	2334 (97.1)	1086 (46.5)	1248 (53.5)	.24
Yes	69 (2.9)	37 (53.6)	32 (46.4)	
Hyperlipemia, n (%)				
No	2181 (90.8)	991 (45.4)	1190 (54.6)	<.001
Yes	222 (9.2)	132 (59.5)	90 (40.5)	
Mpox-related knowledge, median (IQR)	8 (5-11)	8 (5-11)	9 (6 - 12)	<.001
Perceive risk awareness, n (%)				
Low	1173 (48.8)	609 (51.9)	564 (48.1)	<.001
High	1230 (52.2)	514 (41.8)	716 (58.2)	
Mpox risk perception, median (IQR)	3 (2-4)	3 (2-4)	4 (3-5)	<.001

[a] MSM: men who have sex with men.

[b] CNY: Chinese yuan. An exchange rate of CNY 1=US $0.14 has been applied.

Reliability and Validity Assessment of the Mpox Prevention Self-Efficacy Scale

Content Validity of the Scale

The results of the independent samples *t* test revealed statistically significant differences between the high-score group (top 27%) and the low-score group (bottom 27%) across all items (*P*<.001), as described in Table S1 in Multimedia Appendix 1. The critical ratio was 65.957, indicating that the scale possesses strong discriminatory power. Furthermore, Pearson correlation analysis demonstrated that the correlation coefficients between the total questionnaire score and the scores of the 6 items ranged from 0.746 to 0.817 (*P*<.001), suggesting that each item is highly representative of the overall construct.

Construct Validity of the Scale

The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.817, and the Bartlett test yielded a χ^2_{15} value of 3442.5 (*P*<.001), indicating that the data were suitable for factor analysis. EFA of the baseline data extracted a single factor with an eigenvalue greater than 1.00, accounting for 60.02% of the total variance. The factor loadings for each item on the common factor ranged from 0.630 to 0.799 (Table 2).

Table . Factor loadings of the mpox prevention self-efficacy scale for men who have sex with men (n=1225).

Path	Standardized factor loadings	SE	*P* value	CR[a]	AVE[b]
C1_1<--Self_efficacy	0.630			0.845	0.478
C1_2<--Self_efficacy	0.640	0.043	<.001		
C1_3<--Self_efficacy	0.645	0.070	<.001		
C1_4<--Self_efficacy	0.664	0.058	<.001		
C1_5<--Self_efficacy	0.752	0.056	<.001		
C1_6<--Self_efficacy	0.799	0.055	<.001		

[a] CR: composite reliability.

[b] AVE: average variance extracted.

We then conducted a CFA on the baseline data to evaluate the global model fit of the factor structure suggested by the EFA, as described in Figure S1 in Multimedia Appendix 3. The overall model fit was good (χ^2_5=32.1, n=1225; *P*<.001; CFI=0.991; RMSEA=0.067; SRMR=0.02; GFI=0.992; NFI=0.990; IFI=0.991; TLI =0.974). All indices were within acceptable ranges, indicating that the model was well-aligned with the data (Table 3).

Table . Model fit index of the mpox prevention self-efficacy scale for men who have sex with men (n=1225).

Model fitting index	Chi-square (*df*)[a]	RMSEA[b]	SRMR[c]	GFI[d]	CFI[e]	NFI[f]	TLI[g]	IFI[h]
Reference value	<3.000	<0.080	<0.050	>0.900	>0.900	>0.900	>0.900	>0.900
Model value	6.416	0.067	0.0184	0.992	0.991	0.990	0.974	0.991

[a] Chi-square (*df*): χ^2 goodness-of-fit.

[b] RMSEA: root mean square error of approximation.

[c] SRMR: standardized root mean square residual.

[d] GFI: goodness-of-fit index.

[e] CFI: comparative fit index.

[f] NFI: normed fit index.

[g] TLI: Tucker-Lewis index.

[h] IFI: incremental fit index.

Reliability Analysis of the Scale

The Cronbach α coefficient for the mpox prevention self-efficacy scale was 0.859, indicating good internal consistency (Cronbach α>0.80).

Factors Associated With Mpox Prevention Self-Efficacy

Univariate Logistic Regression Analysis

Univariate analysis revealed that higher levels of mpox-related knowledge, greater perceived risk awareness, and increased mpox risk perception were significantly associated with higher mpox prevention self-efficacy. Conversely, older MSM demonstrated lower self-efficacy in mpox prevention. All these differences were statistically significant (Table 4).

Table . Univariate logistic regression analysis of mpox prevention self-efficacy among men who have sex with men in 6 Chinese cities (N=2403), from a survey conducted from October 2023 to March 2024.

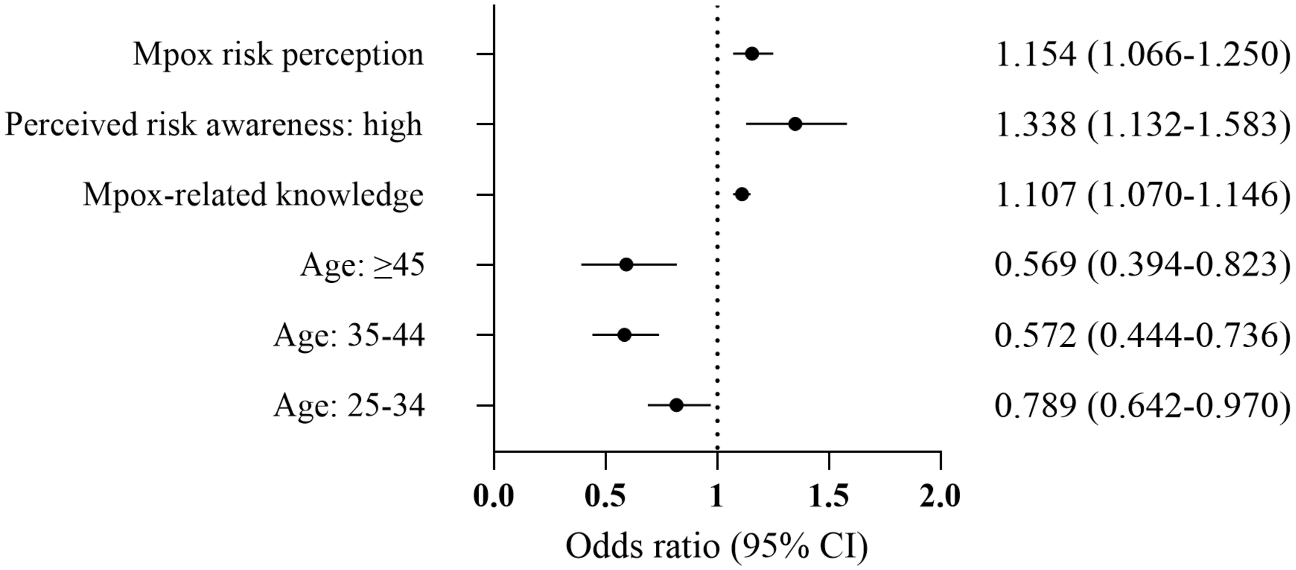
Characteristic	Odds ratio	95% CI	*P* value
Age group, years (reference: 18‑24)			
25‑34	0.784	0.640‑0.960	.02
35‑44	0.560	0.436‑0.718	<.001
≥45	0.509	0.355‑0.732	<.001
Education level (reference: junior high school and below)			
Senior high school	1.350	0.927‑1.967	.12
College or above	1.378	0.994‑1.911	.05
Marital status (reference: unmarried)			
Married	0.759	0.584‑0.985	.04
Divorced or widowed	1.037	0.709‑1.518	.85
Income, CNY[a] (reference: ≤3000 [US $414])			
3001‑6000 (US $414-$828)	1.139	0.913‑1.420	.25
6001‑12,000 (US $828-$1656)	1.392	1.112‑1.741	.004
>12,000 (US $1656)	0.992	0.732‑1.344	.96
Disease-related factor			
Mpox-related knowledge	1.124	1.088‑1.163	<.001
Perceived risk awareness (high versus low)	1.338	1.132‑1.583	.001
Mpox risk perception	1.157	1.070‑1.252	<.001

[a] CNY: Chinese yuan. The exchange rate used is 1 CNY =0.14USD.

Multivariable Logistic Regression Analysis

Variables with statistically significant differences from the univariate analysis were included in the multivariable logistic regression model (Figure 1). In this analysis (Y: mpox prevention self-efficacy, 0=no/low self-efficacy, 1=self-efficacy), the positive factors associated with mpox prevention self-efficacy were mpox-related knowledge (OR 1.107, 95% CI 1.070‑1.146), perceived risk awareness (OR 1.338, 95% CI 1.132‑1.583), and mpox risk perception (OR 1.154, 95% CI 1.066‑1.250). Individuals over the age of 25 years may exhibit lower self-efficacy in mpox prevention (25‑34 years: OR 0.789, 95% CI 0.642‑0.970; 35‑44 years: OR 0.572, 95% CI 0.444‑0.736; 45 years and older: OR 0.569, 95% CI 0.394‑0.823).

Figure 1. Multivariable logistic regression analysis of mpox prevention self-efficacy among MSM in 6 Chinese cities (N=2403), from a survey conducted from October 2023 to March 2024. The categories “18‐24 years old” and “low perceived risk awareness” were used as references in the model. MSM: men who have sex with men.



Discussion

Principal Findings

This study represents one of the earliest investigations into mpox prevention self-efficacy and its influencing factors among MSM. Given that the mpox outbreak was once again declared a Public Health Emergency of International Concern in 2024, concerns about the potential spread of the virus—particularly within the MSM community—have intensified. This highlights the critical need for effective preventive strategies tailored to this high-risk group. Considering the crucial role of self-efficacy in health behaviors and disease prevention, our findings offer valuable public health insights for managing mpox outbreaks and enhancing the psychological and behavioral well-being of the MSM population.

Comparison With Prior Work

The study found that the median mpox prevention self-efficacy score was 23 (IQR 18‐28), with a negatively skewed distribution of self-efficacy scores. This suggests that most participants demonstrated relatively high self-efficacy scores in preventing mpox, though a small proportion reported lower self-efficacy levels. These results indicate that, while the MSM community generally exhibits strong preparedness for mpox prevention, targeted interventions may be necessary for individuals with lower self-efficacy. Establishing and reinforcing high levels of mpox prevention self-efficacy is essential, as it plays a critical role in identifying at-risk populations and implementing targeted interventions [41,42]. Strengthening self-efficacy, particularly among those with lower initial scores, can improve both individual and community-level responses to prevention efforts, thereby contributing to more effective mpox control strategies. Therefore, public health programs should prioritize efforts to improve self-efficacy, focusing on education and risk awareness to increase confidence in preventive actions.

The study found that MSM with an education level of junior high school or below exhibited lower levels of self-efficacy. These findings are consistent with the demographic characteristics observed in previous studies of MSM with mpox [43,44]. Higher education levels may be associated with increased awareness of self-protection measures, which may contribute to higher self-efficacy and a reduction in risky behaviors [45]. This underscores the potential effectiveness of educational interventions aimed at enhancing self-efficacy, particularly for individuals with lower education levels.

An interesting finding of this study is that the proportion of married MSM exhibiting high self-efficacy in mpox prevention was lower compared to their unmarried, divorced, or widowed counterparts, despite the lack of statistical significance (P>.05). This trend may be explained by the fact that unmarried, divorced, or widowed MSM tend to receive less familial support, which may drive them to be more proactive in seeking preventive and self-protective measures to avoid infection [46]. Previous studies have also shown that divorced or widowed MSM tend to have higher levels of mpox-related knowledge compared to married individuals [47], possibly because the latter may be less inclined to seek out information or self-protective behaviors related to mpox [48]. This observation warrants further exploration, as it may suggest that marital status could still influence self-efficacy in mpox prevention, albeit not significantly in this study.

Accurate measurement of self-efficacy is fundamental for developing effective mpox prevention strategies [49]. In this study, the mpox prevention self-efficacy scale was validated and demonstrated strong reliability, validity, and practical utility. Although the General Self-Efficacy Scale, originally developed by Schwarzer and Jerusalem [50], is widely used due to its excellent internal consistency and test-retest reliability, it lacks the specificity required for assessing self-efficacy in the MSM population and in the context of mpox prevention, limiting its applicability in this context. Previous research on mpox predominantly has largely relied on isolated questions to assess self-efficacy, without employing a comprehensive or systematic scale tailored to the MSM community [51]. To fill this gap, this study designed and evaluated the mpox prevention self-efficacy scale specifically tailored for MSM. The scale offers a more precise tool for assessing self-efficacy and its related factors among MSM, providing valuable insights for future research and interventions aimed at improving mpox prevention strategies.

Univariate analysis revealed a significant negative association between age and mpox prevention self-efficacy, a relationship that remained significant in the multivariable regression model. These findings indicate that older MSM may exhibit lower levels of self-efficacy in mpox prevention compared to their younger counterparts (aged 18 - 24 years). This association may be partially attributed to older MSM having reduced access to mpox-related information. Age has been identified as a key determinant of mpox-related knowledge [52], with older individuals potentially having less exposure to relevant health information, which may contribute to lower awareness, confidence, and self-efficacy in preventing mpox. In contrast, younger MSM, being more familiar with using the internet, are more likely to access mpox-related information. Furthermore, previous studies have shown that older MSM are less willing to change their sexual behavior compared to their younger counterparts [37]. Specifically, older MSM are more likely to maintain risky sexual behaviors even after becoming aware of the health risks associated with mpox, highlighting the need for tailored interventions aimed at this population to enhance their self-efficacy in preventing mpox. Therefore, age-specific communication strategies should be considered in mpox prevention efforts, particularly targeting middle-aged and older MSM through effective media channels. Given that MSM extensively use the internet for socializing, seeking sexual partners, and accessing sexual health information [53,54], eHealth technologies have emerged as promising tools for delivering prevention interventions to this population [55]. Research has shown that enhanced behavioral interventions, such as tailored strategies, can lead to reductions in unprotected sex, while advanced eHealth technologies, including social media platforms, are associated with increased health screening in the MSM community [56]. Therefore, eHealth-based interventions, incorporating targeted communication strategies and educational tools, can be effectively implemented on widely used social platforms or websites to enhance mpox prevention self-efficacy, particularly in populations with lower levels of self-efficacy. Tailored interventions focusing on improving self-efficacy among older and less educated MSM may be particularly beneficial, as these groups tend to exhibit lower self-efficacy, which may increase their susceptibility to infection. These findings underscore the importance of demographic factors in informing targeted mpox prevention strategies, especially in guiding surveillance and health education practices.

This study found that mpox-related knowledge, perceived risk awareness, and risk perception were positively correlated with self-efficacy in mpox prevention. Previous research has demonstrated that attitudes towards a disease are closely linked to disease-related knowledge, suggesting that lower levels of knowledge among MSM are associated with more negative attitudes and decreased intentions to engage in mpox prevention behaviors [57]. Mpox perceived risk awareness reflects beliefs about the potential harm or likelihood of mpox in their surroundings, while risk perception represents a subjective judgment of personal susceptibility to the disease. According to social cognitive theory, individuals who recognize the severity of a negative health condition, believe that it can be avoided, maintain a positive attitude toward recommended actions, and have confidence in their ability to successfully execute these actions are more likely to engage in health-related behaviors [22,58]. Therefore, mpox prevention efforts among MSM should include further public health education and training to enhance perceived risk awareness, provide targeted risk communication, and ensure the privacy and confidentiality of services, which are crucial for effective mpox prevention within this population.

This research has several strengths. First, as one of the earliest studies in China examining mpox prevention self-efficacy among MSM, the findings carry significant public health implications for the prevention and control of mpox outbreaks in the country, while also enhancing the psychological and behavioral well-being of the MSM community. Furthermore, the study's multicenter design contributes to the generalizability of the results, as it incorporates a diverse sample from various regions across China.

Limitations

Despite the valuable insights provided by the study, several limitations must be acknowledged. First, participant recruitment was facilitated by government centers for disease control departments and NGOs using a snowball sampling method, which was used to improve the efficiency of identifying and recruiting members of the hidden MSM population [59]. Although this approach is effective for accessing hard-to-reach groups, it inherently introduces selection bias, as the sample composition is influenced by the characteristics of initial recruits (seeds) and the subsequent limited recruitment chains. Consequently, this may restrict the generalizability of our findings. Second, the sensitive nature of some inquiries, combined with self-reported data, may lead to information bias. Third, the cross-sectional design restricts the ability to infer causality, highlighting the need for future longitudinal studies to further explore the temporal relationships between mpox-related knowledge, perceived risk awareness, risk perception, and self-efficacy. Such prospective research could also evaluate the impact of targeted interventions aimed at enhancing self-efficacy in mpox prevention among MSM.

Conclusion

This study examined the mpox prevention self-efficacy of MSM and its influencing factors, based on the design and evaluation of the mpox prevention self-efficacy scale. The findings demonstrated significant associations between mpox prevention self-efficacy and sociodemographic characteristics, as well as mpox-related factors. These results underscore the importance of targeted interventions to enhance mpox prevention self-efficacy, particularly through efforts to increase mpox-related knowledge, perceived risk awareness, and risk perception. Special attention should be given to middle-aged and older MSM, who may experience a decline in self-efficacy. Future research should adopt longitudinal designs to better understand the temporal relationships between self-efficacy and related factors, thereby providing more robust evidence to guide the development of effective mpox prevention strategies. Ultimately, these efforts are crucial for improving mpox prevention outcomes and promoting the overall health of the MSM community.

Acknowledgments

The authors would like to express their gratitude to Shanghai Qingai Health Center (Shanghai), Trans China (Yunnan), Consultation Center of AIDS Aid and Health Service (Liaoning), Xinjiang Rong'ai Health Consulting Co Ltd (Xinjiang), Lingnan Partners Community Support Center (Guangdong), and Xi'an Zhitong AIDS Prevention and Health Center (Shaanxi) for their contributions to data collection. We also extend our thanks to all the participants involved in this study. This study was funded by Key discipline projects of Shanghai Three-Year Action Plan for Public Health under Grant (GWVI-11.1-29), Public Health Research Project from Shanghai Disease Control and Prevention Administration (2024GKM30), Changning District Medical and Health Research Special Project (CNKW2022Y08), Tongren Hospital Talent Launch Project (TR2023rc11, TR2024RC10), Science and Technology Commission Shanghai Municipality (number 20JC1410204) for the Seroepidemiological Study of Novel Coronavirus Pneumonia in Key Populations, national health commission of the PRC (grant number Z155080000004), and the Research Fund of Shanghai Tongren Hospital, Shanghai Jiaotong University School of Medicine (TRYJ2024LC13, TRYJ2022LC10).

Data Availability

The datasets generated and analyzed during this study are not publicly available due to privacy and ethical considerations, but are available from the corresponding author on reasonable request.

Authors' Contributions

QG and SL performed the analyses, interpreted the data, and developed the manuscript; MT and HX conceptualized and validated data analyses; JZ and GX contributed to the data collection and provided critical insights into the methodology; JC reviewed and edited the manuscript; YC, FH, and YW designed and supervised this study, and reviewed and edited the manuscript. All authors have made a significant contribution to this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Independent samples *t* test discrimination results of the mpox prevention self-efficacy scale (N=2403).

[DOC File, 18 KB - publichealth_v11i1e68400_app1.doc]

Multimedia Appendix 2

Results of collinearity diagnosis for the multivariable logistic regression analysis of mpox prevention self-efficacy among MSM in 6 Chinese cities (N=2403), from a survey conducted from October 2023 to March 2024.

[DOC File, 18 KB - publichealth_v11i1e68400_app2.doc]

Multimedia Appendix 3

Standardized one-factor structural model of the mpox prevention self-efficacy scale (n=1225).

[PNG File, 101 KB - publichealth_v11i1e68400_app3.png]

References

1. Lu J, Xing H, Wang C, et al. Mpox (formerly monkeypox): pathogenesis, prevention, and treatment. Signal Transduct Target Ther 2023 Dec 27;8(1):458. [doi: 10.1038/s41392-023-01675-2] [Medline: 38148355]
2. Mpox (monkeypox) outbreak. World Health Organization. 2022. URL: https://www.who.int/emergencies/situations/monkeypox-oubreak-2022 [accessed 2024-08-20]
3. WHO Director-General's opening remarks at the media briefing following the IHR emergency committee meeting on the upsurge of mpox 2024-14. World Health Organization. 2024. URL: https://www.who.int/zh/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-following-the-ihr-emergency-committee-meeting-on-the-upsurge-of-mpox-2024---14-august-2024 [accessed 2024-08-20]
4. 2022-24 mpox (monkeypox) outbreak: global trends. World Health Organization. 2024. URL: https://worldhealthorg.shinyapps.io/mpx_global/ [accessed 2024-08-20]

5. Vaughan AM, Cenciarelli O, Colombe S, et al. A large multi-country outbreak of monkeypox across 41 countries in the WHO European Region, 7 March to 23 August 2022. Euro Surveill 2022 Sep;27(36):2200620. [doi: 10.2807/1560-7917.ES.2022.27.36.2200620] [Medline: 36082686]
6. Hu M, Xu C, Wang J. Spatiotemporal analysis of men who have sex with men in Mainland China: social app capture-recapture method. JMIR mHealth uHealth 2020 Jan 24;8(1):e14800. [doi: 10.2196/14800] [Medline: 32012086]
7. Notice of the General Office of the National Health Commission on the issuance of the Technical Guidelines for Monkeypox Prevention and Control (2022 Edition) [Website in Chinese]. Gov.cn. 2022. URL: https://www.gov.cn/zhengce/zhengceku/2022-07/01/content_5698880.htm [accessed 2024-09-22]
8. Chen HJ, Tang HL, Li PL, et al. Survey on the awareness rate of mpox knowledge and related factors among men who have sex with men in China [Article in Chinese]. Zhonghua Liu Xing Bing Xue Za Zhi 2024 Apr 10;45(4):559-565. [doi: 10.3760/cma.j.cn112338-20231030-00257] [Medline: 38678353]
9. Bandura A. Self-efficacy mechanism in human agency. American Psychologist 1982;37(2):122-147. [doi: 10.1037//0003-066X.37.2.122]
10. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Advances in Behaviour Research and Therapy 1978 Jan;1(4):139-161. [doi: 10.1016/0146-6402(78)90002-4]
11. Zhang C, Huang D, Goldsamt LA, et al. Self-efficacy mediates the effect of providing HIV self-testing kits on HIV testing frequency among Chinese men who have sex with men: results from a randomised controlled trial. Sex Health 2023 Apr;20(2):118-125. [doi: 10.1071/SH22110] [Medline: 36848626]
12. Liu H, Lai G, Shi G, Zhong X. The influencing factors of HIV-preventive behavior based on health belief model among HIV-negative MSMs in Western China: a structural equation modeling analysis. Int J Environ Res Public Health 2022 Aug 17;19(16):10185. [doi: 10.3390/ijerph191610185] [Medline: 36011822]
13. DiIorio C, Dudley WN, Kelly M, Soet JE, Mbwara J, Sharpe Potter J. Social cognitive correlates of sexual experience and condom use among 13- through 15-year-old adolescents. J Adolesc Health 2001 Sep;29(3):208-216. [doi: 10.1016/s1054-139x(00)00200-7] [Medline: 11524220]
14. Safren SA, Blashill AJ, Lee JS, et al. Condom-use self-efficacy as a mediator between syndemics and condomless sex in men who have sex with men (MSM). Health Psychol 2018 Sep;37(9):820-827. [doi: 10.1037/hea0000617] [Medline: 29927272]
15. Luo S, Jiao K, Zhang Y, et al. Behavioral intention of receiving monkeypox vaccination and undergoing monkeypox testing and the associated factors among young men who have sex with men in China: large cross-sectional study. JMIR Public Health Surveill 2024 Mar 19;10:e47165. [doi: 10.2196/47165] [Medline: 38502181]
16. Lawrence A. Assessing vaccine intentions, knowledge, self-efficacy, and trust: a cross-sectional study on perceptions of monkeypox vaccination and public health risk awareness in Makurdi, Benue State, Nigeria. Cureus 2024 Oct;16(10):e72131. [doi: 10.7759/cureus.72131] [Medline: 39440166]
17. Tan F, Oka P, Dambha-Miller H, Tan NC. The association between self-efficacy and self-care in essential hypertension: a systematic review. BMC Fam Pract 2021 Feb 22;22(1):44. [doi: 10.1186/s12875-021-01391-2] [Medline: 33618661]
18. Wang SD, Loftus P, Pang RD, Kirkpatrick MG. Impact of self-efficacy on daily intention to not smoke. Addict Behav 2021 Jul;118:106877. [doi: 10.1016/j.addbeh.2021.106877] [Medline: 33714032]
19. Winters M, Malik AA, Omer SB. Attitudes towards Monkeypox vaccination and predictors of vaccination intentions among the US general public. PLoS One 2022;17(12):e0278622. [doi: 10.1371/journal.pone.0278622] [Medline: 36454991]
20. Lee J, Lee EH, Chae D. Self-efficacy instruments for type 2 diabetes self-care: a systematic review of measurement properties. J Adv Nurs 2020 Aug;76(8):2046-2059. [doi: 10.1111/jan.14411] [Medline: 32400902]
21. Takahashi LM, Magalong MG, Debell P, Fasudhani A. HIV and AIDS in suburban Asian and Pacific Islander communities: factors influencing self-efficacy in HIV risk reduction. AIDS Educ Prev 2006 Dec;18(6):529-545. [doi: 10.1521/aeap.2006.18.6.529] [Medline: 17166079]
22. Khumsaen N, Stephenson R. Beliefs and perception about HIV/AIDS, self-efficacy, and HIV sexual risk behaviors among young Thai men who have sex with men. AIDS Educ Prev 2017 Apr;29(2):175-190. [doi: 10.1521/aeap.2017.29.2.175] [Medline: 28467158]
23. Villegas Rodríguez N, Ferrer Lagunas LM, Cianelli Acosta R, Miner S, Lara Campos L, Peragallo N. Knowledge and self efficacy related with HIV prevention among Chilean women. Invest Educ Enferm 2011;29(2):222-229. [Medline: 25284914]
24. Carter G, Brown L, Mahnke B, Ohmit A, Woodward B. Sexual communication self-efficacy and PrEP literacy helps to meet HIV prevention outcomes among Black and Latinx individuals. J Prim Care Community Health 2022;13:21501319221099789. [doi: 10.1177/21501319221099789] [Medline: 35578769]
25. Black DS, Sun P, Rohrbach LA, Sussman S. Decision-making style and gender moderation of the self-efficacy-condom use link among adolescents and young adults: informing targeted STI/HIV prevention programs. Arch Pediatr Adolesc Med 2011 Apr;165(4):320-325. [doi: 10.1001/archpediatrics.2011.17] [Medline: 21464380]
26. Lavallee S, Hynds PD, Brown RS, Majury A. Classification of sub-populations for quantitative risk assessment based on awareness and perception: a cross-sectional population study of private well users in Ontario. Sci Total Environ 2023 Jan 20;857(Pt 3):159677. [doi: 10.1016/j.scitotenv.2022.159677] [Medline: 36302430]

27. Gray-Scholz D, Haney TJ, MacQuarrie P. Out of sight, out of mind? Geographic and social predictors of flood risk awareness. Risk Anal 2019 Nov;39(11):2543-2558. [doi: 10.1111/risa.13357] [Medline: 31173665]
28. Jiao K, Xu Y, Huang S, et al. Mpox risk perception and associated factors among Chinese young men who have sex with men: Results from a large cross-sectional survey. J Med Virol 2023 Aug;95(8):e29057. [doi: 10.1002/jmv.29057] [Medline: 37635623]
29. Wiedemann PM, Schütz H. The precautionary principle and risk perception: experimental studies in the EMF area. Environ Health Perspect 2005 Apr;113(4):402-405. [doi: 10.1289/ehp.7538] [Medline: 15811829]
30. Renner B, Schupp H, Vollmann M, Hartung FM, Schmälzle R, Panzer M. Risk perception, risk communication and health behavior change. Zeitschrift für Gesundheitspsychologie 2008 Jul;16(3):150-153. [doi: 10.1026/0943-8149.16.3.150]
31. Li Y, Peng X, Fu L, et al. Monkeypox awareness and low vaccination hesitancy among men who have sex with men in China. J Med Virol 2023 Feb;95(2):e28567 [FREE Full text] [doi: 10.1002/jmv.28567]
32. Huang X, Lin Z, Qin J, et al. Willingness to accept monkeypox vaccine and its correlates among men who have sex with men in Southern China: a web-based online cross-sectional study. Front Public Health 2024;12:1289918. [doi: 10.3389/fpubh.2024.1289918] [Medline: 38384873]
33. Zheng M, Qin C, Qian X, et al. Knowledge and vaccination acceptance toward the human monkeypox among men who have sex with men in China. Front Public Health 2022;10:997637. [doi: 10.3389/fpubh.2022.997637] [Medline: 36388271]
34. Gao Y, Liu S, Xu H, et al. The Mpox vaccine hesitancy scale for Mpox: links with vaccination intention among men who have sex with men in six cities of China. Vaccines (Basel) 2024 Sep 3;12(9):1009. [doi: 10.3390/vaccines12091009] [Medline: 39340039]
35. Ravelo GJ, Sanchez M, Cyrus E, De La Rosa M, Peragallo N, Rojas P. Associations between gender norms and HIV self-efficacy among Latina immigrants in a farmworker community. Ethn Health 2022 Jan;27(1):27-39. [doi: 10.1080/13557858.2019.1659234] [Medline: 31450962]
36. Cianelli R, Ferrer L, Norr KF, et al. Mano a Mano-Mujer: an effective HIV prevention intervention for Chilean women. Health Care Women Int 2012;33(4):321-341. [doi: 10.1080/07399332.2012.655388] [Medline: 22420675]
37. Chan ZYS, Chong SY, Niaupari S, et al. Receptiveness to monkeypox vaccines and public health communication strategies among gay, bisexual and other men who have sex with men in Singapore: cross-sectional quantitative and qualitative insights. Sex Transm Infect 2024 Aug 19;100(6):362-367. [doi: 10.1136/sextrans-2024-056230] [Medline: 38821878]
38. Bien-Gund CH, Shaw PA, Agnew-Brune C, et al. HIV self-testing and risk behaviors among men who have sex with men in 23 US cities, 2017. JAMA Netw Open 2022 Dec 1;5(12):e2247540. [doi: 10.1001/jamanetworkopen.2022.47540] [Medline: 36534398]
39. Gellman M, Turner J. Encyclopedia of Behavioral Medicine 2013. [doi: 10.1007/978-1-4419-1005-9]
40. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q 1984;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
41. Villegas N, Cianelli R, Gonzalez-Guarda R, Kaelber L, Ferrer L, Peragallo N. Predictors of self-efficacy for HIV prevention among Hispanic women in South Florida. J Assoc Nurses AIDS Care 2013;24(1):27-37. [doi: 10.1016/j.jana.2012.03.004] [Medline: 22795758]
42. Sheeran P, Maki A, Montanaro E, et al. The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: a meta-analysis. Health Psychol 2016 Nov;35(11):1178-1188. [doi: 10.1037/hea0000387] [Medline: 27280365]
43. Du M, Zheng M, Liu J, et al. The prevalence of mpox and its association with sexual behavior among Chinese men who have sex with men in early August 2023. J Med Virol 2023 Dec;95(12):e29320. [doi: 10.1002/jmv.29320] [Medline: 38102917]
44. Zheng M, Du M, Yang G, et al. Lower rate of mpox vaccination hesitancy and medical consultation among Chinese men who have sex with men living with HIV in comparison with those living without HIV: a national observational study. Hum Vaccin Immunother 2023 Dec 15;19(3):2290788. [doi: 10.1080/21645515.2023.2290788] [Medline: 38054460]
45. Wei L, Tian J, Guo M, et al. Trajectories of sexual risk behaviors and the associated factors among young men who have sex with men in China. Front Public Health 2022;10:854616. [doi: 10.3389/fpubh.2022.854616] [Medline: 35387185]
46. Wu S, Zhu S, Yan X, et al. Willingness to receive the COVID-19 vaccine among HIV positive men who have sex with men in China: a cross-sectional study. BMC Public Health 2023 Jan 10;23(1):64. [doi: 10.1186/s12889-022-14961-5] [Medline: 36627582]
47. Zheng M, Chen W, Qian X, et al. Awareness of mpox-related knowledge among men who have sex with men in China. BMC Public Health 2023 Mar 30;23(1):600. [doi: 10.1186/s12889-023-15503-3] [Medline: 36997990]
48. Liu Z, Chen Y, Yao T, et al. Factors related to HIV testing frequency in MSM based on the 2011-2018 survey in Tianjin, China: a hint for risk reduction strategy. BMC Public Health 2021 Oct 20;21(1):1900. [doi: 10.1186/s12889-021-11948-6] [Medline: 34670542]
49. Fisher WA, Fisher JD, Harman JJ. The information ‐ motivation ‐ behavioral skills model: a general social psychological approach to understanding and promoting health behavior. In: Social Psychological Foundations of Health and Illness 2009. [doi: 10.1002/9780470753552]

50. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Weinman J, Wright S, Johnston M, editors. Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs: NFER-NELSON; 1995. URL: https://www.religiousforums.com/data/attachment-files/2014/12/22334_285ebbd019ec89856493442b1b6f9154.pdf [accessed 2025-02-11]
51. Bakare D, Salako J, Sogbesan A, Olojede OE, Bakare AA. Assessment of the level of awareness, knowledge, and risk perception of community members about mpox infection in Nigeria. Ann Ib Postgrad Med 2024 Apr 30;22(1):76-87. [Medline: 38939883]
52. Jairoun AA, Al-Hemyari SS, Abdulla NM, et al. Awareness and preparedness of human monkeypox outbreak among university student: Time to worry or one to ignore? J Infect Public Health 2022 Oct;15(10):1065-1071. [doi: 10.1016/j.jiph.2022.08.015] [Medline: 36087548]
53. Holloway IW, Rice E, Gibbs J, Winetrobe H, Dunlap S, Rhoades H. Acceptability of smartphone application-based HIV prevention among young men who have sex with men. AIDS Behav 2014 Feb;18(2):285-296. [doi: 10.1007/s10461-013-0671-1] [Medline: 24292281]
54. Zou H, Fan S. Characteristics of men who have sex with men who use smartphone geosocial networking applications and implications for HIV interventions: a systematic review and meta-analysis. Arch Sex Behav 2017 May;46(4):885-894. [doi: 10.1007/s10508-016-0709-3] [Medline: 27043837]
55. Carpenter KM, Stoner SA, Mikko AN, Dhanak LP, Parsons JT. Efficacy of a web-based intervention to reduce sexual risk in men who have sex with men. AIDS Behav 2010 Jun;14(3):549-557. [doi: 10.1007/s10461-009-9578-2] [Medline: 19499321]
56. Xin M, Viswanath K, Li AYC, et al. The effectiveness of electronic health interventions for promoting HIV-preventive behaviors among men who have sex with men: meta-analysis based on an integrative framework of design and implementation features. J Med Internet Res 2020 May 25;22(5):e15977. [doi: 10.2196/15977] [Medline: 32449685]
57. Opoku MP, Agyei-Okyere E, Nketsia W, Torgbenu EL, Kumi EO. Perceived self-efficacy of students and its influence on attitudes and knowledge about HIV/AIDS in Ghana. Int J Health Plann Manage 2022 Mar;37(2):755-769. [doi: 10.1002/hpm.3371] [Medline: 34697830]
58. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. Health Educ Q 1988;15(2):175-183. [doi: 10.1177/109019818801500203] [Medline: 3378902]
59. Kendall C, Kerr L, Gondim RC, et al. An empirical comparison of respondent-driven sampling, time location sampling, and snowball sampling for behavioral surveillance in men who have sex with men, Fortaleza, Brazil. AIDS Behav 2008 Jul;12(4 Suppl):S97-104. [doi: 10.1007/s10461-008-9390-4] [Medline: 18389357]

Abbreviations

CFA: confirmatory factor analysis
CFI: comparative fit index
EFA: exploratory factor analysis
GFI: goodness-of-fit index
IFI: incremental fit index
MSM: men who have sex with men
NFI: normed fit index
NGO: nongovernmental organization
OR: odds ratio
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
TLI: Tucker-Lewis index

Edited by A Mavragani; submitted 07.11.24; peer-reviewed by B Foroutan, O Oni, S Luo, TF Ojo; revised version received 05.01.25; accepted 05.01.25; published 28.02.25.

Please cite as:
Gao Q, Liu S, Tuerxunjiang M, Xu H, Zhang J, Xu G, Chen J, Cai Y, Hu F, Wang Y
Mpox Prevention Self-Efficacy and Associated Factors Among Men Who Have Sex With Men in China: Large Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e68400
URL: https://publichealth.jmir.org/2025/1/e68400
doi:10.2196/68400

Identifying Preferred Features of Influenza Vaccination Programs Among Chinese Clinicians Practicing Traditional Chinese Medicine and Western Medicine: Discrete Choice Experiment

Liuren Zhang[1,2*], MPH; Linchen Chu[1*], MPH; Maria E Sundaram[3*], PhD; Yi Zhou[1], MPH; Xiu Sun[4], MPH; Zheng Wei[5], MPH; Chuanxi Fu[1], PhD

[1]Institute of Infectious Disease and Vaccine, School of Public Health, Zhejiang Chinese Medical University, Hangzhou, China
[2]The Third People's Hospital of Hubei Province, Wuhan, China
[3]Center for Clinical Epidemiology and Population Health, Marshfield Clinic Research Institute, Marshfield, WI, United States
[4]The Third Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China
[5]The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China
[*]these authors contributed equally

Corresponding Author:
Chuanxi Fu, PhD
Institute of Infectious Disease and Vaccine, School of Public Health, Zhejiang Chinese Medical University, Hangzhou, China

Abstract

Background: Achieving high vaccine coverage among clinicians is crucial to curb the spread of influenza. Traditional Chinese medicine (TCM), rooted in cultural symbols and concepts without direct parallels in modern Western medicine, may influence perspectives on vaccination. Therefore, understanding the preferences of TCM clinicians towards influenza vaccines is of great importance.

Objective: To understand preferences for features of influenza vaccination programs and identify the optimal influenza vaccination program among clinicians practicing TCM and Western medicine.

Methods: We conducted a discrete choice experiment with a national sample of 3085 Chinese clinicians from various hospital levels (n=1013 practicing TCM) from January to May 2022. Simulations from choice models using the experimental data generated the coefficients of preference and predicted the uptake rate of different influenza vaccination programs. Clinicians were grouped by vaccine preference classification through a latent class analysis.

Results: All included attributes significantly influenced clinicians’ preferences for choosing an influenza vaccination program. An approximate hypothetical 60% increase of vaccine uptake could be obtained when the attitude of the workplace changed from “no notice” to “encouraging of vaccination”; there was an approximate hypothetical 35% increase of vaccine uptake when vaccination campaign strategies changed from “individual appointment” to “vaccination in a workplace setting.” In the entire sample, about 30% (946/3085) of clinicians preferred free vaccinations, while 26.5% (819/3085) comprehensively considered all attributes, except vaccination campaign strategies, when making a decision about choosing an influenza vaccination program. Clinicians who practiced TCM, worked in tertiary hospital, or had at least a postgraduate degree exhibited a lower preference for free vaccinations. Clinicians who practiced Western medicine, worked in primary hospital, or had at most a bachelor’s degree had a higher preference for vaccinations in workplace settings.

Conclusions: Offering a range of influenza vaccination programs targeting the preferred attributes of different clinician groups could potentially encourage more clinicians, including those practicing TCM, to participate in influenza vaccination programs.

(JMIR Public Health Surveill 2025;11:e63314) doi:10.2196/63314

KEYWORDS

influenza vaccination program; traditional Chinese medicine; clinicians; vaccine; health care worker; hospital-acquired; effectiveness; antiviral; cross-sectional study

Introduction

Annual seasonal influenza epidemics result in 290,000-650,000 deaths worldwide and over 88,000 preventable deaths in China annually [1,2]. During influenza seasons, health care workers (HCWs) are at a high risk of infection, leading to economic loss for the institutions they work for and potentially causing staff shortages [3]. HCWs may be asked or pressured to work despite having an influenza infection, contributing to clinic- or hospital-acquired influenza transmission [4]. Previous studies have indicated that clinicians play a crucial role in influencing vulnerable groups. Clinicians who were vaccinated were more likely to recommend vaccinations to their patients, highlighting the critical role of this group in improving vaccine uptake [5]. Influenza vaccination is the most cost-effective measure to protect high-risk groups against severe influenza-associated diseases and deaths, and the World Health Organization recommends the influenza vaccination in particular for specific target groups, including HCWs [6].

However, despite efforts to promote influenza vaccination among HCWs, uptake rates remain unsatisfactory in China. In the 2019-2020 season, only 61% of Chinese clinicians, who were recruited at a platform for respiratory medical professionals, received vaccines, which was lower than the 87.9% in Finland [7,8]. In the 2021-2022 season, the coverage was even higher in the United States, which had 91.3% clinicians who received vaccines [9]. Barriers to influenza vaccine uptake among HCWs in China include out-of-pocket costs and insufficient workplace regulations [7]. For example, workplaces did not require or encourage clinicians to get vaccinated before the flu season. Most importantly, previous studies have primarily treated Chinese clinicians as practitioners of modern medicine, overlooking the significant presence of clinicians practicing traditional Chinese medicine (TCM) in China [7,10].

Modern Western medicine, originating in Europe, dominates medical practices globally, including in China, and relies on biomedical science, genetics, and advanced medical technologies to diagnose, treat, and prevent illness [11]. In contrast, TCM, rooted in ancient China, follows a distinct philosophy with unique principles, like Yin-Yang and the Five Elements, and employs practices such as acupuncture and herbal remedies [12]. The theoretical and diagnostic foundation of TCM is difficult to interpret through the framework of Western medicine anatomy and physiology; it is based on the philosophy, logic, and beliefs of a distinct civilization, offering a view of health and disease that differs from modern scientific thinking [13]. Ancient texts such as *Huangdi Nei Jing* and *Shanghan Lun* remain core guides for TCM students today [14]. TCM was categorized as a kind of complementary and alternative medicine by the World Health Organization and has been adopted by over 100 countries [15]. In 2020, the United States had over 30,000 licensed acupuncturists, while Germany and Australia also had thriving TCM practitioner communities, with more than 3000 and 4000 practitioners, respectively [12]. In 2020, there were 682,770 registered TCM clinicians, accounting for 16.7% of all clinicians, in mainland China, who provided medical services for over 1 billion outpatients and 35 million hospitalizations [16]. Considering the massive medical work undertaken by TCM clinicians, it is important to understand their influenza vaccination preferences.

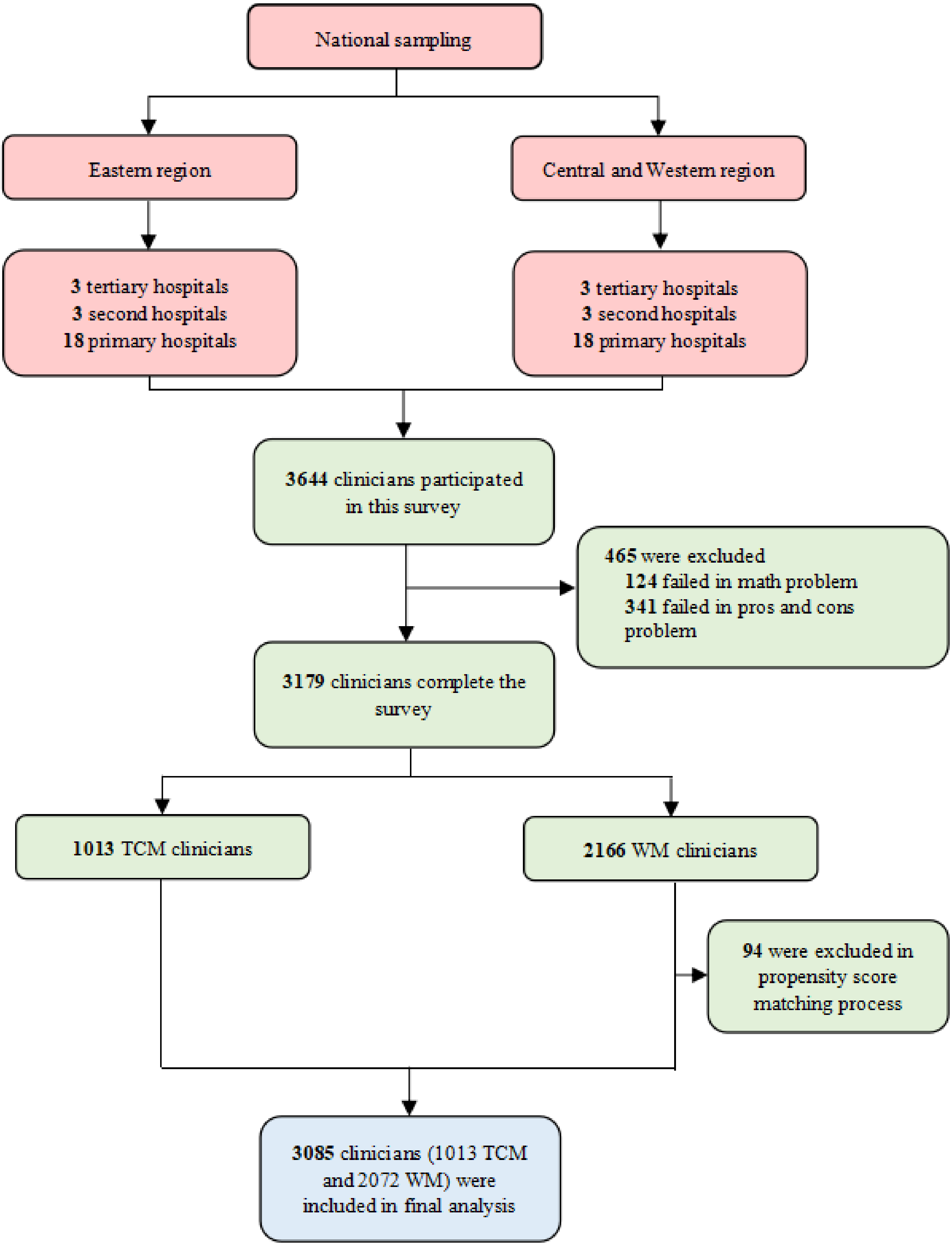
We conducted a nationwide survey using a discrete choice experiment (DCE) among both TCM and Western medicine clinicians from January to May 2022 in China, to capture their stated preferences regarding influenza vaccination programs. This study’s objectives were: (1) to investigate how different characteristics of influenza vaccination programs influenced clinicians’ vaccination decisions, including TCM and Western medicine clinicians; (2) to estimate the expected influenza vaccine coverage and identify the optimal influenza vaccination program among TCM and Western medicine clinicians; and (3) to identify different subgroups that share certain outward preferences using latent class analysis (LCA).

Methods

Participants

Clinicians (both TCM and Western medicine) holding a current practicing license in a hospital or community health service center, as of the date of survey administration, were included as potential participants. Data were collected through stratified cluster sampling between January 10 and May 10, 2022. In China, the coastal region (Eastern) and inland region (Central and Western) were stratified according to geographical location and socioeconomic development. Three tertiary hospitals, 3 secondary hospitals, and 18 primary hospitals were selected as survey spots in both the Eastern region and the Central and Western regions. The names and locations of these tertiary and secondary hospitals are listed in Table S1 in Multimedia Appendix 1. The primary hospitals were selected with the help of local health bureaus. They selected these hospitals through internal networks, and the name of each primary hospital was unknown to the researchers. We confirmed the location of primary hospital clinicians using their internet protocol address when they accessed the online questionnaire. The geographic distribution of all participants is presented in Figure S1 in Multimedia Appendix 1. Each hospital established a messaging group through WeChat (a free messaging and calling app) that included all the clinicians who worked in the hospital. Hospital administrative staff and local health bureaus delivered the online questionnaire link to these WeChat groups. We provided a brief statement for every participant before they answered the questionnaire to introduce the objective of our survey and to obtain informed consent. All participants were anonymized. Each WeChat account was allowed to submit only 1 online questionnaire. In China, real-name registration is required to apply for a phone number, and each phone number can be linked to only 1 WeChat account. Thus, we can reasonably assume that each individual could complete only 1 questionnaire. Participants who completed the questionnaire and correctly answered 2 quality control questions (a simple math problem and a pros-and-cons question where 1 choice is better than the other) were rewarded with a modest monetary incentive. Before the questionnaire was programmed, it was pilot tested by experts as well as 30 clinicians, who were contacted through our personal network, and adapted based on the feedback received. After sampling, we adjusted the Western medicine clinician sample using the propensity score matching method to ensure comparability between TCM clinicians and Western medicine clinicians (Figure 1).

Figure 1. Sample flow chart. TCM: traditional Chinese medicine; WM: Western medicine.



Demographic Variables

We used a 2-part questionnaire with demographic questions and DCE options. The demographic variables assessed are listed in Textbox 1.

Textbox 1. Demographic variables assessed.

- Age in years, divided into groups (≤29, 30–39, 40–49, or ≥50 years)
- Sex (male or female)
- Educational level (junior college degree or lower, bachelor's degree, postgraduate degree, or doctorate)
- Medical license type (traditional Chinese medicine, Western medicine, or traditional Chinese medicine integrated with Western medicine)
- Annual personal income (CNY <¥100,000, ¥100,000-¥200,000, or >¥200,000; US <$13,649.18, $13,649.18-$27,298.35, or >$27,298.35, respectively)
- Length of time practicing medicine, in years (≤3 , 4-9, 10-19, 20-29, or ≥30 years)
- Professional (job) title (resident, fellow, or attending)
- Hospital level (primary, secondary, or tertiary)

Discrete Choice Experiment

A DCE is a technique used widely in the health care sector that elicits people's preferences by prompting them to make specific choices [17]. In this study, participants were asked to select a series of preferable vaccination programs. Each program was defined by a set of attributes, such as cost of the vaccine and risk of mild adverse events, and the variation in each attribute was referred to as a level. By making a series of choices between a vaccination program with one set of attribute levels and another with alternative attribute levels, participants implicitly revealed the degree to which each attribute is more important to them and the value they place on each attribute level.

This DCE was designed according to health-based principles, ranging from technicalities, such as design efficiency, to participant-centered considerations, such as checking if experimental tasks were as clear as possible to participants. In our survey, individuals made 8 choices, and for each choice they needed to choose a preferred influenza vaccination program between 2 alternative influenza vaccination programs. The influenza vaccination programs were described by attributes and levels representing different characteristics of vaccination programs.

Attributes and Levels

Five attributes described the characteristics of vaccination programs in the choice tasks, summarized in Table 1. We selected attributes and levels including: variable cost of vaccine (CNY ¥0; CNY ¥50 [approximately US $7]; or CNY ¥100 [approximately US $14]), variable vaccine effectiveness against infection (20%, 50%, or 80%), variable risk of mild adverse events such as redness, swelling, or pain at the injection site (1%, 3%, or 5%), different vaccination campaign strategies (individual appointment or vaccination in a workplace setting), and different workplace attitudes (no notice or encouragement of vaccination).

Table . Attributes and levels of discrete choice experiment.

Attribute	Levels
1. Cost of vaccine	• CNY ¥0 • CNY ¥50 (approximately $7 USD) • CNY ¥100 (approximately $15 USD)
2. Vaccine effectiveness against infection	• Prevent 20% infection • Prevent 50% infection • Prevent 80% infection
3. Risk of mild adverse events (such as redness, swelling, or pain at the injection site)	• 1% risk • 3% risk • 5% risk
4. Vaccination campaign strategies	• Individual appointment • Vaccination in a workplace setting
5. Workplace attitudes	• No notice • Encouragement of vaccination

We selected attributes and levels related to influenza vaccination based on several sources of evidence: a rapid review of existing research [18-21], one-on-one interviews with 4 TCM clinicians and 4 Western medicine clinicians, and a consultation with subject matter experts. An orthogonal experimental design was used to reduce these choice tasks down to 8 choice tasks with 16 hypothetical vaccination programs. Levels were randomly assigned with equal probabilities and repeatedly recombined, with all combinations being plausible and realistic. Priors were obtained from a pilot study of 30 clinicians. Each clinician answered 8 choice tasks, reporting no understanding problem and respondent fatigue. The full choice tasks, as presented to respondents, are shown in Table S2 in Multimedia Appendix 1.

Statistical Analyses

The rule of thumb, as proposed by Johnson and Orme [22], suggests that the sample size required for the main effects of a DCE analysis depends on the number of choice tasks (*t*), the number of alternatives for each choice task (*a*), and the largest number of levels for any of the attributes (*c*), according to the following equation [23]:

$$N>1000c/(t\times a)$$

For our study, there were 8 choice tasks; each choice task had 2 alternatives, and the largest number of levels was 3 for any of the attributes. Our sample, which exceeded the minimum sample size calculated, of approximately at least 376 clinicians (or 188 of both TCM and Western medicine clinicians), was sufficient for the purposes of our study.

Sex, age groups, and hospital levels were identified as confounders in propensity score matching, and the propensity score of the Western medicine sample and TCM sample was calculated. We then used a nearest-neighbor matching algorithm to match 3 Western medicine clinicians with 1 TCM clinician based on the propensity score. Descriptive statistics were used to summarize characteristics of TCM and Western medicine clinicians. For categorical variables, frequencies were reported, and Pearson χ^2 tests were used to test for differences between the 2 groups. For continuous variables, mean and standard deviation were reported after Shapiro-Wilk tests indicated that the continuous variables were normally distributed. We used *t* tests to compare differences across the 2 groups.

To estimate the relative impact of influenza vaccine attributes in the DCE, we conducted a mixed logit regression model (MLM) to compute preference weight. An MLM is based on the assumption that random error has a normal distribution, taking heterogeneity as well as correlation between the choice task completed by each participant into account. We validated the linear continuous effects of chosen attributes, and the variables of cost, vaccine efficacy, and risk of mild adverse events were considered as continuous, while the vaccination campaign strategies and workplace attitudes were considered as categorized in the analysis. The MLM allowed for the calculation of compensatory effects across any 2 attributes. For example, a 5% decrease of risk of mild adverse events can compensate the negative impact of a 10% reduction of vaccine effectiveness. The willingness to pay (WTP) refers to the compensatory effect between cost and any other attribute. We calculated the expected vaccine coverage for the base case and the change of coverage when the level of one attribute was changed.

Subsequently, we conducted an LCA to classify individuals based on their preference characteristics; an individual was assigned to the class with the highest posterior probability. For example, the LCA presented that 31.1% (644/2072) of Western medicine clinicians preferred free vaccinations over any other attribute. Our choice in the number of subgroups was based on model fit and interpretability of results. We compared classes across demographic characteristics among the estimated classes using χ^2 tests. The preference subgroups were named based on the most preferred attribute levels in each class.

All analyses were based on 2-sided *P* values, with $P<.05$ indicating statistical significance. Stata (Version 16.0, Stata Corp LLC) was used for analysis.

Ethical Considerations

The Zhejiang Chinese Medical University Ethics Committee reviewed and approved this protocol (no. 20221021 - 1). Prior to participation, all study participants were required to sign an informed consent form, thereby confirming their voluntary engagement in the survey process. The study data were anonymized. All participants who completed the questionnaire and successfully submitted it received a reward of CNY ¥20 (US $2.7), distributed via WeChat.

Results

Demographics

A total of 3644 clinicians from 6 tertiary hospitals, 6 secondary hospitals, and 36 primary hospitals were invited to participate in this survey. Among them, 559 (15.3%) were ultimately excluded for failing the basic math tests (n=124, 3.4%) or the pros-and-cons test (n=341, 9.4%) or in the post-hoc during propensity score matching (n=94, 2.6%). Finally, 3085 subjects, consisting of 1013 (32.8%) TCM clinicians and 2072 (67.2%) Western medicine clinicians, were included in the analysis (Figure 1).

The sample matching adjustment achieved the expected distributions of gender, age groups, and hospital level (Figures S2-S4 in Multimedia Appendix 1). Educational attainment showcased a significant discrepancy, with a notable proportion of Western medicine clinicians holding a bachelor’s degree (1106/2072, 53.4%) compared to TCM clinicians (409/1013, 40.4%). Conversely, a higher percentage of TCM clinicians had attained a master’s degree (274/1013, 27%) or a doctorate (110/1013, 10.9%) compared to their Western medicine counterparts. In terms of annual personal income, a higher percentage of TCM clinicians reported incomes of CNY ¥100,000 - ¥200,000 (US$13,649.18-$27,298.35) (345/1013, 34.1%) compared to Western medicine clinicians (570/2072, 27.5%). Conversely, a higher proportion of Western medicine clinicians reported incomes above CNY ¥200,000 (US $27,298.35) (413/2072, 19.9%) compared to TCM clinicians (133/1013, 13.2%). In terms of professional title, 50.4% (511/1013) of TCM clinicians held the title of resident compared to Western medicine clinicians (916/2072, 44.2%). Conversely, a higher proportion of Western medicine clinicians held the title of fellow (827/2072, 39.9%) compared to TCM clinicians (346/1013, 34.2%). For the length of time practicing medicine, there was no significant difference between TCM and Western medicine clinicians (*P*=.44). (Table 2).

Table . Demographic data of clinicians practicing traditional Chinese medicine (TCM) and Western medicine (WM) between January and May 2022 in China.

	TCM clinicians (n=1013), n (%)	WM clinicians (n=2072), n (%)	*P* value
Sex			.31
Male	420 (41.5)	899 (43.4)	
Female	593 (58.5)	1173 (56.6)	
Age group (year)			.63
0-29	319 (31.5)	618 (29.8)	
30-39	431 (42.5)	893 (43.1)	
40-49	183 (18.1)	407 (19.6)	
≥50	80 (7.9)	154 (7.5)	
Hospital level			.55
Primary	295 (29.1)	637 (30.7)	
Secondary	240 (23.7)	498 (24)	
Tertiary	478 (47.2)	937 (45.2)	
Educational level			<.001
Junior college degree or lower	220 (21.7)	584 (28.2)	
Bachelor’s degree	409 (40.4)	1106 (53.4)	
Master’s degree	274 (27)	289 (13.9)	
Doctorate	110 (10.9)	93 (4.5)	
Annual personal income (CNY[a])			<.001
0-100,000	535 (52.8)	1089 (52.6)	
100,000-200,000	345 (34.1)	570 (27.5)	
>200,000	133 (13.2)	413 (19.9)	
Length of time practicing medicine (year)			.44
0-3	287 (28.3)	543 (26.2)	
4–9	331 (32.7)	677 (32.7)	
10–19	242 (23.9)	540 (26.1)	
20–29	106 (10.5)	232 (11.2)	
≥30	47 (4.6)	80 (3.9)	
Professional title			.003
Resident	511 (50.4)	916 (44.2)	
Fellow	346 (34.2)	827 (39.9)	
Attending	156 (15.4)	329 (15.9)	

[a] CNY: Chinese yuan. CNY ¥1=US $0.14.

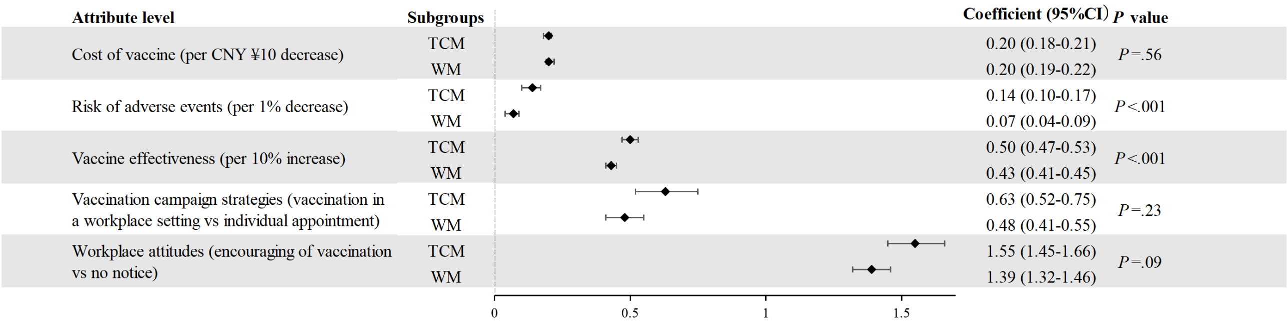
Preference Weight and Compensatory Effects (Including WTP)

All attributes and levels included were significantly associated with clinicians’ preferences for influenza vaccination programs. Both TCM and Western medicine clinicians’ preferences for influenza vaccination programs increased with higher vaccine effectiveness, with the vaccination campaign strategies changing from “individual appointment” to “vaccination in a workplace setting,” or with the workplace attitude changing from “no notice” to “encouraging of vaccination.” Nevertheless, the increase in the risk of mild adverse events and cost could undermine the utility of vaccination towards these 2 groups. The WTPs of TCM clinicians were calculated and showed that each 1% increase in risk of mild adverse events could be compensated by a CNY ¥6.97 (US $1.37) decrease of cost. Similarly, changing the mode of administration from “individual appointment” to “vaccination in a workplace setting” could compensate for the negative effect of a CNY ¥78.84 (US $15.44) increase of cost. Compensatory effects can be calculated between any attributes; for example, a hypothetical 5% increase in risk of mild adverse events could be compensated by a simultaneous hypothetical 13.7% increase in vaccine effectiveness. WTP and the compensatory effect of Western medicine clinicians were calculated in the same way. The preference weights and WTP of attributes and levels of the 2 groups are shown in Table S3 in Multimedia Appendix 1.

Comparison of Preference Weights Among TCM and Western Medicine Clinicians

Significant differences were noted for the coefficients of the risk of mild adverse events and vaccine effectiveness between clinicians practicing TCM and Western medicine. TCM clinicians (odds ratio [OR] 0.14, 95% CI 0.10 - 0.17) expressed a stronger preference for an influenza vaccination program per 1% decrease of mild adverse events risk than Western medicine clinicians (OR 0.07, 95% CI 0.04 - 0.09; *P*<.001). Similarly, TCM clinicians (OR 0.50, 95% CI 0.47 - 0.53) expressed a stronger preference for an influenza vaccination program per 10% increase of vaccine effectiveness than Western medicine clinicians (OR 0.43, 95% CI 0.41 - 0.45; *P*<.001). Other attributes (cost, vaccination campaign strategies, and workplace attitudes) did not present significant differences between the 2 groups (Figure 2).

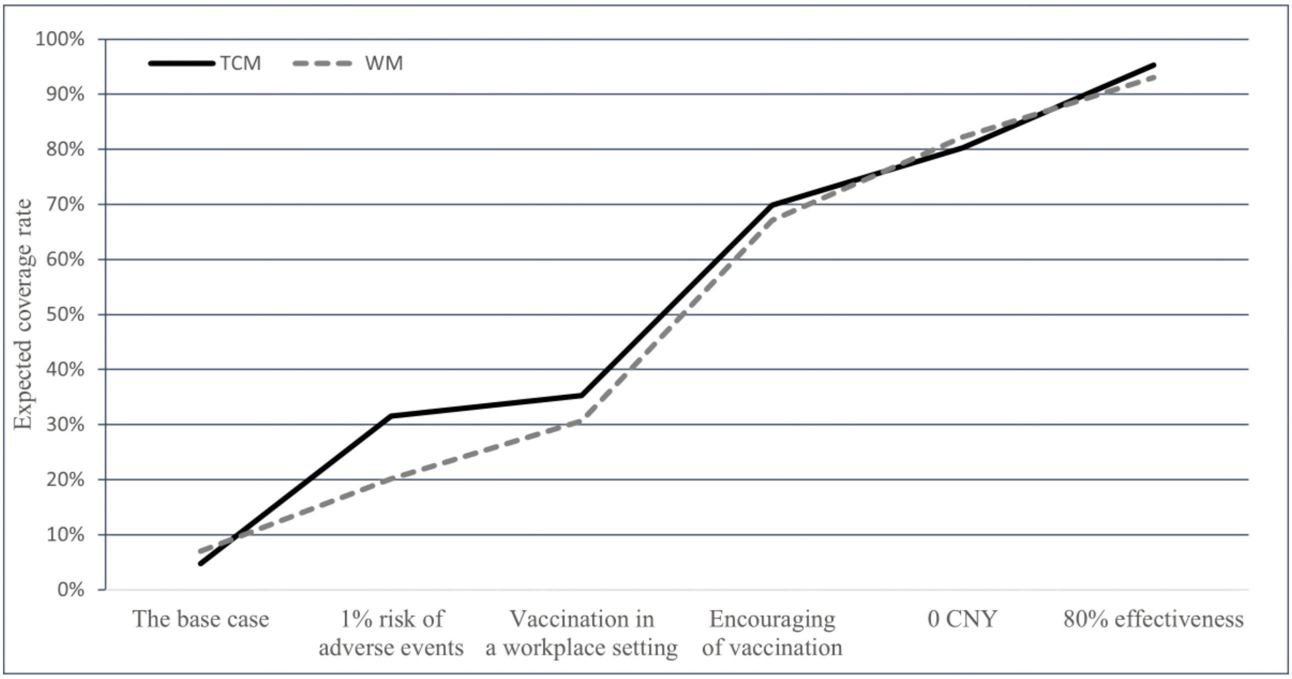
Figure 2. Comparison of preferences in clinicians practicing TCM and WM between January and May 2022 in China. TCM: traditional Chinese medicine; WM: Western medicine.



Expected Influenza Vaccine Coverage Under Various Influenza Vaccination Programs

We set the base case as having a cost of CNY ¥100 (US $13.64), 5% risk of mild adverse events, 20% effectiveness, vaccination by individual arrangement, and no notice from the workplace. The expected coverage of the base case to TCM clinicians was 4.75% and to Western medicine clinicians was 6.99%. Expected coverage increased by approximately 30% with a reduction in risk of mild adverse events from 5% to 1%; greater coverage increases could be achieved by either enhancing effectiveness from 20% to 80% or reducing costs from CNY ¥100 to CNY ¥0. An approximate hypothetical 60% increase could be obtained when the attitude of the workplace changed from "no notice" to "encouraging of vaccination," and there was an approximate hypothetical 35% increase when vaccination campaign strategies changed from "individual appointment" to "vaccination in a workplace setting." The changing of expected coverage due to improvements in the influenza vaccination program is depicted in Figure 3 (additional details available in Table S4 in Multimedia Appendix 1).

Figure 3. Changes in expected coverage rate, as attributes in influenza vaccination programs improve, between January and May 2022 in China. We set the base case as having a cost of CNY ¥100 (US $13.64), 5% risk of mild adverse events, 20% effectiveness, vaccination by individual arrangement, and no notice from the workplace. TCM: traditional Chinese medicine; WM: Western medicine; CNY: Chinese yuan.



Preference of Different Subgroups

Table 3 details the demographic composition variations across various subgroups, while the fitting and categorization process of these subgroups is recorded in Table S5 in Multimedia Appendix 1. In the entire sample, about 30% (946/3085) of clinicians preferred free vaccinations, while approximately 26% (819/3085) comprehensively considered all attributes except for vaccination campaign strategies. Notable trends were observed within 2 specific subgroups: clinicians preferring free vaccinations and those preferring vaccinations in a workplace setting. Among clinicians with different medical licenses, TCM practitioners exhibited a lower preference for free vaccinations (302/1013, 29.8%) compared to Western medicine clinicians (644/2072, 31.1%) and a higher preference for workplace vaccinations (186/1013, 18.4%) compared to Western medicine clinicians (351/2072, 17%) (*P*=.03). Differences were also noted across sex, with female clinicians indicating a lower preference for free vaccinations (533/1766, 30.2%) compared to male clinicians (414/1319, 31.3%) and a higher preference for workplace vaccinations (317/1766, 18%) compared to male clinicians (219/1319, 16.6%) (*P*=.04). Educational level revealed a U-shaped relationship, where the highest preference for free vaccinations was among those with junior college degrees or lower (270/804, 33.6%); however, this preference decreased significantly among those with postgraduate degrees (158/563, 28%; *P*<.001). Conversely, there was an inverted U-shaped relationship between educational level and preference for vaccinations in a workplace setting. A strong monotonic relationship was present between hospital level and clinicians' preference for free vaccinations, with 32.5% (303/932) of primary hospital workers preferring free vaccinations compared to 29.4% (416/1415) of those in tertiary hospitals (*P*<.001). On the other hand, clinicians in tertiary hospitals (262/1415, 18.5%) significantly preferred vaccinations in a workplace setting compared to those in primary (145/932, 15.6%) or secondary hospitals (131/738, 17.7%) (*P*<.001). No significant differences were found based on age group, annual personal income, length of time practicing medicine, or professional title.

Table . Demographic characteristics by influenza vaccination preference from a latent class analysis between January and May 2022 in China.

Demographics and categories	Prefer encouragement of vaccinations in a workplace setting, n (%)	Prefer free vaccinations, n (%)	Prefer higher vaccine effectiveness, n (%)	Consider all attributes except vaccination campaign strategies, n (%)	Prefer vaccinations in a workplace setting, n (%)	*P* value
Medical practicing license						.03
TCM[a] (n=1013)	144 (14.2)	302 (29.8)	113.46 (11.2)	267 (26.4)	186 (18.4)	
WM[b] (n=2072)	300 (14.5)	644 (31.1)	223.78 (10.8)	552 (26.7)	351 (17)	
Age group (year)						.54
0-29 (n=937)	137 (14.6)	292 (31.2)	102 (11)	242 (25.8)	164 (17.5)	
30-39 (n=1324)	188 (14.2)	408 (30.8)	144.32 (10.9)	357 (27)	226 (17.1)	
40-49 (n=590)	87 (14.7)	173 (29.3)	64.31 (10.9)	158 (26.7)	109 (18.4)	
≥50 (n=234)	34 (14.5)	73 (31.2)	24.57 (10.5)	63 (27)	39 (16.8)	
Sex						.04
Male (n=1319)	193 (14.6)	414 (31.3)	142.45 (10.8)	351 (26.6)	219 (16.6)	
Female (n=1766)	253 (14.3)	533 (30.2)	193 (11)	470 (26.6)	317 (18)	
Educational level						<.001
Junior college degree or lower (n=804)	118 (14.7)	270 (33.6)	85.22 (10.6)	211 (26.3)	119 (14.8)	
Bachelor’s degree (n=1515)	218 (14.4)	459 (30.3)	166 (11)	402 (26.6)	270 (17.8)	
Postgraduate degree (n=563)	80 (14.2)	158 (28)	63.62 (11.3)	151 (26.9)	110 (19.6)	
Doctorate (n=203)	28 (14)	59 (29.3)	21.92 (10.8)	54 (26.7)	39 (19.2)	
Annual personal income (CNY[c])						.63
0-100,000 (n=1624)	235 (14.5)	505 (31.1)	175.39 (10.8)	431 (26.5)	277 (17)	
100,000-200,000 (n=915)	131 (14.3)	278 (30.4)	99.74 (10.9)	243 (26.6)	163 (17.8)	
>200,000 (n=546)	78 (14.2)	163 (29.9)	60.61 (11.1)	145 (26.6)	99 (18.2)	
Length of time practicing medicine, in years						.79
0-3 (n=830)	120 (14.5)	262 (31.6)	91.3 (11)	221 (26.6)	135 (16.4)	
4–9 (n=1008)	147 (14.6)	305 (30.3)	107.86 (10.7)	265 (26.3)	182 (18.1)	
10–19 (n=782)	111 (14.2)	236 (30.2)	86 (11.1)	210 (26.8)	139. (17.8)	
20–29 (n=338)	49 (14.4)	103 (30.5)	37.52 (11.1)	91 (27)	58 (17.1)	
≥30 (n=127)	18 (14.3)	39 (30.7)	13.21 (10.4)	34 (26.9)	22 (17.7)	
Professional (job) title						.28
Residents (n=1427)	207 (14.5)	448 (31.4)	154.12 (10.8)	378 (26.5)	240 (16.8)	
Fellows (n=1173)	167 (14.2)	354 (30.2)	129.03 (11)	314 (26.8)	209 (17.8)	

Demographics and categories	Prefer encouragement of vaccinations in a workplace setting, n (%)	Prefer free vaccinations, n (%)	Prefer higher vaccine effectiveness, n (%)	Consider all attributes except vaccination campaign strategies, n (%)	Prefer vaccinations in a workplace setting, n (%)	*P* value
Attending (n=485)	71 (14.6)	144 (29.7)	53.84 (11.1)	128 (26.3)	89 (18.3)	
Hospital level						<.001
Primary (n=932)	133 (14.3)	303 (32.5)	99.72 (10.7)	251 (26.9)	145 (15.6)	
Secondary (n=738)	107 (14.5)	227 (30.7)	80.44 (10.9)	193 (26.2)	131 (17.7)	
Tertiary (n=1415)	204 (14.4)	416 (29.4)	157.07 (11.1)	376 (26.6)	262 (18.5)	

[a] TCM: traditional Chinese medicine.

[b] WM: Western medicine.

[c] CNY: Chinese yuan. CNY ¥1=US $0.14.

Discussion

Principal Findings

We found that (1) TCM clinicians were more sensitive to changes to the risk of mild adverse events and vaccine effectiveness than Western medicine clinicians; (2) decreasing the price of influenza vaccinations, offering convenient vaccinations in workplace settings, and issuing notifications encouraging clinicians to get vaccinated could increase influenza vaccination coverage of both groups; and (3) compared to Western medicine clinicians, a larger proportion of TCM clinicians preferred vaccinations in workplace settings over any other attribute.

The findings indicate a higher probability of vaccines being received when paired with increased vaccine effectiveness and decreased likelihood of vaccine-related adverse events, which aligns with prior research conducted through questionnaire surveys [24]. In our survey, TCM clinicians placed significantly higher preference weights on the risk of mild adverse events and vaccine effectiveness, indicating that fluctuations in the safety and effectiveness of influenza vaccines could have a greater impact on them. A qualitative study shows that patients find vaccination information transmitted by complementary and alternative medicine providers to be more understandable, useful, and trustworthy [25]. A TCM clinician who receives and shares information regarding defective vaccine incidents could cause a chain reaction in the broad population. A systematic review indicated that the provision of free vaccinations, easy access to vaccinations, and modification through educational activities and reminders can effectively increase the influenza vaccine uptake [26]. In the simulation scenario, decreasing the price of the influenza vaccination from CNY ¥100 (US $13.64) to being free resulted in a significant increase in projected coverage. Offering free vaccinations to clinicians can provide additional benefits, such as mitigating absenteeism during the influenza season and preventing hospital-acquired influenza infections. We also found that other workplace-based interventions, such as providing vaccination opportunities at the workplace and promoting vaccination initiatives more broadly, demonstrated a favorable effect on the projected uptake of influenza vaccine. Given the feasibility and efficacy of these measures, we propose that health care facilities should (1) provide free annual influenza vaccinations, (2) furnish accessible vaccination services for employees, and (3) issue notifications encouraging clinicians to avail themselves of vaccination opportunities. In our LCA, the majority of clinicians prioritized free vaccinations, followed by a substantial group who took into account all attributes except vaccination campaign strategies when deciding whether to accept an influenza vaccination program. The results of our survey suggest that a higher proportion of TCM clinicians favored workplace-based vaccinations over any other attribute, whereas a greater number of Western medicine clinicians preferred free vaccinations over any other attribute. This finding can lend support to the creation of a vaccination policy that is more flexible and adaptable. Sex, educational level, and hospital level were all associated with latent classification. Among these considerations, the most significant distinction emerged between clinicians who preferred free vaccinations most and those who preferred vaccinations in a workplace setting most. This differentiation suggests that the other 3 categories (favoring workplace encouragement of vaccinations, prioritizing higher vaccine effectiveness, and considering attributes except vaccination campaign strategies) were evenly distributed across all demographics in our study. As a result, measures like releasing notices in health care workplaces to encourage clinicians to get vaccinated may achieve similar effects across all groups of Chinese clinicians. A previous study showed that women's perceptions of inflation (consumer prices increasing) were higher than men's, but fewer female clinicians preferred free vaccinations over vaccinations in a workplace setting, compared to male clinicians in our study [27]. A policy offering free vaccinations may hold greater appeal for clinicians with lower levels of education (bachelor's degree or below) or those employed in primary health care institutions. Conversely, providing vaccinations in a workplace setting was found to be more favorable among clinicians with higher levels of education (postgraduate degree or doctorate) or those working in tertiary hospitals. In China, clinicians employed in tertiary hospitals serve as the cornerstone of the nation's health care provision, handling extensive medical responsibilities that may make it challenging for them to get vaccinated during their private leisure time. Consequently, providing convenient vaccination services for clinicians should be prioritized as a primary measure for these tertiary hospitals.

Limitations

Our study is subject to several limitations. First, our sample of clinicians did not perfectly mirror the sociodemographic characteristics of the entire nation. Additionally, there may be biases between individuals who completed the survey and those who did not. Second, as a cross-sectional study, potential recall bias is inevitable. Moreover, DCEs are susceptible to hypothetical bias, meaning that responses in surveys may not perfectly align with real-life behaviors [28]. This discrepancy could limit the accuracy of measured preferences. To mitigate this, we designed the experiment based on real-world influenza vaccination programs and ensured that the details and definitions of each attribute were comprehensible to participants.

Conclusions

In summary, while changes in mild adverse event rates and vaccine effectiveness had a greater impact on TCM clinicians, other attributes, such as workplace-provided free vaccinations, issuing notifications to encourage vaccinations, and providing accessible vaccination services, are practical methods to increase the uptake of influenza vaccination programs among both groups. Policymakers can utilize these insights to formulate flexible intervention measures that facilitate greater accessibility of influenza vaccination programs for clinicians, especially TCM clinicians, worldwide.

Acknowledgments

We thank Gang Cai, Limin Chen, Yu Du, Jinglin Han, Conghua Ji, Xia Jin, Fangyou Li, Xikun Li, Peizhen Li, Chunxiao Liu, Jianping Li, Miner Qian, Baoshan Sun, Yongshou Yang, Yuan Lv, Tianxing Zhang for their help with data collection. This work was supported by the National Social Science Fund of China (22BGL316).

Data Availability

The deidentified cross-sectional data and the discrete choice models from the analysis are available from the corresponding author on reasonable request.

Authors' Contributions

LZ and CF conceptualized the study, contributed to the overall design of the survey experiment, analysis, and interpretation of the data, and are the guarantors for the study. LZ cleaned the data, performed the analyses, drew the figures, conducted the literature searches, and wrote the first draft of the manuscript. LZ, LC, YZ, XS, ZW, MES, and CF discussed the results and interpretation of the data. MES and CF reviewed and approved the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables and figures.
[DOCX File, 176 KB - publichealth_v11i1e63314_app1.docx]

References

1. WHO launches new global influenza strategy. World Health Organization. 2019. URL: https://www.who.int/news/item/11-03-2019-who-launches-new-global-influenza-strategy [accessed 2025-01-08]
2. Li L, Liu Y, Wu P, et al. Influenza-associated excess respiratory mortality in China, 2010-15: a population-based study. Lancet Public Health 2019 Sep;4(9):e473-e481. [doi: 10.1016/S2468-2667(19)30163-X] [Medline: 31493844]
3. Gianino MM, Politano G, Scarmozzino A, et al. Cost of sickness absenteeism during seasonal influenza outbreaks of medium intensity among health care workers. Int J Environ Res Public Health 2019 Mar 1;16(5):747. [doi: 10.3390/ijerph16050747] [Medline: 30832264]
4. Chiu S, Black CL, Yue X, et al. Working with influenza-like illness: presenteeism among US health care personnel during the 2014-2015 influenza season. Am J Infect Control 2017 Nov 1;45(11):1254-1258. [doi: 10.1016/j.ajic.2017.04.008] [Medline: 28526310]
5. Blendon RJ, Benson JM, Hero JO. Public trust in physicians--U.S. medicine in international perspective. N Engl J Med 2014 Oct 23;371(17):1570-1572. [doi: 10.1056/NEJMp1407373] [Medline: 25337746]
6. Influenza (seasonal). World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal) [accessed 2025-01-08]
7. Yi H, Yang Y, Zhang L, et al. Improved influenza vaccination coverage among health-care workers: evidence from a web-based survey in China, 2019/2020 season. Hum Vaccin Immunother 2021 Jul 3;17(7):2185-2189. [doi: 10.1080/21645515.2020.1859317] [Medline: 33497309]

8. Hammer CC, Lyytikäinen O, Arifulla D, Toura S, Nohynek H. High influenza vaccination coverage among healthcare workers in acute care hospitals in Finland, seasons 2017/18, 2018/19 and 2019/20. Euro Surveill 2022 Apr;27(17):2100411. [doi: 10.2807/1560-7917.ES.2022.27.17.2100411] [Medline: 35485269]
9. Razzaghi H, Srivastav A, de Perio MA, Laney AS, Black CL. Influenza and COVID-19 vaccination coverage among health care personnel - United States, 2021-22. MMWR Morb Mortal Wkly Rep 2022 Oct 21;71(42):1319-1326. [doi: 10.15585/mmwr.mm7142a2] [Medline: 36264832]
10. Saunier F, Berthelot P, Mottet-Auselo B, et al. Impact of a decision-aid tool on influenza vaccine coverage among HCW in two French hospitals: a cluster-randomized trial. Vaccine 2020 Aug 10;38(36):5759-5763. [doi: 10.1016/j.vaccine.2020.07.011] [Medline: 32684500]
11. Loudon I, editor. Western Medicine: An Illustrated History: Oxford University Press; 1997.
12. Traditional medicine. World Health Organization. 2023. URL: https://www.who.int/news-room/questions-and-answers/item/traditional-medicine [accessed 2025-01-08]
13. Tang JL, Liu BY, Ma KW. Traditional Chinese medicine. Lancet 2008 Dec 6;372(9654):1938-1940. [doi: 10.1016/S0140-6736(08)61354-9] [Medline: 18930523]
14. Cao X. Immunology in China: the past, present and future. Nat Immunol 2008 Apr;9(4):339-342. [doi: 10.1038/ni0408-339] [Medline: 18349809]
15. WHO international standard terminologies on traditional Chinese medicine. World Health Organization. URL: https://www.who.int/publications/b/55892 [accessed 2025-01-08]
16. China's traditional Chinese medicine service capacity has been significantly improved, with the total number of TCM visits reaching 1.2 billion [Article in Chinese]. WeChat. 2023. URL: https://mp.weixin.qq.com/s/eqgzswqtGiQC7NSLgsDgLA [accessed 2025-01-08]
17. Hensher DA, Rose JM, Greene WH. Applied Choice Analysis: Cambridge University Press; 2015.
18. Walsh JM, McPhee SJ. A systems model of clinical preventive care: an analysis of factors influencing patient and physician. Health Educ Q 1992;19(2):157-175. [doi: 10.1177/109019819201900202] [Medline: 1618625]
19. Vasilevska M, Ku J, Fisman DN. Factors associated with healthcare worker acceptance of vaccination: a systematic review and meta-analysis. Infect Control Hosp Epidemiol 2014 Jun;35(6):699-708. [doi: 10.1086/676427] [Medline: 24799647]
20. Wicker S, Marckmann G, Poland GA, Rabenau HF. Healthcare workers' perceptions of mandatory vaccination: results of an anonymous survey in a German University Hospital. Infect Control Hosp Epidemiol 2010 Oct;31(10):1066-1069. [doi: 10.1086/656242] [Medline: 20804415]
21. Pielak KL, McIntyre CC, Tu AW, Remple VP, Halperin B, Buxton JA. Identifying attitudes, beliefs and reported practices of nurses and doctors as immunization providers. J Adv Nurs 2010 Jul;66(7):1602-1611. [doi: 10.1111/j.1365-2648.2010.05326.x] [Medline: 20492025]
22. Johnson R, Orme B. Getting the most from CBC. : Sawtooth Software; 2003 URL: https://sawtoothsoftware.com/resources/technical-papers/getting-the-most-from-cbc [accessed 2025-01-13]
23. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. Patient 2015 Oct;8(5):373-384. [doi: 10.1007/s40271-015-0118-z] [Medline: 25726010]
24. Alshammari TM, Yusuff KB, Aziz MM, Subaie GM. Healthcare professionals' knowledge, attitude and acceptance of influenza vaccination in Saudi Arabia: a multicenter cross-sectional study. BMC Health Serv Res 2019 Apr 15;19(1):229. [doi: 10.1186/s12913-019-4054-9] [Medline: 30992004]
25. Deml MJ, Notter J, Kliem P, et al. "We treat humans, not herds!": a qualitative study of complementary and alternative medicine (CAM) providers' individualized approaches to vaccination in Switzerland. Soc Sci Med 2019 Nov;240:112556. [doi: 10.1016/j.socscimed.2019.112556] [Medline: 31563005]
26. Hollmeyer H, Hayden F, Mounts A, Buchholz U. Review: interventions to increase influenza vaccination among healthcare workers in hospitals. Influenza Other Respir Viruses 2013 Jul;7(4):604-621. [doi: 10.1111/irv.12002] [Medline: 22984794]
27. Plane M, Vermersch G. Inflation and counter-inflationary policy measures: the case of France. Hans-Böckler-Stiftung. 2022. URL: https://nbn-resolving.de/urn:nbn:de:101:1-2023011910431794423230%0A [accessed 2025-01-08]
28. Catalogue of Bias Collaboration, Buckell J, Buchanan J, et al. Hypothetical bias. Catalogue of Bias. URL: https://catalogofbias.org/biases/hypothetical-bias/ [accessed 2025-01-08]

Abbreviations

DCE: discrete choice experiment
HCW: health care worker
MLM: mixed logit regression model
OR: odds ratio
TCM: traditional Chinese medicine
WTP: willingness to pay

Edited by A Mavragani, T Sanchez; submitted 17.06.24; peer-reviewed by N Jiang, S Jimenez-Jorge; revised version received 15.12.24; accepted 19.12.24; published 20.01.25.

Please cite as:
Zhang L, Chu L, Sundaram ME, Zhou Y, Sun X, Wei Z, Fu C
Identifying Preferred Features of Influenza Vaccination Programs Among Chinese Clinicians Practicing Traditional Chinese Medicine and Western Medicine: Discrete Choice Experiment
JMIR Public Health Surveill 2025;11:e63314
URL: https://publichealth.jmir.org/2025/1/e63314
doi:10.2196/63314

Original Paper

Attitudes Toward Psychotherapeutic Treatment and Health Literacy in a Large Sample of the General Population in Germany: Cross-Sectional Study

Rebekka Schröder[1], Dr phil; Tim Hamer[1], MSc; Ralf Suhr[1], PD, Dr med; Lars König[1], Dr rer nat

Stiftung Gesundheitswissen, Berlin, Germany

Corresponding Author:
Lars König, Dr rer nat
Stiftung Gesundheitswissen
Friedrichstr. 134
Berlin, 10117
Germany
Phone: 49 30 4195492 62
Email: lars.koenig@stiftung-gesundheitswissen.de

Abstract

Background: Prevalences of mental disorders are increasing worldwide. However, many people with mental health problems do not receive adequate treatment. An important factor preventing individuals from seeking professional help is negative attitudes toward psychotherapeutic treatment. Although a positive shift in attitudes has been observed in recent years, there is still substantial stigma surrounding psychotherapeutic treatment. First studies have linked higher health literacy with more positive attitudes toward psychotherapy, but more research is needed in this area.

Objective: This study aimed to examine how general and mental health literacy are associated with attitudes toward psychotherapeutic treatment in Germany. Additionally, associations between sociodemographic factors, experience with psychotherapy, and attitudes toward psychotherapy were explored.

Methods: A random sample was drawn from a panel representative of the German-speaking population with internet access in Germany and invited to participate in the study via email. Overall, 2000 individuals aged ≥16 years completed the web-based survey with standardized questionnaires in September and October 2022. Attitudes toward psychotherapy and both general and mental health literacy were assessed using the Questionnaire on Attitudes Towards Psychotherapeutic Treatment (QAPT) with 2 subscales (“positive attitudes” and “non-acceptance of society”), the European Health Literacy Survey instrument (HLS-EU-Q16) and the Mental Health Literacy Tool for the Workplace (MHL-W-G). Associations between the questionnaire scales were assessed with Pearson correlations. Additionally, basic sociodemographic information and information on personal and family experiences with psychotherapy were collected. Pearson correlations (age), ANOVAs (level of education and subjective social status), and *t* tests (experience with psychotherapy, gender, and migration background) were used to analyze how these relate to attitudes toward psychotherapy.

Results: More favorable attitudes toward psychotherapy and lower perceived societal nonacceptance were found in those with higher general (r=0.14, P<.001; r=−0.32, P<.001, respectively) and mental health literacy (r=0.18, P<.001; r=−0.23, P<.001, respectively). Participants with treatment experience for mental health problems ($t_{1260.12}$=−10.40, P<.001, Cohen d=−0.49; $t_{1050.95}$=3.06, P=.002, Cohen d=0.16) and who have relatives with treatment experience ($t_{1912.06}$=−5.66, P<.001, Cohen d=−0.26; t_{1926}=4.77, P<.001, Cohen d=0.22) reported more positive attitudes and higher perceived societal acceptance than those without treatment experience. In terms of sociodemographic differences, being a woman (t_{1992}=−3.60, P<.001, Cohen d=−0.16), younger age (r=−0.11, P<.001), higher subjective social status ($F_{2,1991}$=5.25, P=.005, η^2=.005), and higher levels of education ($F_{2,1983}$=22.27, P<.001, η^2=.021) were associated with more positive attitudes toward psychotherapeutic treatment. Being a man (t_{1994}=5.29, P<.001, Cohen d=0.24), younger age (r=−0.08, P<.001), and lower subjective social status ($F_{2,1993}$=7.71, P<.001, η^2=.008) were associated with higher perceived nonacceptance of psychotherapy.

Conclusions: Positive associations between attitudes toward psychotherapy and both general and mental health literacy were delineated. Future studies should investigate whether targeted health literacy interventions directed at individuals with lower general and mental health literacy might also help to improve attitudes toward psychotherapeutic treatment and help-seeking behavior.

(JMIR Public Health Surveill 2025;11:e67078) doi:10.2196/67078

KEYWORDS

health literacy; mental health literacy; attitude to health; cross-sectional studies; Germany; adults; representative; psychotherapy

Introduction

In Germany, almost 30% of the population are estimated to have a mental disorder [1] but only about 10% have received treatment in the past year [2]. This means that about two-thirds of those with mental health problems do not receive adequate treatment. While the number of individuals not seeking help for mental disorders has decreased in recent years, the majority of individuals with mental disorders still do not seek treatment [3]. With a global rise in the prevalence of mental disorders [4], it becomes even more important to understand the factors that influence whether an individual seeks help or not.

Psychotherapy is considered effective for treating many mental disorders [5]. However, there are several reasons that prevent individuals from seeking psychotherapy. These encompass structural challenges such as long waiting periods and underprovision, especially in rural areas [2,6,7], as well as individual reasons often related to stigma against mental disorders and psychotherapy [8,9]. Specifically, in self-reports, about one-third of respondents report negative attitudes and stigma toward psychotherapy, with more negative attitudes in male individuals and in individuals with lower education [10]. Similarly, almost a third of the respondents believed that professional care was worse than or equal to no help when faced with serious emotional problems in a large European study [8]. In recent years, attitudes toward professional help, particularly toward psychotherapeutic and psychiatric treatment, have improved both on a global level and in Germany [11,12]. In addition, perceived public stigma toward people with mental disorders has substantially decreased in the past 30 years [13]. At the same time, the individual wish for social distance toward people with depression slightly decreased but increased toward people with schizophrenia [11,14]. Furthermore, in many non-Western countries, there are still substantial negative attitudes toward individuals with mental illnesses [15,16]. Critically, it has been shown that negative attitudes toward help seeking and stigmatizing attitudes toward people with mental illness predict help-seeking behavior [7,17-20]. In summary, this means that although some progress has been made in recent years, there is still considerable stigma associated with both mental illness and its treatment, which needs to be addressed to increase the number of individuals who receive treatment in addition to reducing structural barriers.

Health literacy refers to “people’s knowledge, motivation and competencies to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life concerning health care, disease prevention and health promotion to maintain or improve quality of life during the life course” [21]. Low health literacy has been associated with several negative health-related outcomes, including higher rates of hospitalization, lower medical treatment adherence, decreased quality of life, fewer health-promoting behaviors, and higher mortality [22-27]. Mental health literacy was developed from general health literacy and applies the concept to mental health. It is defined as “understanding how to obtain and maintain positive mental health; understanding mental disorders and their treatments; decreasing stigma related to mental disorders; and, enhancing help-seeking efficacy” [28]. High mental health literacy has a number of positive correlates such as mental well-being, physical activity, life satisfaction, healthy eating habits, and self-efficacy [29-31].

First studies have linked increased health literacy with lower stigma and more positive attitudes toward mental disorders and psychotherapy [12,19,30,32-34]. For example, Svensson and Hansson [32] have shown that a higher degree of mental health literacy is related to less stigma and a lower wish for social distance toward individuals with depression and to a lesser extent also toward individuals with schizophrenia. However, this pattern of results was not found in all studies. In a Swiss community sample of young adults, no direct association between mental health literacy and the wish for social distance was delineated. Instead, it was shown that increased mental health literacy may also lead to more stigma by enhancing stereotypes of individuals with mental disorders (eg, perceived dangerousness) due to higher endorsement of biogenetic causal explanation of mental disorders [35]. It has been demonstrated in diverse cultural contexts that specific interventions may enhance both general and mental health literacy [36-41]. Moreover, it has been shown in longitudinal studies that more positive treatment attitudes can predict more frequent mental health service use in the future [19,20]. Along with significantly reducing structural barriers, indirectly improving attitudes toward psychotherapeutic treatment via increased health literacy may therefore be an important starting point to ensure that more people with mental health problems receive professional help.

Although some studies have started to investigate the relationship between general and mental health literacy and attitudes toward psychotherapeutic treatments, there are still substantial gaps in the literature as well as some limitations of previous studies that need to be addressed in order to get a better understanding of how to increase the number of individuals seeking psychotherapeutic treatment. First, in past investigations, mental health literacy was often assessed via unlabeled vignettes [32,35] describing the symptoms of a specific disorder (eg, depression or schizophrenia), and high mental health literacy was assigned to those individuals who correctly identified these vignettes. While this covers a small aspect of the broad definition of mental health literacy, that is, the recognition of mental disorders, other important aspects, such as promoting positive mental health, are not included [28]. This limitation in previous methodological approaches is why we decided to use an instrument with a broader scope to assess mental health literacy in the present investigation [42,43]. Second, the literature on general health literacy and how it relates to treatment attitudes is relatively sparse and there is a need to replicate and extend previous findings [34]. Third, there is some evidence to suggest that there are substantial interindividual differences in attitudes toward psychotherapy in the community [19,34]. However, the specific determinants of these differences need to be explored more in depth to identify potentially vulnerable segments of the population that could particularly benefit from interventions. For example, it has been shown that gender, age, level of education, and social status play an important role in accessing psychiatric and psychotherapeutic services and might therefore also be related to attitudes [2,10]. In addition, having a migration background or being part of an ethnic minority has been associated with access to mental health care [44-48]. Finally, negative attitudes might be improved by hearing about the positive experiences of others or making these experiences themselves [19,49-51], which is why we decided to include 2 questions on personal and family experience with psychotherapy.

To address these previous limitations and gaps in the literature, in this study, we aimed to explore the relationship between attitudes toward psychotherapy and both general and mental health literacy with comprehensive questionnaires and in a large sample. Furthermore, we aimed at exploring how additional factors including basic sociodemographic factors and experience with psychotherapy relate to attitudes toward psychotherapy in order to inform the future development of target group-specific interventions with the aim of reducing individual barriers to seeking professional help.

Methods

Ethical Considerations

The ethics committee of the Berlin Medical Association had no ethical or professional objections to the study protocol (reference Eth-39/22). Participants gave informed consent to take part in the study. They were not directly compensated for their participation by the independent nonprofit foundation Stiftung Gesundheitswissen. Data collection was conducted by the market research institute forsa (forsa Gesellschaft für Sozialforschung und statistische Analysen mBH), which was commissioned and financially compensated for this work by the nonprofit foundation Stiftung Gesundheitswissen. Importantly, Stiftung Gesundheitswissen did not have any influence on the data collected and the direction of the results. Forsa provided only anonymized data to the Stiftung Gesundheitswissen.

Data Acquisition

Data collection was conducted by forsa using the forsa.omninet panel. Forsa.omninet is a representative panel for the German-speaking population with internet access aged 14 years and older in Germany and currently has around 100,000 participants. Panelists are recruited exclusively offline, that is, by telephone within the framework of forsa.omniTel, a telephone omnibus survey of forsa, in which randomly selected individuals aged 14 years and older are interviewed daily. The composition of the panel is continuously monitored on the basis of key characteristics (eg, region, age, and gender) and recruitment is adjusted accordingly. In addition, approximately 1000 new individuals are recruited monthly. There is no possibility to apply for participation in the panel so the selection of the participants is solely controlled by a random sampling procedure. Inclusion criteria for this investigation were the age of 16 years or older; sufficient German knowledge; and being able to use a computer, laptop, or mobile device for the survey. The only exclusion criterion was not fulfilling the inclusion criteria. With the aim of collecting data from at least 2000 individuals, 3927 panelists aged 16 years and older were randomly drawn from the panel and invited to take part in the survey by email. Exactly 2000 participants completed the study, equaling a response rate of 50.9% (2000/3927). The invitation email contained brief information on the overall topic of the survey as well as the incentive. If necessary, the selected panelists were reminded about their participation in 2 further emails. Recruitment and data collection took place in September and October 2022 with a standardized questionnaire presented in the form of computer-assisted web interviews. The total questionnaire consisted of 155 items, of which not all are relevant to this investigation. For economic reasons, data collection for different research objectives was combined. Results are published in separate and independent publications [31,52].

To minimize the effects of nonresponse and potential bias due to systematically missing data from certain population segments, survey weights were calculated by forsa using an iterative proportional fitting approach with the following weight variables and combinations: (1) gender × age (16-29 years, 30-45 years, 46-64 years, or ≥65 years) × region (West Germany and Berlin, or East Germany) and (2) federal state. This procedure resulted in a single weighting factor for each individual applied before further statistical analyses. The weighting was informed by the population update of the German Federal Statistical Office (as of December 31, 2020).

Measures

Attitudes Toward Psychotherapy

Attitudes toward psychotherapy were assessed with the German version of the Questionnaire on Attitudes Towards Psychotherapeutic Treatment (QAPT) in its revised 2-factorial form [44,53,54]. The questionnaire consists of 11 items answered on a 4-point Likert scale ranging from "do not agree" to "agree." The items address how the respondents perceive psychotherapeutic treatments. The 2 subscales are "positive attitudes" (6 items) and "non-acceptance of society" (5 items). The positive attitude scale reflects the individuals' beliefs about the positive effects of psychotherapeutic treatment and the competence of psychotherapists, whereas the nonacceptance scale reflects the fear of stigma and social isolation from others. The second scale is originally labeled "acceptance of society" ("Akzeptanz der Gesellschaft" in German [54]). However, as the items of this scale are negatively coded, that is, higher values indicate lower acceptance, we decided to label the scale "nonacceptance of society" for easier interpretability. For each scale, a mean value is calculated across all items. The mean value reaches from 1 to 4 with higher values indicating more positive attitudes and higher nonacceptance (eg, lower acceptance), respectively.

Health Literacy

General health literacy was assessed with the German translation of the short version of the European Health Literacy Survey instrument (HLS-EU-Q16 [55-57]). This 16-item questionnaire addresses the subjective difficulty in accessing, understanding, appraising, and applying information in the fields of health care, disease prevention, and health promotion. Respondents are asked to answer each item on a 4-point Likert scale with the options “very easy,” “fairly easy,” “fairly difficult,” and “very difficult.” An overall sum score is calculated for the dichotomized items (eg, 1=“fairly easy” and “very easy,” 0=“fairly difficult” and “very difficult”). The sum score ranges from 0 to 16, with higher values indicating higher health literacy. Groups of individuals with varying degrees of health literacy were determined using this sum score: inadequate health literacy (scores 0-8), problematic health literacy (scores 9-12), and adequate health literacy (scores 13-16).

Mental Health Literacy

Mental health literacy was assessed with the German version of the Mental Health Literacy Tool for the Workplace (MHL-W-G [42,43]). Instructions were slightly adapted so that they also apply to people who are not currently working. Specifically, participants were instructed to imagine they worked if this was not the case. The instrument presents 4 case vignettes, in each of which current symptoms and the circumstances of one individual are described. The vignettes differ in the nature of the mental health issue, its impact on workplace performance, and the gender of the individual. Importantly, the nature of the mental health issue is not explicitly labeled. For each vignette, the same 4 questions are answered covering important aspects of mental health literacy, namely the ability to recognize specific mental disorders, knowledge and beliefs about risk factors and prevention, knowledge and attitudes to facilitate help seeking, and knowledge and beliefs about mental health interventions. After reading each of the vignettes, participants are asked to rate their knowledge about what might be happening, how they could prevent the situation from worsening, what to say or do in the situation, and which resources and services might be helpful. Answers are given on a 5-point Likert scale ranging from “very low” to “very high”. A sum score is calculated by adding up the ratings of each of the 4 questions in each of the 4 vignettes (ranging from a minimum of 16 to a maximum of 80 points), with higher scores reflecting higher levels of self-reported mental health literacy. So far, no generally accepted cut-off scores have been developed to compare groups of individuals with varying levels of mental health literacy using this questionnaire, which is why we do not present stratified data here.

Experience with Psychotherapy

In addition, the participants were presented with 2 simple yes-or-no-questions to assess whether they have ever received any treatments for mental disorders themselves and whether any of their close family members have ever received treatment for mental disorders.

Sociodemographic Information

Participants provided basic sociodemographic information including gender (men, women) and age.

In addition, participants were asked to provide their highest level of formal education. Participants were then categorized into 3 groups: low, middle, and high level of education. The low level of education is equivalent to no formal education or basic secondary school (ohne Haupt-/Volksschulabschluss, Haupt-/Volksschulabschluss); the middle level of education is equivalent to intermediate secondary school (Mittlere Reife, Realschulabschluss, Fachschulreife, Abschluss der Polytechnischen Oberschule, Fachhochschulreife, Abschluss einer Fachoberschule); and the high level of education is equivalent to most advanced secondary school, for example, grammar schools to obtain a general or specialized university entrance qualification or a university degree (Abitur, allgemeine oder fachgebundene Hochschulreife, Fach-/Hochschulstudium).

Subjective social status was assessed with the German version of the MacArthur scale [58,59], in which the participants rank themselves relative to other members of society. They are presented with a metaphor, in which a ladder represents the social structure of society with the highest ladder rung representing the individuals in society with the highest status and the bottom rung presenting the individuals in society with the lowest status. The respondents are asked to identify the rung on which they place themselves. Three categories of subjective social status were determined according to the respondents’ answers, that is, low subjective social status (scores 1-4), middle subjective social status (scores 5-7), and high subjective social status (scores 8-10).

Migration status was assessed with a simple yes-or-no-question. The participants were asked to indicate if they have a migration background. They were instructed to answer with yes if they or one of their parents were not born in Germany.

Statistical Analysis

Statistical analyses were performed with the software SPSS (version 29.0.2.0; IBM). All analyses reported in the main manuscript were conducted with the weighted dataset (details on weighting can be found in the “data acquisition” paragraph). Results for the unweighted dataset can be found in the Multimedia Appendix 1. Cronbach α was calculated to determine the internal consistency of the attitudes and nonacceptance subscales of the QAPT as well as the HLS-EU-Q16 and MHL-W-G total scores. The resulting reliability coefficients were interpreted using the following rule of thumb: >.90 excellent, >.80 good, >.70 acceptable, >.60 questionable, >.50 poor, and <.50 unacceptable [60]. For all inferential analyses, a threshold of α=.05 was used as the significance level and 2-sided *P* values are reported. Associations between the 2 attitudes toward psychotherapy scales and both mental and general health literacy were assessed with Pearson correlations, including 95% CI. Likewise, for the associations between the 2 attitudes toward psychotherapy scales and the metric sociodemographic measure of age, Pearson correlations were calculated. For each of the dichotomous sociodemographic measures and the 2 experiences with treatment for mental disorder questions (ie, gender, migration status, personal experience, and the experience of a close family member) separate 2-tailed *t* tests were performed for the 2 attitudes toward psychotherapy scales. Effect sizes for the *t* tests are reported as Cohen *d*, with the values of Cohen *d*=0.20, Cohen *d*=0.50, and Cohen *d* = 0.80 corresponding to small, medium, and large effects, respectively [61]. If the Levene test for equality of variances indicated that variances were not homogeneous, degrees of freedom were adjusted accordingly. For each categorial sociodemographic measure with more than 2-factor levels (ie, level of education and subjective social status), separate one-way ANOVAs were conducted. These included the sociodemographic measures as predictors and the attitudes toward psychotherapy scales as the dependent variables. Effect sizes for the ANOVAs are reported as η^2, with the values η^2=.01, η^2=.06, and η^2=.14 corresponding to small, medium, and large effects, respectively [62]. Significant main effects in the ANOVAs were followed up with Bonferroni-corrected pairwise *t* tests. Uncorrected *P* values and corrected α-thresholds are reported for these post hoc *t* tests. Participants could opt out of any question, resulting in a missing value for that item. Individuals with missing data for one item were excluded from all analyses including this item, for example, a scale or mean sum score was not calculated for an individual if a specific item relevant to that scale or mean sum score was missing, but it was calculated for other scales of the same questionnaire that do not contain that item if all other items were complete.

Results

Sample Characteristics and Descriptive Statistics

A sample of 2000 individuals was recruited. The sample characteristics can be found in Table 1 (cumulative percentages may not add up to precisely 100% and sample sizes may vary due to weighting and rounding). Descriptive statistics for the QAPT, HLS-EU-Q16, and MHL-W-G questionnaires across the entire sample are in Table 2. Detailed descriptive statistics for the QAPT attitudes and nonacceptance subscales are in Tables 3 and 4, respectively, and detailed descriptive statistics for each QAPT item can be found in Multimedia Appendix 1.

Table 1. Characteristics of the weighted and unweighted sample.

Variable	Unweighted sample (n=2000), n (%)	Weighted sample (n=2000), n (%)
Gender		
Men	957 (47.9)	980 (49.0)
Women	1043 (52.1)	1020 (51.0)
Age group (years)		
16-29	461 (23.1)	355 (17.8)
30-45	433 (21.6)	475 (23.7)
46-64	599 (29.9)	655 (32.7)
>65	507 (25.4)	515 (25.8)
Level of education		
Low	349 (17.4)	349 (17.5)
Middle	847 (42.4)	861 (43.0)
High	792 (39.6)	782 (39.1)
Missing	12 (0.6)	8 (0.4)
Social status		
Low	312 (15.6)	315 (15.7)
Middle	1397 (69.8)	1389 (69.4)
High	291 (14.5)	296 (14.8)
Migration background		
No	1871 (93.5)	1876 (93.8)
Yes	129 (6.5)	124 (6.2)
Own experience with psychotherapy		
No	1376 (68.8)	1379 (68.9)
Yes	603 (30.1)	596 (29.8)
Missing	21 (1.1)	25 (1.2)
Family experience with psychotherapy		
No	1062 (53.1)	1053 (52.6)
Yes	870 (43.5)	879 (43.9)
Missing	68 (3.4)	68 (3.4)
General health literacy categories		
Inadequate	192 (9.6)	190 (9.5)
Problematic	540 (27.0)	545 (27.3)
Adequate	1255 (62.7)	1252 (62.6)
Missing	13 (0.7)	13 (0.6)

Table 2. Descriptive statistics of the QAPT[a], HLS-EU-Q16[b], and MHL-W-G[c] questionnaires.

Variable	Sample, n	Score, mean (SD)	Score, 95% CI of the mean	Score, median (range)
QAPT				
Positive attitudes toward psychotherapy	1994	3.28 (0.54)	3.26-3.31	3.33 (1-4)
Nonacceptance of society	1996	2.06 (0.61)	2.03-2.08	2.00 (1-4)
HLS-EU-Q16	1987	12.92 (3.03)	12.79-13.05	14.00 (0-16)
MHL-W-G	1994	52.03 (11.02)	51.55-52.51	52.00 (16-80)

[a] QAPT: Questionnaire on Attitudes Towards Psychotherapeutic Treatment.

[b] HLS-EU-Q16: European Health Literacy Survey instrument.

[c] MHL-W-G: Mental Health Literacy Tool for the Workplace (German version).

Table 3. Descriptive statistics for the QAPT[a] attitudes subscale across the sociodemographic groups.

Variable	Score, mean (SD)	Score, 95% CI of the mean	Score, median
Gender			
Men	3.24 (0.55)	3.20-3.27	3.33
Women	3.33 (0.52)	3.29-3.36	3.33
Level of education			
Low	3.20 (0.56)	3.14-3.25	3.17
Middle	3.23 (0.54)	3.19-3.27	3.17
High	3.38 (0.50)	3.25-3.42	3.50
Social status			
Low	3.22 (0.55)	3.16-3.28	3.17
Middle	3.28 (0.52)	3.25-3.31	3.33
High	3.35 (0.57)	3.30-3.43	3.50
Migration background			
No	3.29 (0.54)	3.25-3.31	3.33
Yes	3.25 (0.53)	3.15-3.34	3.33
Own experience with psychotherapy			
No	3.21 (0.54)	3.18-3.23	3.17
Yes	3.46 (0.48)	3.42-3.50	3.50
Family experience with psychotherapy			
No	3.22 (0.56)	3.19-3.25	3.17
Yes	3.36 (0.50)	3.32-3.39	3.33

[a] QAPT: Questionnaire on Attitudes Towards Psychotherapeutic Treatment.

Table 4. Descriptive statistics for the QAPT[a] nonacceptance subscale across the sociodemographic groups.

Variable	Score, mean (SD)	Score, 95% CI of the mean	Score, median
Gender			
Men	2.13 (0.60)	2.09-2.17	2.20
Women	1.99 (0.61)	1.95-2.02	2.00
Level of education			
Low	2.05 (0.61)	1.99-2.12	2.00
Middle	2.05 (0.60)	2.01-2.09	2.00
High	2.06 (0.61)	2.02-2.10	2.00
Social status			
Low	2.16 (0.59)	2.09-2.22	2.20
Middle	2.05 (0.60)	2.02-2.09	2.00
High	1.96 (0.63)	1.89-2.04	1.80
Migration background			
No	2.06 (0.60)	2.03-2.08	2.00
Yes	2.07 (0.64)	1.96-2.19	2.00
Own experience with psychotherapy			
No	2.08 (0.59)	2.05-2.11	2.00
Yes	1.99 (0.64)	1.94-2.04	2.00
Family experience with psychotherapy			
No	2.12 (0.60)	2.08-2.15	2.20
Yes	1.98 (0.61)	1.94-2.02	2.00

[a] QAPT: Questionnaire on Attitudes Towards Psychotherapeutic Treatment.

Reliability Analyses

The reliability analyses revealed an internal consistency of α=.83 (good) for the attitudes subscale and α=.74 (acceptable) for the nonacceptance subscale of the QAPT. The internal consistency was α=.80 (good) for the HLS-EU-Q16 and α=.92 (excellent) for the MHL-W-G.

Attitudes Toward Psychotherapy and General and Mental Health Literacy

The attitudes subscale of the QAPT correlated negatively with the nonacceptance subscale questionnaire (*r*=−0.17, 95% CI −0.21 to −0.13; *P*<.001). General health literacy correlated positively with the attitudes subscale (*r*=0.14, 95% CI 0.10-0.18; *P*<.001) and negatively with the nonacceptance subscale (*r*=−0.32, 95% CI −0.36 to −0.28; *P*<.001). Mental health literacy correlated positively with the attitudes subscale (*r*=0.18, 95% CI 0.13-0.22; *P*<.001) and negatively with the nonacceptance subscale (*r*=−0.23, 95% CI −0.27 to −0.19; *P*<.001).

Attitudes Toward Psychotherapy and Experience with Psychotherapy

Participants who have received treatment for a mental disorder reported significantly more positive attitudes toward psychotherapy than participants without treatment experience ($t_{1260.12}$=−10.40, *P*<.001, Cohen *d*=−0.49). Likewise, participants who have received treatment for a mental disorder reported significantly less nonacceptance regarding psychotherapy than participants without treatment experience ($t_{1050.95}$=3.06, *P*=.002, Cohen *d*=0.16).

Similarly, participants who have a close family member who received treatment for a mental disorder reported significantly more positive attitudes toward psychotherapy than participants without family treatment experience ($t_{1912.06}$=−5.66, *P*<.001, Cohen *d*=−0.26). Likewise, participants who have a close family member who received treatment for a mental disorder reported significantly less nonacceptance regarding psychotherapy than participants without family treatment experience (t_{1926}=4.77, *P*<.001, Cohen *d*=0.22).

Attitudes Toward Psychotherapy and Sociodemographic Measures

Age correlated negatively with the attitudes subscale (*r*=−0.11, 95% CI −0.16 to −0.07, *P*<.001) and negatively with the nonacceptance subscale (*r*=−0.08, 95% CI −0.12 to −0.03, *P*<.001).

There were significant gender differences for both subscales. Women reported higher positive attitudes toward psychotherapy than men (t_{1992}=−3.60, *P*<.001, Cohen *d*=−0.16). Women also reported lower nonacceptance than men (t_{1994}=5.29, *P*<.001, Cohen *d*=0.24).

Subjective social status had a significant effect on both the attitude subscale ($F_{2,1991}$=5.25, P=.005, η^2=.005) and the nonacceptance subscale ($F_{2,1993}$=7.71, P<.001, η^2=.008). Post hoc Bonferroni-corrected *t* tests for the attitudes subscale (Bonferroni-corrected α=.017) revealed significant differences between the low and high subjective social status groups (t_{607}=−3.08, P=.002, Cohen d=−0.25). Differences between the middle and high (t_{1678}=−2.34, P=.02, Cohen d=−0.15) and the low and middle subjective social status groups (t_{1698}=−1.82, P=.07, Cohen d=−0.11) were not significant after Bonferroni correction. Descriptively, positive attitudes were lowest in individuals with low subjective social status, followed by middle and high subjective social status. Post hoc Bonferroni-corrected *t* tests for the nonacceptance subscale (Bonferroni-corrected α=.017) revealed significant differences between the low and middle (t_{1699}=2.72, P=.007, Cohen d=0.17) and low and high subjective social status groups (t_{608}=3.88, P<.001, Cohen d=0.32). The difference between the middle and high subjective social status groups was not significant after Bonferroni correction (t_{1679}=2.32, P=.02, Cohen d=0.15). Descriptively, nonacceptance was highest in individuals with low subjective social status, followed by middle and high subjective social status.

Level of education had a significant effect on the attitudes subscale ($F_{2,1983}$=22.27, P<.001, η^2=.021). Post hoc Bonferroni-corrected *t* tests for the attitudes subscale (Bonferroni-corrected α=.017) revealed significant differences between the low and high educational levels (t_{1126}=−5.52, P<.001, Cohen d=−0.36) and the middle and high educational levels ($t_{1635.23}$=−5.82, P<.001, Cohen d=−0.29), but not between the low and middle educational levels (t_{1204}=−1.01, P=.32, Cohen d=−0.06). Descriptively, positive attitudes were highest in individuals with high educational levels, followed by middle and low educational levels.

There was no significant effect of level of education on the nonacceptance subscale ($F_{2,1984}$=0.02, P=.99, η^2<.001). Descriptively, participants with low, middle, and high levels of education reached similar scores.

There were no significant differences for the attitude subscale between individuals with and without a migration background (t_{1992}=0.73, P=.46, Cohen d=0.07). Likewise, there were no significant differences for the nonacceptance subscale between individuals with and without a migration background (t_{1994}=−0.29, P=.77, Cohen d=−0.03).

Results of the inferential analyses for the unweighted dataset can be found in the Multimedia Appendix 1. Importantly, weighting had no effect on the inferential conclusions except for one analysis. The post hoc *t* test for the difference between the low and high subjective social status groups did not reach significance in the unweighted dataset after the Bonferroni correction.

Discussion

Overview

This study investigated the associations between attitudes toward psychotherapeutic treatment and both general and mental health literacy in a large and representative sample of the general population in Germany. In addition, attitudes toward psychotherapeutic treatment were analyzed concerning interindividual differences in terms of experience with mental health treatment and sociodemographic factors.

Principal Findings

Overall, participants expressed largely positive attitudes toward psychotherapeutic treatment and reported relatively low perceived nonacceptance. However, significant interindividual differences were observed and are discussed below.

Importantly, while our results demonstrate several statistically significant results, the practical significance of these findings has to be considered before interpretation. All effect sizes (Cohen d and η^2) reported here fall into the small to medium range [61,62]. In addition, some of the CI of the means overlap. This means that the practical impact of our findings in real-world settings may be limited. Nonetheless, these findings point to potentially vulnerable sociodemographic groups in need of specific interventions to promote positive attitudes toward psychotherapy. The small effect sizes observed here suggest that while tailored interventions for specific groups may be more effective due to a better fit between the specific needs of a sociodemographic group and the intervention, there is also a potential overall benefit of broader general-population interventions [63,64].

The questionnaire scales used in this study reached acceptable to excellent internal consistency coefficients speaking for a high degree of interrelatedness of the scale items and their overall reliability which is an important prerequisite for the validity of the results [60]. Furthermore, weighting did not have a marked effect on the conclusions drawn from the statistical findings which highlights the generalizability of the findings.

General and Mental Health Literacy

Both higher general and higher mental health literacy were associated with more positive attitudes and lower nonacceptance of psychotherapeutic treatment. For mental health literacy, this finding aligns well with previous research [12,32,33,35]. Of note, these previous studies mostly used the simple identification of vignettes as a measure of mental health literacy. Here, we used a much broader instrument for mental health literacy more closely mirroring the definition of mental health literacy by also covering aspects of support, prevention, and use of resources [28,53]. For the first time, we show here that by including these broader aspects of mental health literacy positive associations with positive attitudes and negative associations with nonacceptance can be obtained, significantly extending previous observations in studies using vignette identification as a proxy for mental health literacy. Another novel contribution of our investigation is that we replicated initial findings of the associations between the QAPT and MHL-W-G scales previously observed only in a small sample of working individuals in a substantially larger general population sample with similar effect sizes [43]. Notably, here, we also add to the literature by showing a positive association between general health literacy and the attitudes measure, which to our knowledge was previously only reported in one study [34]. Here, we replicate these initial findings in a larger sample and generalize them to another measure of general health literacy, namely the widely used HLS-EU-Q16. This suggests that individuals with higher levels of both general and mental health literacy are more likely to exhibit positive attitudes toward psychotherapy and perceive societal acceptance to be higher. Consequently, they might also be more likely to seek treatment when faced with mental problems [19]. Crucially, however, it has to be noted that our cross-sectional study design does not allow to infer causal associations between the investigated constructs (also refer to limitations below; [65]). Therefore, we cannot unequivocally conclude whether inadequate or problematic health literacy is indeed the antecedent and origin of less positive attitudes toward psychotherapy. We recommend future studies with appropriate study designs to test for causal effects, for example, randomized controlled trials. In addition, other potentially confounding variables should be carefully assessed and controlled for.

The present findings should also be interpreted in the context of mental health stigma in general. Although some progress in reducing stigma has been made in the past years substantial stigma against people with mental health problems remains [13,14]. Crucially, individuals with more stigmatizing attitudes show less active help-seeking behavior [17]. The promotion of health literacy with targeted interventions along with far-reaching antistigma campaigns could therefore also lead to higher rates of individuals seeking treatment [66]. Importantly, antistigma interventions may be more effective when focusing on psychosocial instead of biogenetic causes for mental disorders [35]. In addition to these interventions focusing on individual barriers, structural barriers to psychotherapeutic treatment (although not the focus of the present research) should not be disregarded and must be addressed as well. For example, waiting periods should be significantly reduced and provision should be increased, especially in rural areas [2,6,7].

Experience with Psychotherapy

Attitudes were more positive, and nonacceptance was lower in participants who have received treatment for mental disorders in the past or those who have a close relative with treatment experience. This is consistent with previous research [8,35]. It is noteworthy, that the treatment experience of a close family member may already improve the treatment attitudes of their relatives although with lower effect sizes than personal experience. Speaking openly about mental health treatment experiences and thus providing information about treatments might therefore be a valuable path to lowering psychotherapy stigma [49]. However, as no causal conclusions can be drawn from our study design, it is possible that positive treatment attitudes are not the result of personal or family treatment experience but that those with a priori more positive treatment attitudes are more likely to seek treatment.

Sociodemographic Factors

Women exhibited significantly more positive attitudes and reported lower perceived societal nonacceptance (ie, higher acceptance) of psychotherapeutic treatment compared to men. This finding aligns with previous research [10,67] and may be attributed to masculinity ideals and self-stigma, preventing men from seeking help for mental health problems [68,69].

Regarding age, more positive attitudes toward psychotherapeutic treatment were found in younger individuals, consistent with some [8] but not all previous investigations [10,70]. However, surprisingly, younger individuals also displayed higher levels of nonacceptance compared to older individuals, suggesting a discrepancy between perceived societal acceptance and individual attitudes. This pattern of results might be interpreted as higher awareness of the persisting mental health stigma in the community in younger individuals, although a generational shift toward less stigma is slowly taking place [14]. Notably, in our data, correlations with age were relatively small for both subscales. Integrating our results with previous literature, an overall heterogeneous picture emerges [8,10,70]: age seems to have no or a very small effect on attitudes toward psychotherapy and how it is perceived to be accepted by society. Future interventions to improve help-seeking for mental problems should therefore address all age groups.

Attitudes toward psychotherapy were also related to the subjective social status of the participants. Specifically, positive attitudes increased, and nonacceptance decreased with higher subjective social status, although the post hoc test between the high and low social status groups for the attitudes subscale narrowly missed significance in the unweighted dataset (refer to Multimedia Appendix 1). Social status has been identified as a major determinant of help-seeking behavior in the past [71,72]. However, more research is needed to determine the role of social status in attitudes toward psychotherapy. Critically, social status also plays a role in access to psychotherapy. Individuals with lower social status are less likely to be offered treatment than individuals with higher social status suggesting that structural barriers are also higher for these individuals than those with higher social status [73]. Potential financial barriers to access to psychotherapeutic treatment should also be considered but do not apply to the majority of our participants as the German public health insurance that insures about 90% of the population covers psychotherapeutic treatment for many diagnoses [74-76]. Treatment attitudes and treatment seeking should be carefully addressed in help-seeking interventions for individuals with lower social status and treatment should be provided equally to individuals of all social status groups. As general and mental health literacy were found to be low in individuals with lower subjective social status, these groups might particularly benefit from interventions [31,77]. There is first evidence to suggest that individuals with low social status might particularly benefit from culturally appropriate and tailored interventions that are based on scientific theories and incorporate different modalities, for example, using print materials and offering direct contact with the interventionists [63,64,78].

Higher educational levels were associated with more positive attitudes toward psychotherapy. However, there was no association between education and perceived acceptance of psychotherapy. In the literature, there is heterogeneous evidence concerning the role of education in attitudes toward mental health help-seeking [8,10,53]. Therefore, more research in this area is recommended to determine the role of education in help-seeking attitudes. Our results carefully suggest that programs directed at individuals with lower levels of education could be particularly beneficial.

Migration background was not related to neither positive attitudes nor perceived nonacceptance in our sample. This result is surprising given previous evidence of more negative attitudes toward psychotherapeutic treatment, and simultaneous higher endorsement of pharmacotherapy, in people with a migration background [44-46]. However, as we only included individuals with sufficient knowledge of the German language, it cannot be ruled out that our sample of individuals with a migration background was not representative of all individuals with a migration background in Germany (refer to *Limitations* section). Future studies could try to overcome this problem by providing translated questionnaires to reach a broader segment of the population with a migration background. In addition, substantial cultural differences have been delineated in the past, which is why we recommend investigating more thoroughly how the cultural backgrounds of the individuals with migration backgrounds are associated with their endorsement of psychotherapeutic treatment [44,79,80], but our data does not allow more fine-grained analyses. Notably, however, interventions to increase the use of mental health care may be more effective when tailored to the specific needs of minority groups and different cultures [47,63,64,78].

Overall, the findings presented here imply that more interventions to improve attitudes toward psychotherapy might be necessary and might be achieved by explicitly targeting both general and health literacy in vulnerable segments of the community. Our findings suggest that male individuals, those with higher age, lower social status, and lower levels of education might particularly benefit from interventions to promote positive attitudes toward psychotherapy. Of note, while a large number of interventions have already been developed and evaluated in the past, systematic reviews of the literature show substantial heterogeneity in the efficacy of these interventions and call for more high-quality evaluations [36-39]. In particular, more effort needs to be put into interventions in different cultures and across ethnic minority groups [40,41,80]. Notably, there are first promising results, for example concerning digital mental health interventions that are both cost-effective and easy to implement [39]. Web-based interventions appear to be generally efficacious when they include certain active components such as interactivity, target-group specificity, evidence-based content, and a sequential modular concept [81], for example as part of an e-learning platform [82,83]. While potential improvements in health literacy may not directly translate to increased help-seeking behavior, further research is needed to explore this connection [81].

Limitations

The findings presented here have to be interpreted in the light of some limitations. First, both attitudes and health literacy were assessed via self-reports. Future studies could focus on more applied measures rather than self-reports for a broader scope and higher ecological validity. Second, 6.2% of the participants in the weighted sample reported having a migration background. This is substantially lower than recent census data from Germany [84], suggesting that our sample may not be representative in terms of the proportion of individuals with a migration background. Third, giving our participants the option to opt-out from specific items might have resulted in systematically missing data which could have affected the generalizability of the results. However, the number of individuals using that option was relatively low, with the highest number (3.4%) for the item asking about family experience with psychotherapy which we speculate might have been the result of lack of knowledge. Importantly, the proportions of missing values were smaller for the main questionnaires used here (eg, 0.3% and 0.2% for the QAPT scales, refer to Table 2). As these proportions are substantially lower than the recommended threshold of 5%, no imputation analyses for missing values were conducted and all analyses were conducted with all available data [85]. Fourth, we relied on data collected on the web which might have resulted in lower data quality [86,87]. However, with this interview mode, a substantially larger sample size could be achieved, simultaneously reducing error variance. In addition, by working with the forsa.omninet panel a few issues common to web-based surveys could be mitigated. For example, participants were randomly selected and could not apply to take part in the study, people who rarely use the internet were also represented due to the panel recruitment via telephone, and multiple registrations of the same individuals were prevented. However, the sample is only representative of the population with internet access as data collection took place exclusively on the internet, which is why results may not be generalizable to individuals without internet access. Finally, as noted above, the present investigation used a cross-sectional study design. Consequently, the results cannot be interpreted in terms of causal relationships [65], which is why further research is needed with more sophisticated study designs including longitudinal and experimental studies that control for potentially confounding factors.

Summary and Conclusion

In this study, positive attitudes toward psychotherapy and perceived societal acceptance of psychotherapeutic treatment were investigated in a large general population sample in Germany. Our results highlight a positive association between these factors and both general and mental health literacy. Individuals who have been treated for mental problems or who have relatives with treatment experience reported more favorable attitudes and higher perceived acceptance of society than those without treatment experience. In terms of sociodemographic differences, it was shown that being a woman, younger age, higher social status, and higher levels of education were associated with more positive attitudes toward psychotherapeutic treatment. Being a woman, older age, and higher subjective social status were associated with higher perceived societal acceptance of psychotherapy. Since previous studies suggest that target group-specific interventions have the potential to improve health literacy, future studies should focus on establishing whether tailored interventions designed to increase general or mental health literacy and aimed at individuals with deficits in this area may improve attitudes toward seeking help for mental problems.

Acknowledgments

This research was funded by the independent, nonprofit foundation Stiftung Gesundheitswissen. forsa Gesellschaft für Sozialforschung und statistische Analysen mBH was responsible for the recruitment of study participants and data acquisition.

Data Availability

The datasets generated and analyzed during this study are available from the Stiftung Gesundheitswissen on reasonable request.

Authors' Contributions

R Schröder contributed to methodology, formal analysis, writing - original draft, review, and editing. TH managed validation, writing, review, and editing. R Suhr handled supervision, writing reviews, and editing. LK contributed to conceptualization; methodology; supervision; writing, review, and editing.

Conflicts of Interest

All authors are employees of the independent, nonprofit foundation Stiftung Gesundheitswissen.

Multimedia Appendix 1
Supplementary analyses and tables.
[PDF File (Adobe PDF File), 273 KB - publichealth_v11i1e67078_app1.pdf]

References

1. Jacobi F, Höfler M, Siegert J, Mack S, Gerschler A, Scholl L, et al. Twelve-month prevalence, comorbidity and correlates of mental disorders in Germany: the mental health module of the German health interview and examination survey for adults (DEGS1-MH). Int J Methods Psychiatr Res 2014;23(3):304-319 [FREE Full text] [doi: 10.1002/mpr.1439] [Medline: 24729411]
2. Inanspruchnahme psychiatrischer und psychotherapeutischer Leistungen – individuelle Determinanten und regionale Unterschiede [Article in German]. Robert Koch-Institut. 2017. URL: https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsJ/Focus/JoHM_04_2017_Psychiatrische_Psychotherapeutische_Leistungen.html [accessed 2025-01-11]
3. Brandstetter S, Dodoo-Schittko F, Speerforck S, Apfelbacher C, Grabe H, Jacobi F, et al. Trends in non-help-seeking for mental disorders in Germany between 1997-1999 and 2009-2012: a repeated cross-sectional study. Soc Psychiatry Psychiatr Epidemiol 2017;52(8):1005-1013. [doi: 10.1007/s00127-017-1384-y] [Medline: 28456871]
4. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022;9(2):137-150 [FREE Full text] [doi: 10.1016/S2215-0366(21)00395-3] [Medline: 35026139]
5. Huhn M, Tardy M, Spineli LM, Kissling W, Förstl H, Pitschel-Walz G, et al. Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: a systematic overview of meta-analyses. JAMA Psychiatry 2014;71(6):706-715. [doi: 10.1001/jamapsychiatry.2014.112] [Medline: 24789675]
6. Singer S, Maier L, Paserat A, Lang K, Wirp B, Kobes J, et al. Wartezeiten auf einen Psychotherapieplatz vor und nach der Psychotherapiestrukturreform. Psychotherapeut 2022;67(2):176-184. [doi: 10.1007/s00278-021-00551-0]
7. Tomczyk S, Schomerus G, Stolzenburg S, Muehlan H, Schmidt S. Ready, willing and able? an investigation of the theory of planned behaviour in help-seeking for a community sample with current untreated depressive symptoms. Prev Sci 2020;21(6):749-760 [FREE Full text] [doi: 10.1007/s11121-020-01099-2] [Medline: 32140825]
8. ten Have M, de Graaf R, Ormel J, Vilagut G, Kovess V, Alonso J, et al. Are attitudes towards mental health help-seeking associated with service use? Results from the European Study of Epidemiology of Mental Disorders. Soc Psychiatry Psychiatr Epidemiol 2010;45(2):153-163 [FREE Full text] [doi: 10.1007/s00127-009-0050-4] [Medline: 19381427]
9. Osman N, Michel C, Schimmelmann BG, Schilbach L, Meisenzahl E, Schultze-Lutter F. Influence of mental health literacy on help-seeking behaviour for mental health problems in the Swiss young adult community: a cohort and longitudinal case-control study. Eur Arch Psychiatry Clin Neurosci 2023;273(3):649-662 [FREE Full text] [doi: 10.1007/s00406-022-01483-9] [Medline: 36088495]
10. Petrowski K, Hessel A, Körner A, Weidner K, Brähler E, Hinz A. Attitudes toward psychotherapy in the general population. Psychother Psychosom Med Psychol 2014;64(2):82-85. [doi: 10.1055/s-0033-1361155] [Medline: 24515849]

11. Angermeyer MC, Matschinger H, Schomerus G. Attitudes towards psychiatric treatment and people with mental illness: changes over two decades. Br J Psychiatry 2013;203(2):146-151. [doi: 10.1192/bjp.bp.112.122978] [Medline: 23787060]
12. Angermeyer MC, van der Auwera S, Carta MG, Schomerus G. Public attitudes towards psychiatry and psychiatric treatment at the beginning of the 21st century: a systematic review and meta-analysis of population surveys. World Psychiatry 2017;16(1):50-61 [FREE Full text] [doi: 10.1002/wps.20383] [Medline: 28127931]
13. Angermeyer MC, Matschinger H, Carta MG, Schomerus G. Changes in the perception of mental illness stigma in Germany over the last two decades. Eur Psychiatry 2014;29(6):390-395. [doi: 10.1016/j.eurpsy.2013.10.004] [Medline: 24321774]
14. Schomerus G, Schindler S, Sander C, Baumann E, Angermeyer MC. Changes in mental illness stigma over 30 years - improvement, persistence, or deterioration? Eur Psychiatry 2022;65(1):e78 [FREE Full text] [doi: 10.1192/j.eurpsy.2022.2337] [Medline: 36328960]
15. Ghanean H, Nojomi M, Jacobsson L. A community study on attitudes to and knowledge of mental illness in Tehran. Open J Psychiatry 2015;05(01):26-30.
16. Atilola O. Level of community mental health literacy in sub-Saharan Africa: current studies are limited in number, scope, spread, and cognizance of cultural nuances. Nord J Psychiatry 2015;69(2):93-101. [doi: 10.3109/08039488.2014.947319] [Medline: 25151995]
17. Schnyder N, Panczak R, Groth N, Schultze-Lutter F. Association between mental health-related stigma and active help-seeking: systematic review and meta-analysis. Br J Psychiatry 2017;210(4):261-268. [doi: 10.1192/bjp.bp.116.189464] [Medline: 28153928]
18. Villatoro AP, DuPont-Reyes MJ, Phelan JC, Link BG. 'Me' vs. 'Them': how mental illness stigma influences adolescent help-seeking behaviors for oneself and recommendations for peers. Stigma Health 2022;7(3):300-310 [FREE Full text] [doi: 10.1037/sah0000392] [Medline: 36776352]
19. Bonabi H, Müller M, Ajdacic-Gross V, Eisele J, Rodgers S, Seifritz E, et al. Mental health literacy, attitudes to help seeking, and perceived need as predictors of mental health service use: a longitudinal study. J Nerv Ment Dis 2016;204(4):321-324 [FREE Full text] [doi: 10.1097/NMD.0000000000000488] [Medline: 27015396]
20. Mojtabai R, Evans-Lacko S, Schomerus G, Thornicroft G. Attitudes toward mental health help seeking as predictors of future help-seeking behavior and use of mental health treatments. Psychiatr Serv 2016;67(6):650-657. [doi: 10.1176/appi.ps.201500164] [Medline: 26876662]
21. Sørensen K, van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
22. Zheng M, Jin H, Shi N, Duan C, Wang D, Yu X, et al. The relationship between health literacy and quality of life: a systematic review and meta-analysis. Health Qual Life Outcomes 2018;16(1):201 [FREE Full text] [doi: 10.1186/s12955-018-1031-7] [Medline: 30326903]
23. Miller TA. Health literacy and adherence to medical treatment in chronic and acute illness: a meta-analysis. Patient Educ Couns 2016;99(7):1079-1086 [FREE Full text] [doi: 10.1016/j.pec.2016.01.020] [Medline: 26899632]
24. Fan Z, Yang Y, Zhang F. Association between health literacy and mortality: a systematic review and meta-analysis. Arch Public Health 2021;79(1):119 [FREE Full text] [doi: 10.1186/s13690-021-00648-7] [Medline: 34210353]
25. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]
26. Ranjbaran S, Chollou KM, Babazadeh T. Assessment of health literacy and health promoting behaviors among the urban adult population. Ethiop J Health Sci 2022;32(5):985-992 [FREE Full text] [doi: 10.4314/ejhs.v32i5.14] [Medline: 36262704]
27. Rezakhani Moghaddam H, Ranjbaran S, Babazadeh T. The role of e-health literacy and some cognitive factors in adopting protective behaviors of COVID-19 in Khalkhal residents. Front Public Health 2022;10:916362 [FREE Full text] [doi: 10.3389/fpubh.2022.916362] [Medline: 35942262]
28. Kutcher S, Wei Y, Coniglio C. Mental health literacy: past, present, and future. Can J Psychiatry 2016;61(3):154-158 [FREE Full text] [doi: 10.1177/0706743715616609] [Medline: 27254090]
29. Bjørnsen HN, Espnes GA, Eilertsen MB, Ringdal R, Moksnes UK. The relationship between positive mental health literacy and mental well-being among adolescents: implications for school health services. J Sch Nurs 2019;35(2):107-116 [FREE Full text] [doi: 10.1177/1059840517732125] [Medline: 28950750]
30. Suwanwong C, Jansem A, Intarakamhang U, Prasittichok P, Tuntivivat S, Chuenphittayavut K, et al. Modifiable predictors of mental health literacy in the educational context: a systematic review and meta-analysis. BMC Psychol 2024;12(1):378 [FREE Full text] [doi: 10.1186/s40359-024-01878-4] [Medline: 38965633]
31. König L, Hamer T, Suhr R. Die psychische Gesundheitskompetenz der Bevölkerung in Deutschland [Article in German]. Präv Gesundheitsf 2023;19(4):521-526. [doi: 10.1007/s11553-023-01079-8]
32. Svensson B, Hansson L. How mental health literacy and experience of mental illness relate to stigmatizing attitudes and social distance towards people with depression or psychosis: a cross-sectional study. Nord J Psychiatry 2016;70(4):309-313. [doi: 10.3109/08039488.2015.1109140] [Medline: 26643359]

33. Cheng H, Wang C, McDermott RC, Kridel M, Rislin JL. Self-stigma, mental health literacy, and attitudes toward seeking psychological help. Jour of Counseling & Develop 2018;96(1):64-74. [doi: 10.1002/jcad.12178]
34. Fleary SA, Joseph PL, Gonçalves C, Somogie J, Angeles J. The relationship between health literacy and mental health attitudes and beliefs. Health Lit Res Pract 2022;6(4):e270-e279 [FREE Full text] [doi: 10.3928/24748307-20221018-01] [Medline: 36350237]
35. Doll CM, Michel C, Betz LT, Schimmelmann BG, Schultze-Lutter F. The important role of stereotypes in the relation between mental health literacy and stigmatization of depression and psychosis in the community. Community Ment Health J 2022;58(3):474-486 [FREE Full text] [doi: 10.1007/s10597-021-00842-5] [Medline: 34037914]
36. Visscher BB, Steunenberg B, Heijmans M, Hofstede JM, Devillé W, van der Heide I, et al. Evidence on the effectiveness of health literacy interventions in the EU: a systematic review. BMC Public Health 2018;18(1):1414 [FREE Full text] [doi: 10.1186/s12889-018-6331-7] [Medline: 30594180]
37. Walters R, Leslie SJ, Polson R, Cusack T, Gorely T. Establishing the efficacy of interventions to improve health literacy and health behaviours: a systematic review. BMC Public Health 2020;20(1):1040 [FREE Full text] [doi: 10.1186/s12889-020-08991-0] [Medline: 32605608]
38. Ma KKY, Anderson JK, Burn A. Review: school-based interventions to improve mental health literacy and reduce mental health stigma - a systematic review. Child Adolesc Ment Health 2023;28(2):230-240. [doi: 10.1111/camh.12543] [Medline: 35005830]
39. Yeo G, Reich SM, Liaw NA, Chia EYM. The effect of digital mental health literacy interventions on mental health: systematic review and meta-analysis. J Med Internet Res 2024;26:e51268 [FREE Full text] [doi: 10.2196/51268] [Medline: 38421687]
40. Fernández-Gutiérrez M, Bas-Sarmiento P, Albar-Marín MJ, Paloma-Castro O, Romero-Sánchez JM. Health literacy interventions for immigrant populations: a systematic review. Int Nurs Rev 2018;65(1):54-64. [doi: 10.1111/inr.12373] [Medline: 28449363]
41. Meherali S, Punjani NS, Mevawala A. Health literacy interventions to improve health outcomes in low- and middle-income countries. Health Lit Res Pract 2020;4(4):e251-e266 [FREE Full text] [doi: 10.3928/24748307-20201118-01] [Medline: 33313935]
42. Moll S, Zanhour M, Patten SB, Stuart H, MacDermid J. Evaluating mental health literacy in the workplace: development and psychometric properties of a vignette-based tool. J Occup Rehabil 2017;27(4):601-611. [doi: 10.1007/s10926-017-9695-0] [Medline: 28120136]
43. Wulf IC, Ruhle SA. Psychische Gesundheitskompetenz am Arbeitsplatz [Article in German]. Diagnostica 2020;66(1):50-61. [doi: 10.1026/0012-1924/a000237]
44. Schulz W, Shin MA, Schmid-Ott G. Attitudes towards psychotherapy in South Korea and Germany : a cross-cultural comparative study. Nervenarzt 2018;89(1):51-57. [doi: 10.1007/s00115-016-0252-4] [Medline: 27909737]
45. Brandl EJ, Dietrich N, Mell N, Winkler JG, Gutwinski S, Bretz HJ, et al. Attitudes towards psychopharmacology and psychotherapy in psychiatric patients with and without migration background. BMC Psychiatry 2020;20(1):176 [FREE Full text] [doi: 10.1186/s12888-020-02585-1] [Medline: 32303265]
46. Verhulp EE, Stevens GWJM, Pels TVM, van Weert CMC, Vollebergh WAM. Lay beliefs about emotional problems and attitudes toward mental health care among parents and adolescents: Exploring the impact of immigration. Cultur Divers Ethnic Minor Psychol 2017;23(2):269-280. [doi: 10.1037/cdp0000092] [Medline: 27709973]
47. Lee-Tauler SY, Eun J, Corbett D, Collins PY. A systematic review of interventions to improve initiation of mental health care among racial-ethnic minority groups. Psychiatr Serv 2018;69(6):628-647. [doi: 10.1176/appi.ps.201700382] [Medline: 29716446]
48. Lehti V, Kieseppä V, Gissler M, Suvisaari J, Markkula N. Psychotherapy use among migrants: a register-based longitudinal study. J Epidemiol Community Health 2024;79(1):49-55 [FREE Full text] [doi: 10.1136/jech-2024-222330] [Medline: 39147569]
49. Braun-Koch K, Rief W, Teige-Mocigemba S. Changing attitudes towards psychotherapy via social observations: are similarities more important than discrepancies? BMC Psychol 2022;10(1):286 [FREE Full text] [doi: 10.1186/s40359-022-00952-z] [Medline: 36461123]
50. Braun-Koch K, Rief W. Maintenance vs. change of negative therapy expectation: an experimental investigation using video samples. Front Psychiatry 2022;13:836227 [FREE Full text] [doi: 10.3389/fpsyt.2022.836227] [Medline: 35444567]
51. Albarracin D, Shavitt S. Attitudes and attitude change. Annu Rev Psychol 2018;69:299-327. [doi: 10.1146/annurev-psych-122216-011911] [Medline: 28841390]
52. König L, Kuhlmey A, Suhr R. Digital health literacy of the population in Germany and its association with physical health, mental health, life satisfaction, and health behaviors: nationally representative survey study. JMIR Public Health Surveill 2024;10:e48685 [FREE Full text] [doi: 10.2196/48685] [Medline: 38381497]
53. Schmid-Ott G, Reibold S, Ernst G, Niederauer H, Künsebeck H, Schulz W, et al. Development of a questionnaire to assess attitudes towards psychotherapeutic treatment. Dermatol Psychosom 2004;4(4):187-193. [doi: 10.1159/000075905]

54. Ditte D, Schulz W, Schmid-Ott G. Attitude towards psychotherapy in the Russian population and in the population with a Russian/Soviet cultural background in Germany. A pilot study. Nervenarzt 2006;77(1):64-72. [doi: 10.1007/s00115-005-1889-6] [Medline: 15776258]
55. Pelikan JM, Röthlin F, Ganahl K. Measuring comprehensive health literacy in general populations: validation of instrument, indices and scales of the HLS-EU study. Bethesda, MD: Hyatt Regency Bethesda; 2014.
56. Sørensen K, van den Broucke S, Pelikan JM, Fullam J, Doyle G, Slonska Z, et al. Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). BMC Public Health 2013;13:948 [FREE Full text] [doi: 10.1186/1471-2458-13-948] [Medline: 24112855]
57. Jordan S, Hoebel J. Health literacy of adults in Germany: findings from the German Health Update (GEDA) study. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015;58(9):942-950. [doi: 10.1007/s00103-015-2200-z] [Medline: 26227894]
58. Hoebel J, Müters S, Kuntz B, Lange C, Lampert T. Measuring subjective social status in health research with a German version of the MacArthur Scale. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015;58(7):749-757. [doi: 10.1007/s00103-015-2166-x] [Medline: 25986532]
59. Adler NE, Epel ES, Castellazzo G, Ickovics JR. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. Health Psychol 2000;19(6):586-592. [doi: 10.1037//0278-6133.19.6.586] [Medline: 11129362]
60. Gliem J, Gliem R. Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. Indiana University. URL: https://hdl.handle.net/1805/344 [accessed 2025-01-11]
61. Cohen J. A power primer. Psychol Bull 1992;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
62. Ellis PD. The essential guide to effect sizes: statistical power, meta-analysis, and the interpretation of research results 1st Edition. Cambridge, England: Cambridge University Press; 2010.
63. Noar S, Benac C, Harris M. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
64. Lustria MLA, Noar SM, Cortese J, van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
65. Wang X, Cheng Z. Cross-sectional studies: strengths, weaknesses, and recommendations. Chest 2020;158(1S):S65-S71. [doi: 10.1016/j.chest.2020.03.012] [Medline: 32658654]
66. Hansson L, Stjernswärd S, Svensson B. Changes in attitudes, intended behaviour, and mental health literacy in the Swedish population 2009-2014: an evaluation of a national antistigma programme. Acta Psychiatr Scand 2016;134 Suppl 446:71-79. [doi: 10.1111/acps.12609] [Medline: 27426648]
67. Nam SK, Chu HJ, Lee MK, Lee JH, Kim N, Lee SM. A meta-analysis of gender differences in attitudes toward seeking professional psychological help. J Am Coll Health 2010;59(2):110-116. [doi: 10.1080/07448481.2010.483714] [Medline: 20864437]
68. Yousaf O, Popat A, Hunter MS. An investigation of masculinity attitudes, gender, and attitudes toward psychological help-seeking. Psychol. Men Masculinity 2015;16(2):234-237. [doi: 10.1037/a0036241]
69. Levant RF, Stefanov DG, Rankin TJ, Halter MJ, Mellinger C, Williams CM. Moderated path analysis of the relationships between masculinity and men's attitudes toward seeking psychological help. J Couns Psychol 2013;60(3):392-406. [doi: 10.1037/a0033014] [Medline: 23668728]
70. Pepin R, Segal DL, Coolidge FL. Intrinsic and extrinsic barriers to mental health care among community-dwelling younger and older adults. Aging Ment Health 2009;13(5):769-777. [doi: 10.1080/13607860902918231] [Medline: 19882416]
71. Choi N, Miller MJ. Social class, classism, stigma, and college students‘ attitudes toward counseling. Couns. Psychol 2018;46(6):761-785. [doi: 10.1177/0011000018796789]
72. Eisenberg D, Golberstein E, Gollust SE. Help-seeking and access to mental health care in a university student population. Med Care 2007;45(7):594-601. [doi: 10.1097/MLR.0b013e31803bb4c1] [Medline: 17571007]
73. Niemeyer H, Knaevelsrud C. Socioeconomic status and access to psychotherapy. J Clin Psychol 2023;79(4):937-953. [doi: 10.1002/jclp.23449] [Medline: 36251952]
74. Epping J, Muschik D, Geyer S. Social inequalities in the utilization of outpatient psychotherapy: analyses of registry data from German statutory health insurance. Int J Equity Health 2017;16(1):147 [FREE Full text] [doi: 10.1186/s12939-017-0644-5] [Medline: 28814311]
75. Sozialleistungen: Angaben zur Krankenversicherung (Ergebnisse des Mikrozensus) [Article in German]. Fachserie 13 Reihe 1.1. Wiesbaden: Statistisches Bundesamt; 2019. URL: https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00014767/2130110119004_Rev2015.pdf [accessed 2025-01-11]
76. Sareen J, Jagdeo A, Cox BJ, Clara I, ten Have M, Belik S, et al. Perceived barriers to mental health service utilization in the United States, Ontario, and the Netherlands. Psychiatr Serv 2007;58(3):357-364. [doi: 10.1176/ps.2007.58.3.357] [Medline: 17325109]
77. Schaeffer D, Berens E, Vogt D. Health literacy in the German population. Dtsch Arztebl Int 2017;114(4):53-60 [FREE Full text] [doi: 10.3238/arztebl.2017.0053] [Medline: 28211318]

78. Stormacq C, Wosinski J, Boillat E, van den Broucke S. Effects of health literacy interventions on health-related outcomes in socioeconomically disadvantaged adults living in the community: a systematic review. JBI Evid Synth 2020;18(7):1389-1469. [doi: 10.11124/JBISRIR-D-18-00023] [Medline: 32813388]
79. Balkir Neftçi N, Barnow S. One size does not fit all in psychotherapy: understanding depression among patients of turkish origin in Europe. Noro Psikiyatr Ars 2016;53(1):72-79 [FREE Full text] [doi: 10.5152/npa.2016.12666] [Medline: 28360770]
80. Selkirk M, Quayle E, Rothwell N. A systematic review of factors affecting migrant attitudes towards seeking psychological help. J Health Care Poor Underserved 2014;25(1):94-127. [doi: 10.1353/hpu.2014.0026] [Medline: 24509015]
81. Brijnath B, Protheroe J, Mahtani KR, Antoniades J. Do web-based mental health literacy interventions improve the mental health literacy of adult consumers? Results from a systematic review. J Med Internet Res 2016;18(6):e165 [FREE Full text] [doi: 10.2196/jmir.5463] [Medline: 27323907]
82. König L, Suhr R. The effectiveness of publicly available web-based interventions in promoting health app use, digital health literacy, and media literacy: pre-post evaluation study. J Med Internet Res 2023;25:e46336 [FREE Full text] [doi: 10.2196/46336] [Medline: 38048146]
83. König L, Marbach-Breitrück E, Engler A, Suhr R. The development and evaluation of an e-Learning course that promotes digital health literacy in school-age children: pre-post measurement study. J Med Internet Res 2022;24(5):e37523 [FREE Full text] [doi: 10.2196/37523] [Medline: 35576572]
84. Bevölkerung mit Migrationshintergrund - Ergebnisse des Mikrozensus 2021: Fachserie 1 Reihe 2.2 [Article in German]. Endergebnisse. 2021. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Publikationen/Downloads-Migration/migrationshintergrund-endergebnisse-2010220217004.pdf?__blob=publicationFile [accessed 2025-01-11]
85. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol 2017;17(1):162 [FREE Full text] [doi: 10.1186/s12874-017-0442-1] [Medline: 29207961]
86. Zhang X, Kuchinke L, Woud ML, Velten J, Margraf J. Survey method matters: online/offline questionnaires and face-to-face or telephone interviews differ. Comput. Hum. Behav 2017;71:172-180. [doi: 10.1016/j.chb.2017.02.006]
87. Heerwegh D, Loosveldt G. Face-to-face versus web surveying in a high-internet-coverage population: differences in response quality. Public Opin. Q 2008;72(5):836-846.

Abbreviations

HLS-EU-Q16: European Health Literacy Survey instrument
MHL-W-G: Mental Health Literacy Tool for the Workplace (German version)
QAPT: Questionnaire on Attitudes Towards Psychotherapeutic Treatment

Edited by A Mavragani; submitted 04.10.24; peer-reviewed by C Timm, S Ranjbaran, J Rosario; comments to author 04.11.24; revised version received 18.12.24; accepted 05.01.25; published 24.01.25.

Please cite as:
Schröder R, Hamer T, Suhr R, König L
Attitudes Toward Psychotherapeutic Treatment and Health Literacy in a Large Sample of the General Population in Germany: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e67078
URL: https://publichealth.jmir.org/2025/1/e67078
doi:10.2196/67078
PMID:39854696

Original Paper

Nationwide Trends in Screen Time and Associated Risk Factors by Family Structures Among Adolescents, 2008-2022: Nationwide Cross-Sectional Study

Seokjun Kim[1,2*], MD; Hyesu Jo[2,3*], MS; Yejun Son[2,4*], MS; Min Kyung Shin[5*], MD; Kyeongmin Lee[2,3], MS; Jaeyu Park[2,4], PhD; Hayeon Lee[2,6], PhD; Lee Smith[7], PhD; Elena Dragioti[8], PhD; Guillaume Fond[9], MD, PhD; Laurent Boyer[9], MD, PhD; Guillermo F López Sánchez[10], PhD; Mark A Tully[11], PhD; Masoud Rahmati[9,12,13], PhD; Damiano Pizzol[14,15], PhD; Selin Woo[1,2], PhD; Dong Keon Yon[1,2,3,4,16], MD, PhD

[1]Department of Medicine, Kyung Hee University College of Medicine, Seoul, Republic of Korea
[2]Center for Digital Health, Medical Science Research Institute, Kyung Hee University College of Medicine, Seoul, Republic of Korea
[3]Department of Regulatory Science, Kyung Hee University, Seoul, Republic of Korea
[4]Department of Precision Medicine, Kyung Hee University College of Medicine, Seoul, Republic of Korea
[5]Department of Medicine, Wonkwang University College of Medicine, Iksan, Republic of Korea
[6]Department of Electronics and Information Convergence Engineering, Kyung Hee University, Yongin, Republic of Korea
[7]Centre for Health, Performance and Wellbeing, Anglia Ruskin University, Cambridge, United Kingdom
[8]Research Laboratory Psychology of Patients, Families, and Health Professionals, Department of Nursing, School of Health Sciences, University of Ioannina, Ioannina, Greece
[9]CEReSS-Health Service Research and Quality of Life Center, Assistance Publique-Hôpitaux de Marseille, Aix-Marseille University, Marseille, France
[10]Division of Preventive Medicine and Public Health, Department of Public Health Sciences, School of Medicine, University of Murcia, Murcia, Spain
[11]School of Medicine, Ulster University, Londonderry, United Kingdom
[12]Department of Physical Education and Sport Sciences, Faculty of Literature and Human Sciences, Lorestan University, Khoramabad, Iran
[13]Department of Physical Education and Sport Sciences, Faculty of Literature and Humanities, Vali-E-Asr University of Rafsanjan, Rafsanjan, Iran
[14]Health Unit, Eni, San Donato Milanese, Italy
[15]Health Unit, Eni, Maputo, Mozambique
[16]Department of Pediatrics, Kyung Hee University Medical Center, Kyung Hee University College of Medicine, Seoul, Republic of Korea
[*]these authors contributed equally

Corresponding Author:
Dong Keon Yon, MD, PhD
Department of Medicine
Kyung Hee University College of Medicine
23 Kyungheedae-ro
Dongdaemun-gu
Seoul, 02447
Republic of Korea
Phone: 82 269352476
Fax: 82 5044780201
Email: yonkkang@gmail.com

Abstract

Background: Although understanding long-term trends in adolescent screen time and the influence of family structure is essential, there is a lack of research addressing these issues comprehensively.

Objective: This study aimed to conduct comprehensive investigations into adolescent screen time before and during the COVID-19 pandemic, with a particular focus on family structures.

Methods: This study used nationwide, large-scale data from the Korea Youth Risk Behavior Web-Based Survey from South Korea. We aimed to indicate the changes in adolescent screen time over 15 years from 2008 to 2022. Weighted linear regression was used to analyze annual trends in screen time before and during the pandemic, and stratified analyses were conducted to examine associated risk factors across different family structures.

Results: This study used data from a total of 836,972 individuals (n=403,456, 48.2% women), with an age range of 12-18 years. The analysis revealed an overall increase in screen time prepandemic (β=8.06, 95% CI 7.74-8.39), with a notable increase observed at the onset of the pandemic (β=162.06, 95% CI 159.49-164.64). Among diverse family structures, the orphanage group showed the most substantial increase in screen time during the pandemic (β_{diff}=221.90, 95% CI 159.62-284.17). Risk factors associated with screen time during the pandemic varied by family structure. Notably, the nuclear family group presented distinct screen time–related risk factors, including grade, region of residence, physical activity frequency, sadness and despair, and the highest education level of parents.

Conclusions: There has been a notable increase in average screen time among adolescents since the onset of the pandemic, with the orphanage group exhibiting a pronounced trend. The risk factors associated with screen time during the pandemic varied for each family structure. Findings from this study suggest that the implementation of individualized measures tailored to each family structure should be adopted to effectively address the increased issue of adolescent screen time since the pandemic.

(JMIR Public Health Surveill 2025;11:e57962) doi:10.2196/57962

KEYWORDS

adolescents; family type; pandemic; screen time; South Korea; sedentary activity; risk factor; mobile phone

Introduction

There was a declaration of COVID-19 as a global pandemic by the World Health Organization in March 2020 [1]. With the proclamation, countries worldwide have advocated for indoor living, remote work, and web-based learning to prevent the spread of COVID-19. For example, the National Health Service in England advised individuals to remain at home and participate in remote work unless necessary, while the Chinese government implemented restrictions on public transport and postponed the reopening of schools following the vacation period [2]. Globally, universities and educational institutions were closed, prompting a rapid digital transition in the education system [3,4]. The increase in indoor time since the pandemic has likely induced a concurrent increase in screen time [5,6]. Throughout the pandemic, social distancing and economic hardships restricted daily activities and social interactions, negatively impacting individuals' mental health worldwide [7,8]. Among preschool children, psychological resilience has been identified as a factor that amplifies the positive outcomes associated with the use of digital technologies [9]. The need for discussing how the utilization of digital technologies evolves in response to mental health changes brought about by the pandemic is increasingly evident. In such circumstances, specific focus is placed on the increased screen time of adolescents, due to the association with adolescent depression and mental health [10,11]. Global guidelines, such as Canada's 24-Hour Movement Guidelines for Children and Youth, recommend limiting recreational screen time to 2 hours or less per day, highlighting the potential risks of excessive screen use [12]. With the onset of the COVID-19 pandemic, there is growing global attention to screen time among adolescents.

In South Korea, concerns about excessive screen time among adolescents predate the pandemic. The "Shutdown Law," implemented in 2011, restricted web-based gaming for those younger than 16 years old during late-night hours to prevent gaming addiction but was abolished in 2021 [13]. A recent study highlighted a significant increase in screen time among Korean adolescents, with teenagers spending an average of 303.66 minutes daily during the pandemic, a substantial rise from 133.24 minutes in the prepandemic period [14]. The screen time of adolescents appears to be influenced by their family structure, likely due to variations in parental monitoring, restrictions on screen use, and the promotion of outdoor activities [15]. However, studies on this topic are limited by small sample sizes, a narrow focus on specific family structures, and a lack of consideration for the significant changes brought about by the COVID-19 pandemic [15,16]. This underscores the need for a comprehensive analysis of long-term trends in adolescent screen time, particularly about diverse family structures, both before and during the pandemic.

Therefore, this study investigated the 15-year trends in adolescent screen time and potential risk factors based on family structures, using a nationwide dataset of Korean adolescents surveyed between 2008 and 2022. The primary aim of this study was to assess the changes in adolescent screen time trends and associated risk factors in each family structure since the onset of the pandemic.

Methods

Study Design and Population

This study used the Korea Youth Risk Behavior Web-Based Survey (KYRBS). KYRBS is a large-scale, nationwide survey covering health-related behaviors of Korean adolescents [17]. The KYRBS is designed to collect data from middle and high school students (grades 7 to 12) aged 12 to 18 years enrolled in public and private schools throughout South Korea. To achieve national representativeness, the survey uses a stratified, multistage cluster sampling approach [18]. Schools are chosen as the primary sampling units based on regional and school-type stratifications, followed by a random selection of classes within these schools [18]. This methodology ensures a balanced representation of students from both urban and rural areas. Administered annually by the Korea Disease Control and Prevention Agency, the survey has consistently maintained an average response rate of 95.7% over the past 15 years. Students voluntarily completed a web-based anonymous questionnaire following a standardized protocol in a computer lab in each school [19].

To ensure the data's representativeness, sampling weights were determined and applied. The weighting process considers the probability of selection at each stage of sampling (schools and classes), adjusts for nonresponse, and incorporates poststratification adjustments to align with the overall student population distributions by grade, sex, and school type for each survey year. To avoid overrepresentation, extreme weights were capped using interquartile range techniques. These adjustments ensure that the results reliably represent the health behaviors of the adolescent population across South Korea. In this study, an initial cohort of 992,702 respondents completed the questionnaires from 2008 to 2022. Following the exclusion of participants with missing data, the final dataset comprised 836,972 samples (84.3% of the initial raw data). The exclusion of missing data was conducted systematically to maintain dataset robustness and minimize bias.

Ethical Considerations

The study protocol was approved by the Institutional Review Board of the Korea Disease Control and Prevention Agency (2014-06EXP-02-P-A), and all participants provided written informed consent. Additionally, the study complied with the Population Health Promotion Act 19 (117058) as mandated by the Korean government. This study was conducted following the principles of the Declaration of Helsinki.

Endpoints

We aimed to investigate trends in screen time and explore associated risk factors, with a focus on the diverse family structures of adolescents. The COVID-19 pandemic period was defined as 2020-2022, starting with the first observed COVID-19 case in South Korea on January 20, 2020, and aligning with widespread public health measures [20,21]. The prepandemic years were grouped into 6-year intervals, each spanning two consecutive years, as follows: 2008-2009, 2010-2011, 2012-2013, 2014-2015, 2016-2017, and 2018-2019. This classification allows for a clear comparison of trends before and during the pandemic.

Screen time was defined as the average daily time spent on the internet or using smartphones for noneducational purposes. Participants were asked: “What was your average daily internet or smartphone use on weekdays in the past 30 days?” and “What was your average daily internet or smartphone use on weekends in the past 30 days?” The overall average screen time was calculated by combining weekday and weekend use, weighted to reflect their relative proportions within a week [19].

Family structures were categorized into 4 groups: nuclear family, living with relatives, living alone, and orphanage. A nuclear family was defined as adolescents living with one or both parents, regardless of biological relationships. The “living with relatives” category included adolescents residing with extended family members, such as grandparents or uncles or aunts, excluding parents. “Living alone” referred to those managing households independently without the presence of any family members. The “orphanage” category encompassed adolescents living in child welfare facilities. These classifications are based on established definitions in prior studies on family dynamics and adolescent health [22].

Covariate Definitions

A total of 11 covariates were included in the analytical models: sex, BMI group (underweight, normal, overweight, and obese), grade (7th to 9th [middle-school] and 10th to 12th [high-school]), region of residence (urban and rural) [23,24], smoking status, alcoholic consumption, school performance (low, middle-low, middle, middle-high, and high), sexual experience, physical activity frequency (lower activity, moderate activity, and higher activity), sadness and despair, and highest education level of parents (high school or lower and college or higher). BMI was calculated using self-reported height and weight data. Conforming to the 2017 Korean National Growth Charts, BMI was categorized into the following four groups: underweight (<5%), normal (5% to 84 %), overweight (85% to 94%), and obese (≥95%) [19]. School performance was categorized into the following five groups according to the self-reports of students: low (<20%), middle-low (20% to 39%), middle (40% to 59%), middle-high (60% to 79%), and high (≥80%). Physical activity frequency was classified into the following three groups based on the engagement in vigorous aerobic and resistance training more than three days per week: lower activity (neither activity is done for more than three days per week), moderate activity (either one activity), and higher activity (both activities). The definitions of variables were extracted from peer-reviewed literature [19].

Statistical Analysis

We used a weighted linear regression model to compute β coefficient and β difference (β_{diff}) along with a 95% CI [25]. This strategy was selected for trend analysis and associated factor analysis. In the analysis of screen time trends, the β coefficient expressed the annual trend of screen time among adolescents in prepandemic or pandemic. β_{diff} was used to indicate the trend difference between the divided year groups based on the onset of the pandemic. In the analysis of screen time-associated factors, the β coefficient represents the correlation between screen time and each biological, social, and familial factor. β_{diff} was also used to reveal the difference in correlation between the year groups divided by the onset of the pandemic. In addition, the associations between screen time and each variable were recalculated after adjusting for potential influencing factors, including age, sex, BMI group, grade, region of residence, smoking status, alcoholic consumption, school performance, sexual experience, physical activity frequency, sadness and despair, and highest education level of parents. Statistical significance, as applied in all the above analyses, was determined by a 2-sided P value less than .05. Statistical analysis in this study was conducted using SAS software (version 9.4; SAS Institute Inc).

Results

This study was conducted using KYRBS spanning from 2008 to 2022. A total of 836,972 individuals (n=403,456, 48.20% women) with an average age of 15.01 (SD 1.75) years were finally selected for the analysis following the missing data handling. Baseline characteristics of the raw data are indicated in Table 1.

Table 1. General characteristics of participants in KYRBS[a], 2008-2022 (n=836,972).

	Total	Prepandemic						Pandemic
		2008-2009	2010-2011	2012-2013	2014-2015	2016-2017	2018-2019	2020-2022
Overall, n (%)	836,972 (100)	118,457 (14.15)	121,079 (14.47)	111,432 (13.31)	109,654 (13.1)	115,169 (13.76)	101,771 (12.16)	159,410 (19.05)
Age (years), mean (SD)	15.01 (1.75)	15.03 (1.75)	15.08 (1.75)	14.90 (1.76)	14.95 (1.75)	15.00 (1.74)	15.00 (1.78)	15.10 (1.74)
Sex, n (%)								
Men	433,516 (51.8)	63,372 (53.5)	63,168 (52.17)	58,694 (52.67)	56,442 (51.47)	58,685 (50.96)	51,467 (50.57)	81,688 (51.24)
Women	403,456 (48.2)	55,085 (46.5)	57,911 (47.83)	52,738 (47.33)	53,212 (48.53)	56,484 (49.04)	50,304 (49.43)	77,722 (48.76)
BMI group[b], n (%)								
Underweight	65,467 (7.82)	10,586 (8.94)	10,485 (8.66)	8649 (7.76)	8390 (7.65)	7972 (6.92)	6850 (6.73)	12,535 (7.86)
Normal	608,013 (72.64)	89,246 (75.34)	91,583 (75.64)	83,935 (75.32)	81,264 (74.11)	82,875 (71.96)	71,846 (70.6)	107,264 (67.29)
Overweight	68,622 (8.2)	8068 (6.81)	8428 (6.96)	8496 (7.62)	8849 (8.07)	10,046 (8.72)	9315 (9.15)	15,420 (9.67)
Obese	69,758 (8.33)	6281 (5.3)	6708 (5.54)	7146 (6.41)	7783 (7.1)	10,878 (9.45)	10,875 (10.69)	20,087 (12.6)
Grade, n (%)								
7th	140,196 (16.75)	19,698 (16.63)	20,046 (16.56)	18,516 (16.62)	17,531 (15.99)	18,718 (16.25)	16,875 (16.58)	28,812 (18.07)
8th	145,377 (17.37)	20,953 (17.69)	21,030 (17.37)	19,272 (17.29)	18,982 (17.31)	18,990 (16.49)	17,324 (17.02)	28,826 (18.08)
9th	148,190 (17.71)	21,475 (18.13)	21,644 (17.88)	19,804 (17.77)	19,702 (17.97)	19,492 (16.92)	17,771 (17.46)	28,302 (17.75)
10th	132,805 (15.87)	18,780 (15.85)	19,190 (15.85)	17,531 (15.73)	17,055 (15.55)	19,018 (16.51)	15,835 (15.56)	25,396 (15.93)
11th	135,789 (16.22)	19,028 (16.06)	19,731 (16.3)	17,996 (16.15)	17,843 (16.27)	19,633 (17.05)	16,415 (16.13)	25,143 (15.77)
12th	134,615 (16.08)	18,523 (15.64)	19,438 (16.05)	18,313 (16.43)	18,541 (16.91)	19,318 (16.77)	17,551 (17.25)	22,931 (14.38)
Region of residence, n (%)								
Urban	541,928 (64.75)	71,150 (60.06)	75,778 (62.59)	74,454 (66.82)	72,549 (66.16)	76,221 (66.18)	67,554 (66.38)	104,222 (65.38)
Rural	295,044 (35.25)	47,307 (39.94)	45,301 (37.41)	36,978 (33.18)	37,105 (33.84)	38,948 (33.82)	34,217 (33.62)	55,188 (34.62)
Smoking status, n (%)								
No	798,876 (95.45)	102,201 (86.28)	111,177 (91.82)	108,595 (97.45)	105,948 (96.62)	113,177 (98.27)	99,929 (98.19)	157,849 (99.02)
Yes	38,096 (4.55)	16,256 (13.72)	9902 (8.18)	2837 (2.55)	3706 (3.38)	1992 (1.73)	1842 (1.81)	1561 (0.98)
Alcoholic consumption, n (%)								
No	693,100 (82.81)	90,006 (75.98)	94,937 (78.41)	91,157 (81.81)	91,744 (83.67)	97,895 (85)	86,087 (84.59)	141,274 (88.62)
Yes	143,872 (17.19)	28,451 (24.02)	26,142 (21.59)	20,275 (18.2)	17,910 (16.33)	17,274 (15)	15,684 (15.41)	18,136 (11.38)
School performance[c], n (%)								
Low	95,889 (11.46)	15,897 (13.42)	15,551 (12.84)	14,929 (13.4)	12,301 (11.22)	11,507 (9.99)	10,056 (9.88)	15,648 (9.82)

	Total	Prepandemic						Pandemic
		2008-2009	2010-2011	2012-2013	2014-2015	2016-2017	2018-2019	2020-2022
Middle-low	199,758 (23.87)	30,782 (25.99)	31,193 (25.76)	28,149 (25.26)	25,943 (23.66)	25,751 (22.36)	22,255 (21.87)	35,685 (22.39)
Middle	235,371 (28.12)	31,917 (26.94)	32,365 (26.73)	29,912 (26.84)	30,137 (27.48)	32,783 (28.47)	29,891 (29.37)	48,366 (30.34)
Middle-high	203,302 (24.29)	26,665 (22.51)	28,639 (23.65)	26,227 (23.54)	27,243 (24.84)	29,369 (25.5)	25,768 (25.32)	39,391 (24.71)
High	102,652 (12.26)	13,196 (11.14)	13,331 (11.01)	12,215 (10.96)	14,030 (12.79)	15,759 (13.68)	13,801 (13.56)	20,320 (12.75)
Sexual experience, n (%)								
No	751,024 (89.73)	93,933 (79.3)	97,202 (80.28)	98,844 (88.7)	104,074 (94.91)	109,715 (95.26)	96,079 (94.41)	151,177 (94.84)
Yes	85,948 (10.27)	24,524 (20.7)	23,877 (19.72)	12,588 (11.3)	5580 (5.09)	5454 (4.74)	5692 (5.59)	8233 (5.16)
Physical activity frequency[d], n (%)								
Lower activity	502,005 (59.98)	71,137 (60.05)	72,357 (59.76)	65,536 (58.81)	61,683 (56.25)	64,707 (56.18)	59,804 (58.76)	106,781 (66.99)
Moderate activity	207,881 (24.84)	31,427 (26.53)	32,204 (26.6)	30,804 (27.64)	30,540 (27.85)	31,667 (27.5)	25,261 (24.82)	25,978 (16.3)
Higher activity	127,086 (15.18)	15,893 (13.42)	16,518 (13.64)	15,092 (13.54)	17,431 (15.9)	18,795 (16.32)	16,706 (16.42)	26,651 (16.72)
Sadness and despair, n (%)								
No	585,417 (69.94)	72,615 (61.3)	77,976 (64.4)	76,928 (69.04)	81,804 (74.6)	86,103 (74.76)	73,455 (72.18)	116,536 (73.1)
Yes	251,555 (30.06)	45,842 (38.7)	43,103 (35.6)	34,504 (30.96)	27,850 (25.4)	29,066 (25.24)	28,316 (27.82)	42,874 (26.9)
Highest educational level of parents, n (%)								
High school or lower	424,689 (50.74)	49,612 (41.88)	56,851 (46.95)	55,905 (50.17)	60,585 (55.25)	67,515 (58.62)	52,916 (52)	81,305 (51)
College or higher	412,283 (49.26)	68,845 (58.12)	64,228 (53.05)	55,527 (49.83)	49,069 (44.75)	47,654 (41.38)	48,855 (48)	78,105 (49)
Family structure, n (%)								
Nuclear family	396,035 (47.32)	67,622 (57.09)	64,799 (53.52)	50,829 (45.61)	52,840 (48.19)	52,794 (45.84)	40,550 (39.84)	66,601 (41.78)
Living with relatives	14,831 (1.77)	3160 (2.67)	2862 (2.36)	2328 (2.09)	1843 (1.68)	1729 (1.50)	1269 (1.25)	1640 (1.03)
Living alone	6889 (0.82)	645 (0.54)	629 (0.52)	535 (0.48)	611 (0.56)	772 (0.67)	1322 (1.3)	2375 (1.49)
Orphanage	2786 (0.33)	398 (0.34)	431 (0.36)	372 (0.33)	417 (0.38)	379 (0.33)	327 (0.32)	462 (0.29)

[a] KYRBS: Korea Youth Risk Behavior Web-Based Survey.

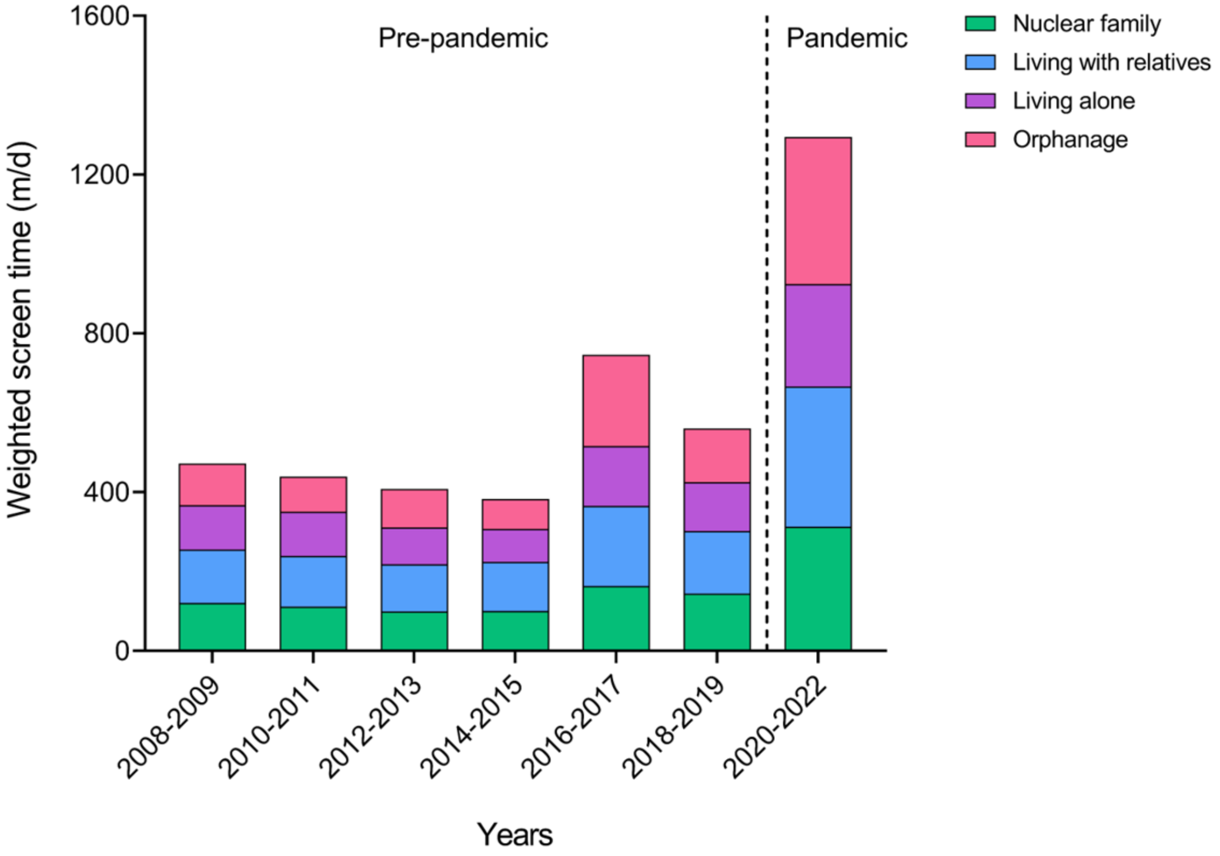
[b] BMI was divided into four groups according to the 2017 Korean National Growth Charts: underweight (<5%), normal (5% to 84%), overweight (85% to 94%), and obese (≥95%).

[c] School performance was divided into five groups: low (<20%), middle-low (20% to 39%), middle (40% to 59%), middle-high (60% to 79%), and high (≥80%).

[d] Physical activity frequency was divided into three groups based on the engagement in vigorous aerobic and resistance training more than three days per week: lower activity (neither activity is done for more than three days per week), moderate activity (either one activity), and higher activity (both activities).

The average weekly screen time and the trend change before and during the pandemic for each demographic over 15 years are exhibited in Multimedia Appendix 1 and Tables S1-S4 in Multimedia Appendix 2. Particularly, the trend in average screen time for each family structure is illustrated in Figure 1. Screen time showed a consistent increase from 2008 to 2022. Starting at 119.80 minutes per day (m/d; 95% CI 118.63-120.98) in 2008-2009, it temporarily decreased to 99.64 m/d (95% CI 98.45-100.83) in 2012-2013, then increased to 165.68 m/d (95% CI 163.27-168.08) in 2016-2017, and finally reached 306.80 m/d (95% CI 304.30-309.30) in 2020-2022, reflecting a sharp increase during the pandemic.

Figure 1. Trends in weighted screen time among adolescents in South Korea by family type, 2008-2022.



The pandemic had a significant impact on screen time, as detailed in Tables S1-S4 in Multimedia Appendix 2. The analysis showed a modest increase in screen time during the prepandemic period (β=8.06, 95% CI 7.74-8.39) and a substantial rise during the pandemic (β=162.06, 95% CI 159.49-164.64), with a significant difference in growth rate ($β_{diff}$=149.43, 95% CI 132.79-166.07). Among family structures, students in orphanages exhibited the largest increase ($β_{diff}$=221.90, 95% CI 159.62-284.17).

Table 2 highlights associations between screen time and biological, social, and familial factors. Women ($β_{diff}$=74.91, 95% CI 71.52-78.30) and higher BMI ($β_{diff}$=2.11, 95% CI 0.80-3.42) were identified as significant biological vulnerabilities. Among social factors, smoking ($β_{diff}$=79.09, 95% CI 63.13-95.06), alcohol consumption ($β_{diff}$=54.37, 95% CI 50.40-58.35), and sexual experience ($β_{diff}$=46.73, 95% CI 40.78-52.68) had the strongest associations with increased screen time. Academic achievement ($β_{diff}$=28.05, 95% CI 27.04-29.06]), rural residence ($β_{diff}$=–11.61, 95% CI –15.49 to –7.73), sadness and despair ($β_{diff}$ 10.71, 95% CI 8.34-13.07), higher physical activity ($β_{diff}$=5.35, 95% CI 3.95-6.75), and higher grade ($β_{diff}$=3.99, 95% CI 1.52-6.46) showed varying impacts (Figures 2 and 3).

Table 2. Factors associated with weighted average screen time among adolescents before and during the COVID-19 pandemic in KYRBS[a].

Factors	Unadjusted model			Adjusted model[b]		
	Prepandemic screen time, β (95% CI)	Pandemic screen time, β (95% CI)	β_{diff} (95% CI)	Prepandemic screen time, β (95% CI)	Pandemic screen time, β (95% CI)	β_{diff} (95% CI)
Biological factors						
Sex						
Men	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Women	–5.51 (–6.82 to –4.19)[c]	61.18 (57.75 to 64.62)[c]	66.69 (63.01 to 70.38)[c]	–9.12 (–10.33 to –7.90)[c]	65.79 (62.63 to 68.95)[c]	74.91 (71.52 to 78.30)[c]
BMI group[d]						
Underweight	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Obese	5.13 (4.65 to 5.61)[c]	2.61 (1.31 to 3.91)[c]	–2.52 (–3.90 to –1.13)[c]	3.51 (3.04 to 3.98)[c]	5.62 (4.40 to 6.85)[c]	2.11 (0.80 to 3.42)[c]
Social factors						
Grade						
7th	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
12th	–0.09 (–0.40 to 0.22)	9.43 (8.18 to 10.68)[c]	9.52 (8.24 to 10.81)[c]	–5.65 (–6.31 to –4.98)[c]	–1.66 (–4.04 to 0.72)	3.99 (1.52 to 6.46)[c]
Region of residence						
Rural	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Urban	–0.79 (–2.09 to 0.51)	–16.20 (–21.03 to –11.38)[c]	–15.41 (–20.41 to –10.41)[c]	1.08 (–0.13 to 2.29)	–10.53 (–14.22 to –6.84)[c]	–11.61 (–15.49 to –7.73)[c]
Smoking status						
No	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Yes	9.10 (7.27 to 10.93)[c]	174.49 (158.20 to 190.79)[c]	165.39 (148.99 to 181.79)[c]	6.46 (4.55 to 8.38)[c]	85.55 (69.70 to 101.41)[c]	79.09 (63.13 to 95.06)[c]
Alcoholic consumption						
No	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Yes	14.79 (13.75 to 15.83)[c]	95.80 (91.51 to 100.08)[c]	81.01 (76.60 to 85.42)[c]	13.89 (12.85 to 14.93)[c]	68.26 (64.42 to 72.10)[c]	54.37 (50.40 to 58.35)[c]
School performance[e]						
Low	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
High	10.15 (9.84 to 10.46)[c]	43.33 (42.33 to 44.33)[c]	33.18 (32.14 to 34.23)[c]	9.29 (8.97 to 9.60)[c]	37.34 (36.38 to 38.30)[c]	28.05 (27.04 to 29.06)[c]
Sexual experience						
No	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Yes	–0.31 (–1.55 to 0.94)	88.53 (82.21 to 94.86)[c]	88.84 (82.40 to 95.28)[c]	0.92 (–0.30 to 2.14)	47.65 (41.83 to 53.47)[c]	46.73 (40.78 to 52.68)[c]
Physical activity frequency[f]						

Factors	Unadjusted model			Adjusted model[b]		
	Prepandemic screen time, β (95% CI)	Pandemic screen time, β (95% CI)	β_{diff} (95% CI)	Prepandemic screen time, β (95% CI)	Pandemic screen time, β (95% CI)	β_{diff} (95% CI)
Lower activity	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Higher activity	–3.96 (–4.47 to –3.44)[c]	–11.69 (–13.17 to –10.22)[c]	–7.74 (–9.30 to –6.17)[c]	–7.37 (–7.84 to –6.91)[c]	–2.02 (–3.34 to –0.71)[c]	5.35 (3.95 to 6.75)[c]
Sadness and despair						
No	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Yes	6.91 (6.17 to 7.66)[c]	44.01 (41.61 to 46.42)[c]	37.10 (34.58 to 39.62)[c]	7.84 (7.11 to 8.57)[c]	18.55 (16.29 to 20.80)[c]	10.71 (8.34 to 13.07)[c]
Familial factor						
Highest education level of parents						
High school or lower	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
College or higher	–14.11 (–14.88 to –13.34)[c]	–41.55 (–43.79 to –39.30)[c]	–27.43 (–29.81 to –25.06)[c]	–11.22 (–11.97 to –10.46)[c]	–26.31 (–28.45 to –24.17)[c]	–15.10 (–17.36 to –12.83)[c]

[a] KYRBS: Korea Youth Risk Behavior Web-Based Survey.

[b] Adjustment for age, sex, BMI group, grade, region of residence, smoking status, alcoholic consumption, school performance, sexual experience, physical activity frequency, sadness and despair, and highest education level of parents.

[c] Indicate a significant difference (*P*<.05).

[d] BMI was divided into four groups according to the 2017 Korean National Growth Charts: underweight (<5%), normal (5% to 84%), overweight (85% to 94%), and obese (≥95%). BMI was calculated as weight in kilograms divided by height in meters squared.

[e] School performance was divided into five groups: low (<20%), middle-low (20% to 39%), middle (40% to 59%), middle-high (60% to 79%), and high (≥80%).

[f] Physical activity frequency was divided into three groups based on the engagement in vigorous aerobic and resistance training more than three days per week: lower activity (neither activity is done for more than three days per week), moderate activity (either one activity), and higher activity (both activities).

Figure 2. β differences of biological and familial factors in weighted average screen time among adolescents before and during the COVID-19 pandemic in KYRBS. KYRBS: Korea Youth Risk Behavior Web-Based Survey.

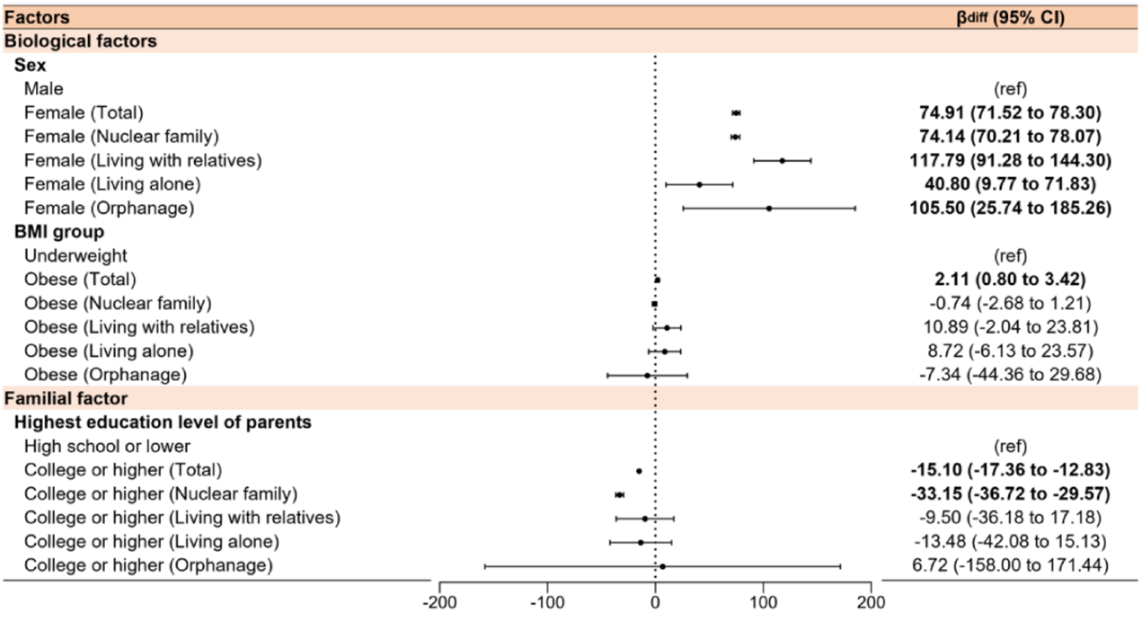
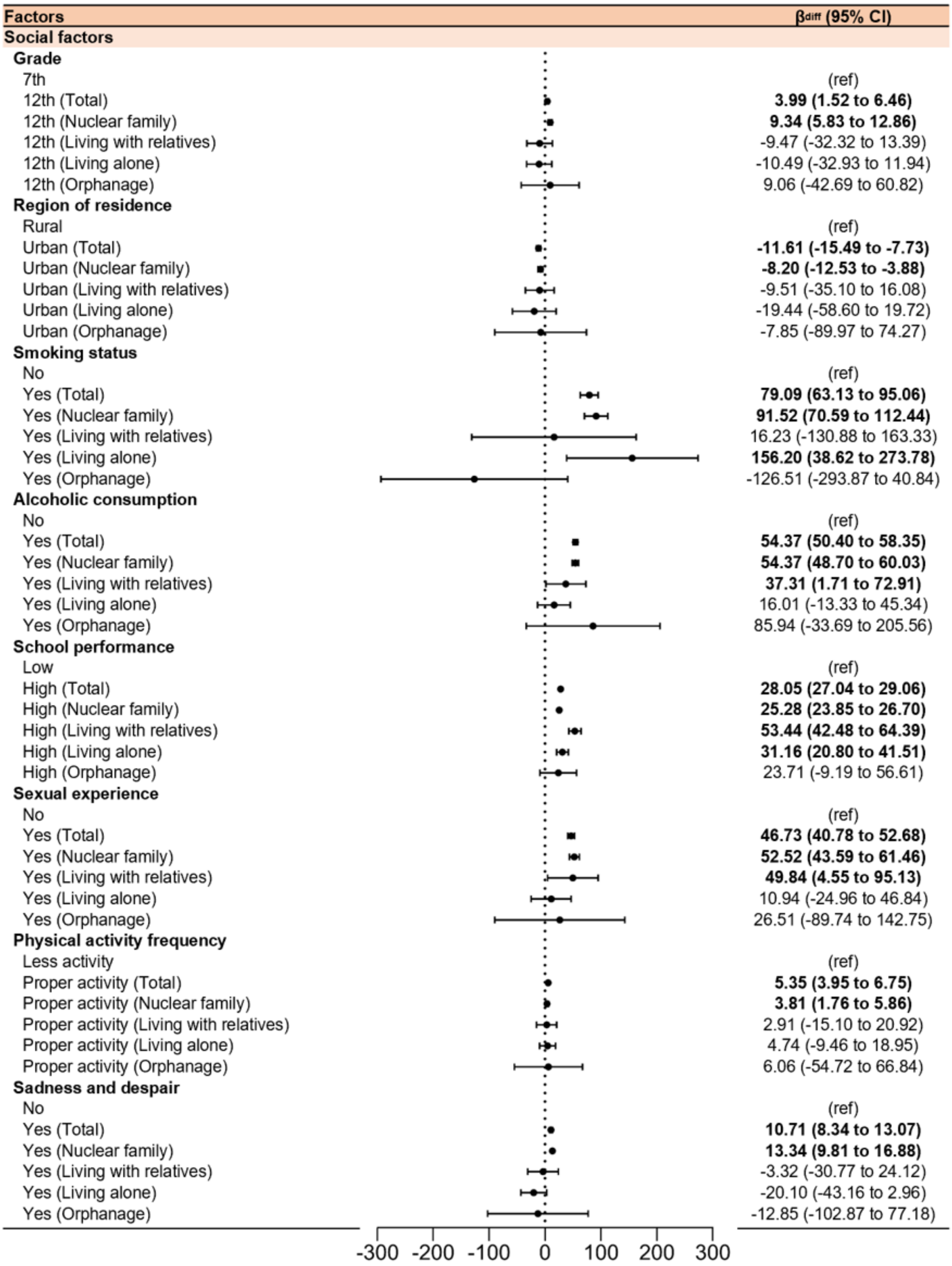


Figure 3. β differences of social factors in weighted average screen time among adolescents before and during the COVID-19 pandemic in KYRBS. KYRBS: Korea Youth Risk Behavior Web-Based Survey.



Factors	$β_{diff}$ (95% CI)
Social factors	
Grade	
7th	(ref)
12th (Total)	**3.99 (1.52 to 6.46)**
12th (Nuclear family)	**9.34 (5.83 to 12.86)**
12th (Living with relatives)	-9.47 (-32.32 to 13.39)
12th (Living alone)	-10.49 (-32.93 to 11.94)
12th (Orphanage)	9.06 (-42.69 to 60.82)
Region of residence	
Rural	(ref)
Urban (Total)	**-11.61 (-15.49 to -7.73)**
Urban (Nuclear family)	**-8.20 (-12.53 to -3.88)**
Urban (Living with relatives)	-9.51 (-35.10 to 16.08)
Urban (Living alone)	-19.44 (-58.60 to 19.72)
Urban (Orphanage)	-7.85 (-89.97 to 74.27)
Smoking status	
No	(ref)
Yes (Total)	**79.09 (63.13 to 95.06)**
Yes (Nuclear family)	**91.52 (70.59 to 112.44)**
Yes (Living with relatives)	16.23 (-130.88 to 163.33)
Yes (Living alone)	**156.20 (38.62 to 273.78)**
Yes (Orphanage)	-126.51 (-293.87 to 40.84)
Alcoholic consumption	
No	(ref)
Yes (Total)	**54.37 (50.40 to 58.35)**
Yes (Nuclear family)	**54.37 (48.70 to 60.03)**
Yes (Living with relatives)	**37.31 (1.71 to 72.91)**
Yes (Living alone)	16.01 (-13.33 to 45.34)
Yes (Orphanage)	85.94 (-33.69 to 205.56)
School performance	
Low	(ref)
High (Total)	**28.05 (27.04 to 29.06)**
High (Nuclear family)	**25.28 (23.85 to 26.70)**
High (Living with relatives)	**53.44 (42.48 to 64.39)**
High (Living alone)	**31.16 (20.80 to 41.51)**
High (Orphanage)	23.71 (-9.19 to 56.61)
Sexual experience	
No	(ref)
Yes (Total)	**46.73 (40.78 to 52.68)**
Yes (Nuclear family)	**52.52 (43.59 to 61.46)**
Yes (Living with relatives)	**49.84 (4.55 to 95.13)**
Yes (Living alone)	10.94 (-24.96 to 46.84)
Yes (Orphanage)	26.51 (-89.74 to 142.75)
Physical activity frequency	
Less activity	(ref)
Proper activity (Total)	**5.35 (3.95 to 6.75)**
Proper activity (Nuclear family)	**3.81 (1.76 to 5.86)**
Proper activity (Living with relatives)	2.91 (-15.10 to 20.92)
Proper activity (Living alone)	4.74 (-9.46 to 18.95)
Proper activity (Orphanage)	6.06 (-54.72 to 66.84)
Sadness and despair	
No	(ref)
Yes (Total)	**10.71 (8.34 to 13.07)**
Yes (Nuclear family)	**13.34 (9.81 to 16.88)**
Yes (Living with relatives)	-3.32 (-30.77 to 24.12)
Yes (Living alone)	-20.10 (-43.16 to 2.96)
Yes (Orphanage)	-12.85 (-102.87 to 77.18)

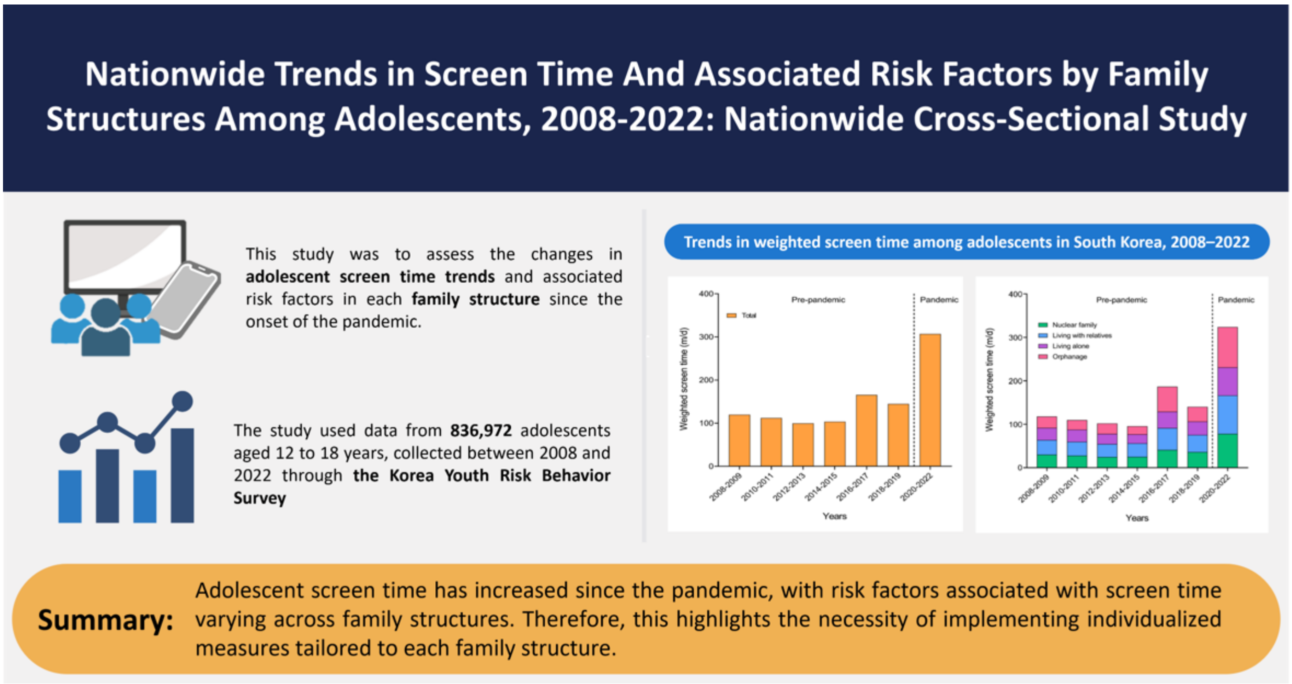
Adjusted models presented in Tables S5-S8 in Multimedia Appendix 2 explored associations across family structures. Before the pandemic, men in nuclear families (β=–10.50, 95% CI –11.92 to –9.08) and those living with relatives (β=–15.14, 95% CI –21.58 to –8.71) had lower screen time. In contrast, women living alone (β=14.80, 95% CI 3.57-26.03) showed higher screen time. During the pandemic, screen time significantly increased for women across all family structures: nuclear families ($β_{diff}$=74.14, 95% CI 70.21-78.07), living with relatives ($β_{diff}$ 117.79, 95% CI 91.28-144.30), living alone ($β_{diff}$ 40.80, 95% CI 9.77-71.83), and in orphanages ($β_{diff}$ 105.50, 95% CI 25.74-185.26; Figure 3).

Discussion

Key Findings

This study stands out as the initial comprehensive long-term trend analysis of adolescent screen time and associated factors across various family structures, encompassing the COVID-19 era (Figure 4). There was an overall increase in screen time from 2008 to 2022. Specifically, throughout the pandemic spanning from 2020 to 2022, there was a substantial increase in adolescent screen time compared to the prepandemic. Among various family structures, students within the orphanage group exhibited the most significant increase in screen time since the pandemic onset. All examined factors exhibited statistically significant alterations in their association with adolescent screen time during the pandemic. Particularly, stratified by family structure, it was consistently observed that women had more vulnerability to increased screen time during the pandemic regardless of family structure. Meanwhile, the nuclear family group showed distinctive screen time-associated factors, including grade, region of residence, physical activity frequency, sadness and despair, and highest education level of parents, which were not present in the other family groups.

Figure 4. Graphical conclusion.



Nationwide Trends in Screen Time And Associated Risk Factors by Family Structures Among Adolescents, 2008-2022: Nationwide Cross-Sectional Study

This study was to assess the changes in **adolescent screen time trends** and associated risk factors in each **family structure** since the onset of the pandemic.

The study used data from **836,972** adolescents aged 12 to 18 years, collected between 2008 and 2022 through **the Korea Youth Risk Behavior Survey**

Trends in weighted screen time among adolescents in South Korea, 2008–2022

Summary: Adolescent screen time has increased since the pandemic, with risk factors associated with screen time varying across family structures. Therefore, this highlights the necessity of implementing individualized measures tailored to each family structure.

Comparison With Previous Studies

Earlier investigations have predominantly focused on the changes in adolescent screen time trends during the COVID-19 pandemic. Studies conducted in countries such as Canada (n=29,027), Germany (n=1711), the United States (n=5412), China (n=1680), and South Korea (n=913,191) consistently reported significant increases in screen time following the onset of the pandemic [14,26-28]. Many of these studies also examined factors influencing screen time, including depression and mental health [10,11]. However, most were limited by relatively small sample sizes, restricted age ranges of participants, or insufficient analysis of trends across specific demographic subgroups, such as those defined by family structure [26-29].

Additionally, while existing studies from Norway (n=4509) and France (n=3720) have explored the association between adolescent screen time and family structure, they often lack diversity in the types of family structures considered and fail to account for the profound shifts introduced by the pandemic [15,16]. These limitations highlight the importance of considering a wider range of family structures and the influence of contextual factors like the pandemic when analyzing screen time trends [27,28].

In contrast, this study addresses these gaps by analyzing a significantly larger sample size (n=836,972) collected over 15 years, allowing for a more robust examination of long-term trends. Furthermore, the stratification of adolescents by diverse family structures—nuclear families, living with relatives, living alone, and orphanages—enables a unique comparative analysis of disparities in screen time trends across these groups. This approach not only provides a more comprehensive understanding of the changes in adolescent screen time but also sheds light on the distinct risk factors associated with different family structures during the pandemic.

Plausible Mechanism

The overall elevation in average screen time of adolescents spanning from 2008 to 2022 could be elucidated by the rise in smartphone penetration among adolescents. The prevalence of smartphones among adolescents in South Korea increased by 15.3%, rising from 79.3% in 2015 to 94.6% in 2022 [30]. This upward trend would have played a role in the elevation of screen time among Korean adolescents over the 15 years. Particularly, a substantial increase in screen time with the onset of the pandemic is likely due to the increased duration of indoor activities. During the pandemic, there was an increase in time spent indoors at home and a decrease in outdoor activities, including walking, cycling, and commuting [31]. Students were restricted from the absence of the school community covering a range of cultural, sports, social, and emotional experiences. Instead, there was an increase in internet use time for video watching and gaming with the onset of the pandemic, suggesting an influence on the rapid rise in screen time [32]. Fundamentally, the rise in screen time could be attributed to the shift in the pandemic lifestyles of students: decreased outdoor activities and various web-based engagements.

Across all family structures, a marked increase in screen time was evident following the pandemic. However, there was a specific increase in screen time among students belonging to the orphanage group. This suggests a potential influence of the presence of parents or relatives on screen time during the pandemic. The Royal College of Paediatrics and Child Health emphasized the significance of the roles of guardians in managing adolescent screen time. They recommended careful parental consideration when introducing technology to adolescents and setting individualized regulations for their screen time [16]. This detailed assistance, initially inadequate for the orphanage group, became further deficient during the pandemic due to the obstruction of support from teachers in schools. It is likely to have contributed to the substantial screen time adjustment failure observed in the orphanage group.

We analyzed the association between various biological, social, and familial factors and screen time according to the family structure. Among the diverse family structures, the nuclear family group exhibited several unique screen time-associated factors from the onset of the pandemic that were not observed in other family structures. In particular, the region of residence indicated that students residing in urban regions were less susceptible to increased screen time during the pandemic compared to their counterparts in rural regions. This finding is likely linked to parental abilities associated with the region of residence. Parents have been identified as having heightened concerns related to screen-related risks, such as cyberbullying and exposure to inappropriate or harmful content [33]. The perceived difference in screen time between regions is likely because parents residing in urban regions exhibited higher levels of parenting compared to those living in rural regions [34]. Particularly, the increase in indoor time during the pandemic induced an increase in the time spent with parents, indicating an intensified influence of parenting on adolescents. Therefore, it may be that students residing in urban regions were less vulnerable to increased screen time during the pandemic than students residing in rural regions.

Clinical and Policy Implications

This study offers a reminder of the recent rapid increase in screen time among adolescents. Considering the rise in screen time since the pandemic, it is imperative for clinical professionals and policymakers to promptly propose relevant regulation solutions. As a noticeable increase in screen time has been observed among adolescents belonging to the orphanage group following the pandemic, officials should implement more intensive management strategies for this particular group. Heightened management is also required for women adolescents, who exhibited vulnerability to increased screen time across all family structures during the pandemic. We additionally suggest the development of a customized screen time management strategy tailored to the specific family structures of each adolescent. Providing individualized solutions targeting the risk factors associated with each family structure can contribute to a more effective reduction in screen time. Specifically, focusing on the distinct risk factors (grade, region of residence, physical activity frequency, sadness and despair, and highest education level of parents) within the nuclear family group could prompt targeted management strategies for this particular cohort. Given the physical and mental health and socioemotional well-being concerns associated with excessive screen time [10,35,36], appropriate efforts from clinical, social, and political professionals are needed to normalize the rapidly increased screen time among adolescents since the pandemic. To address this, it is essential to expand international health expenditure for pandemic response and preparedness [37]. Thus, securing long-term and sustainable funding is crucial to address the rise in adolescents' digital device use and the associated mental health challenges [38].

In particular, investments in digital education programs, enhanced mental health services, and community support networks are vital to mitigate the physical and mental health impacts of increased screen time among adolescents during the pandemic. Additionally, policy and educational initiatives should focus on leveraging the already identified positive aspects of increased screen time during restricted social interactions, such as its potential for information exchange, education, and communication [39]. Pandemic response funds should be prioritized to address these issues and support adolescents in growing up in a healthy digital environment. Appropriate allocation of funds and policy support will play a critical role in addressing the challenges of increased adolescent screen time and alleviating health disparities exacerbated by the pandemic.

Limitations

This study has several limitations. First, variations in survey questions from 2008 to 2022 affected data comparability. Specifically, while smartphone use time was the focus in 2017 and from 2020 to 2022, other years emphasized internet use time, excluding total screen time such as television viewing. Nevertheless, to the best of our knowledge, this is the first study to analyze long-term data spanning 15 years, providing comprehensive insights into prepandemic and pandemic-era trends in adolescent screen time. Second, the self-response methodology in KYRBS may not fully reflect the characteristics of the actual population. However, KYRBS has been historically acknowledged as a credible dataset [40,41], and the weight-based analysis conducted with over 800,000 survey respondents further enhances the representativeness and validity of the screen time trends observed in this study. Third, this study is limited to Korean adolescents, restricting its representation of global trends. Despite this, the focus on a specific demographic allowed us to examine the nuanced relationships between screen time and family structures in the Korean context, providing insights relevant to similar societies. Finally, the cross-sectional design prevents establishing causal relationships between screen time and associated factors, highlighting the need for longitudinal and comprehensive studies. However, the study emphasizes the need for longitudinal research and offers a foundational understanding of screen time trends and influencing factors across diverse family structures.

Conclusions

This study examined adolescent screen time trends from 2008 to 2022 across various family structures, revealing a consistent increase over time, with the steepest rise observed during the COVID-19 pandemic. Screen time patterns varied significantly by family structure, with orphaned adolescents experiencing the most pronounced increase, while other groups, such as those in nuclear families, also showed distinct trends influenced by factors like parental education and region of residence. These findings underscore the importance of tailored interventions, including family-type–specific workshops, institutional programs, and digital literacy resources, to address the diverse needs of adolescents. Reducing regional disparities and promoting alternative activities are also critical for mitigating the negative impacts of excessive screen time. Future research on broader structural and cultural influences on digital engagement is necessary to develop comprehensive and sustainable strategies for managing adolescent screen time.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2024-00460379) and the Ministry of Science and ICT, South Korea, under the Information Technology Research Center) support program (IITP-2024-RS-2024-00438239) supervised by the Institute for Information and Communications Technology Planning and Evaluation. This research was supported by grants from the NRF funded by the Korean government (RS-2024-00465719). The funders had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Data Availability

The datasets generated or analyzed during this study are available from the Korea Disease Control and Prevention Agency repository [42].

Authors' Contributions

DKY had full access to all of the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. All authors approved the final version before submission. SK, HJ, YS, and MKS drafted the manuscript and conducted statistical analyses. KL, JP, HL, LS, ED, GF, LB, GFLS, MAT, MR, and DP reviewed and edited the manuscript. SW and DKY supervised the study and are the guarantors for this study. SK, HJ, YS, and MKS contributed equally as the first authors. SW and DKY contributed equally as corresponding authors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The trend in the average screen time (m/d) of adolescents before and during the COVID-19 pandemic, weighted mean (95% CI), in the Korea Youth Risk Behavior Web-Based Survey.

[DOCX File , 35 KB - publichealth_v11i1e57962_app1.docx]

Multimedia Appendix 2

Trends and factors associated with adolescent screen time before and during the COVID-19 pandemic by household type in the Korea Youth Risk Behavior Web-Based Survey.

[DOCX File , 100 KB - publichealth_v11i1e57962_app2.docx]

References

1. Cha Y, Jung W, Seo M, Rahmati M. The emerging pandemic recent: SARS-CoV-2. Life Cycle 2023;3:e2. [doi: 10.54724/lc.2023.e2]
2. Lee SW, Yuh WT, Yang JM, Cho Y, Yoo IK, Koh HY, et al. Nationwide results of COVID-19 contact tracing in South Korea: individual participant data from an epidemiological survey. JMIR Med Inform 2020;8(8):e20992 [FREE Full text] [doi: 10.2196/20992] [Medline: 32784189]
3. Maqsood A, Abbas J, Rehman G, Mubeen R. The paradigm shift for educational system continuance in the advent of COVID-19 pandemic: mental health challenges and reflections. Curr Res Behav Sci 2021;2:100011 [FREE Full text] [doi: 10.1016/j.crbeha.2020.100011] [Medline: 38620741]

4. Rahmat TE, Raza S, Zahid H, Abbas J, Sobri FAM, Sidiki SN. Nexus between integrating technology readiness 2.0 index and students' e-library services adoption amid the COVID-19 challenges: implications based on the theory of planned behavior. J Educ Health Promot 2022;11:50 [FREE Full text] [doi: 10.4103/jehp.jehp_508_21] [Medline: 35372596]
5. Trott M, Driscoll R, Irlado E, Pardhan S. Changes and correlates of screen time in adults and children during the COVID-19 pandemic: a systematic review and meta-analysis. EClinicalMedicine 2022;48:101452 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101452] [Medline: 35615691]
6. Abbas J, Aman J, Nurunnabi M, Bano S. The impact of social media on learning behavior for sustainable education: evidence of students from selected universities in Pakistan. Sustainability 2019;11(6):1683. [doi: 10.3390/su11061683]
7. Abbas J. The impact of coronavirus (SARS-CoV2) epidemic on individuals mental health: the protective measures of Pakistan in managing and sustaining transmissible disease. Psychiatr Danub 2020;32(3-4):472-477 [FREE Full text] [doi: 10.24869/psyd.2020.472] [Medline: 33370755]
8. Abbas J, Al-Sulaiti K, Lorente D, Shah S, Shahzad U. Economic Growth and Environmental Quality in a Post-Pandemic World. United Kingdom: Routledge; 2023.
9. Meng Q, Yan Z, Abbas J, Shankar A, Subramanian M. Human–computer interaction and digital literacy promote educational learning in pre-school children: mediating role of psychological resilience for kids’ mental well-being and school readiness. Int J Hum Comput Interact 2023;41(1):16-30. [doi: 10.1080/10447318.2023.2248432]
10. Boers E, Afzali MH, Newton N, Conrod P. Association of screen time and depression in adolescence. JAMA Pediatr 2019;173(9):853-859 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.1759] [Medline: 31305878]
11. Tandon PS, Zhou C, Johnson AM, Gonzalez ES, Kroshus E. Association of children's physical activity and screen time with mental health during the COVID-19 pandemic. JAMA Netw Open 2021;4(10):e2127892 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.27892] [Medline: 34596669]
12. Wise J. Screen time: two hour daily limit would improve children’s cognition, study finds. BMJ 2018;362:k4070. [doi: 10.1136/bmj.k4070]
13. Yang Q, Wang H, Wu H, Li W, Zhang Y, Yao Y, et al. Effect of online game policy on smartphone game play time, addiction, and emotion in rural adolescents of China. BMC Psychiatry 2023;23(1):814 [FREE Full text] [doi: 10.1186/s12888-023-05325-3] [Medline: 37940898]
14. Oh J, Kim M, Rhee SY, Rahmati M, Koyanagi A, Smith L, et al. National trends in the prevalence of screen time and its association with biopsychosocial risk factors among Korean adolescents, 2008-2021. J Adolesc Health 2024;74(3):504-513. [doi: 10.1016/j.jadohealth.2023.10.021] [Medline: 38085201]
15. Langøy A, Smith ORF, Wold B, Samdal O, Haug EM. Associations between family structure and young people's physical activity and screen time behaviors. BMC Public Health 2019;19(1):433 [FREE Full text] [doi: 10.1186/s12889-019-6740-2] [Medline: 31023280]
16. Ashton JJ, Beattie RM. Screen time in children and adolescents: is there evidence to guide parents and policy? Lancet Child Adolesc Health 2019;3(5):292-294. [doi: 10.1016/S2352-4642(19)30062-8] [Medline: 30853301]
17. Park S, Yon H, Ban CY, Shin H, Eum S, Lee SW, et al. National trends in alcohol and substance use among adolescents from 2005 to 2021: a Korean serial cross-sectional study of one million adolescents. World J Pediatr 2023;19(11):1071-1081. [doi: 10.1007/s12519-023-00715-9] [Medline: 36977821]
18. Lee JH, Son Y, Park J, Lee H, Choi Y, Lee M, et al. Comparison of national trends in physical activity among adolescents before and during the COVID-19 pandemic: a nationally representative serial study in South Korea. Heliyon 2024;10(21):e40004. [doi: 10.1016/j.heliyon.2024.e40004] [Medline: 39553655]
19. Oh J, Lee M, Lee H, Yang H, Park J, Rahmati M, et al. Hand and oral hygiene practices of South Korean adolescents before and during the COVID-19 pandemic. JAMA Netw Open 2023;6(12):e2349249 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.49249] [Medline: 38147331]
20. Park S, Son Y, Lee H, Lee H, Lee J, Kang J, et al. Sex-specific trends in the prevalence of osteoarthritis and rheumatoid arthritis from 2005 to 2021 in South Korea: nationwide cross-sectional study. JMIR Public Health Surveill 2024;10:e57359 [FREE Full text] [doi: 10.2196/57359] [Medline: 39486025]
21. Hong S, Son Y, Lee M, Lee JH, Park J, Lee H, et al. Association between sociodemographic factors and vaccine acceptance for influenza and SARS-CoV-2 in South Korea: nationwide cross-sectional study. JMIR Public Health Surveill 2024;10:e56989. [doi: 10.2196/56989] [Medline: 39629701]
22. Park S, Yim Y, Lee M, Lee H, Park J, Lee JH, et al. Longitudinal trends in depression, suicidal ideation, and suicide attempts by family structure in South Korean adolescents, 2009-2022: a nationally representative serial study. Asian J Psychiatr 2024;98:104122. [doi: 10.1016/j.ajp.2024.104122] [Medline: 38959550]
23. Yoo IK, Marshall DC, Cho JY, Yoo HW, Lee SW. N-Nitrosodimethylamine-contaminated ranitidine and risk of cancer in South Korea: a nationwide cohort study. Life Cycle 2021;1:e1. [doi: 10.54724/lc.2021.e1]
24. Yoo HW, Jin HY, Yon DK, Effenberger M, Shin YH, Kim SY, et al. Non-alcoholic fatty liver disease and COVID-19 susceptibility and outcomes: a Korean nationwide cohort. J Korean Med Sci 2021;36(41):e291 [FREE Full text] [doi: 10.3346/jkms.2021.36.e291] [Medline: 34697932]
25. Lee SW. Regression analysis for continuous independent variables in medical research: statistical standard and guideline of life cycle committee. Life Cycle 2022;2:e3. [doi: 10.54724/lc.2022.e3]

26. Madigan S, Eirich R, Pador P, McArthur BA, Neville RD. Assessment of changes in child and adolescent screen time during the COVID-19 pandemic: a systematic review and meta-analysis. JAMA Pediatr 2022;176(12):1188-1198 [FREE Full text] [doi: 10.1001/jamapediatrics.2022.4116] [Medline: 36342702]
27. Schmidt SCE, Anedda B, Burchartz A, Eichsteller A, Kolb S, Nigg C, et al. Physical activity and screen time of children and adolescents before and during the COVID-19 lockdown in Germany: a natural experiment. Sci Rep 2020;10(1):21780 [FREE Full text] [doi: 10.1038/s41598-020-78438-4] [Medline: 33311526]
28. Nagata JM, Cortez CA, Cattle CJ, Ganson KT, Iyer P, Bibbins-Domingo K, et al. Screen time use among US adolescents during the COVID-19 pandemic: findings from the Adolescent Brain Cognitive Development (ABCD) study. JAMA Pediatr 2022;176(1):94-96 [FREE Full text] [doi: 10.1001/jamapediatrics.2021.4334] [Medline: 34724543]
29. Xiao S, Yan Z, Zhao L. Physical activity, screen time, and mood disturbance among Chinese adolescents during COVID-19. J Psychosoc Nurs Ment Health Serv 2021;59(4):14-20. [doi: 10.3928/02793695-20201104-04] [Medline: 33180950]
30. Kwon S, Kim R, Lee JT, Kim J, Song S, Kim S, et al. Association of smartphone use with body image distortion and weight loss behaviors in Korean adolescents. JAMA Netw Open 2022;5(5):e2213237 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.13237] [Medline: 35594044]
31. Park S, Kim HJ, Kim S, Rhee SY, Woo HG, Lim H, et al. National trends in physical activity among adults in South Korea before and during the COVID-19 pandemic, 2009-2021. JAMA Netw Open 2023;6(6):e2316930 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.16930] [Medline: 37273204]
32. Major D, Fazekas-Pongor V, Pártos K, Tabák AG, Ungvari ZI, Eörsi D, et al. Effect of school lockdown due to the COVID-19 pandemic on screen time among adolescents in hungary: a longitudinal analysis. Front Public Health 2023;11:1233024 [FREE Full text] [doi: 10.3389/fpubh.2023.1233024] [Medline: 38098820]
33. Kimball HG, Fernandez F, Moskowitz KA, Kang M, Alexander LM, Conway KP, et al. Parent-perceived benefits and harms associated with internet use by adolescent offspring. JAMA Netw Open 2023;6(10):e2339851 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.39851] [Medline: 37883086]
34. Han J, Hao Y, Cui N, Wang Z, Lyu P, Yue L. Parenting and parenting resources among Chinese parents with children under three years of age: rural and urban differences. BMC Prim Care 2023;24(1):38 [FREE Full text] [doi: 10.1186/s12875-023-01993-y] [Medline: 36726105]
35. Mayor S. Added screen time linked to more depression in adolescents, study finds. BMJ 2019;366:l4691. [doi: 10.1136/bmj.l4691] [Medline: 31311792]
36. Stiglic N, Viner RM. Effects of screentime on the health and well-being of children and adolescents: a systematic review of reviews. BMJ Open 2019;9(1):e023191 [FREE Full text] [doi: 10.1136/bmjopen-2018-023191] [Medline: 30606703]
37. Global Burden of Disease 2021 Health Financing Collaborator Network. Global investments in pandemic preparedness and COVID-19: development assistance and domestic spending on health between 1990 and 2026. Lancet Global Health 2023;11(3):e385-e413 [FREE Full text] [doi: 10.1016/S2214-109X(23)00007-4] [Medline: 36706770]
38. Abbas J. Crisis management, transnational healthcare challenges and opportunities: the intersection of COVID-19 pandemic and global mental health. Res Globalization 2021;3:100037 [FREE Full text] [doi: 10.1016/j.resglo.2021.100037]
39. Yu S, Abbas J, Draghici A, Negulescu OH, Ain NU. Social media application as a new paradigm for business communication: the role of COVID-19 knowledge, social distancing, and preventive attitudes. Front Psychol 2022;13:903082 [FREE Full text] [doi: 10.3389/fpsyg.2022.903082] [Medline: 35664180]
40. Kim Y, Choi S, Chun C, Park S, Khang YH, Oh K. Data resource profile: the Korea youth risk behavior web-based survey (KYRBS). Int J Epidemiol 2016;45(4):1076-1076e. [doi: 10.1093/ije/dyw070] [Medline: 27380796]
41. Bae J, Joung H, Kim JY, Kwon KN, Kim YT, Park SW. Test-retest reliability of a questionnaire for the Korea youth risk behavior web-based survey. J Prev Med Public Health 2010;43(5):403-410 [FREE Full text] [doi: 10.3961/jpmph.2010.43.5.403] [Medline: 20959711]
42. Korea Disease Control and Prevention Agency (KDCA). URL: https://www.kdca.go.kr/yhs/ [accessed 2025-02-22]

Abbreviations

KYRBS: Korea Youth Risk Behavior Web-Based Survey

Edited by W Goedel; submitted 01.03.24; peer-reviewed by J Abbas, D Carvalho, YH Kim, S Kim; comments to author 12.01.25; revised version received 17.01.25; accepted 27.01.25; published 10.03.25.

Please cite as:
Kim S, Jo H, Son Y, Shin MK, Lee K, Park J, Lee H, Smith L, Dragioti E, Fond G, Boyer L, López Sánchez GF, Tully MA, Rahmati M, Pizzol D, Woo S, Yon DK
Nationwide Trends in Screen Time and Associated Risk Factors by Family Structures Among Adolescents, 2008-2022: Nationwide Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e57962
URL: https://publichealth.jmir.org/2025/1/e57962
doi:10.2196/57962
PMID:

Association of Frailty Index at 66 Years of Age with Health Care Costs and Utilization Over 10 Years in Korea: Retrospective Cohort Study

Jieun Jang[1], PhD; Anna Kim[2], PhD; Mingee Choi[3], PhD; Ellen P McCarthy[4,5], MPH, PhD; Brianne Olivieri-Mui[6], MPH, PhD; Chan Mi Park[4], MD, MPH; Jae-Hyun Kim[7], PhD; Jaeyong Shin[3,8], MD, MPH, PhD; Dae Hyun Kim[4,5], MD, MPH, SCD

[1]Department of Preventive Medicine, College of Medicine, Dongguk University, Gyeongju, Republic of Korea

[2]School of Economics, Yonsei University, Seoul, Republic of Korea

[3]Department of Preventive Medicine, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

[4]Frailty Research Center, Hinda and Arthur Marcus Institute for Aging Research, Hebrew SeniorLife, Harvard Medical School, Boston, MA, United States

[5]Division of Gerontology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States

[6]Department of Health Sciences, Bouvé College of Health Sciences, Northeastern University, Boston, MA, United States

[7]Department of Health Administration, College of Health Science, Dankook University, Cheonan, Republic of Korea

[8]Department of Policy Analysis and Management, College of Human Ecology, Cornell University, Ithaca, NY, United States

Corresponding Author:
Jaeyong Shin, MD, MPH, PhD
Department of Preventive Medicine, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

Abstract

Background: The long-term economic impact of frailty measured at the beginning of elderhood is unknown.

Objective: The objective of our study was to examine the association between an individual's frailty index at 66 years of age and their health care costs and utilization over 10 years.

Methods: This retrospective cohort study included 215,887 Koreans who participated in the National Screening Program for Transitional Ages at 66 years of age between 2007 - 2009. Frailty was categorized using a 39-item deficit accumulation frailty index: robust (<0.15), prefrail (0.15 - 0.24), and frail (≥0.25). The primary outcome was total health care cost, while the secondary outcomes were inpatient and outpatient health care costs, inpatient days, and number of outpatient visits. Generalized estimating equations with a gamma distribution and identity link function were used to investigate the association between the frailty index and health care costs and utilization until December 31, 2019.

Results: The study population included 53.3% (n=115,113) women, 32.9% (n=71,082) with prefrailty, and 9.7% (n=21,010) with frailty. The frailty level at 66 years of age was associated with higher cumulative total costs (robust to frail: $19,815 to $28.281; *P*<.001), inpatient costs (US $11,189 to US $16,627; *P*<.001), and outpatient costs (US $8,625 to US $11,654; *P*<.001) over the next 10 years. In the robust group, a one-year increase in age was associated with increased total health care costs (mean change per beneficiary per year: US $206.2; SE: $1.2; *P*<.001), inpatient costs (US $126.8; SE: $1.0; *P*<.001), and outpatient costs (US $74.4; SE: $0.4; *P*<.001). In the frail group, the increase in total health care costs was greater compared to the robust group (difference in mean cost per beneficiary per year: US $120.9; SE: $5.3; *P*<.001), inpatient costs (US $102.8; SE: $5.22; *P*<.001), and outpatient costs (US $15.6; SE: $1.5; *P*<.001). Similar results were observed for health care utilization (*P*<.001). Among the robust group, a one-year increase in age was associated with increased inpatient days (mean change per beneficiary per year: 0.9 d; *P*<.001) and outpatient visits (2.1 visits; *P*<.001). In the frail group, inpatient days increased annually compared to the robust group (difference in the mean inpatient days per beneficiary per year: 1.5 d; *P*<.001), while outpatient visits increased to a lesser extent (difference in the mean outpatient visits per beneficiary per year: −0.2 visits; *P*<.001).

Conclusions: Our study demonstrates the potential utility of assessing frailty at 66 years of age in identifying older adults who are more likely to incur high health care costs and utilize health care services over the subsequent 10 years. The long-term high health care costs and utilization associated with frailty and prefrailty warrants public health strategies to prevent and manage frailty in aging populations.

(JMIR Public Health Surveill 2025;11:e50026) doi:10.2196/50026

KEYWORDS

frailty index; health care costs; health care utilization; elderly; Korea; frailty; aging; utilization; older adults; sociodemographic; cost; prevention

Introduction

Health care systems face the challenge of managing increasing health care costs in aging populations [1,2]. South Korea experiences one of the fastest rates of population aging in the world, with 40% of its population expected to be aged over 65 years by 2050 [3]. The country is already experiencing a deficit in health insurance budget [4]. Proactive identification of individuals who are likely to incur high health care costs and utilization is critical for developing strategies to control health care expenditures in the aging society.

Frailty is defined by a decline in physiological reserves across multiple organ systems, leading to increased susceptibility to poor health outcomes following stressors [5]. Previous studies have found that frailty was associated with increased health care costs and utilization [6-10]. Moreover, individuals with frailty accounted for 43.9% of preventable health care expenditures [11]. One of the established models for measuring frailty is the deficit accumulation frailty model, expressed as a frailty index (FI), which is calculated by dividing the number of deficits present by the total number of deficits considered [12,13]. The FI can range from 0 to 1, with higher scores indicating greater frailty [13]. It can be derived from an existing database that contains information on standardized health assessments.

Since 2007, all Koreans who turn 66 years old have been invited to participate in a comprehensive health evaluation as part of the National Screening Program for Transitional Ages at government-approved clinics, hospitals, and public health facilities [14]. This examination assesses lifestyle, medical history, functional status, and laboratory tests, providing data to quantify a deficit accumulation FI on a national level. We have recently shown that higher FI scores at age 66 were associated with faster development of age-related chronic diseases over the subsequent 10 years among nearly 1 million Koreans [15]. The availability of standardized health assessments through the National Screening Program for Transitional Ages provides a unique opportunity to identify the FI across a large cohort. However, it is uncertain whether the FI at age 66 can predict long-term health care costs and utilization.

We conducted a nationwide cohort study of Koreans who participated in the National Screening Program for Transitional Ages in 2007 - 2009 to examine the association between the FI at age 66 and subsequent health care costs and utilization over 10 years. South Korea offers a unique context for research in frailty due to its nationwide single-payer health insurance system, which ensures both universal coverage, and provides detailed and standardized claims data for the entire population. This system facilitates long-term follow-up of health care costs and utilization, enabling population-level analyses that are rarely feasible in other health care settings. In this study, we hypothesized that a higher FI at age 66 would be associated with a greater increase in health care costs and utilization over 10 years.

Methods

Data Sources

We accessed the National Screening Program for Transitional Ages database (2007 - 2009), which was linked to the Korean National Health Insurance database (2004 - 2019), through the Korean National Health Insurance Corporation research program. The dataset included a 35% (n=435,572) random sample of adults who reached the age of 66 years between 2007-2009. The screening program database includes information on lifestyle indicators, medical history, functional status, and laboratory measurements. The Korean National Health Insurance database includes ICD-10 diagnostic codes, sociodemographic variables, health service claims, health care utilization, and long-term care insurance claims [16].

Study Population

We included individuals with complete sociodemographic information who participated in the screening program (n=222,480) through the 435,572 enrollees of the Korean National Health Insurance aged 66 years between 2007 - 2009. We excluded those with (1) duplicate records (n=1460), (2) insufficient data (<80% of the necessary items) for calculating frailty (n=5119), or (3) death within the month of examination or the month following the examination (n=14). Our final cohort included 215,887 enrollees (Figure S1 in Multimedia Appendix 1). This cohort included both community-dwelling older adults and long-term care residents.

Measurement of Frailty and Other Characteristics

The detailed procedure for constructing the FI has been described previously [15]. Briefly, we followed the standard procedure established by Searle et al [17]. The variables were selected as health deficits if they met the following criteria: (1) associated with health status, (2) prevalence increased with age, (3) did not saturate too early, and (4) covered a range of organ systems. We calculated the FI (range: 0 to 1; higher scores indicated greater frailty) based on 39 health-deficit items in the following health domains: medical history (15 items), biometric or laboratory measures (8 items), physical health (2 items), psychological health (8 items), and disability (6 items) (the definition of each item is provided in Table S1 in Multimedia Appendix 1) [15]. Frailty categories were defined using the previously used cut off points [15,18,19]: robust (<0.15), prefrail (0.15 to <0.25), and frail (≥0.25). We also assessed the presence of chronic conditions using ICD-10 diagnosis codes (ie, 1 inpatient or 2 outpatient diagnoses) from the previous year. Additionally, the following characteristics were obtained from the screening examination: sex, annual income level (quantiles), insurance status (employee insurance, self-employed insurance, or medical aid for low income), residential area (capital, metropolitan, or rural areas), and examination year (2007, 2008, or 2009).

Outcome Measurements and Follow Up

The primary outcome was total health care costs per beneficiary per year. These health care costs were the sum of reimbursements from the Korean National Health Insurance and beneficiaries' cost-sharing for inpatient and outpatient care. Secondary outcomes were inpatient costs, outpatient costs, inpatient days, and the number of outpatient visits per beneficiary per year. Follow up began on the day after the screening examination and lasted until the following, depending on whichever occurred earlier: date of death, 10 years from the screening examination, or December 31, 2019. To account for variations in inflation, we adjusted the costs to 2007 Korean Won (KRW) using an annual conversion factor (converted into US dollars at an exchange rate of 1 USD=1200 KRW). The extreme cost values were truncated at the 1st and 99th percentiles (ie, replacing values smaller than the 1st percentile with the 1st percentile value, and values larger than the 99th percentile with the 99th percentile).

Statistical Analysis

We used the χ^2 test to compare baseline characteristics by frailty category. The cumulative annual costs (total health care costs, inpatient costs, and outpatient costs) were compared using one-way ANOVA. Additionally, we compared the annual rates of health care utilization (inpatient days and number of outpatient visits) over 10 years, according to the frailty category. To investigate the association between the FI at age 66 years and health care costs and utilization over 10 years, generalized estimating equations were applied with a gamma distribution and the identity link function [20,21]. This was done to model right-skewed and over-dispersed health care cost and utilization data, accounting for repeated annual measures of costs and utilization within each beneficiary. Subgroup analyses were conducted by sex and by health insurance type. We tested the homogeneity of the interaction term between time and frailty categories across sex and health insurance type [22,23]. All models were adjusted for sex, annual income, insurance status, residential area, and examination years.

$$Yit=B0+B1*Prefrailit+B2*frailit+B3*Timeit+B4*Timeit*Prefrailit+B5*Timeit*Frailit+B6*Xi+eit$$

- Y_{iT}: Dependent variable for an individual participant i observed for time T

- $T=t–c$; where time T is defined as year t minus the calendar year in which an individual participant i underwent a medical examination year C^i (2007, 2008, or 2009)

- Time_t: Time in years [0 (67 years), 1 (68 years), 2 (69 years),…, 9 (76 years)]

- Prefrail_{iT}: Dummy variable (1 for prefrail group, 0 otherwise)

- X_i: Independent variables

The analyses were performed using SAS Enterprise Guide (version 7.15; SAS Institute,) and STATA (version 15; Stata Corporation). A two-sided *P*-value <.05 was considered statistically significant.

Ethical Considerations

This study was exempted from review by the Institutional Review Board of Ajou University Health System (AJIRB-MED-EXP-20 - 127) as the data used in this study were deidentified and secondary analysis was performed using public data. No compensation was provided to the participants.

Results

The study population included 123,795 (57.3%) women, 71,082 individuals with prefrailty (32.9%), and 21,010 individuals with frailty (9.7%) (Table 1). Individuals with greater frailty were more likely to be women (robust vs frail: 44.8% vs 68.6%) and medical aid recipients (robust vs frail: 1.8% vs 7.8%). Over the 10-years follow up, 20,189 (9.4%) individuals died. The survival rate to age 76 years was higher in the robust group than in the prefrail and frail groups (robust: 91.8%; pre-frail: 90.3%; frail: 85.2%) (Table S2 in Multimedia Appendix 1).

Figure 1 shows the cumulative health care costs over 10 years by frailty category in individuals aged 66. Individuals in the frail group incurred the highest cumulative total health care costs (US $28,281), followed by the prefrail (US $23,793), and robust groups (US $19,815; *P*<.001). Compared to the robust group, the frail group was associated with higher cumulative inpatient costs (US $11,189 vs US $16,627) and outpatient costs (US $8,625 vs US $11,654) for the subsequent 10 years.

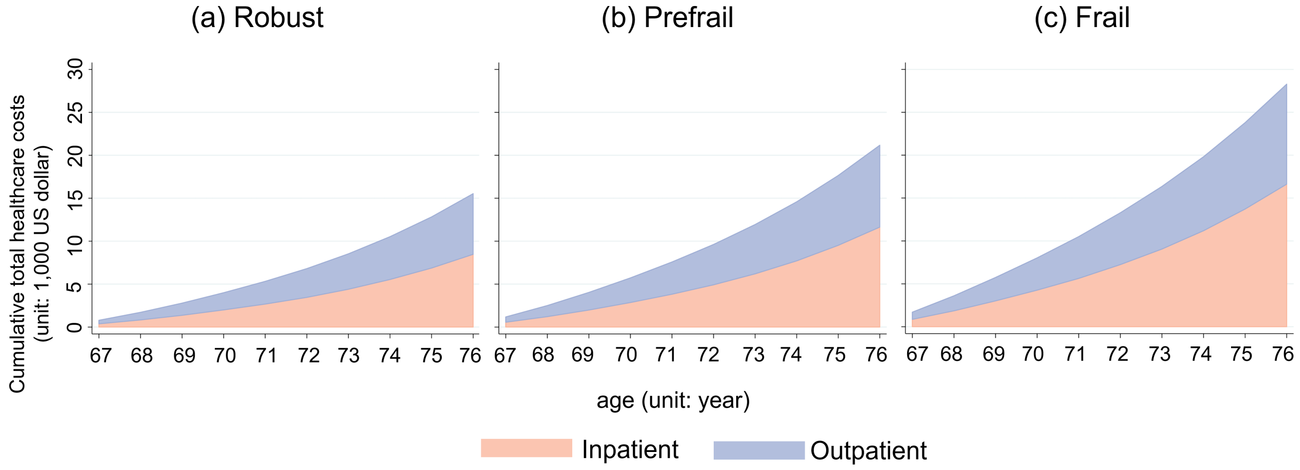
Table . Characteristics of Korean adults who participated in the National Screening Program for Transitional Ages at 66 years of age.

Characteristics	Total (N=215,887), n (%)	Frailty category[a], n (%)			*P* value[b]
		Robust (n=123,795)	Prefrail (n=71,082)	Frail (n=21,010)	
Gender					<.001
Men	100,774 (46.68)	68,390 (55.24)	25,780 (36.27)	6604 (31.43)	
Women	115,113 (53.32)	55,405 (44.76)	45,302 (63.73)	14,406 (68.57)	
Annual income (US $)					<.001
Quartile 1 (lowest)	46,306 (21.45)	26,944 (21.77)	14,525 (20.43)	4837 (23.02)	
Quartile 2	33,002 (15.29)	19,065 (15.40)	10,716 (15.08)	3221 (15.33)	
Quartile 3	55,973 (25.93)	32,026 (25.87)	18,624 (26.20)	5323 (25.34)	
Quartile 4 (highest)	80,606 (37.34)	45,760 (36.96)	27,217 (38.29)	7629 (36.31)	
Insurance status					<.001
Employee insurance	67,302 (31.17)	37,738 (30.48)	22,880 (32.19)	6684 (31.81)	
Self-employed insurance	141,955 (65.75)	83,769 (67.67)	45,492 (64.00)	12,694 (60.42)	
Medical aid for low income	6630 (3.07)	2288 (1.85)	2710 (3.81)	1632 (7.77)	
Residential area					<.001
Capital area	77,728 (36)	44,030 (35.57)	25,886 (36.42)	7812 (37.18)	
Metropolitan area	53,811 (24.93)	30,909 (24.97)	17,487 (24.60)	5415 (25.77)	
Rural area	84,348 (39.07)	48,856 (39.47)	27,709 (38.98)	7783 (37.04)	
Examination year					<.001
2007	53,907 (24.97)	29,875 (24.13)	18,297 (25.74)	5735 (27.30)	
2008	86,427 (40.03)	48,638 (39.29)	29,050 (40.87)	8739 (41.59)	
2009	75,553 (35.00)	45,282 (36.58)	23,735 (33.39)	6536 (31.11)	

[a] Frailty categories were defined as robust (frailty index <0.15, prefrail (0.15 to <0.25), and frail (≥0.25).

[b] P values were calculated using the χ^2 test for categorical variables.

Figure 1. Cumulative growth in total health care costs according to frailty index at age 66 years (A) Robust; (B) Prefrail; (C) Frail categories. The X-axis represents age (years) and the Y-axis represents cumulative total health care costs (US $). The pink area represents inpatient costs and blue area represents outpatient costs.



Over a period of 10 years, the annual growth in total health care costs, inpatient costs, and outpatient costs was greater in the frail group than in the prefrail and robust groups (Figure 2). Multivariable analyses (Table 2; Table S3 in Multimedia Appendix 1) showed that the frail group had higher mean total health care costs (difference: $827.2, SE: $20.0; *P*<.001), inpatient costs ($432.2, SE: $16.3; *P*<.001), and outpatient costs ($395.0, SE: $7.9; *P*<.001) than the robust group at baseline.

In the robust group, a one-year increase in age was associated with increased total health care costs (mean change per beneficiary per year: \$206.2; SE: \$1.2; *P*<.001), inpatient costs (\$126.8; SE: \$1.0; *P*<.001), and outpatient costs (\$74.4; SE: \$0.4; *P*<.001). In the frail group, there were greater increases in total health care costs each year compared to the robust group (difference in the mean cost per beneficiary per year: \$120.9; SE: \$5.3; *P*<.001), inpatient costs (\$102.8; SE: \$5.2; *P*<.001), and outpatient costs (\$15.6; SE: \$1.5; *P*<.001). The prefrail group had higher mean total health care costs, inpatient costs, and outpatient costs than the robust group at baseline, with the annual change per one-year increase in age falling between the robust and frail groups (Table 2).

Figure 2. Trends in annual health care costs per beneficiary per year according to frailty index at age 66 years. The X-axis represents age (years) and the Y-axis is annual health care costs per beneficiary per year (US dollar). The navy diamond line represents frail individuals, the blue triangle line represents pre-frail individuals, and the sky-blue square line represents robust individuals.

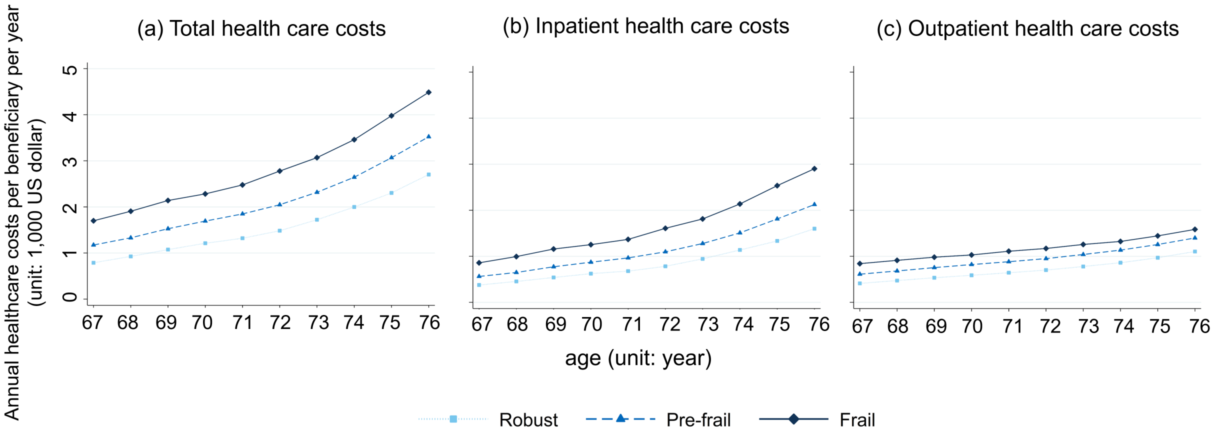


Table . Association between frailty at 66 years of age and annual health care costs and health care utilization over 10 years.[a] results of the generalized estimating equations models that examined the association between frailty at 66 years of age and annual health care costs and health care utilization over 10 years.

Characteristics	Total health care costs		Inpatient costs		Outpatient costs		Number of inpatient days		Number of outpatient visits	
	β (SE)[b]	*P* value[c]	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value
Frailty category[d]										
Robust	Ref.[e]		Ref.		Ref.		Ref.		Ref.	
Pre-frail	354 (8.4)	<.001	169.4 (6.4)	<.001	180.6 (3.5)	<.001	1.7 (0)	<.001	16.7 (0)	<.001
Frail	827.2 (20)	<.001	432.2 (16.3)	<.001	395 (7.9)	<.001	5 (0)	<.001	36.0 (0.1)	<.001
Year (Ref: robust)										
Per 1-year increase	206.2 (1.2)	<.001	126.8 (1)	<.001	74.4 (0.4)	<.001	0.9 (0)	<.001	2.1 (0)	<.001
Year * Frailty category										
Year * Prefrail	52 (2.4)	<.001	38.7 (2.1)	<.001	13.1 (0.7)	<.001	0.3 (0)	<.001	0.1 (0)	<.001
Year * Frail	120.9 (5.3)	<.001	102.8 (5.2)	<.001	15.6 (1.5)	<.001	0.9 (0)	<.001	-0.2 (0)	<.001

[a] Generalized estimation equation models were used for the analysis. The models were adjusted for the examination year, sex, annual income, insurance status, and residential area.

[b] β (SE) represents the regression co-efficient (β) and its standard error (SE).

[c] P values indicate the significance level of the comparisons.

[d] Frailty categories were defined as robust (frailty index<0.15), pre-frail (0.15 to <0.25), and frail (≥0.25).

[e] Ref. denotes the reference category used for comparisons.

The examination of health care utilization over 10 years showed that the frail group had a greater increase in inpatient days and the outpatient visits over 10 years than the prefrail or robust groups (Figure S2 and Figure S3 in Multimedia Appendix 1). The frail group had a greater mean number of inpatient days (5 d; SE: 0 d; *P*<.001) and outpatient visits (36 visits; SE: 0.1; *P*<.001) than the robust group at baseline (Table 2). In the robust group, a one-year increase in age was associated with increases in inpatient days (mean change per beneficiary per year: 0.9 d; SE: 0 d; *P*<.001) and outpatient visits (mean change per beneficiary per year: 2.1 visits; SE: 0 visits; *P*<.001). In the frail group, the number of inpatient days increased more each year compared to the robust group (difference in the mean inpatient days per beneficiary per year: 0.9 d; SE: 0 d; *P*<.001), whereas the number of outpatient visits increased to a lesser degree (difference in the mean outpatient visits per beneficiary per year: −0.2 visits; SE: 0 visits; *P*<.001). The prefrail group had a greater mean number of inpatient days and outpatient visits than the robust group at baseline, with more increases in both inpatient days and outpatient visits annually (Table 2).

In subgroup analyses, men with frailty had a greater increase in total health care costs than women with frailty (*P*-for-interaction:<.001) (Table S4 in Multimedia Appendix 1). Frail people with low-income medical aid were more likely to incur higher total health care costs than those with other insurance types and frailty groups; however, the observed difference was not significant (*P*-for-interaction: .74) (Table S5 in Multimedia Appendix 1).

Discussion

Principal Findings

Using a nationwide Korean cohort, we found that the frailty level at 66 years of age was associated with higher health care costs and health care utilization over the subsequent 10 years. The growth in annual health care costs and health care utilization, particularly inpatient days, was greater in individuals with frailty than in prefrail or robust individuals. Furthermore, prefrail individuals, who made up 32.9% of the population, had higher health care costs and utilization than robust individuals. Given the high prevalence of prefrailty and associated long-term health care costs and utilization, our findings suggest the importance of identifying frailty and prefrailty to control health care costs and utilization in aging populations.

Our findings are consistent with previous studies demonstrating that frailty is associated with increased use of health care resources [24-26]. A population-based cohort study showed that the association between frailty onset and increasing self-reported health care costs was prominent in inpatient care and informal nursing care [6]. A meta-analysis of 7 cohorts of community-dwelling older adults found that health care costs of prefrail and frail older adults were higher than robust individuals [27]. In addition, those who were frail faced a greater risk of hospitalization, skilled nursing facility stays [28,29], emergency department visits [30], and institutionalization [31], compared to robust individuals. Our study contributes to the existing literature by examining 10-year trajectories of health care costs and utilization in a nationally representative cohort of Koreans aged 66 years. The choice of this time point in the beginning of elderhood removes the effect of chronologic age and emphasizes the importance of early identification, prevention, and management of frailty and prefrailty. Early identification can provide the opportunity to proactively address the needs of these individuals to lower health care costs and utilization in the future [32,33]. Frailty may be prevented or delayed by interventions such as physical activity, nutrition, and comprehensive geriatric assessment [34]. Therefore, by linking frailty with long-term health care costs and utilization, our findings expand the understanding of frailty's economic impact beyond short-term or disease-specific analyses commonly found in the literature.

Previous research suggested that incorporating frailty into a diagnosis-based model, such as the hierarchical condition category method (which is used to predict Medicare health care costs), may improve the accuracy of cost projections [35]. However, most prediction models for health care costs used in Korea do not account for frailty [36]. Although health care cost prediction models have been developed using the National Screening Program for Transitional Ages database, they have been designed for the entire population rather than specifically for older adults [37]. Our findings suggest that frailty metrics could enhance existing health care cost prediction models, providing a more nuanced understanding of health care needs in aging populations.

In our study, the associations between frailty and health care costs and utilization were stronger in men than women. Li et al [38] and colleagues found that frailty or worsening frailty had a stronger association with increased hospitalization and outpatient costs in men than in women. However, other studies did not find an interaction between frailty and sex on catastrophic health expenditures [39] or health care utilization (outpatient visit, inpatient visit, and inpatient length of stay) [40]. There are several explanations for higher health care costs among frail men than among frail women in our study. Men are generally more likely to develop serious health conditions that can be costly to manage, such as coronary heart disease, cancer, cerebrovascular disease, emphysema, cirrhosis of the liver, kidney disease, and atherosclerosis [41]. In addition, men are more likely to engage in risky health behaviors such as smoking and drinking and may be less likely to seek health care services for health issues, leading to delayed treatment [42]. These findings suggest that taking into account the differences in health risks and needs between frail men and frail women may help to improve health outcomes and reduce costs in older adults.

Limitations

Our study had important limitations. First, we were unable to assess costs and utilization of noncovered services including the costs of outpatient prescription medications. Second, selection bias is possible due to nonparticipation of otherwise eligible individuals in the National Screening Program for Transitional Ages and deaths, which may affect the longitudinal cost trajectories. However, we previously reported no major differences in characteristics between nonparticipants and participants in the program [15]. Third, the association between frailty and health care costs and utilization may be subject to the choice of frailty definitions. Both the deficit accumulation FI and frailty phenotype have been associated with increased health care costs in previous studies [8,43]. Fourth, causality may not be inferred from our observational data. Finally, our findings may not be generalizable to other countries with different health care systems and financing structures.

Conclusion

Our study demonstrates the potential utility of assessing frailty at 66 years of age to identify older adults who are more likely to incur high health care costs and utilize health care services in the subsequent 10 years. The long-term high health care costs and utilization associated with frailty and prefrailty call for public health approaches to prevent and manage frailty in aging populations.

Acknowledgments

This study used customized data provided by the Korean National Health Insurance Service and was supported by faculty research grant 6-2021-0112 from Yonsei University College of Medicine (JS), grant KSPM-2021-01 from the Korean Society for Preventive Medicine and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute funded by grant HI21C1249 from the Ministry of Health & Welfare, Korea (JJ), and by Mid-Career Investigator Award K24AG073527 from the National Institute on Aging (DHK). The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

An earlier version of this study was presented at the Gerontological Society of America Annual Scientific Meeting on November 6, 2022 in Indianapolis, Indiana, and a poster was presented at the American Geriatrics Society Annual Scientific Meeting on May 4, 2022 in Long Beach, California.

Data Availability

The datasets generated and analyzed during this study are not publicly available. We obtained permission to use the datasets for the present study.

Authors' Contributions

Conceptualization: JJ, JS, DHK
Data curation: JJ, JS, DHK, AK
Formal analysis: JJ, DHK, JS, AK
Project administration: JJ, JS, DHK
Resources: JJ, JS, DHK
Supervision: JS, DHK
Writing – original draft: JJ, JS, DHK
Writing – review and editing: JJ, JS, DHK, AK

Conflicts of Interest

JJ reported receiving grant funding from the Korea Health Industry Development Institute during the conduct of the study. JS reported receiving grant funding from Yonsei University College of Medicine and the Korean Society for Preventive Medicine during the conduct of the study and being the chief executive officer and founder of Evertri, a company unrelated to the work that was submitted. DHK reported receiving grant funding from the National Institutes of Health during the conduct of the study and personal fees from Alosa Health and Village MD outside the submitted work. Generative AI was not used in any portion of the manuscript writing. No other disclosures were reported.

Multimedia Appendix 1
Analysis of Frailty index, healthcare utilization trends, and cost variations over time.
[DOCX File, 229 KB - publichealth_v11i1e50026_app1.docx]

References

1. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 8;388(10053):1545-1602. [doi: 10.1016/S0140-6736(16)31678-6] [Medline: 27733282]
2. Osborn R, Doty MM, Moulds D, Sarnak DO, Shah A. Older Americans were sicker and faced more financial barriers to health care than counterparts in other countries. Health Aff (Millwood) 2017 Dec;36(12):2123-2132. [doi: 10.1377/hlthaff.2017.1048]

3. World social report 2023: Leaving no one behind In an ageing world. United Nations. 2023. URL: https://www.un.org/development/desa/dspd/world-social-report/2023-2.html
4. National health insurance statistical yearbook 2021. Health Insurance Review and Assessment service. 2023.
5. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
6. Hajek A, Bock JO, Saum KU, et al. Frailty and healthcare costs-longitudinal results of a prospective cohort study. Age Ageing 2018 Mar 1;47(2):233-241. [doi: 10.1093/ageing/afx157] [Medline: 29036424]
7. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019 Oct 12;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6] [Medline: 31609228]
8. Sirven N, Rapp T. The cost of frailty in France. Eur J Health Econ 2017 Mar;18(2):243-253. [doi: 10.1007/s10198-016-0772-7] [Medline: 26914932]
9. Bock JO, König HH, Brenner H, et al. Associations of frailty with health care costs--results of the ESTHER cohort study. BMC Health Serv Res 2016 Apr 14;16(1):128. [doi: 10.1186/s12913-016-1360-3] [Medline: 27074800]
10. Shi SM, Olivieri-Mui B, McCarthy EP, Kim DH. Changes in a frailty index and association with mortality. J Am Geriatr Soc 2021 Apr;69(4):1057-1062. [doi: 10.1111/jgs.17002] [Medline: 33377190]
11. Figueroa JF, Joynt Maddox KE, Beaulieu N, Wild RC, Jha AK. Concentration of potentially preventable spending among high-cost medicare subpopulations: an observational study. Ann Intern Med 2017 Nov 21;167(10):706-713. [doi: 10.7326/M17-0767] [Medline: 29049488]
12. Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. ScientificWorldJournal 2001 Aug 8;1:323-336. [doi: 10.1100/tsw.2001.58] [Medline: 12806071]
13. Kaskirbayeva D, West R, Jaafari H, et al. Progression of frailty as measured by a cumulative deficit index: a systematic review. Ageing Res Rev 2023 Feb;84:101789. [doi: 10.1016/j.arr.2022.101789] [Medline: 36396032]
14. Kim HS, Shin DW, Lee WC, Kim YT, Cho B. National screening program for transitional ages in Korea: a new screening for strengthening primary prevention and follow-up care. J Korean Med Sci 2012 May;27 Suppl(Suppl):S70-S75. [doi: 10.3346/jkms.2012.27.S.S70] [Medline: 22661875]
15. Jang J, Jung H, Shin J, Kim DH. Assessment of frailty index at 66 years of age and association with age-related diseases, disability, and death over 10 years in Korea. JAMA Netw Open 2023 Mar 1;6(3):e2248995. [doi: 10.1001/jamanetworkopen.2022.48995] [Medline: 36862415]
16. Lee J, Lee JS, Park SH, Shin SA, Kim K. Cohort profile: The National Health Insurance Service-National Sample Cohort (NHIS-NSC), South Korea. Int J Epidemiol 2017 Apr 1;46(2):e15-e15. [doi: 10.1093/ije/dyv319] [Medline: 26822938]
17. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC Geriatr 2008 Sep 30;8(1):1-10. [doi: 10.1186/1471-2318-8-24] [Medline: 18826625]
18. Rockwood K, Mitnitski A. Frailty in relation to the accumulation of deficits. J Gerontol A Biol Sci Med Sci 2007 Jul;62(7):722-727. [doi: 10.1093/gerona/62.7.722] [Medline: 17634318]
19. Shi SM, McCarthy EP, Mitchell S, Kim DH. Changes in predictive performance of a frailty index with availability of clinical domains. J Am Geriatr Soc 2020 Aug;68(8):1771-1777. [doi: 10.1111/jgs.16436] [Medline: 32274807]
20. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. Biometrics 1988 Dec;44(4):1049-1060. [Medline: 3233245]
21. Hanley JA, Negassa A, Edwardes MDD, Forrester JE. Statistical analysis of correlated data using generalized estimating equations: an orientation. Am J Epidemiol 2003 Feb 15;157(4):364-375. [doi: 10.1093/aje/kwf215] [Medline: 12578807]
22. Fleiss JL. The statistical basis of meta-analysis. Stat Methods Med Res 1993;2(2):121-145. [doi: 10.1177/096228029300200202] [Medline: 8261254]
23. Rothman K. Spreadsheets for the analysis of epidemiologic data. RTI Health Solutions. 2002. URL: https://www.drugepi.org/dope/software#Episheet
24. Petermann-Rocha F, Lyall DM, Gray SR, et al. Associations between physical frailty and dementia incidence: a prospective study from UK Biobank. Lancet Healthy Longev 2020 Nov;1(2):e58-e68. [doi: 10.1016/S2666-7568(20)30007-6] [Medline: 36094146]
25. Veronese N, Cereda E, Stubbs B, et al. Risk of cardiovascular disease morbidity and mortality in frail and pre-frail older adults: results from a meta-analysis and exploratory meta-regression analysis. Ageing Res Rev 2017 May;35:63-73. [doi: 10.1016/j.arr.2017.01.003] [Medline: 28143778]
26. Cheng MH, Chang SF. Frailty as a risk factor for falls among community dwelling people: evidence from a meta-analysis. J Nurs Scholarsh 2017 Sep;49(5):529-536. [doi: 10.1111/jnu.12322] [Medline: 28755453]
27. Chi J, Chen F, Zhang J, et al. Impacts of frailty on health care costs among community-dwelling older adults: A meta-analysis of cohort studies. Arch Gerontol Geriatr 2021;94:104344. [doi: 10.1016/j.archger.2021.104344] [Medline: 33516075]
28. Ensrud KE, Kats AM, Schousboe JT, et al. Frailty phenotype and healthcare costs and utilization in older women. J Am Geriatr Soc 2018 Jul;66(7):1276-1283. [doi: 10.1111/jgs.15381] [Medline: 29684237]
29. Ensrud KE, Kats AM, Schousboe JT, et al. Frailty phenotype and healthcare costs and utilization in older men. J Am Geriatr Soc 2020 Sep;68(9):2034-2042. [doi: 10.1111/jgs.16522]

30. Kojima G. Frailty as a predictor of emergency department utilization among community-dwelling older people: a systematic review and meta-analysis. J Am Med Dir Assoc 2019 Jan;20(1):103-105. [doi: 10.1016/j.jamda.2018.10.004] [Medline: 30470576]
31. Kojima G. Frailty as a predictor of nursing home placement among community-dwelling older adults: a systematic review and meta-analysis. J Geriatr Phys Ther 2018;41(1):42-48. [doi: 10.1519/JPT.0000000000000097] [Medline: 27341327]
32. Holroyd-Leduc J, Resin J, Ashley L, et al. Giving voice to older adults living with frailty and their family caregivers: engagement of older adults living with frailty in research, health care decision making, and in health policy. Res Involv Engagem 2016;2(1):23. [doi: 10.1186/s40900-016-0038-7] [Medline: 29062523]
33. Abbasi M, Rolfson D, Khera AS, Dabravolskaj J, Dent E, Xia L. Identification and management of frailty in the primary care setting. CMAJ 2018 Sep 24;190(38):E1134-E1140. [doi: 10.1503/cmaj.171509] [Medline: 30249759]
34. Walston J, Buta B, Xue QL. Frailty screening and interventions: considerations for clinical practice. Clin Geriatr Med 2018 Feb;34(1):25-38. [doi: 10.1016/j.cger.2017.09.004] [Medline: 29129215]
35. Johnston KJ, Wen H, Joynt Maddox KE. Relationship of a claims-based frailty index to annualized medicare costs: a cohort study. Ann Intern Med 2020 Apr 21;172(8):533-540. [doi: 10.7326/M19-3261] [Medline: 32252070]
36. Lee S, Moon Y. A study of the financial projection of health expenditures of the aged of national health insurance-focused on the healthy ageing of EU. J Crit Soc Pol 2018 Feb 28;58:53-93. [doi: 10.47042/ACSW.2018.02.58.53]
37. Choi Y, An J, Ryu S, Kim J. Development and evaluation of machine learning-based high-cost prediction model using health check-up data by the national health insurance service of Korea. IJERPH 2022;19(20):13672. [doi: 10.3390/ijerph192013672]
38. Li CY, Al Snih S, Chou LN, et al. Frailty transitions predict healthcare use and Medicare payments in older Mexican Americans: a longitudinal cohort study. BMC Geriatr 2020 Jun 1;20(1):189. [doi: 10.1186/s12877-020-01583-y] [Medline: 32487037]
39. Fan L, Hou XY, Liu Y, Chen S, Wang Q, Du W. Catastrophic health expenditure associated with frailty in community-dwelling Chinese older adults: a prospective cohort analysis. Front Public Health 2021;9:718910. [doi: 10.3389/fpubh.2021.718910] [Medline: 34568260]
40. Fan L, Tian Y, Wang J, et al. Frailty predicts increased health care utilization among community-dwelling older adults: a longitudinal study in China. J Am Med Dir Assoc 2021 Sep;22(9):1819-1824. [doi: 10.1016/j.jamda.2021.01.082] [Medline: 33662331]
41. Case A, Paxson C. Sex differences in morbidity and mortality. Demography 2005 May;42(2):189-214. [doi: 10.1353/dem.2005.0011] [Medline: 15986983]
42. Harris CR, Jenkins M. Gender differences in risk assessment: Why do women take fewer risks than men? Judgm decis mak 2006 Jul;1(1):48-63. [doi: 10.1017/S1930297500000346]
43. Comans TA, Peel NM, Hubbard RE, Mulligan AD, Gray LC, Scuffham PA. The increase in healthcare costs associated with frailty in older people discharged to a post-acute transition care program. Age Ageing 2016 Mar;45(2):317-320. [doi: 10.1093/ageing/afv196] [Medline: 26769469]

Abbreviations

KRW: Korean Won

Edited by A Mavragani; submitted 16.06.23; peer-reviewed by LK Chen, P Heyn; revised version received 29.11.24; accepted 29.11.24; published 27.01.25.

Please cite as:
Jang J, Kim A, Choi M, McCarthy EP, Olivieri-Mui B, Park CM, Kim JH, Shin J, Kim DH
Association of Frailty Index at 66 Years of Age with Health Care Costs and Utilization Over 10 Years in Korea: Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e50026
URL: https://publichealth.jmir.org/2025/1/e50026
doi:10.2196/50026

Impact of the National Essential Public Health Service Package on Blood Pressure Control in Chinese People With Hypertension: Retrospective Population-Based Longitudinal Study

Feiran Wei[1*], MD; You Ge[1,2*], MD; Han Li[2], MPH; Yuan Liu[3], MSc

[1]School of Public Health, Southeast University, Nanjing, China
[2]Department of Infectious Disease, The Second Hospital of Nanjing, Nanjing, China
[3]Institute for Chronic Disease Management, Jining No. 1 People’s Hospital, 0802 Huoju, Jining, China
[*]these authors contributed equally

Corresponding Author:
Yuan Liu, MSc
Institute for Chronic Disease Management, Jining No. 1 People’s Hospital, 0802 Huoju, Jining, China

Abstract

Background: The National Essential Public Health Service Package (NEPHSP) was launched in 2009 to tackle poor blood pressure control in Chinese people with hypertension; however, it’s effect is still unclear.

Objective: In a retrospective population-based longitudinal study, we aimed to evaluate effect of the NEPHSP on blood pressure control.

Methods: A total of 516,777 patients registered in the NEPHSP were included. The blood pressure control data were assessed based on the Residence Health Record System dataset. We longitudinally evaluated the effects of the NEPHSP on blood pressure control by analyzing changes in blood pressure at quarterly follow-ups. Both the degree and trend of the blood pressure changes were analyzed. We conducted stratified analysis to explore effects of the NEPHSP on blood pressure control among subgroups of participants with specific characteristics.

Results: The mean baseline systolic blood pressure (SBP) and diastolic blood pressure (DBP) were 147.12 (SD 19.88) mm Hg and 85.11 (SD 11.79) mm Hg, respectively. The control rates of baseline SBP and DBP were 39.79% (205,630/516,777) and 69.21% (357,685/516,777). Compared to baseline, the mean SBP decreased in each quarter by 5.06 mm Hg (95% CI −5.11 to −5.00; *P*<.001), 6.69 mm Hg (95% CI; −6.74 to −6.63; *P*<.001), 10.30 mm Hg (95% CI −10.34 to −10.23; *P*<.001), and 6.63 mm Hg (95% CI −6.68 to −6.57; *P*<.001). The SBP control rates increased in each quarter to 53.12% (274,493/516,777; β coefficient=0.60, 95% CI 0.59-0.61; *P*<.001), 56.61% (292,537/516,777; β coefficient=0.76, 95% CI 0.75-0.77; *P*<.001), 63.4% (327,648/516,777; β coefficient=1.08, 95% CI 1.07-1.09; *P*<.001), and 55.09% (284,711/516,777; β coefficient=0.69, 95% CI 0.68-0.70; *P*<.001). Compared to baseline, the mean DBP decreased in each quarter by 1.75 mm Hg (95% CI −1.79 to −1.72; *P*<.001), 2.64 mm Hg (95% CI −2.68 to −2.61; *P*<.001), 4.20 mm Hg (95% CI −4.23 to −4.16; *P*<.001), and 2.64 mm Hg (95% CI −2.68 to −2.61; *P*<.001). DBP control rates increased in each quarter to 78.11% (403,641/516,777; β coefficient=0.52, 95% CI 0.51-0.53; *P*<.001), 80.32% (415,062/516,777; β coefficient=0.67, 95% CI 0.66-0.68; *P*<.001), 83.17% (429,829/516,777; β coefficient=0.89, 95% CI 0.88-0.90; *P*<.001), and 79.47% (410,662/516,777; β coefficient=0.61, 95% CI 0.60-0.62; *P*<.001). The older age group had a larger decrease in their mean SBP (β coefficient=0.87, 95% CI 0.85-0.90; *P*<.001) and a larger increase in SBP control rates (β coefficient=0.054, 95% CI 0.051-0.058; *P*<.001). The participants with cardiovascular disease (CVD) had a smaller decrease in their mean SBP (β coefficient=−0.38, 95% CI −0.41 to −0.35; *P*<.001) and smaller increase in SBP control rates (β coefficient=−0.041, 95% CI −0.045 to −0.037; *P*<.001) compared to the blood pressure of participants without CVD.

Conclusions: The NEPHSP was effective in improving blood pressure control of Chinese people with hypertension. Blood pressure control of older individuals and those with CVD need to be intensified.

(JMIR Public Health Surveill 2025;11:e65783) doi:10.2196/65783

KEYWORDS

hypertension; primary care; public health; blood pressure control; cardiovascular disease

Introduction

The Global Burden of Disease Study reported that cardiovascular diseases (CVDs) are the leading noncommunicable disease contributing to the increase in disease-adjusted life years in China [1-3]. Hypertension, the major modifiable risk factor for CVDs, is prevalent in China [4-12]. According to the data of a recent survey conducted nationwide, the prevalence of hypertension was 54.7% among Chinese adults over 18 years old, and reportedly, merely 60.1% of Chinese adults with hypertension accepted antihypertensive treatments [13]. The low hypertension treatment rate has led to a suboptimal control of blood pressure among Chinese adults with hypertension [13]. Merely 7.2% of Chinese adults with hypertension had their blood pressures under control [13].

To tackle the poor hypertension control and growing burden of CVDs, the Chinese government launched the National Essential Public Health Service Package (NEPHSP) in 2009 [14]. The NEPHSP was a set of public health services which were available for all community-dwelling residents [14-18]. The public health services contain 4 modules, which are screening, monitoring, regular follow-up, and individualized interventions for hypertension control. The 4 modules were provided by primary health care professionals [14-18]. NEPHSP registration has been continuously expanded since it was launched in 2009 [14,18]. A total of 109 million Chinese adults diagnosed with hypertension registered with the NEPHSP from 2009 to 2019 [14]. The government investment increased from 15 to 84 RMB (US $2.05 to US $11.46) per person annually [15,18]. Because of the registration expansion and investment increase for the NEPHSP, the proportion of Chinese adult with hypertension who accepted antihypertensive treatment increased by 59.58% from 2009 to 2019 [14,18].

Although the registration expansion and investment increase for the NEPHSP improved the hypertension treatment rate, the NEPHSP lacks quality-oriented evaluations, which leads to an uncertainty whether the NEPHSP is effective for blood pressure control [14-18]. Furthermore, research demonstrating the effectiveness of the NEPHSP for blood pressure control in Chinese adults with hypertension is lacking [14-18]. To fulfill the research gap, this study aimed to explore effect of the NEPHSP on blood pressures control of Chinese adult with hypertension.

Methods

Study Design

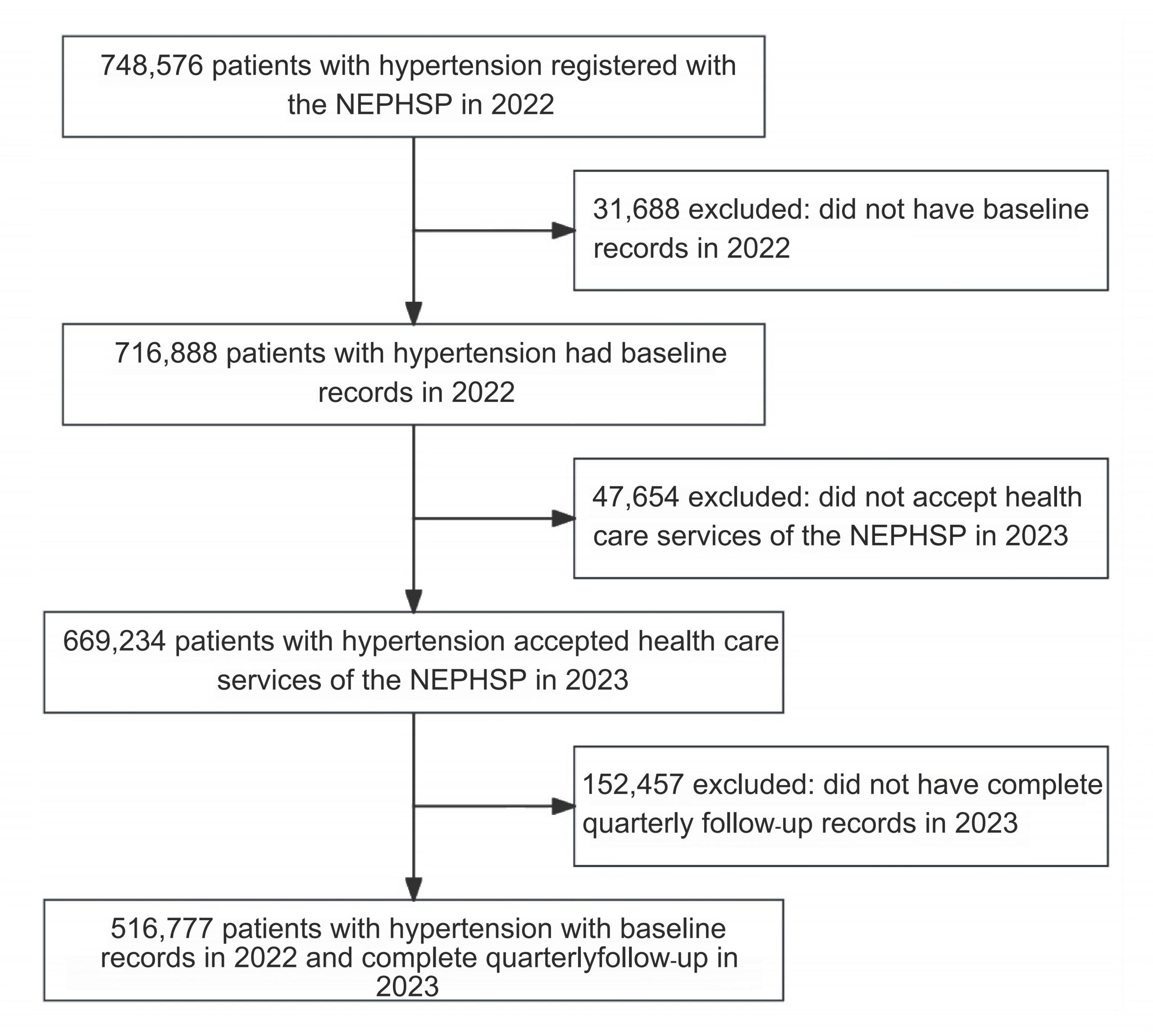
This study was designed as longitudinal study to evaluate effect of the NEPHSP intervention on blood pressure control. The longitudinal study observed changes in the health status and clinical indicators of individuals who accepted treatment in a period of time, in order to explore the association between treatment and disease control [19-22]. Longitudinal studies have been commonly used in clinical research to evaluate the effects of treatment [19-22]. In this longitudinal study, we repetitively measured the blood pressure in participants who accepted hypertension care services offered by the NEPHSP in each quarter of 2023. All participants had a recorded blood pressure measurement prior to joining the NEPHSP in 2022. With the baseline outcomes in 2022 as the control, the degree and trend of the blood pressure changes after receiving NEPHSP in 2023 were analyzed longitudinally.

The data in this study were extracted from the Residence Health Record System, which is an electronic health record system built to recording demographic and health information of people registering with the NEPHSP [16,17]. This study used data of the Residence Health Record System built in Jining City, which is a national pilot city of the Chronic Disease Comprehensive Prevention and Control programs and is located in eastern China.

Study Participants

We analyzed data from the NEPHSP participants from 1593 primary care institutions in the 9 districts in Jining City. Eligibility criteria were (1) individuals registering with the NEPHSP as a patient with hypertension, (2) individuals who had baseline records of health examinations in 2022, (3) individuals who accepted health care services offered by the NEPHSP in 2023, and (4) individuals who had records for quarterly follow-ups conducted in 2023. The flowchart of participant inclusion is described in Figure 1.

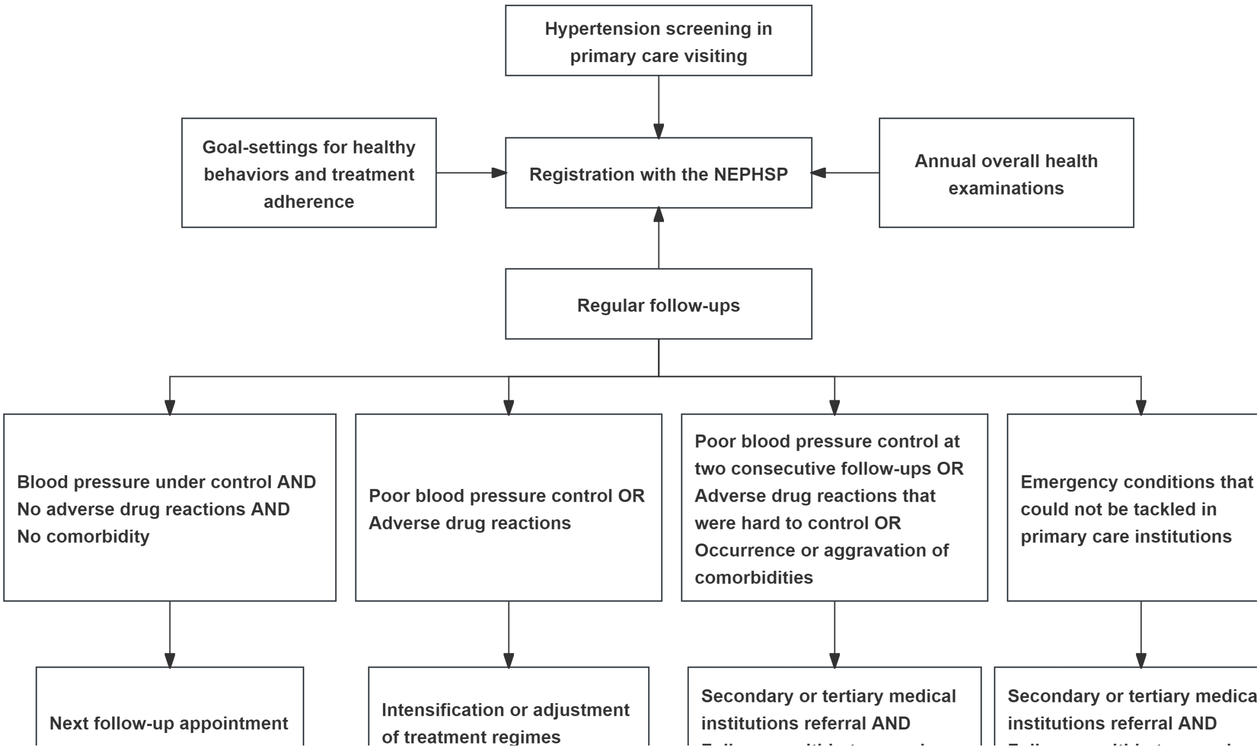
Figure 1. The flowchart of the participant inclusion. NEPHSP: National Essential Public Health Service Package.



Intervention

The NEPHSP was divided into 4 dimensions: screening, routine follow-ups, individualized interventions, and annual health examinations. The explicit flowchart of the NEPHSP is described in Figure 2.

Figure 2. Flowchart of the NEPHSP. NEPHSP: National Essential Public Health Service Package.



The screening was conducted at primary care institutions when community residents had an appointment with primary health care professionals. The criteria for a hypertension diagnosis were a systolic blood pressure (SBP) over 140 mm Hg or a diastolic blood pressure (DBP) over 90 mm Hg for 3 separate measurements on different days. The residents were invited to register with the NEPHSP if they were diagnosed with hypertension.

The routine follow-up contained 5 facets: (1) assessing emergency conditions, (2) evaluating hypertensive symptoms, (3) measuring blood pressure, (4) evaluating healthy behaviors, and (5) evaluating medication treatment regimens and compliance. Assessing emergency conditions was the first part of the routine follow-up. The participants would be referred to secondary or tertiary care settings if there were emergency conditions that were hard to tackle in primary care institutions. If no emergency conditions were identified, the remaining parts would be conducted. The regular follow-ups were provided quarterly.

The individualized interventions contained 3 components: medication treatment adjustment, referral channel to superior medical institutions, and healthy behavior changes. According to health conditions that were evaluated in the regular follow-ups, primary care professionals would make decisions on medication treatment adjustments for the participants. If the health issues could not be handled in primary care settings, the primary care professionals would transfer the patient to superior medical institutions. The explicit processes and criteria of medication treatment adjustment and referral were divided into 3 dimensions. First, if blood pressure was under control, there were no adverse drug reactions, and there was not an occurrence of comorbidities or an aggravation of comorbidities, the next follow-up appointments were made for the participants. Second, if blood pressures were uncontrolled or adverse drug reactions were identified, the primary health care professionals would adjust the antihypertensive medication regimens. The conditions of the participants would be re-evaluated within 2 weeks. Third, if the patients had refractory uncontrolled blood pressure, adverse drug reactions that could not be tackled in primary care settings, or an occurrence and aggravation of comorbidities, the primary health care professionals would refer the participants to secondary or tertiary care settings. The conditions of the participants would be re-evaluated within 2 weeks.

Changes in health behavior were available for all participants, and goal-setting was used to formulate their health behaviors. The primary health care professionals developed shared goals with the participants. The targets for changes in health behavior included medication treatment compliance, dietary changes and weight control, control of smoking and alcohol drinking, and exercise engagement. The achievements of the health behavior changes were evaluated at the routine follow-ups. Annual health examinations were available for all patients. The health examinations contained overall physical and bioclinical examinations, as well as lifestyle and medication treatment evaluations.

Outcome Variables

We evaluated the impact of the NEPHSP on blood pressure control in Chinese patients with hypertension through changes in SBP and DBP and changes in the control rates of SBP and DBP. The control rates of SBP and DBP were the proportion of study participants whose SBP and DBP values were in the normal range, respectively. The normal range of SBP and DBP in the Chinese hypertension guideline is defined as <140 mm Hg and <90 mm Hg, respectively [23].

We used the mean blood pressure values at baseline in 2022 and at 4 quarterly follow-up visits in 2023 to depict trends in blood pressure for the patients before and after receiving the NEPHSP services. The baseline values of SBP and DBP were averaged by the blood pressure measurements of the health examinations in 2022. The mean follow-up values of SBP and DBP were averaged by blood pressure measurements of the routine follow-ups. We extracted the blood pressure data from the Residence Health Record System, which recorded the SBP and DBP values measured in the annual health examinations and the regular follow-ups of the NEPHSP participants.

Statistical Analysis

Baseline characteristics of the participants were statistically described. The quantitative variables were described as mean and SD, while categorical variables were described as counts and percentage. The generalized estimating equation (GEE) model was used to analyze the repetitively measured blood pressure data that contained a continuous time variable [24]. The GEE model was capable of analyzing numerical and categorical data that were repetitively measured [24]. Both the trend and degree of the blood pressure changes were analyzed. First, the variable of time was set as the continuous variable to analyze the trend of the blood pressure changes in the GEE models. Linear trend tests were conducted to examine whether the changes in the blood pressure values with the continuous time variation were linear. The linear trend test were reflected by *P* for trend [25,26]. If the changes in the blood pressure values for the continuous time variation were linear, we used the coefficients of GEE models to reflect linear trends of blood pressure changes. Second, we also set the variable of time as the categorical variable to analyze the degree of blood pressure changes in the GEE models. The baseline blood pressure values in 2022 were chosen as the reference group. All statistical analyses were carried out by R software version 4.0.4 (R Core Team). The analysis was performed with 2-tailed tests at an α level of .05.

The covariates were first selected based on evidence of previous studies that had proven the factors that could impact blood pressure [27-34]. Subsequently, we also conducted univariate analyses to confirm whether the covariates had an impact on blood pressure. The following participant characteristics were adjusted as covariates: age, sex, BMI, waist circumference, exercise engagement, alcohol drinking and smoking status, fasting plasma glucose, total cholesterol, triglycerides, low density lipoprotein, high density lipoprotein, baseline SBP and DBP values, diagnosis of cardiovascular and renal diseases, and duration since hypertension diagnosis. The results of univariate analyses and corresponding information of each covariate in the GEE models are described in Multimedia Appendix 1 and Multimedia Appendix 2.

Sensitivity and Stratified Analyses

A series of stratified and sensitivity analyses were conducted to identify whether the changes in blood pressure were different by subgroup. The subgroups were predefined as an older age group (≥65 years old) and younger age group (<65 years old), a male group and female group, a CVD diagnosis group and no CVD diagnosis group, and a baseline controlled SBP group and baseline uncontrolled SBP group. We fit GEE models to assess the changes in blood pressure separately for those subgroups. We then tested the interactive effects between the time variables (continuous time variable) and the subgroup indicators using the total sample, to formally evaluate whether the changes in blood pressure differed between the subgroups.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and approved by the Committee on Human Research of the Jining No. 1 People’s Hospital (2023 Ethical Approval No. KYLL-204609 - 176). The need for informed consent was waived due to the retrospective nature of this study.

Results

Baseline Characteristics

The baseline characteristics are accessible in Table 1. A total of 516,777 participants were involved in this study. The mean age was 68.59 (SD 9.57) years, 58.56% (n=302,109) of the participants were female, and 66.94% (n=345,923) of the participants were over 65 years old. The mean values for baseline SBP and DBP were 157.12 (SD 19.88) mm Hg and 85.11 (SD 11.79) mm Hg, respectively. The mean values of baseline SBP were above the normal criteria [19]. The control rates of baseline SBP and DBP were 39.79% (205,630/516,777) and 69.21% (357,685/516,777), respectively.

Table . Baseline characteristics.

Characteristics	Values
Age (years), mean (SD)	68.39 (9.57)
Sex, n/N (%)	
Male	214,668/516,777 (41.54)
Female	302,109/516,777 (58.56)
DBP[a] (mm Hg), mean (SD)	85.11 (11.79)
DBP control rate, n/N (%)	357685/516,777 (69.21)
Cardiovascular disease diagnosis, n/N (%)	124409/516,777 (24.07)
SBP[b] (mm Hg), mean (SD)	
All participants	147.12 (19.88)
Older participants	149.81 (20.37)
Younger participants	141.67 (17.64)
Participants with cardiovascular disease	149.14 (18.48)
Participants without cardiovascular disease	144.47 (20.02)
SBP control rate, n/N (%)	
All participants	205,630/516,777 (39.79)
Older participants	114,649/345,923 (33.14)
Younger participants	90,981/170,854 (53.25)
Participants with cardiovascular disease	48,552/124,409 (39.02)
Participants without cardiovascular disease	157,078/392,368 (40.03)

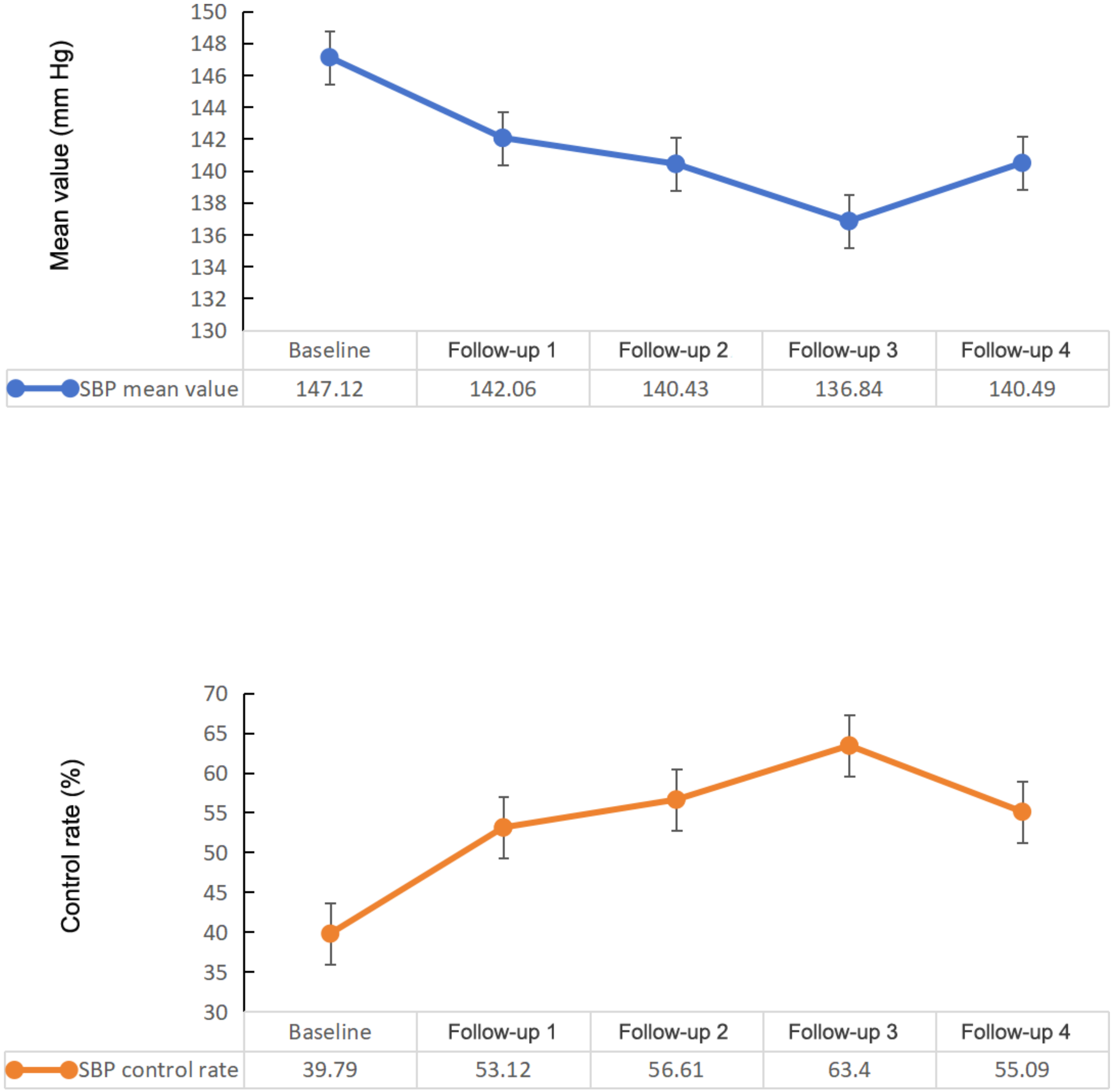
[a] DBP: diastolic blood pressure.

[b] SBP: systolic blood pressure.

Changes in SBP Control

The changes in SBP control are described in Figure 3. The SBP values had a linear change trend with continuous time variation (β coefficient=−1.85, 95% CI −1.86 to −1.93; *P*<.001).

Figure 3. Changes in SBP control. SBP: systolic blood pressure.



The SBP values showed a decreasing trend (Table 2, Figure 3). Compared to baseline, SBP decreased in each quarter by 5.06 mm Hg, 6.69 mm Hg, 10.60 mm Hg, and 6.63 mm Hg (Table 3).

Table . Slopes of the blood pressure changes from baseline to the fourth follow-up.

Variable	β coefficient (95% CI)	*P* value
SBP[a]	−1.85 (−1.86 to −1.93)	<.001
SBP control rate	0.185 (0.183-0.187)	<.001
DBP[b]	−0.77 (−0.78 to −0.76)	<.001
DBP control rate	0.169 (0.166-0.171)	<.001

[a] SBP: systolic blood pressure.

[b] DBP: diastolic blood pressure.

Table . Degree of change in systolic blood pressure (SBP) control.

Appointment	SBP		SBP control rate	
	β coefficient (95% CI)	*P* value	β coefficient (95% CI)	*P* value
Follow-up 1	−5.00 (−5.11 to −5.00)	<.001	0.18 (0.183-0.187)	<.001
Follow-up 2	−6.69 (−6.74 to −6.63)	<.001	0.60 (0.59-0.61)	<.001
Follow-up 3	−10.60 (−10.54 to −10.63)	<.001	0.76 (0.75-0.77)	<.001
Follow-up 4	−6.63 (−6.68 to −6.57)	<.001	1.08 (1.07-1.09)	<.001

SBP control rates for the NEPHSP participants showed an increasing trend (Figure 3, Table 2). Compared to baseline, SBP control rates increased in each quarter to 53.12% (274,493/516,777), 56.61% (292,537/516,777), 63.4% (327,648/516,777), and 55.09% (284,711/516,777; Table 3).

Changes in DBP Control

The changes in DBP control are described in Figure 4. The DBP values had a linear trend with continuous time variation (β coefficient=−1.85, 95% CI −1.86 to −1.93; *P*<.001). The DBP values of the NEPHSP participants showed a decreasing trend (Table 2, Figure 4). Compared to baseline DBP, DBP decreased in each quarter by 1.75 mm Hg, 2.64 mm Hg, 4.20 mm Hg, and 2.64 mm Hg (Table 4).

Figure 4. Changes in DBP control. DBP: diastolic blood pressure.

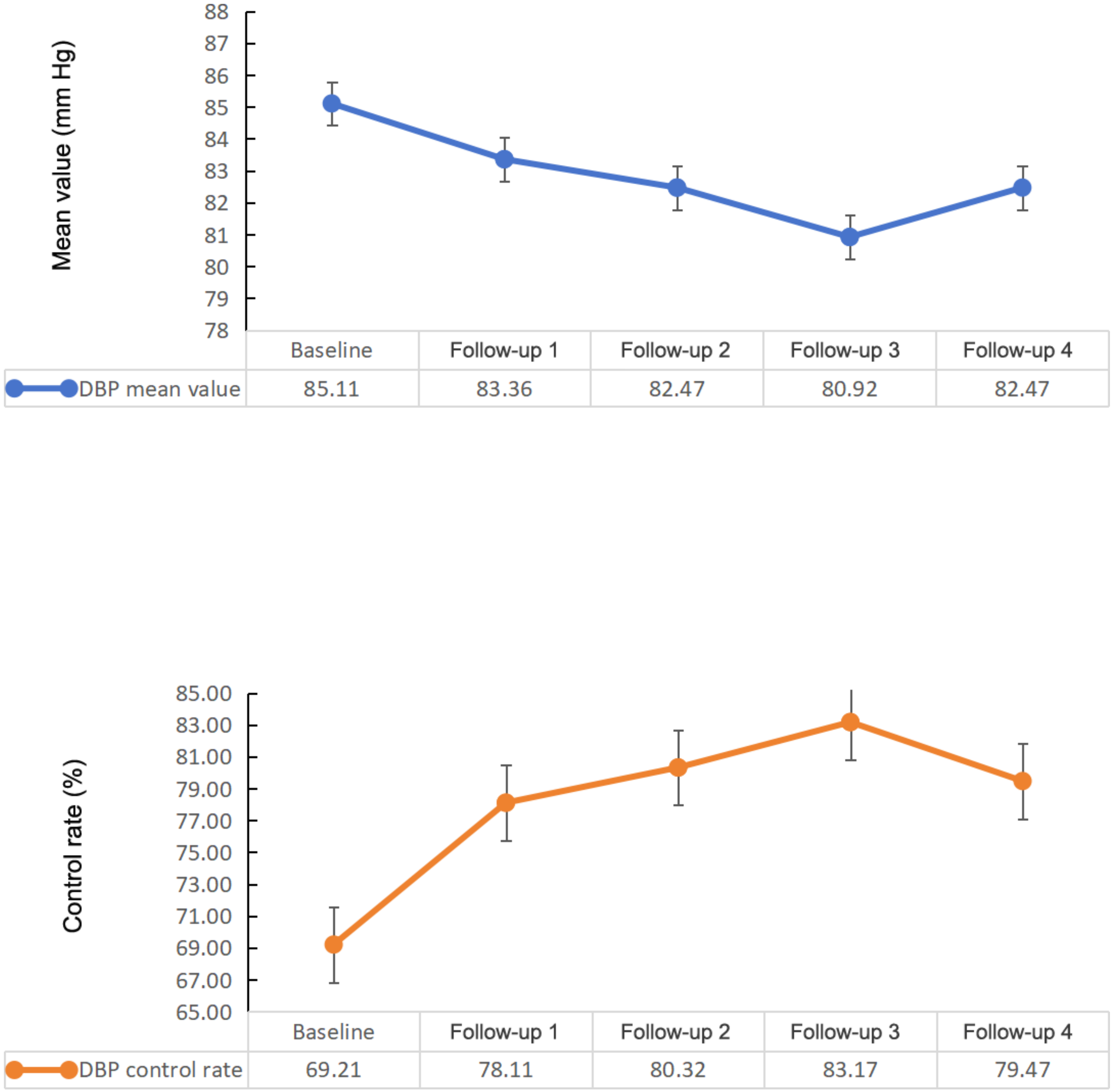


Table . Degree of change in diastolic blood pressure (DBP) control.

Appointment	DBP		DBP control rate	
	β coefficient (95% CI)	*P* value	β coefficient (95% CI)	*P* value
Follow-up 1	−1.75 (−1.79 to −1.72)	<.001	0.62 (0.61-0.63)	<.001
Follow-up 2	−2.64 (−2.68 to −2.61)	<.001	0.67 (0.66-0.68)	<.001
Follow-up 3	−4.20 (−4.63 to −4.16)	.01	0.89 (0.88-0.90)	<.001
Follow-up 4	−2.64 (−2.68 to −2.61)	<.001	0.61 (0.60-0.62)	<.001

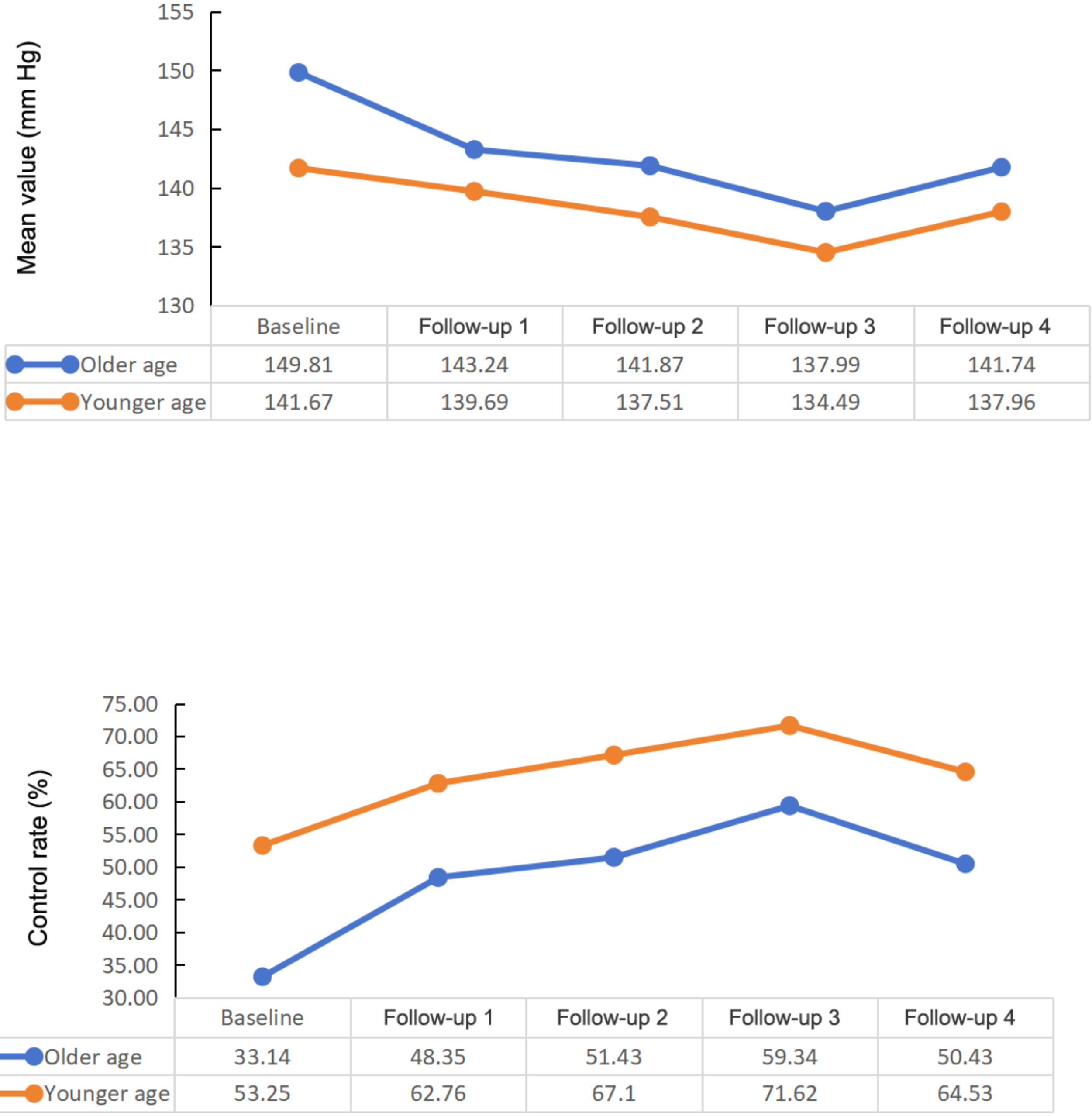
DBP control rates for the NEPHSP participants showed an increasing trend (Table 2, Figure 4). Compared to baseline, DBP control rates increased in each quarter to 78.11% (403,641/516,777), 80.32% (415,062/516,777), 83.17% (429,829/516,777), and 79.47% (410,662/516,777; Table 4).

Sensitivity and Stratified Analyses

Our sensitivity analysis stratified by the predefined subgroups showed that a decrease in blood pressure and an increase in blood pressure control rates were consistent across subgroups of age, sex, and a diagnosis of baseline CVD. The parameters of the sensitivity analysis are accessible in Multimedia Appendix 3. The changes in SBP were converse between the baseline controlled SBP and uncontrolled SBP groups. The baseline controlled SBP group had an increased control of their SBP, while the baseline uncontrolled SBP group had a decreased control of their SBP. The SBP control rates decreased in the baseline controlled SBP group, while the SBP control rates increased in the baseline uncontrolled SBP group.

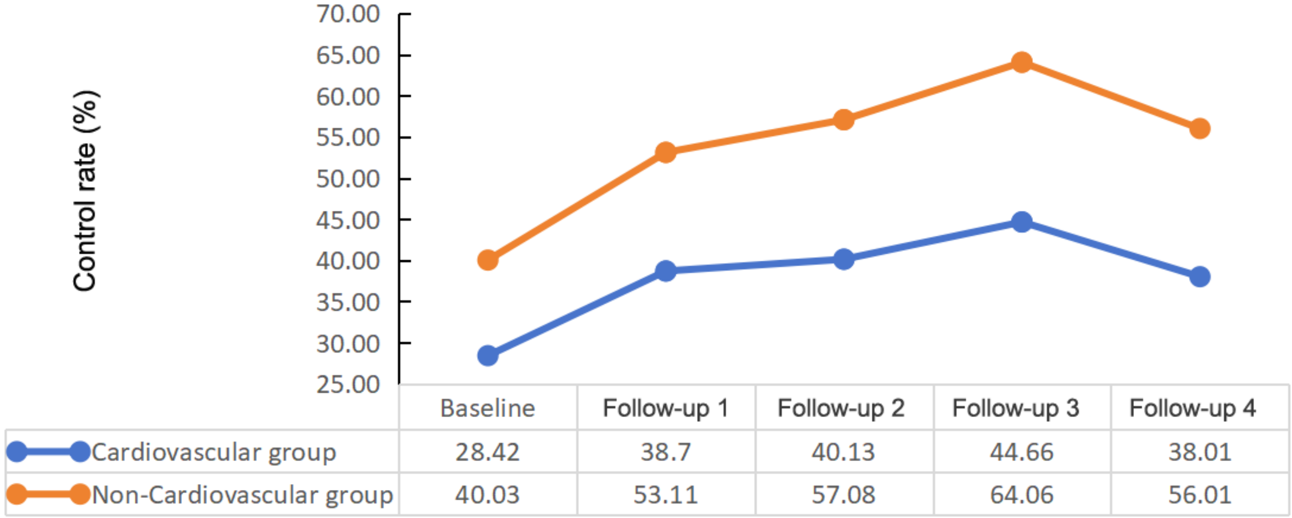
The changes in SBP control in the older and younger age groups are described in Figure 5. Compared with younger age group, the older age group showed a larger decrease in SBP and a larger increase in SBP control rates. The SBP of the older age group were higher than the younger age group. The SBP control rates of the older age group were significantly lower than the younger age group (*P*<.001).

Figure 5. Systolic blood pressure control of the age groups.



The changes in SBP control of the participants with and without CVD were described in Figure 6. The participants with CVD had a smaller decrease in SBP and smaller increase in SBP control rates than the participants without CVD. The SBP control rates of the participants with CVD were significantly lower than the participants without CVD (*P*<.001).

Figure 6. Systolic blood pressure control of the cardiovascular disease group.



The stratified analysis by sex demonstrated that the decrease in SBP values and the increase in SBP control rates were greater in the female group than in male group; however, the differences were relatively small.

Discussion

Principal Findings

To our knowledge, this is the first study that longitudinally explored effect of the NEPHSP on blood pressure control. The study findings demonstrated that the NEPHSP was effective for improving SBP and DBP control. The NEPHSP decreased the values of SBP and DBP and increased the control rates of SBP and DBP.

The routine follow-up that monitored blood pressure and the individualized interventions of the NEPHSP might contribute to the improvement of SBP and DBP control. The routine follow-ups generated clinically crucial data that could recognize prioritization of the participant subgroups who had poorly controlled blood pressures [35]. Previous studies also reported that the regular population review that enables the formulation of customized interventions for improving blood pressure control is an effective strategy for hypertension management [35-48]. Individualized interventions were formulated according to the evaluations of routine follow-ups to identify appropriate strategies to improve blood pressure control in participants with poorly controlled blood pressure [35,36,47].

Although the NEPHSP improved the control of blood pressure, the mean values of SBP were still slightly higher than the normal criteria. This might be caused by therapeutic inertia [33]. Therapeutic inertia was defined as a failure of health care professionals to initiate or intensify treatment appropriately during visits [33,49]. Previous research demonstrated that the failure to appropriately initiate or intensify treatment was introduced by the perspective of health care professionals that near-target blood pressures were acceptable and the overconcern about adverse effects caused by treatment intensification [47,48,50,51]. Because of the doctor-related factors leading to therapeutic inertia, participants with near-target blood pressures might not accept intensification of antihypertensive treatments, which could contribute to the slightly higher values of blood pressure [47,48,50,51]. Although therapeutic inertia is associated with poor blood pressure control and cardiovascular events, therapeutic inertia is common in hypertension management in primary care settings [47,48,51]. Given this situation, interventions for handling therapeutic inertia are important for improving hypertension management in primary care settings [33,47,48,51]. However, in terms of the NEPHSP, no component for tackling therapeutic inertia were embedded. The relevant components addressing therapeutic inertia can be considered for inclusion in the NEPHSP [33].

The subgroup analysis showed that the older participants had greater blood pressure improvement compared to the younger participants. This finding might be relevant to the referral mechanism of the NEPHSP. Compared with younger patients with hypertension, older patients with hypertension have more complex health conditions, such as frailty, comorbidities, and polypharmacy [34,52-58]. These complex health conditions can introduce challenges for blood pressure control in primary care settings, which are positively associated with poor blood pressure control [34,52-58]. To improve poor blood pressure control in older patients with complex health conditions in primary care settings, a referral mechanism to superior medical institutions is important [34,46,47,57-59]. In terms of the NEPHSP, the older participants with complex health conditions and poor blood pressure control were recognized via the regular follow-up evaluations. Based on the health condition evaluations of the regular follow-ups, the older participants could accept more referrals to superior medical institutions than the younger participants. Through the referral to superior medical institutions, the older participants could receive more intensive antihypertensive treatments than the treatments in primary care settings [34,46,47,57-59]. Because of the intensive antihypertensive treatments provided in superior medical institutions, the older participants had more blood pressure control improvement than the younger participants [34,46,47,57-59].

Although older participants had a greater improvement in blood pressure control than younger participants, the SBP of the older participants were still higher than the younger participants.

Moreover, the SBP control rates for the older participants were significantly lower than the younger participants. These study findings might be associated with concerns about adverse events caused by blood pressure–lowering pharmacotherapy [33,34,45-58]. However, previous robust research evidence demonstrated that decreasing the blood pressures of older adults to normal range is safe and effective for reducing risks of cardiovascular events [59,60]. Consequently, older people with hypertension should have a similar attitude toward blood pressure control as younger people with hypertension [59,60]. Given this situation, strategies to intensify blood pressure control in older participants engaging in the NEPHSP are necessary [59,60].

The improvement in the blood pressure control for the participants with CVD was less than the participants without CVD, while blood pressure control for people with hypertension and CVD needs to be more intensive than people with hypertension but without CVD [60]. Similar to the older participants, the patients with hypertension with CVD also required more referrals to superior medical institutions to achieve optimal blood pressure control [61-66]. Compared to patients with hypertension without CVD, the patients with CVD demand more intensive blood pressure control to prevent cardiovascular events, and blood pressure control in patients with hypertension and CVD is challenging in primary care settings [61-66]. The main challenges are potential adverse drug effects caused by polypharmacy and emergent cardiovascular events caused by intensive antihypertensive treatments [61-66]. Given this situation, referral to superior medical institutions is necessary for patients with hypertension and CVD to achieve optimal blood pressure control via intensive and safe antihypertensive treatments [61-66].

Although the NEPHSP had the algorithm to evaluate health conditions of the participants with comorbidities and subsequent referral mechanism to superior medical settings, the improvement of blood pressure control in the participants with CVD was still less than the participants without CVD. The lesser improvement in blood pressure control in the participants with CVD could be caused by the health evaluations of the regular follow-ups. The health condition evaluations might fail to accurately assess the health conditions of the participants with CVD, as the evaluations did not have systematic tools such as checklists, which could lead to failure to refer the participants who needed referral [61-66]. Furthermore, no explicit criteria for referral of participants with CVD were predefined, which could lead to confusion for the primary health care professionals when making the decisions on referral [61-66]. Given that the blood pressure control in people with hypertension and CVD needs to be more intensive than people without CVD, strategies for intensifying blood pressure control in the NEPHSP participants with CVD should be considered [61-66].

The subgroup analysis showed that the participants with controlled SBP at baseline had an increase in SBP and a decrease in SBP control rates, which were converse to the SBP results of all study participants. These results could be associated with the health evaluation algorithm of the regular follow-ups of the NEPHSP. According to the health evaluation algorithm, the participants who had controlled blood pressures that were evaluated in the regular follow-ups would not accept specific interventions. Because specific interventions were not provided, the participants with controlled SBP at baseline could have an increase in SBP and a decrease in SBP control rates. Although SBP of the participants was controlled at baseline, the increase of SBP values within the normal range could also increase CVD events [58]. Given this situation, the health evaluation algorithm and interventions of the NEPHSP for the participants with controlled SBP at baseline were suggested to be reformulated [58-61].

Blood lipid and glucose control could impact hypertension management. Previous studies reported that the triglyceride-glucose index (TyG), which is calculated by multiplying fasting triglyceride by fasting glucose, had a positive relationship with blood pressure control [31,32,61,62]. The increase of the TyG could contribute to poor blood pressures control [31,32,61,63]. A linear dose-response relationship between changes in the TyG and the change in blood pressure was identified [31,32,62,63]. To achieve optimal blood pressure control, blood lipid and glucose should be concurrently managed in people with hypertension [31,32,62,63]. However, in terms of the NEPHSP, no treatment targets for blood lipid and glucose were predefined for the study participants, which might cause ignorance toward blood lipid and glucose control. The ignorance of blood lipid and glucose control might cause a high TyG, which could contribute to the near-target blood pressure levels in this study [31,32,62,63]. Given this situation, further studies are required to confirm whether blood lipid and glucose control impact blood pressure control of the NEPHSP participants, in order to clarify the target for improving blood pressure control in this population.

Limitations

This study had several limitations. Although the effects of the NEPHSP on blood pressure control were modest, this study could not confirm the factors impacting blood pressure control of the NEPHSP. The group divisions for subgroup analysis were general and not explicit. Further studies that include a gradient group division to explore factors impacting hypertension control of NEPHSP are needed. Randomization for group division was not applied in the subgroup analysis, which could lead to an uneven distribution of confounding factors. Further studies that use randomization or matched group division are need to confirm the factors impacting hypertension control of the NEPHSP. This study explored baseline characteristics that could impact hypertension control of the NEPHSP, and future studies exploring the impact of the factor trajectories on hypertension control are needed. In addition, this study evaluated NEPHSP in a pilot city, and further nationwide studies are needed.

The blood pressure data in this study were the clinical records obtained from routine blood pressure measurements that were conducted via the Riva-Rocci and Korotkoff techniques [64-66]. The Riva-Rocci and Korotkoff techniques are still considered a cornerstone in evaluation of blood pressure levels in clinical trials [64-66]. However, the techniques have limitations, including that they can provide only a limited number of blood pressure values and can not reflect blood pressure fluctuations over 24 hours [64-66]. As a result, potential influences of blood pressure fluctuations over 24 hours in our study findings were uncertain [64-66]. Further studies that use advanced devices such as ambulatory blood pressure monitoring to explore impacts of 24-hour blood pressure fluctuations on effects of the NEPHSP on blood pressure control are required. Despite the uncertainty of impacts of 24-hour blood pressure fluctuations, our study findings also provide preliminary evidence for the control of blood pressure by the NEPHSP. First, our study was conducted at population level. The large scale of the study participants might reduce the impact of 24-hour blood pressure fluctuations on the changes in blood pressure levels to some degree. Second, the follow-up time was one year for this study, which could offset the impact of blood pressure fluctuations over 24 hours.

This study was designed as a self-controlled longitudinal study. We compared blood pressure levels before the participants accepted health services from the NEPHSP versus after accepting the health services. The changes in blood pressure levels could preliminarily reflect the effect of the NEPHSP on blood pressure control for the participants with hypertension. However, the data monitoring blood pressure changes were only available for the NEPHSP participants and were inaccessible for people with hypertension who did not register with the NEPHSP. Because of the inaccessibility of blood pressure monitoring data for this population, this study was unable to compare blood pressure control between the NEPHSP participants and people who did not register with the NEPHSP. Given this situation, further cohort and control trials are required to confirm advantages of the NEPHSP.

Conclusion

The NEPHSP was effective for improving blood pressure control of Chinese people with hypertension. However, the SBP of the study participants was near-target, which could imply therapy inertia in the NEPHSP participants. A suggested strategy to tackle potential therapy inertia was to intensify blood pressure control of the NEPHSP. The effects of the NEPHSP on blood pressure control in older participants and participants with CVD were modest. Embedding intensive treatment modules into the NEPHSP was suggested to improve blood pressure control of older participants and participants with CVD. Further studies confirming effects of the NEPHSP on blood pressure control for people with specific characteristics, such as older age and comorbidities, are required.

Acknowledgments

We are grateful to all the staff who participated in this research survey.

Data Availability

The data that support the findings of this study are available on reasonable request from the corresponding author. The data are not publicly available due to their containing information that could compromise the privacy of the participants.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Univariate regression analysis of covariates.
[XLSX File, 11 KB - publichealth_v11i1e65783_app1.xlsx]

Multimedia Appendix 2
Corresponding information of each covariate in the generalized estimating equation regressions.
[XLSX File, 12 KB - publichealth_v11i1e65783_app2.xlsx]

Multimedia Appendix 3
Subgroup analysis.
[XLSX File, 10 KB - publichealth_v11i1e65783_app3.xlsx]

References

1. GBD 2019 Viewpoint Collaborators. Five insights from the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1135-1159. [doi: 10.1016/S0140-6736(20)31404-5] [Medline: 33069324]
2. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222. [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
3. Lewington S, Lacey B, Clarke R, et al. The burden of hypertension and associated risk for cardiovascular mortality in China. JAMA Intern Med 2016 Apr;176(4):524-532. [doi: 10.1001/jamainternmed.2016.0190] [Medline: 26975032]

4. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1223-1249. [doi: 10.1016/S0140-6736(20)30752-2] [Medline: 33069327]
5. Roth GA, Mensah GA, Johnson CO, et al. Global burden of chronic respiratory diseases and risk factors, 1990-2019: an update from the Global Burden of Disease Study 2019. EClinMed 2023;59:101956. [doi: 10.1016/j.eclinm.2023.101956] [Medline: 37125402]
6. Yang ZJ, Liu J, Ge JP, et al. Prevalence of cardiovascular disease risk factor in the Chinese population: the 2007-2008 China National Diabetes and Metabolic Disorders Study. Eur Heart J 2012 Jan;33(2):213-220. [doi: 10.1093/eurheartj/ehr205] [Medline: 21719451]
7. Li D, Lv J, Liu F, et al. Hypertension burden and control in mainland China: analysis of nationwide data 2003-2012. Int J Cardiol 2015 Apr 1;184:637-644. [doi: 10.1016/j.ijcard.2015.03.045] [Medline: 25771229]
8. Feng XL, Pang M, Beard J. Health system strengthening and hypertension awareness, treatment and control: data from the China Health and Retirement Longitudinal Study. Bull World Health Organ 2014 Jan 1;92(1):29-41. [doi: 10.2471/BLT.13.124495] [Medline: 24391298]
9. He J, Gu D, Chen J, et al. Premature deaths attributable to blood pressure in China: a prospective cohort study. Lancet 2009 Nov 21;374(9703):1765-1772. [doi: 10.1016/S0140-6736(09)61199-5] [Medline: 19811816]
10. Yang G, Wang Y, Zeng Y. Rapid health transition in China, 1990-2010: findings from the Global Burden of Disease Study 2010. Lancet 2013 Jun 8;381(9882):1987-2015. [doi: 10.1016/S0140-6736(13)61097-1.] [Medline: 23746901]
11. Wang J, Zhang L, Wang F, Liu L, Wang H, China National Survey of Chronic Kidney Disease Working Group. Prevalence, awareness, treatment, and control of hypertension in China: results from a national survey. Am J Hypertens 2014 Nov;27(11):1355-1361. [doi: 10.1093/ajh/hpu053] [Medline: 24698853]
12. Li W, Gu H, Teo KK, et al. Hypertension prevalence, awareness, treatment, and control in 115 rural and urban communities involving 47 000 people from China. J Hypertens 2016 Jan;34(1):39-46. [doi: 10.1097/HJH.0000000000000745] [Medline: 26630211]
13. Lu J, Lu Y, Wang X, et al. Prevalence, awareness, treatment, and control of hypertension in China: data from 17 million adults in a population-based screening study (China PEACE Million Persons Project). Lancet 2017 Dec 9;390(10112):2549-2558. [doi: 10.1016/S0140-6736(17)32478-9] [Medline: 29102084]
14. You L, Zhao J, Chen X. National essential public health services programs over the past decade research report two: progress and achievements of the implementation of National Essential Public Health Services Programs over the past decade. Chin Gen Pract 2022;25:3209-3220. [doi: 10.12114/j.issn.1007-9572.2022.0407]
15. Fang G, Yang D, Wang L, Wang Z, Liang Y, Yang J. Experiences and challenges of implementing universal health coverage with China’s National Basic Public Health Service Program: literature review, regression analysis, and insider interviews. JMIR Public Health Surveill 2022;8(7):e31289. [doi: 10.2196/31289] [Medline: 35867386]
16. Li X, Lu J, Hu S, et al. The primary health-care system in China. Lancet 2017 Dec 9;390(10112):2584-2594. [doi: 10.1016/S0140-6736(17)33109-4] [Medline: 29231837]
17. Li X, Krumholz HM, Yip W, et al. Quality of primary health care in China: challenges and recommendations. Lancet 2020 Jun 6;395(10239):1802-1812. [doi: 10.1016/S0140-6736(20)30122-7] [Medline: 32505251]
18. Xiong S, Jiang W, Meng R, et al. Factors associated with the uptake of national essential public health service package for hypertension and type-2 diabetes management in China’s primary health care system: a mixed-methods study. Lancet Reg Health West Pac 2023 Feb;31:100664. [doi: 10.1016/j.lanwpc.2022.100664] [Medline: 36879777]
19. Liu J, Colditz GA. Optimal design of longitudinal data analysis using generalized estimating equation models. Biom J 2017 Mar;59(2):315-330. [doi: 10.1002/bimj.201600107] [Medline: 27878852]
20. Newman AB. An overview of the design, implementation, and analyses of longitudinal studies on aging. J Am Geriatr Soc 2010 Oct;58 Suppl 2(Suppl 2):S287-S291. [doi: 10.1111/j.1532-5415.2010.02916.x] [Medline: 21029055]
21. Schneeweiss S, Rassen JA, Brown JS, et al. Graphical depiction of longitudinal study designs in health care databases. Ann Intern Med 2019 Mar 19;170(6):398-406. [doi: 10.7326/M18-3079] [Medline: 30856654]
22. Hund H, Gerth S, Lossnitzer D, Fegeler C. Longitudinal data driven study design. Stud Health Technol Inform 2014;205(373-7):373-377. [Medline: 25160209]
23. Wang JG, Chia YC, Chen CH, et al. What is new in the 2018 Chinese hypertension guideline and the implication for the management of hypertension in Asia? J Clin Hypertens (Greenwich) 2020 Mar;22(3):363-368. [doi: 10.1111/jch.13803] [Medline: 31955513]
24. Huang Y, Pan J. Penalized joint generalized estimating equations for longitudinal binary data. Biom J 2022 Jan;64(1):57-73. [doi: 10.1002/bimj.202000336] [Medline: 34587284]
25. Lee DH, Keum N, Hu FB, et al. Predicted lean body mass, fat mass, and all cause and cause specific mortality in men: prospective US cohort study. BMJ 2018 Jul 3;362:k2575. [doi: 10.1136/bmj.k2575] [Medline: 29970408]
26. Zhu J, Chen C, Lu L, Shikany JM, D’Alton ME, Kahe K. Folate, vitamin B6, and vitamin B12 status in association with metabolic syndrome incidence. JAMA Netw Open 2023 Jan 3;6(1):e2250621. [doi: 10.1001/jamanetworkopen.2022.50621] [Medline: 36630134]

27. Zhang W, He K, Zhao H, et al. Association of body mass index and waist circumference with high blood pressure in older adults. BMC Geriatr 2021 Apr 19;21(1):260. [doi: 10.1186/s12877-021-02154-5] [Medline: 33874889]
28. Bowling CB, Sloane R, Pieper C, et al. Association of sustained blood pressure control with multimorbidity progression among older adults. J Am Geriatr Soc 2020 Sep;68(9):2059-2066. [doi: 10.1111/jgs.16558] [Medline: 32501546]
29. Karmali KN, Lloyd-Jones DM, van der Leeuw J, et al. Blood pressure-lowering treatment strategies based on cardiovascular risk versus blood pressure: a meta-analysis of individual participant data. PLoS Med 2018 Mar;15(3):e1002538. [doi: 10.1371/journal.pmed.1002538] [Medline: 29558462]
30. Czernichow S, Castetbon K, Salanave B, et al. Determinants of blood pressure treatment and control in obese people: evidence from the general population. J Hypertens 2012 Dec;30(12):2338-2344. [doi: 10.1097/HJH.0b013e3283593010] [Medline: 23032140]
31. Ren Y, Cheng L, Song Y, et al. Ideal cardiovascular health index and high-normal blood pressure in elderly people: evidence based on real-world data. Sci Rep 2024 May 3;14(1):10166. [doi: 10.1038/s41598-024-60906-w] [Medline: 38702348]
32. Aljuraiban GS, Alharbi FJ, Aljohi AO, et al. Triglyceride-glucose index (TyG index) in association with blood pressure in adults: a retrospective study. Int J Gen Med 2024;17:3395-3402. [doi: 10.2147/IJGM.S469147] [Medline: 39130487]
33. Masoli JAH, Mensah E, Rajkumar C. Age and ageing cardiovascular collection: blood pressure, coronary heart disease and heart failure. Age Ageing 2022 Aug 2;51(8):afac179. [doi: 10.1093/ageing/afac179] [Medline: 35934320]
34. Wong MCS, Wang HHX, Cheung CSK, et al. Factors associated with multimorbidity and its link with poor blood pressure control among 223,286 hypertensive patients. Int J Cardiol 2014 Nov 15;177(1):202-208. [doi: 10.1016/j.ijcard.2014.09.021] [Medline: 25499379]
35. Glynn LG, Murphy AW, Smith SM, Schroeder K, Fahey T. Interventions used to improve control of blood pressure in patients with hypertension. Cochrane Database Syst Rev 2010 Mar 17;17(3):CD005182. [doi: 10.1002/14651858.CD005182.pub4] [Medline: 20238338]
36. Jaffe MG, Lee GA, Young JD, Sidney S, Go AS. Improved blood pressure control associated with a large-scale hypertension program. JAMA 2013 Aug 21;310(7):699-705. [doi: 10.1001/jama.2013.108769] [Medline: 23989679]
37. Jafar TH, Gandhi M, de Silva HA, et al. A community-based intervention for managing hypertension in rural South Asia. N Engl J Med 2020 Feb 20;382(8):717-726. [doi: 10.1056/NEJMoa1911965] [Medline: 32074419]
38. Schwalm JD, McCready T, Lopez-Jaramillo P, et al. A community-based comprehensive intervention to reduce cardiovascular risk in hypertension (HOPE 4): a cluster-randomised controlled trial. Lancet 2019 Oct 5;394(10205):1231-1242. [doi: 10.1016/S0140-6736(19)31949-X] [Medline: 31488369]
39. Anand TN, Joseph LM, Geetha AV, Prabhakaran D, Jeemon P. Task sharing with non-physician health-care workers for management of blood pressure in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Glob Health 2019 Jun;7(6):e761-e771. [doi: 10.1016/S2214-109X(19)30077-4] [Medline: 31097278]
40. Schlaich MP, Kiuchi MG, Nolde JM, Lugo-Gavidia LM, Carnagarin R. Therapeutic inertia in hypertension management - status quo in primary care. J Hypertens 2021 Jun 1;39(6):1107-1108. [doi: 10.1097/HJH.0000000000002830] [Medline: 33967214]
41. Phillips LS, Branch WT, Cook CB, et al. Clinical inertia. Ann Intern Med 2001 Nov 6;135(9):825-834. [doi: 10.7326/0003-4819-135-9-200111060-00012] [Medline: 11694107]
42. Milman T, Joundi RA, Alotaibi NM, Saposnik G. Clinical inertia in the pharmacological management of hypertension: a systematic review and meta-analysis. Medicine (Baltimore) 2018 Jun;97(25):e11121. [doi: 10.1097/MD.0000000000011121] [Medline: 29924011]
43. Lebeau JP, Cadwallader JS, Vaillant-Roussel H, et al. General practitioners' justifications for therapeutic inertia in cardiovascular prevention: an empirically grounded typology. BMJ Open 2016 May 13;6(5):e010639. [doi: 10.1136/bmjopen-2015-010639] [Medline: 27178974]
44. Kerr EA, Zikmund-Fisher BJ, Klamerus ML, Subramanian U, Hogan MM, Hofer TP. The role of clinical uncertainty in treatment decisions for diabetic patients with uncontrolled blood pressure. Ann Intern Med 2008 May 20;148(10):717-727. [doi: 10.7326/0003-4819-148-10-200805200-00004] [Medline: 18490685]
45. Masoli JAH, Delgado J, Pilling L, Strain D, Melzer D. Blood pressure in frail older adults: associations with cardiovascular outcomes and all-cause mortality. Age Ageing 2020 Aug 24;49(5):807-813. [doi: 10.1093/ageing/afaa028] [Medline: 32133525]
46. Boyd C, Smith CD, Masoudi FA, et al. Decision making for older adults with multiple chronic conditions: executive summary for the American Geriatrics Society Guiding Principles on the Care of Older Adults With Multimorbidity. J Am Geriatr Soc 2019 Apr;67(4):665-673. [doi: 10.1111/jgs.15809] [Medline: 30663782]
47. Benetos A, Petrovic M, Strandberg T. Hypertension management in older and frail older patients. Circ Res 2019 Mar 29;124(7):1045-1060. [doi: 10.1161/CIRCRESAHA.118.313236] [Medline: 30920928]
48. Anker D, Santos-Eggimann B, Zwahlen M, et al. Blood pressure control and complex health conditions in older adults: impact of recent hypertension management guidelines. J Hum Hypertens 2021 Mar;35(3):280-289. [doi: 10.1038/s41371-020-0334-4] [Medline: 32346124]

49. Bowling CB, Lee A, Williamson JD. Blood pressure control among older adults with hypertension: narrative review and introduction of a framework for improving care. Am J Hypertens 2021 Apr 2;34(3):258-266. [doi: 10.1093/ajh/hpab002] [Medline: 33821943]
50. Zhang WL, Cai J. STEP to blood pressure management of elderly hypertension: evidence from Asia. Hypertens Res 2022 Apr;45(4):576-582. [doi: 10.1038/s41440-022-00875-7] [Medline: 35277670]
51. Blood Pressure Lowering Treatment Trialists' Collaboration. Age-stratified and blood-pressure-stratified effects of blood-pressure-lowering pharmacotherapy for the prevention of cardiovascular disease and death: an individual participant-level data meta-analysis. Lancet 2021 Sep 18;398(10305):1053-1064. [doi: 10.1016/S0140-6736(21)01921-8] [Medline: 34461040]
52. Curfman G, Bauchner H, Greenland P. Treatment and control of hypertension in 2020: the need for substantial improvement. JAMA 2020 Sep 22;324(12):1166-1167. [doi: 10.1001/jama.2020.13322] [Medline: 32902571]
53. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA 2005 Aug 10;294(6):716-724. [doi: 10.1001/jama.294.6.716] [Medline: 16091574]
54. Valderas JM, Starfield B, Roland M. Multimorbidity’s many challenges: a research priority in the UK. BMJ 2007 Jun 2;334(7604):1128. [doi: 10.1136/bmj.39226.427095.3A] [Medline: 17540918]
55. Kennard L, O’Shaughnessy KM. Treating hypertension in patients with medical comorbidities. BMJ 2016 Feb 16;352:i101. [doi: 10.1136/bmj.i101] [Medline: 26884124]
56. Sarkar C, Dodhia H, Crompton J, et al. Hypertension: a cross-sectional study of the role of multimorbidity in blood pressure control. BMC Fam Pract 2015 Aug 7;16:98. [doi: 10.1186/s12875-015-0313-y] [Medline: 26248616]
57. Paulsen MS, Andersen M, Thomsen JL, et al. Multimorbidity and blood pressure control in 37 651 hypertensive patients from Danish general practice. J Am Heart Assoc 2012 Dec 31;2(1):e004531. [doi: 10.1161/JAHA.112.004531] [Medline: 23525411]
58. Messerli FH, Hofstetter L, Rimoldi SF, Rexhaj E, Bangalore S. Risk factor variability and cardiovascular outcome: JACC review topic of the week. J Am Coll Cardiol 2019 May 28;73(20):2596-2603. [doi: 10.1016/j.jacc.2019.02.063] [Medline: 31118154]
59. Mehlum MH, Liestøl K, Kjeldsen SE, et al. Blood pressure variability and risk of cardiovascular events and death in patients with hypertension and different baseline risks. Eur Heart J 2018 Jun 21;39(24):2243-2251. [doi: 10.1093/eurheartj/ehx760] [Medline: 29365085]
60. Mezue K, Goyal A, Pressman GS, Matthew R, Horrow JC, Rangaswami J. Blood pressure variability predicts adverse events and cardiovascular outcomes in SPRINT. J Clin Hypertens (Greenwich) 2018 Sep;20(9):1247-1252. [doi: 10.1111/jch.13346] [Medline: 29984884]
61. Stevens SL, Wood S, Koshiaris C, et al. Blood pressure variability and cardiovascular disease: systematic review and meta-analysis. BMJ 2016 Aug 9;354:i4098. [doi: 10.1136/bmj.i4098] [Medline: 27511067]
62. Wang Y, Yang W, Jiang X. Association between triglyceride-glucose index and hypertension: a meta-analysis. Front Cardiovasc Med 2021;8:644035. [doi: 10.3389/fcvm.2021.644035] [Medline: 34136539]
63. Xu AR, Jin Q, Shen Z, Zhang J, Fu Q. Association between the risk of hypertension and triglyceride glucose index in Chinese regions: a systematic review and dose-response meta-analysis of a regional update. Front Cardiovasc Med 2023;10:1242035. [doi: 10.3389/fcvm.2023.1242035] [Medline: 37583585]
64. Parati G, Bilo G, Mancia G. Blood pressure measurement in research and in clinical practice: recent evidence. Curr Opin Nephrol Hypertens 2004 May;13(3):343-357. [doi: 10.1097/00041552-200405000-00013] [Medline: 15073495]
65. Sewell K, Halanych JH, Russell LB, et al. Blood pressure measurement biases in clinical settings, Alabama, 2010-2011. Prev Chronic Dis 2016 Jan 7;13:E01. [doi: 10.5888/pcd13.150348] [Medline: 26741995]
66. Atherosclerosis, Hypertension, and Obesity in the Young Committee of the American Heart Association Council on Cardiovascular Disease in the Young, Alpert B, McCrindle B, et al. Recommendations for blood pressure measurement in human and experimental animals; part 1: blood pressure measurement in humans. Hypertension 2006 Jul;48(1):e3. [doi: 10.1161/01.HYP.0000229661.06235.08] [Medline: 16769991]

Abbreviations

CVD: cardiovascular disease
DBP: diastolic blood pressure
GEE: generalized estimating equation
NEPHSP: National Essential Public Health Service Package
SBP: systolic blood pressure
TyG: triglyceride-glucose index

Edited by A Mavragani, T Sanchez; submitted 26.08.24; peer-reviewed by W Wang, Y Ren; revised version received 17.12.24; accepted 19.12.24; published 06.02.25.

Please cite as:
Wei F, Ge Y, Li H, Liu Y
Impact of the National Essential Public Health Service Package on Blood Pressure Control in Chinese People With Hypertension: Retrospective Population-Based Longitudinal Study
JMIR Public Health Surveill 2025;11:e65783
URL: https://publichealth.jmir.org/2025/1/e65783
doi:10.2196/65783

Investigating the Trajectories of Poor Vision in Children and Adolescents in Wuhan, China From 2016 to 2019: Prospective Cohort Study

Lijuan Xu[1*], MM; Hanjia Li[2*], MPH; Fang Li[3*], MPH; Tinghui Zhang[2], BS; Jingyan Yan[2], MPH; Hong Yan[2], PhD; Lu He[1], MM; Bin Yu[2,4], PhD

[1]Physical Examination Center, Renmin Hospital of Wuhan University, Wuhan, China

[2]Department of Epidemiology and Health Statistics, School of Public Health, Wuhan University, 115 Donghu Rd, Wuchang District, Wuhan, China

[3]Wuhan Center for Disease Control and Prevention, Wuhan, China

[4]Population and Health Research Center, Wuhan University, Wuhan, China

[*]these authors contributed equally

Corresponding Author:
Bin Yu, PhD
Department of Epidemiology and Health Statistics, School of Public Health, Wuhan University, 115 Donghu Rd, Wuchang District, Wuhan, China

Abstract

Background: Poor vision is a challenging public health problem among children and adolescents globally and in China. It is well-recognized that early onset of poor vision and progressing to moderate and severe poor vision will increase the risk of irreversible blinding complications. To achieve the national goal of poor vision control and prevention, it is essential to investigate and understand the development of poor vision among children and adolescents in China.

Objective: This study aims to investigate the progression of poor vision among children and adolescents in Wuhan, China, based on a prospective cohort and to provide scientific evidence for the development and implementation of effective poor vision prevention and control programs.

Methods: Data were derived from a 4-year prospective cohort (2016 - 2019) of primary and middle school students (N=108,585) in Wuhan, China. Vision condition was measured using the standard logarithmic visual acuity charts. A group-based trajectory model was used to identify trajectories of poor vision overall and by gender and region.

Results: The mean age of the study subjects was 11.13 (SD 3.33) years, 200,110 (53.91%) were male and the majority (354404, 95.48%) were from urban areas. The prevalence of poor vision was 58.51% in 2016, 58.95% in 2017, 53.83% in 2018, and 54.79% in 2019. Group-based trajectory model identified 3 groups, including normal vision group (NVG) (27.4%), vision decline group (VDG) (17.8%), and moderate poor vision group (MPVG) (54.8%). A higher proportion of girls (57.8%) were in the MPV group compared to boys (50.5%), and the VDG showed greater changes in girls compared to boys. Furthermore, urban students (55.3%) had a higher proportion of MPV compared to rural students (47.5%), while urban students (17.2%) had a smaller proportion in the VDG compared to rural students (24%). Further analyses showed that as age increased, the likelihood of being categorized in the NVG decreased (β=−.417, *P*<.001), while the likelihood of being in the VDG (β=.058, *P*<.001) increased. Compared with boys, girls were more likely to be categorized in the VDG (β=.597, *P*<.001) and MPV group (β=.362, *P*<.001). Rural students were less likely than urban students to be categorized in the VDG (β=−.311, *P*<.001).

Conclusions: The prevalence of poor vision among children and adolescents in Wuhan has remained high over the years, with a slight decrease in recent years. The study identified three groups: normal vision, vision decline, and moderate poor vision. Girls and students from urban areas were more likely to have moderate poor vision, while boys and rural students had a higher proportion of vision decline. These findings provide valuable information for implementing poor vision prevention and control policies in the region.

(JMIR Public Health Surveill 2025;11:e53028) doi:10.2196/53028

KEYWORDS

children and adolescents; poor vision; group-based trajectory model; myopia; gender difference; China; normal vision group; vision decline group; moderate poor vision group; prevalence

Introduction

With the increasing prevalence of poor vision among children and adolescents in recent years, vision problems have been recognized as a major public health challenge globally in the 21st century [1] and in China [2]. Childhood and adolescence are well-accepted as the most vulnerable periods for poor visual development. Visual acuity reaches its full development by approximately 8 years of age, while visual cortex plasticity begins to diminish from the age of 2 years onward [1]. Poor vision can exert significant psychological, educational, and social impacts on children and adolescents, including psychological distress [3], perceived barriers [4], and poor academic achievement [5]. One systematic review and meta-analysis indicated that approximately one-third of children and adolescents were affected by myopia worldwide, and the number was projected to 740 million by 2050 [6] . Another review reported the prevalence of poor vision in children and adolescents as 7.26%, 7.34%, and 2.91% based on uncorrected visual acuity, presenting visual acuity and best corrected visual acuity, respectively, with refractive errors being the most common cause [7].

Compared with other countries, children and adolescents in East Asia reported a higher prevalence of poor vision. For example, data from a literature review indicated that the estimated prevalence of myopia among adolescents in East Asia was 69%, far greater than 16.7% in White adolescents and 14.3% in Hispanic or Latino adolescents [8]. Further, it is well-known that children and adolescents in China experience a significant burden of poor vision. Findings from a systematic review showed that studies from China reported a much higher prevalence of poor vision than those from other countries (13.33% - 41.17% in China vs 1.34% - 29.42% in other countries) [7]. Data from Chinese government indicated that the myopia rate among children and adolescents was 53.6% in 2018,decreased to 50.2% in 2019, and reversed to 52.7% in 2020 [9]. Additionally, data from a review showed that the prevalence of vision impairment in China increased from 19% at age 6 years to 66.9% at age 17 years [10]. To achieve the national goal of controlling and preventing poor vision, it is essential to investigate and understand its development among children and adolescents in China.

It is well-recognized that progression to moderate and severe poor vision increases the risk of irreversible blinding complications, such as retinal detachment, glaucoma, and myopic maculopathy [11]. Previous studies have shown that individuals with early-onset poor vision are at a higher risk of progressing to severe poor vision and vision impairment [12-14]. Therefore, it is valuable to identify the trends of poor vision progression among children and adolescents. However, to the best of our knowledge, most studies among children and adolescents in China have been cross-sectional, limiting the ability to examine the incidence and progression of poor vision [15,16]. Based on a prospective cohort, this study aimed to investigate the progression of poor vision among children and adolescents in China. Findings of the study will provide evidence for the future design and implementation of effective interventions and prevention programs for poor vision.

Methods

Participants and Sampling

Study data were derived from a prospective cohort of students in primary and middle schools in Wuchang District, Wuhan City, China. The cohort was developed based on a screening project that included all primary and middle schools in the Wuchang District. All students who were currently studying in the selected schools, with oral or written informed consent collected from the parents or guardians, were recruited for the study. Students who were absent from school on the test day or had severe eye diseases or conditions that may influence the testing results (eg, infections, surgeries) were excluded. Enrollment proportions were 91.10% in 2016, 91.43% in 2017, 95% in 2018, and 96.88% in 2019. The screening team from the Physical Examination Center, Renmin Hospital of Wuhan University visited each primary or middle school on a scheduled day. Schools prepared a classroom for the testing in advance, and students in that school entered the room one class at a time until all students were tested. The screening team completed the screening test for all students in the primary or middle schools in Wuchang District; the process was repeated every year between 2016 and 2019. Wuhan is one of the largest metropolitan cities in central China, with a population of 15 million and a per capita GDP (gross domestic product) of $25,000. Wuchang District is one of the 13 districts of Wuhan and ranks among the largest, with a population of 1.27 million [17]. We aimed to investigate the vision status among students and provide evidence and recommendations for decision-makers. Data from 2016 to 2019 were used in this study, with 89,872, 91,060, 93,391, and 96,849 students enrolled in each year, respectively.

Ethical Considerations

The original study was a free governmental screening project for which the informed consent was waived from the parents or guardians. This study involved secondary data analysis of screening data and was approved by the Medical Ethics Committee at Wuhan University School of Medicine (WHU-LFMD-IRB2023031). Informed consent for the secondary data analysis was waived, and all data were was deidentified. No compensation was provided to the participants.

Measurements

Vision Status Assessment

Participants' visual acuity was measured by the standard logarithmic visual acuity charts, which has been widely used in poor vision screening in ophthalmology clinics and schools for over 20 years in China [18]. Participants were asked to stand at a distance of 5 m from the visual acuity chart and identify the direction of the "E optotype." Visual acuity was recorded using the 5-point method; this is an original visual acuity recording method in China, which stipulates normal vision as ≥5.0 and no light perception as 0. The participant's visual acuity is the recorded value of visual acuity corresponding to the smallest row for which they can correctly identify more than half the directions of the "E" optotype [18]. All inspectors received professional training.

The screening results were categorized based on the requirements of the technical standard for physical examination for students [19]; normal vision was defined as 5.0 in both left and right eyes. In cases where the visual acuity of left and right eyes differed, the lower value was used for classification. Mild poor vision was defined as a visual acuity of 4.9, moderate poor vision as <4.9 but >4.5, and severe poor vision as ≤4.5.

Demographic Variables

Demographic variables, including age (in years), sex (male or female), and place of origin (urban or rural), were provided by the schools where the children and adolescents were enrolled.

Statistical Analysis

Descriptive analyses (eg, frequency, proportion, mean, standard deviation) were used to summarize the sample characteristics. The group-based trajectory model (GBTM) was used to analyze the longitudinal data, explore the heterogeneity within the study sample, and further identify different developmental trajectories for poor vision. The underlying assumption of the GBTM is that the overall study sample consists of several latent subgroups with distinct developmental trajectories or patterns. Findings from the GBTM could help researchers and decision-makers to have a deeper understanding of the variability within the population and how different individuals progress over time. The GBTM has been widely used in various areas, including behaviors, mental health [20], and medical conditions [21]. It is an application of finite mixture modeling, which can approximate different developmental trajectories with a finite number of groups through maximum likelihood estimation [22]. The model accommodates missing data for certain time points, as demonstrated in previous studies [23,24]. Here, $Y_i=\{y_{i1},y_{i2},...,y_{it}\}$ denotes individual i's longitudinal sequence of the k^{th} measured over T=4 time periods, π_g denotes the probability that individual i belongs to g^{th} group, $P^g(Y_i)$ represents the conditional probability of Y_i for a given membership in group g. The probability of the individual i's longitudinal sequence is expressed by Equation 1.

$$(1) P(Y_i)=\sum_{g=1}^{G}\pi_g P^g(Y_i)$$

Furthermore, GBTM assumes that the random variables $y_{it}(t=1,2,...,T)$ are independent of each other when within groups. Therefore, we can use Equation 2 to definite $P^g(Y_i)$.

$$(2) P^g(Y_i)=\prod_{t=1}^{T}P^g(y_{it})$$

We tested models with varying numbers of trajectory groups and orders of polynomial function, and then selected the best model based on four criteria [25-30] : (a) a lower absolute value of Bayesian Information Criteria , (b) the average posterior probability (Avepp) for class membership >0.7 for all groups, (c) all groups included at least 5% of the total sample, and (d) a simplified and rational model. Following the establishment of the foundational model, the incorporation of age as a time-dependent covariate, along with initial age, sex, and place of origin as independent covariates, was undertaken to ascertain the impact of these variables on the progression of visual acuity. All analyses were implemented using SAS (version 9.4; SAS Institute) and R software (version 4.2.1; R Foundation for Statistical Computing).

Results

Characteristics of the Study Sample

Results in Table 1 show that among the 89,872 children and adolescents in Wave 1 (2016), 48,431 (53.89%) were male, 83,102 (92.47%) were from urban areas, and the mean age of the sample was 11.24 (SD 3.33) years. Among the 91,060 participants in Wave 2 (2017), 49,056 (53.87%) were male, 85,826 (94.25%) were from urban areas, and the mean age of the sample was 11.16 (SD 3.32) years. Among the 93,391 participants in Wave 3 (2018), 50,350 (53.91%) were male, 90,503 (96.91%) were from urban areas, and the mean age of the sample was 11.09 (SD 3.33) years. In Wave 4 (2019), 52,273 (53.97%) were male, 94,973 (98.06%) were from urban areas, and the mean age of the sample was 11.04 (SD 3.33) years.

Table . Characteristics of the children and adolescents in Wuhan, China, between Wave 1 in 2016 - Wave 4 in 2019.

Variables	Number of participants between Wave I (2016) and Wave IV (2019)			
	Wave 1 (n=89,872)	Wave 2 (n=91,060)	Wave 3 (n=93,391)	Wave 4 (96,849)
Sex, n (%)				
Male	48,431 (53.89)	49,056 (53.87)	50,350 (53.91)	52,273 (53.97)
Female	41,441 (46.11)	42,004 (46.13)	43,041 (46.09)	44,576 (46.03)
Age (years), mean (SD)	11.24 (3.33)	11.16 (3.32)	11.09 (3.33)	11.04 (3.33)
Grade, n (%)				
Grade 1	9161 (10.19)	9350 (10.27)	10,460 (11.2)	10,659 (11.01)
Grade 2	8878 (9.88)	9224 (10.13)	9253 (9.91)	10,649 (11)
Grade 3	8663 (9.64)	8966 (9.85)	9140 (9.79)	9170 (9.47)
Grade 4	7788 (8.67)	8609 (9.45)	8844 (9.47)	9190 (9.49)
Grade 5	7953 (8.85)	7672 (8.43)	8667 (9.28)	8878 (9.17)
Grade 6	7690 (8.56)	7972 (8.75)	7774 (8.32)	8698 (8.98)
Grade 7	7777 (8.65)	7757 (8.52)	7762 (8.31)	7735 (7.99)
Grade 8	7665 (8.53)	7713 (8.47)	7640 (8.18)	7715 (7.97)
Grade 9	7671 (8.54)	7513 (8.25)	7552 (8.09)	7479 (7.72)
Grade 10	5795 (6.45)	5722 (6.28)	5692 (6.1)	6005 (6.2)
Grade 11	5715 (6.36)	5718 (6.28)	5625 (6.02)	5701 (5.89)
Grade 12	5116 (5.69)	4844 (5.32)	4979 (5.33)	4944 (5.11)
Region, n (%)				
Rural	6770 (7.53)	5234 (5.75)	2888 (3.09)	1876 (1.94)
Urban	83,102 (92.47)	85,826 (94.25)	90,503 (96.91)	94,973 (98.06)

Prevalence of Poor Vision

Results in Table 2 show that the prevalence of poor vision in the study sample was 58.51% in 2016, 58.95% in 2017, 53.83% in 2018, and 54.79% in 2019. The prevalence of poor vision from 2016 to 2019 ranged from 12.71% to 12.32% for mild poor vision, from 20.46% to 21.05% for moderate poor vision, and from 25.34% to 21.42% for severe poor vision. Additional information by gender and region is presented in Table 2.

Table . Prevalence of poor vision among children and adolescents in Wuhan, China, from 2016 to 2019.

Vision condition	Number of participants between 2016 and 2019			
	2016 (N=82,477)	2017 (N=90,396)	2018 (N=91,548)	2019 (N=93,068)
Overall, n (%)				
Normal	34,224 (41.49)	37,107 (41.05)	42,266 (46.17)	42,079 (45.21)
Mild poor vision	10,480 (12.71)	10,389 (11.49)	9211 (10.06)	11,467 (12.32)
Moderate poor vision	16,873 (20.46)	18,145 (20.07)	17,622 (19.25)	19,589 (21.05)
Severe poor vision	20,900 (25.34)	24,755 (27.39)	22,449 (24.52)	19,933 (21.42)
Males, n (%)				
Normal	19,737 (44.13)	21,303 (43.77)	24,142 (48.88)	23,969 (47.67)
Mild poor vision	5422 (12.12)	5382 (11.06)	4838 (9.8)	5928 (11.79)
Moderate poor vision	9075 (20.29)	9596 (19.72)	9320 (78.87)	10,402 (20.69)
Severe poor vision	10,491 (23.46)	12,387 (15.45)	11,087 (22.45)	9980 (19.85)
Females, n (%)				
Normal	14,487 (38.37)	15,804 (37.87)	18,124 (42.99)	18,110 (42.32)
Mild poor vision	5058 (13.4)	5007 (12)	4373 (10.37)	5539 (12.94)
Moderate poor vision	7798 (20.66)	8549 (20.49)	8302 (19.69)	9187 (21.47)
Severe poor vision	10,409 (27.57)	12,368 (29.64)	11362 (26.95)	9953 (23.26)
Urban students, n (%)				
Normal	30,985 (40.76)	34,530 (40.54)	40,703 (45.91)	41,251 (45.19)
Mild poor vision	9616 (12.65)	9894 (11.62)	8932 (10.07)	11,252 (12.33)
Moderate poor vision	15,567 (20.48)	17,081 (20.06)	16,997 (19.17)	19,179 (21.01)
Severe poor vision	19,852 (26.11)	23,665 (27.79)	22,034 (24.85)	19,596 (21.47)
Urban males, n (%)				
Normal	17,757 (43.37)	19,727 (43.23)	23,159 (48.54)	23,450 (47.62)
Mild poor	4940 (12.07)	5106 (11.19)	4681 (9.81)	5817 (11.81)
Moderate poor vision	8307 (20.29)	8991 (19.7)	8997 (18.86)	10,162 (20.64)
Severe poor vision	9938 (24.27)	11,811 (25.88)	10,879 (22.8)	9816 (19.93)
Urban females, n (%)				
Normal	13,228 (37.71)	14,803 (37.44)	17,544 (42.84)	17,801 (42.35)
Mild poor vision	4676 (13.33)	4788 (12.11)	4251 (10.38)	5435 (12.93)
Moderate poor vision	7260 (20.7)	8090 (20.46)	8000 (19.54)	9017 (21.45)
Severe poor vision	9914 (28.26)	11854 (29.98)	11,155 (27.24)	9780 (23.27)
Rural students, n (%)				
Normal	3239 (50.16)	2577 (49.31)	1563 (54.23)	828 (46.26)
Mild poor vision	864 (13.38)	495 (9.47)	279 (9.68)	215 (12.01)
Moderate poor vision	1306 (20.23)	1064 (20.36)	625 (21.69)	410 (22.91)
Severe poor vision	1048 (16.23)	1090 (20.86)	415 (14.4)	337 (18.83)
Rural males, n (%)				
Normal	1980 (52.34)	1576 (51.96)	983 (58.82)	519 (50.19)
Mild poor vision	482 (12.74)	276 (9.1)	157 (9.4)	111 (10.74)
Moderate poor vision	768 (20.3)	605 (19.95)	323 (19.33)	240 (23.21)
Severe poor vision	553 (14.62)	576 (18.99)	208 (12.45)	164 (15.86)
Rural, females, n (%)				

Vision condition	Number of participants between 2016 and 2019			
	2016 (N=82,477)	2017 (N=90,396)	2018 (N=91,548)	2019 (N=93,068)
Normal	1259 (47.08)	1001 (45.64)	580 (47.89)	309 (40.87)
Mild poor vision	382 (14.29)	219 (9.99)	122 (10.07)	104 (13.76)
Moderate poor vision	538 (20.12)	459 (20.93)	302 (24.94)	170 (22.49)
Severe poor vision	495 (18.51)	514 (23.44)	207 (17.09)	173 (22.88)

Trajectories of Poor Vision

Figure 1 shows the three distinct trajectories that were identified based on the lowest absolute Bayesian Information Criteria value (−361460.5). Group 1 (n=29752, 27.4%; Avepp=0.76) was characterized by a low degree of poor vision with a reduction in first 2 years and was named the normal vision group (NVG). Group 2 (n=19328, 17.8%; Avepp=0.72) was characterized by a rapid increase in poor vision and was named the vision decline group (VDG). Group 3 (n=59505, 54.8%; Avepp=0.93) was characterized by a stable level of moderate poor vision and was named the moderate poor vision group (MPV).

Figure 1. Trajectories of poor vision among children and adolescents in Wuhan, China, 2016 - 2019. Dotted lines represent the actual value, solid lines represent the predicted value. NVG: normal vision group, VDG: vision decline group, MPV: moderate poor vision group.

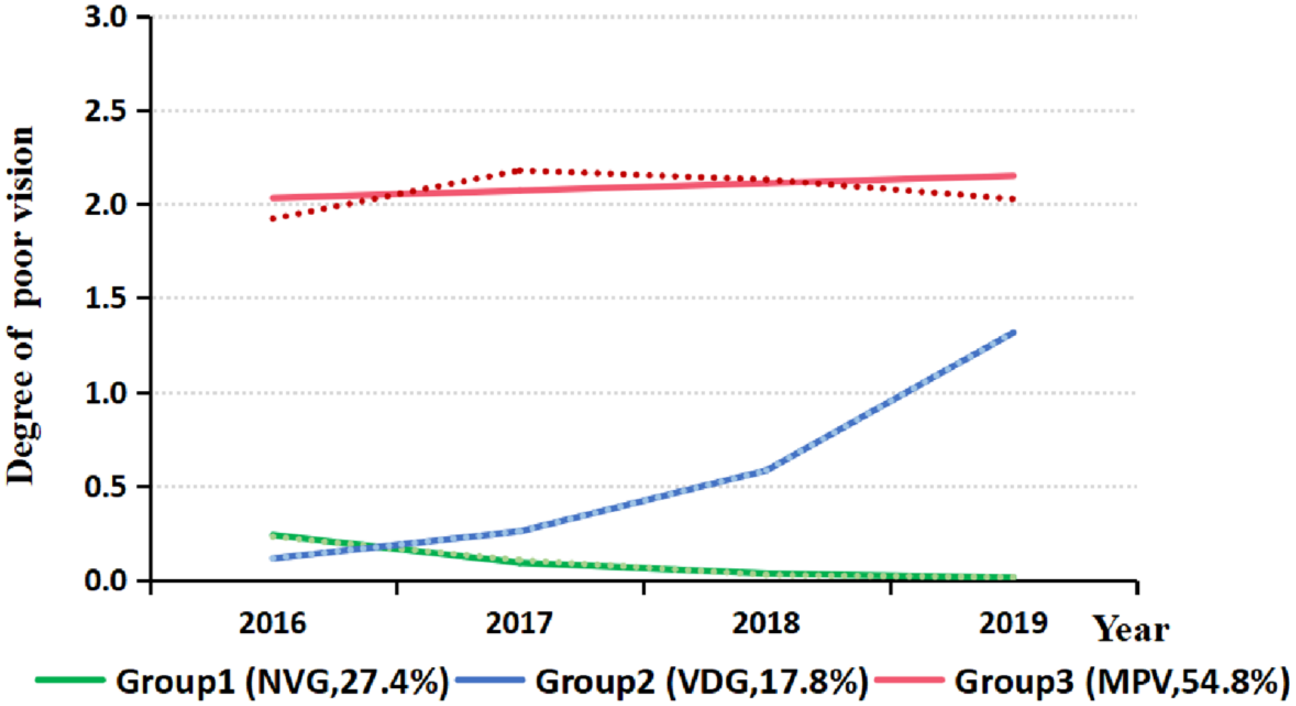


Figure 2 shows the trajectories of poor vision by sex and region, with a total of three trajectories identified. Compared to males, females had a higher proportion in Group 3 (57.8% vs 50.5%), and a lower proportion in Group 2 (17.8% vs 27.9%). Compared to rural areas, students in the urban areas had a higher proportion in Group 3 (55.3% vs 47.5%) and a lower proportion in Group 2 (17.2% vs 24.5%).

Figure 2. Trajectories of poor vision among children and adolescents in Wuhan, China, by sex and region between 2016-2019. Dotted lines represent the actual value, solid lines represent the predicted value. NVG: normal vision group; VDG: vision decline group; MPV: moderate poor vision group.

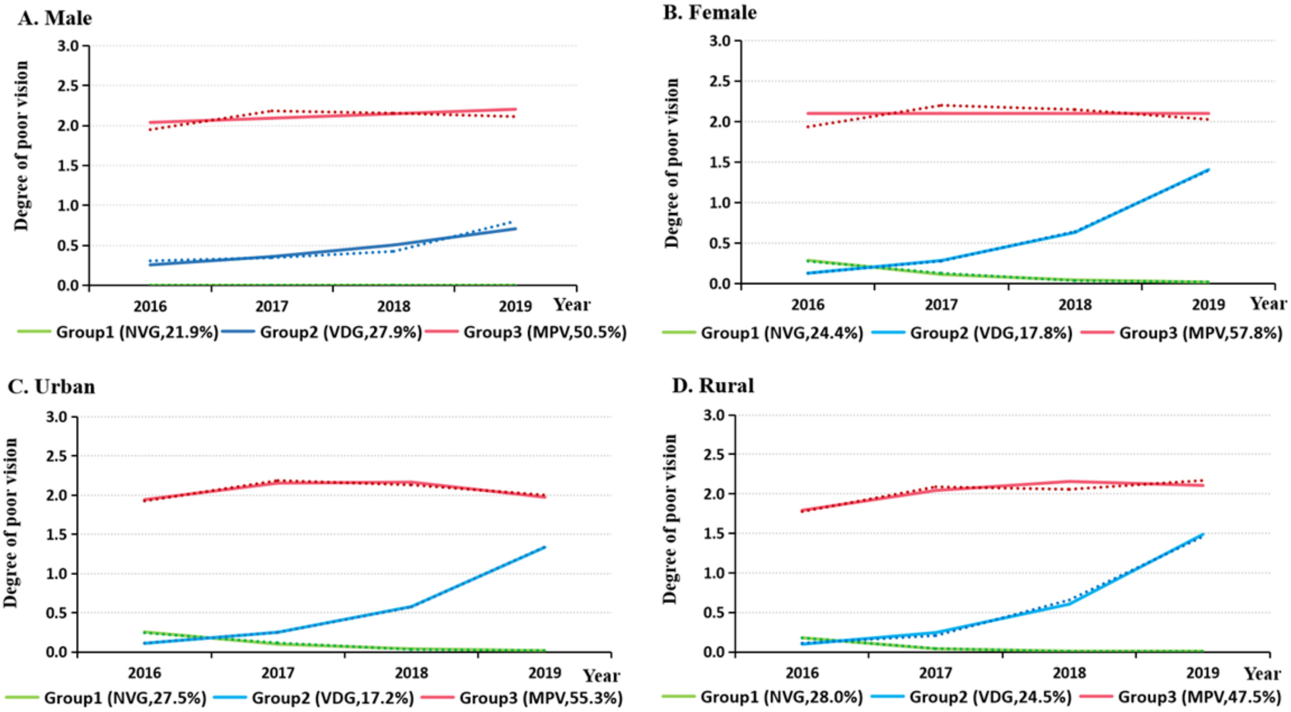
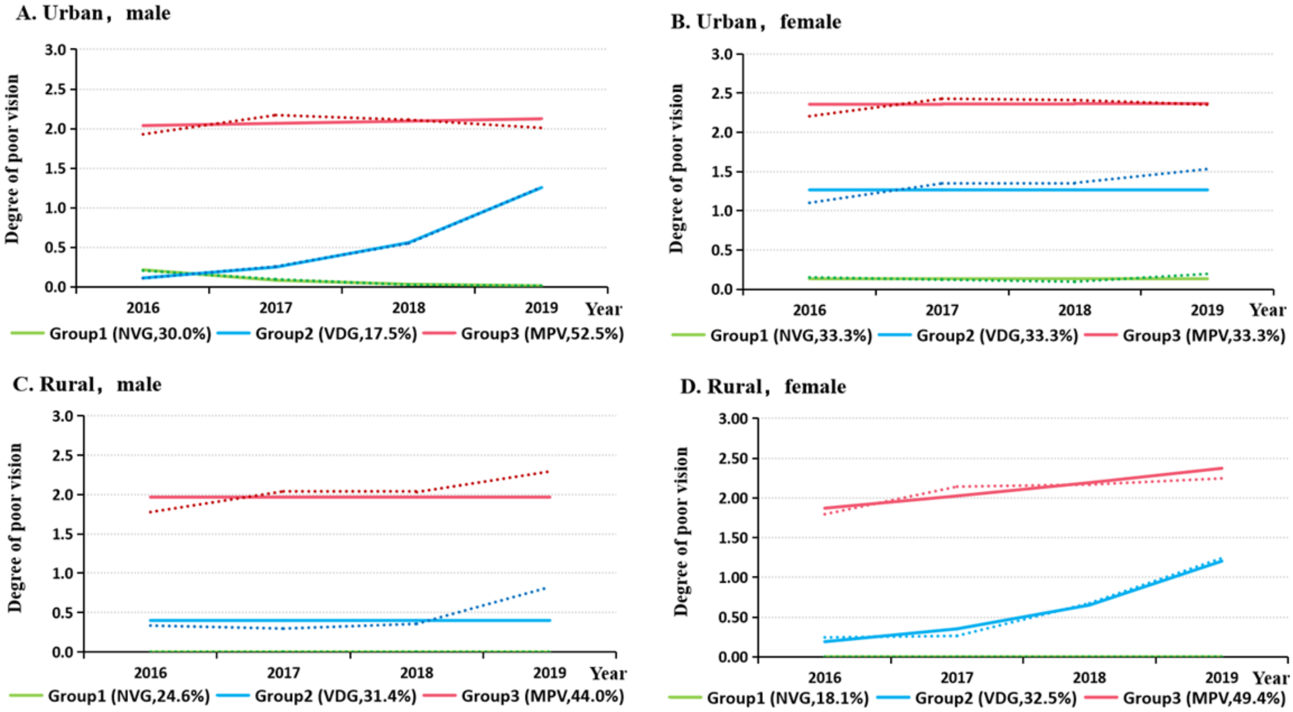


Figure 3 shows the trajectories of poor vision by the interaction of sex and region. The highest proportion in Group 3 was found among urban females (56.8%), while the lowest proportion was found in rural males (43.6%). The highest proportion in Group 2 was found among rural males (32.8%), followed by urban females (27.2%), rural females (25.8%), and urban males (17.5%).

Figure 3. Trajectories of poor vision among children and adolescents in Wuhan, China, 2016 - 2019, across sex and regions. Dotted lines represent the actual value, solid lines represent the predicted value. NVG: normal vision group, VDG: vision decline group, MPV: moderate poor vision group.



Factors Influencing Trajectory Membership and Progression

Data in Table 3 show that with increasing age, students were less likely to be categorized into Group 1 (β=−0.417, *P*<.001), and more likely to be categorized into Group 2 (β=.058,<.001). For group membership (Group 1 was used as reference), compared to boys, girls were more likely to be categorized into Group 2 (β=.597, *P*<.001) and Group 3 (β=.362, *P*<.001). Additionally, compared to urban students, students from rural area were less likely to be categorized into Group 2 (β=−.311, *P*<.001). Further, with increasing initial enrollment age, students were more likely to be categorized into Group 1 (β=.369, *P*<.001) and Group 3 (β=.096, *P*<.001).

Table . Factors influencing trajectory membership and progression of poor vision among children and adolescents in Wuhan, China between 2016 and 2019.

Group	Variables	β	*t* test (*df*)[a]	*P* value
Trajectories progression				
Group 1[b]	Age	−.417	23.730 (85,820)	<.001
Group 2[c]	Age	.058	58.250 (85,820)	<.001
Group 3[d]	Age	.002	0.208 (85,820)	.84
Group membership (ref: Group 1)				
Group 2	Sex, (ref:Boys)			
	Girls	.597	22.709 (49,894)	<.001
	Region (ref:Urban)			
	Rural	−.311	−6.973 (5,092)	<.001
	Initial age	.369	33.711 (84,929)	<.001
Group 3	Sex (ref:Boys)			
	Females	.362	9.338 (49,994)	<.001
	Region (ref:Urban)			
	Rural	−.007	−0.113 (5,092)	.910
	Initial age	.096	5.391 (84,929)	<.001

[a] 2-tailed t test was applied in the analysis.

[b] Group 1: normal vision group (NVG).

[c] Group 2: vision decline group (VDG).

[d] Group 3: moderate poor vision group (MPV).

Discussion

Principal Findings

This study investigated the progression of poor vision by analyzing data from a prospective cohort of primary and middle school students in Wuhan, China. To the best of our knowledge, this is the first study to use longitudinal data to identify the progression pattern of poor vision among children and adolescents. The findings of the study deepen our understanding of the development of poor vision and provide valuable evidence for devising and implementing intervention programs.

Population-based research indicates that poor vision in most school children results from uncorrected refractive errors [31,32]. Visual acuity has been widely used as a proxy measure for refractive error in children in previous studies [33-35] and is useful for estimating the prevalence of myopia in large populations that undergo vision screening when refraction is not feasible [36]. In this study, we found that the prevalence of poor vision among children and adolescents was 58.51% in 2016, 58.95% in 2017, 53.83% in 2018, and 54.79% in 2019, showing a decline in the latter two years. It is well-recognized that Chinese students are under greater academic pressure, spend more time on near-work activities and less time on outdoor activities, which may contribute to the high rates of poor vision [28]. The declines observed in the latter two years may be associated with the nationwide policy change regarding myopia prevention and control [29].

In this study, we identified 3 different development trajectories of poor vision among children and adolescents. Notably, more than half of the participants were categorized into the progression pattern of MPV during 2016 and 2019, amounting to 49,880 primary and middle school students in Wuchang district. If this pattern was applied to Wuhan City, approximately 589,135 primary and middle students would be fall into the MPV category. Meanwhile, our results show that 17.8% of the students enrolled were categorized into the VDG, which was characterized by a progression from normal vision to MPV. This finding suggests that the magnitude of vision loss is increasing every year. A study conducted by Ma et al [10] also showed that the prevalence rate of vision impairment was associated with grade and age in a nonlinear pattern. The underlying mechanisms of the rapid increasing pattern from normal to moderate poor vision may include high academic stress, long time of study, frequent use of electronic devices, lack of outdoor physical activities, among others [28,37-40]. These findings also highlight the urgent need for effective intervention and prevention programs to control poor vision development; potential actions include encouraging students to play outdoor sports, reducing their academic stress, and limiting time spent on electronic devices [41,42]. Further, we found that 27.4% of the participants maintained relatively normal vision during the study period. Future studies should investigate the characteristics of this group and explore the potential protective factors against poor vision.

This study also suggests that being girls are at a higher risk of poor vision than boys, with the average poor vision rate being 59.61% in girls compared to 53.89% in boys, consistent with previous studies [26,43]. Moreover, a study on nationally representative samples conducted by Li and Mo [39] reported that 64.49% of girls were myopic while 54.44% of the boys. This trend was further confirmed in the trajectory analysis, where a greater proportion of females were classified into the MPV group compared to males. Although fewer girls belonged to the VDG, female children and adolescents experienced a greater decline in vision than males. The findings of this study indicated that girls had a greater risk of being assigned to both the VDG and the MPV than to the NVG; this result may be attributed to the fact that girls may spend less time on outdoor activities and more time on near-work tasks than boys [28,44,45].

Further, study findings suggest that urban students had a higher risk of poor vision than youth from rural areas, with a higher prevalence of poor vision and a greater proportion in the MPV category. Compared to urban students, rural children, and adolescents were less likely to be classified into the VDG than the NVG. Previous studies have similarly demonstrated higher rates of poor vision among urban students than those in rural areas [46]. This phenomenon could result from higher academic pressure, more time spent using electronic devices, and less outdoor activities among urban students [47,48].

Limitations

Although this study included a large population, it had several limitations. First, the results of the standard for logarithmic visual acuity charts were used as the sole criterion to determine the visual acuity of participants, and the cycloplegic refraction was not measured, which may have introduced measurement errors. We used the term poor vision instead of myopia throughout the paper to avoid the overestimation. Second, this study was only performed in a single district in Wuhan, and caution should be exercised when generalizing the findings to other areas. Third, there is a lack of data on demographic variables and risk factors for poor vision, that limited further analysis.

Conclusions

This study identified 3 groups with distinct poor vision progression patterns, including normal vision, visual decline, and MPV. Furthermore, the prevalence of poor vision was higher in girls than boys and higher in urban than rural students. Therefore, particular attention should be paid to girls and urban students for the prevention and control of poor vision among primary and secondary school students in the future.

Acknowledgments

The study was supported by Hubei Provincial Outstanding Young Talent Project for Public Health (PI: Dr. Bin Yu), Fundamental Research Funds for the Cross Innovation Talent Project of Renmin Hospital of Wuhan University (PI: Lijuan Xu), and the Funding for Scientific Research Projects from Wuhan Municipal Health Commission (WY22A07, PI: Dr. Fang Li).

Data Availability

The datasets generated during and/or analyzed during this study are not publicly available due to the surveillance policies from the government but are available from the corresponding author on reasonable request.

Authors' Contributions

Data curation: LX, HL
Statistical analysis: HL, FL, TZ, JY
Supervision: HY
Writing – original draft: LX, HL and FL
Writing – review & editing: BY, LH
Author Lu He is the co-corresponding author of this paper and can be reached at: Dr. Lu He, Physical Examination Center, Renmin Hospital of Wuhan University, Wuhan, Hubei Zhang Road (formerly Ziyang Road) Wuchang District No. 99 Jiefang Road 238, Wuhan, China, 86 13545211012; rm003619@whu.edu.cn.

Conflicts of Interest

None declared.

References

1. The Lancet Child Adolescent Health. Vision for the future. Lancet Child Adolesc Health 2021 Mar;5(3):155. [doi: 10.1016/S2352-4642(21)00029-8] [Medline: 33609489]
2. Pan CW, Wu RK, Li J, Zhong H. Low prevalence of myopia among school children in rural China. BMC Ophthalmol 2018 Jun 11;18(1):140. [doi: 10.1186/s12886-018-0808-0] [Medline: 29890943]

3. Pinquart M, Pfeiffer JP. Psychological well-being in visually impaired and unimpaired individuals: a meta-analysis. Brit J Vis Imp 2011 Jan;29(1):27-45. [doi: 10.1177/0264619610389572]
4. Greguol M, Gobbi E, Carraro A. Physical activity practice among children and adolescents with visual impairment--influence of parental support and perceived barriers. Disabil Rehabil 2015;37(4):327-330. [doi: 10.3109/09638288.2014.918194] [Medline: 24828394]
5. Horvat-Gitsels LA, Cortina-Borja M, Rahi JS. Do adolescents with impaired vision have different intentions and ambitions for their education, career and social outcomes compared to their peers? Findings from the Millennium Cohort Study. Br J Ophthalmol 2023 Dec 18;108(1):159-164. [doi: 10.1136/bjo-2021-320972] [Medline: 36307166]
6. Liang J, Pu Y, Chen J, et al. Global prevalence, trend and projection of myopia in children and adolescents from 1990 to 2050: a comprehensive systematic review and meta-analysis. Br J Ophthalmol 2024 Sep 24:bjo-2024-325427. [doi: 10.1136/bjo-2024-325427] [Medline: 39317432]
7. Yekta A, Hooshmand E, Saatchi M, et al. Global prevalence and causes of visual impairment and blindness in children: a systematic review and meta-analysis. J Curr Ophthalmol 2022;34(1):1-15. [doi: 10.4103/joco.joco_135_21] [Medline: 35620376]
8. Rudnicka AR, Kapetanakis VV, Wathern AK, et al. Global variations and time trends in the prevalence of childhood myopia, a systematic review and quantitative meta-analysis: implications for aetiology and early prevention. Br J Ophthalmol 2016 Jul;100(7):882-890. [doi: 10.1136/bjophthalmol-2015-307724]
9. Transcript of the national health commission’s press conference. China Government Network, Propaganda. 2022. URL: http://www.nhc.gov.cn/xcs/s3574/202207/16833fe2630e49338195edd979b6d386.shtml [accessed 2023-06-05]
10. Ma Y, Zhang X, He F, et al. Visual impairment in rural and migrant Chinese school-going children: prevalence, severity, correction and associations. Br J Ophthalmol 2022 Feb;106(2):275-280. [doi: 10.1136/bjophthalmol-2020-317072] [Medline: 33127829]
11. Hu Y, Ding X, Guo X, Chen Y, Zhang J, He M. Association of age at myopia onset with risk of high myopia in adulthood in a 12-year follow-up of a Chinese cohort. JAMA Ophthalmol 2020 Nov 1;138(11):1129-1134. [doi: 10.1001/jamaophthalmol.2020.3451] [Medline: 32940622]
12. Chua SYL, Sabanayagam C, Cheung YB, et al. Age of onset of myopia predicts risk of high myopia in later childhood in myopic Singapore children. Ophthalmic Physiol Opt 2016 Jul;36(4):388-394. [doi: 10.1111/opo.12305] [Medline: 27350183]
13. Williams KM, Hysi PG, Nag A, Yonova-Doing E, Venturini C, Hammond CJ. Age of myopia onset in a British population-based twin cohort. Ophthalmic Physiol Opt 2013 May;33(3):339-345. [doi: 10.1111/opo.12042] [Medline: 23510367]
14. Gwiazda J, Hyman L, Dong LM, et al. Factors associated with high myopia after 7 years of follow-up in the Correction of Myopia Evaluation Trial (COMET) Cohort. Ophthalmic Epidemiol 2007;14(4):230-237. [doi: 10.1080/01658100701486459] [Medline: 17896302]
15. Li SM, Wei S, Atchison DA, et al. Annual Incidences and Progressions of Myopia and High Myopia in Chinese Schoolchildren Based on a 5-Year Cohort Study. Invest Ophthalmol Vis Sci 2022 Jan 3;63(1):8. [doi: 10.1167/iovs.63.1.8] [Medline: 34989760]
16. Wang SK, Guo Y, Liao C, et al. Incidence of and factors associated with myopia and high myopia in Chinese children, based on refraction without cycloplegia. JAMA Ophthalmol 2018 Sep 1;136(9):1017-1024. [doi: 10.1001/jamaophthalmol.2018.2658] [Medline: 29978185]
17. Communiqué of the seventh national population census of Wuhan city [Chinese]. Wuhan Bureau of Statistics, Census Column. 2021. URL: http://tjj.wuhan.gov.cn/ztzl_49/pczl/202109/t20210916_1779157.shtml [accessed 2023-06-05]
18. Standard no.GB/T 11533-2011 [Chinese]. National Standards Full Text Public System. 2011. URL: https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=A9F9E03A346211223DE34421A85CA1C8 [accessed 2023-06-05]
19. Ministry of Health of the PRC. Technical Standard for Physical Examination for Students: GB/T26343 — 2010: Standards Press of China; 2011. URL: https://www.codeofchina.com/standard/GBT26343-2010.html [accessed 2023-06-06]
20. Wei L, Tian J, Guo M, et al. Trajectories of sexual risk behaviors and the associated factors among young men who have sex with men in China. Front Public Health 2022;10:854616. [doi: 10.3389/fpubh.2022.854616] [Medline: 35387185]
21. Beltrame TB, de Oliveira AC, Manera F, Fernandes RC, Castilho BM, Höfelmann DA. Trajectory of blood pressure levels and weight gain in pregnant women: A group-based approach. Nutr Metab Cardiovasc Dis 2023 Oct;33(10):2019-2027. [doi: 10.1016/j.numecd.2023.06.006] [Medline: 37500346]
22. Haviland A, Nagin DS, Rosenbaum PR, Tremblay RE. Combining group-based trajectory modeling and propensity score matching for causal inferences in nonexperimental longitudinal data. Dev Psychol 2008 Mar;44(2):422-436. [doi: 10.1037/0012-1649.44.2.422] [Medline: 18331133]
23. Le Brocque RM, Dow BL, McMahon H, et al. The course of posttraumatic stress in Children: examination of symptom trajectories and predictive factors following admission to pediatric intensive care. Pediatr Crit Care Med 2020 Jul;21(7):e399-e406. [doi: 10.1097/PCC.0000000000002316] [Medline: 32224826]
24. Crothers A, Haeusler GM, Slavin MA, et al. Examining health-related quality of life in pediatric cancer patients with febrile neutropenia: factors predicting poor recovery in children and their parents. EClinicalMedicine 2021 Oct;40:101095. [doi: 10.1016/j.eclinm.2021.101095] [Medline: 34746716]

25. Kiadaliri A, Englund M. Variability in end-of-life healthcare use in patients with osteoarthritis: a population-based matched cohort study. Osteoarthr Cartil 2021 Oct;29(10):1418-1425. [doi: 10.1016/j.joca.2021.07.001]
26. Schuster AK, Krause L, Kuchenbäcker C, et al. Prevalence and time trends in myopia among children and adolescents. Dtsch Arztebl Int 2020 Dec 11;117(50):855-860. [doi: 10.3238/arztebl.2020.0855] [Medline: 33612155]
27. Yotsukura E, Torii H, Inokuchi M, et al. Current prevalence of myopia and association of myopia with environmental factors among schoolchildren in Japan. JAMA Ophthalmol 2019 Nov 1;137(11):1233-1239. [doi: 10.1001/jamaophthalmol.2019.3103] [Medline: 31415060]
28. Wang J, Ying GS, Fu X, et al. Prevalence of myopia and vision impairment in school students in Eastern China. BMC Ophthalmol 2020 Jan 2;20(1):2. [doi: 10.1186/s12886-019-1281-0] [Medline: 31898504]
29. The notice on the issue of implementation plan for comprehensive prevention and control of myopia in children and adolescents. Ministry of Education of the People’s Republic of China. 2018. URL: http://www.moe.gov.cn/srcsite/A17/moe_943/s3285/201808/t20180830_346672.html [accessed 2024-11-06]
30. Yang J, Huang A, Chen D. Characteristics of physical developmental trajectories of infants and children aged 0-3 years based on cluster trajectory model and influencing factors. Chinese Journal of Reproductive 2023;34(1):25-34.
31. Fan DSP, Lam DSC, Lam RF, et al. Prevalence, incidence, and progression of myopia of school children in Hong Kong. Invest Ophthalmol Vis Sci 2004 Apr;45(4):1071-1075. [doi: 10.1167/iovs.03-1151] [Medline: 15037570]
32. Goh PP, Abqariyah Y, Pokharel GP, Ellwein LB. Refractive error and visual impairment in school-age children in Gombak District, Malaysia. Ophthalmology 2005 Apr;112(4):678-685. [doi: 10.1016/j.ophtha.2004.10.048] [Medline: 15808262]
33. Tay MT, Au Eong KG, Ng CY, Lim MK. Myopia and educational attainment in 421,116 young Singaporean males. Ann Acad Med Singap 1992 Nov;21(6):785-791. [Medline: 1295418]
34. Au Eong KG, Tay TH, Lim MK. Education and myopia in 110,236 young Singaporean males. Singapore Med J 1993 Dec;34(6):489-492. [Medline: 8153707]
35. Xiang F, He M, Zeng Y, Mai J, Rose KA, Morgan IG. Increases in the prevalence of reduced visual acuity and myopia in Chinese children in Guangzhou over the past 20 years. Eye (Lond) 2013 Dec;27(12):1353-1358. [doi: 10.1038/eye.2013.194] [Medline: 24008929]
36. Sun HP, Li A, Xu Y, Pan CW. Secular trends of reduced visual acuity from 1985 to 2010 and disease burden projection for 2020 and 2030 among primary and secondary school students in China. JAMA Ophthalmol 2015 Mar;133(3):262-268. [doi: 10.1001/jamaophthalmol.2014.4899] [Medline: 25429523]
37. Shen DY, Li XH. Influencing factors of myopia among primary and secondary school students in Shenzhen. Chinese J School Health 2020;41:583-587.
38. Bian H. Investigation on the current situation of myopia among pupils in Baotou City and analysis of related factors. Int Eye Sci 2021;21:1834-1838.
39. Li H, Mo J. Myopia and risk factors among Chinese adolescents. Modern preventive medicine (14):2552-2557 [FREE Full text]
40. Guan H, Yu NN, Wang H, et al. Impact of various types of near work and time spent outdoors at different times of day on visual acuity and refractive error among Chinese school-going children. PLoS One 2019;14(4):e0215827. [doi: 10.1371/journal.pone.0215827] [Medline: 31026279]
41. Morgan IG, French AN, Ashby RS, et al. The epidemics of myopia: aetiology and prevention. Prog Retin Eye Res 2018 Jan;62:134-149. [doi: 10.1016/j.preteyeres.2017.09.004] [Medline: 28951126]
42. Li SM, Ran AR, Kang MT, et al. Effect of text messaging parents of school-aged children on outdoor time to control myopia: a randomized clinical trial. JAMA Pediatr 2022 Nov 1;176(11):1077-1083. [doi: 10.1001/jamapediatrics.2022.3542] [Medline: 36155742]
43. Yin C, Gan Q, Xu P, et al. Dietary patterns and associations with myopia in Chinese children. Nutrients 2023 Apr 18;15(8):1946. [doi: 10.3390/nu15081946]
44. Guo L, Yang J, Mai J, et al. Prevalence and associated factors of myopia among primary and middle school-aged students: a school-based study in Guangzhou. Eye (Lond) 2016 Jun;30(6):796-804. [doi: 10.1038/eye.2016.39] [Medline: 26965016]
45. Yin Y, Qiu C, Qi Y. Myopia in Chinese adolescents: its influencing factors and correlation with physical activities. Comput Math Methods Med 2022 Aug 24;2022:1-10. [doi: 10.1155/2022/4700325]
46. Huang J, Zhang Y, Du K, et al. Prevalence of poor vision and eyeglass wear among primary and high school students in urban and rural western China: a comparative study. Chin J Public 2020;36(6):849-852. [doi: 10.11847/zgggws1122103]
47. Pan CW, Ramamurthy D, Saw SM. Worldwide prevalence and risk factors for myopia. Ophthalmic Physiol Opt 2012 Jan;32(1):3-16. [doi: 10.1111/j.1475-1313.2011.00884.x] [Medline: 22150586]
48. Pan CW, Dirani M, Cheng CY, Wong TY, Saw SM. The age-specific prevalence of myopia in Asia: a meta-analysis. Optom Vis Sci 2015 Mar;92(3):258-266. [doi: 10.1097/OPX.0000000000000516] [Medline: 25611765]

Abbreviations

Avepp: average posterior probability
GBTM: Group-based trajectory model

GDP: gross domestic product
MPV: moderate poor vision
NVG: normal vision group
VDG: vision decline group

Edited by A Mavragani; submitted 23.09.23; peer-reviewed by L Marković, W Wang; revised version received 01.01.25; accepted 02.01.25; published 18.02.25.

Please cite as:
Xu L, Li H, Li F, Zhang T, Yan J, Yan H, He L, Yu B
Investigating the Trajectories of Poor Vision in Children and Adolescents in Wuhan, China From 2016 to 2019: Prospective Cohort Study
JMIR Public Health Surveill 2025;11:e53028
URL: https://publichealth.jmir.org/2025/1/e53028
doi:10.2196/53028

Incidence of Deliberate Self-Harm in Hong Kong Before and During the COVID-19 Pandemic: Population-Wide Retrospective Cohort Study

Luke Y F Luk[1,2*], BSc; Christie J Y Ching[3*], MPH; Tsz Fung Yip[1,2], BEng; Sunny C L Chan[3], MStat; Catherine Y T Lam[4]; Elizabeth T C Lam[5]; Serena R B Yue[4], MSc; Hoi Ching Pang[4]; Janet Y H Wong[6], PhD; Carlos K H Wong[2,7,8], PhD; Chak Kwan Tong[9], MBBS; Tafu Yamamoto[10], MBChB; Timothy H Rainer[3,11], MD; Abraham K C Wai[3,11,12], MD; Joshua W K Ho[1,2], PhD

[1]School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China (Hong Kong)
[10]Department of Accident & Emergency, Yan Chai Hospital, Hong Kong SAR, China (Hong Kong)
[11]Department of Accident & Emergency, Queen Mary Hospital, Hong Kong SAR, China (Hong Kong)
[12]Department of Accident & Emergency, The University of Hong Kong-Shenzhen Hospital, Shenzhen, China
[2]Laboratory of Data Discovery for Health Limited (D24H), Hong Kong Science Park, Hong Kong SAR, China (Hong Kong)
[3]Department of Emergency Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, G06, G/F, Jockey Club Building for Interdisciplinary Research, 5 Sassoon Road, Pokfulam, Hong Kong SAR, China (Hong Kong)
[4]Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China (Hong Kong)
[5]Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China (Hong Kong)
[6]School of Nursing & Health Sciences, Hong Kong Metropolitan University, Hong Kong SAR, China (Hong Kong)
[7]Department of Family Medicine and Primary Care, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China (Hong Kong)
[8]Department of Pharmacology and Pharmacy, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China (Hong Kong)
[9]Intensive Care Unit, Tuen Mun Hospital, Hong Kong SAR, China (Hong Kong)
[*]these authors contributed equally

Corresponding Author:
Abraham K C Wai, MD
Department of Emergency Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, G06, G/F, Jockey Club Building for Interdisciplinary Research, 5 Sassoon Road, Pokfulam, Hong Kong SAR, China (Hong Kong)

Abstract

Background: COVID-19 ended on May 5, 2023, and since then Hong Kong reported increased mental distress, which was speculated to be from the policies implemented during the pandemic. Despite this, longitudinal surveillance of deliberate self-harm (DSH) incidences throughout the pandemic in Hong Kong remained insufficient.

Objective: The objective of this study was to outline the changes in DSH incidences before and during the COVID-19 pandemic in Hong Kong, with respect to sex, age, and co-occurring mental health issues.

Methods: A quasi-experiment was conducted using an interrupted time series design to estimate the impact of the pandemic on DSH-related emergency department (ED) visits. This design enabled the estimation of DSH-related ED visits based on prepandemic data from 2016 to 2019, assuming the pandemic had not occurred, and allowed for a comparison with observed DSH-related ED visits during the pandemic. The descriptive results were reported as the observed monthly DSH-related ED visits and observed incidence ratios during the pandemic. Afterwards, a negative binomial model was fitted to the prepandemic data (2016 - 2019) and adjusted for temporal trends, seasonality, and population variation to estimate the expected monthly DSH-related ED visits and adjusted incidence ratios (aIRs).

Results: Between January 2016 and December 2022, a total of 31,893 DSH episodes were identified. Initial descriptive analysis showed a significant difference in demographic characteristics (sex) and clinical characteristics (death within 28 d, diagnoses of co-occurring mental health issues, public assistance pay code, and triage level). Subsequent interrupted time-series analysis demonstrated significantly increasing trends in comparison with the prepandemic period. As reported in the aIRs among young adult males (aIR in 2020=1.34, *P*=.002; 2021: aIR=1.94, *P*<.001; and 2022: aIR=2.53, *P*<.001), adult males (aIR in 2020=1.58, *P*<.001; 2021: aIR=2.64, *P*<.001; and 2022: aIR=3.13, *P*<.001), adult females (aIR in 2020=1.13, *P*=.01; 2021: aIR=1.52, *P*<.001; and 2022: aIR=1.64, *P*<.001), and older male adults (aIR in 2020=1.53, *P*<.001; 2021: aIR=2.37, *P*<.001; and 2022: aIR=3.01, *P*<.001).

Conclusions: The average annual DSH-related ED visits increased during the pandemic period. Therefore, there is a need to raise awareness for such vulnerable groups in Hong Kong to prepare for postpandemic spillover.

(JMIR Public Health Surveill 2025;11:e57500) doi:10.2196/57500

KEYWORDS

emergency department; COVID-19; deliberate self-harm; mental health; self-harm; self-injury; self-violence; Hong Kong; SARS-Cov-2

Introduction

During the early phase of COVID-19 pandemic, Hong Kong experienced a significant increase in mental health distress within the community [1-3]. Surveys indicated that an alarming 53.9% (299/555) of participants exhibited moderate to very high levels of psychological distress [4]. In response, the Hong Kong Chief Executive allocated HK $300 million (the conversion rate at the latest revision date of the Chief Executive's 2020 policy address [November 25, 2020] was US $0.129 per unit HKD) to raise mental health awareness [5]. An advisory committee subsequently developed a comprehensive, community-based initiative. This initiative included an education and health promotion program called "Shall We Talk," emotional support workshop training for school staff, mental health first aid training for service professionals, and the deployment of suicide bereavement liaison officers. Despite these efforts, there were ongoing concerns about the sustainability of these resources [6].

The association between stress and deliberate self-harm (DSH) was well-documented in the literature [7-11]. However, recent local studies on DSH risk factors, conducted before the pandemic, identified factors such as sex, age, social welfare status, and co-occuring mental health disorders (ie, depression, bipolar disorder, personality disorder, and substance misuse) [12]. The literature from other countries on this topic is limited and sometimes contradictory, with mixed definitions of self-harm and varying conclusions about youth as a risk factor [13-17]. This highlights the need for longitudinal surveillance to better understand the pre-existing risk factors of DSH.

One of the major challenges in emergency departments (EDs) is reducing waiting times without compromising the quality of health care, which includes preventing readmission and reducing health care costs. The Hospital Authority in Hong Kong has pledged to limit the waiting time for emergency (triage category 2) patients to 15 minutes, and for urgent (triage category 3) patients to 30 minutes [18]. However, by May 2022, over 1500 patients were still waiting for admissions in ED [19]. This underscores the necessity for effective health policies to manage DSH-related ED visits.

The aim of this study was to describe the impact of COVID-19 on DSH incidences in Hong Kong through a longitudinal analysis of DSH-related ED visits before and during the pandemic. This study was stratified by sex, age, and co-occurring mental health issues. The findings from this study justify the need for further surveillance to monitor a possible spill-over effect post the COVID-19 pandemic.

Methods

Study Overview and Settings

This territory-wide cohort study analyzed data from January 2020 to December 2022, which was further categorized into pandemic months (January 2020 to December 2022) and the prepandemic months (January 2016 to December 2019).

Data Sources and International Classification of Diseases, Ninth Revision, Clinical Modification Codes

Data were sourced from the Clinical Data Analysis and Reporting System (CDARS), an administrative clinical database managed by the Hospital Authority of the Hong Kong Special Administrative Region of China [20], encompassing ED data from 18 public hospitals.

DSH and mental health issues were identified using both nursing codes for self-inflicted injury and the *ICD-9-CM* (*International Classification of Diseases, Ninth Revision, Clinical Modification*) diagnosed by clinicians (Table S1 in Multimedia Appendix 1).

In addition, yearly and midyearly population counts for sex and age groups were obtained from the government in Hong Kong [21].

Outcome Measures

This study defined DSH according the National Institute for Clinical Excellence guidelines, which describe it as "self-poisoning or injury, irrespective of the apparent purpose of the act" [22]. This guideline was adopted due to the absence of patients' intent in the medical records from CDARS.

Other key definitions included age groups, public assistance pay codes, triage categories, and mental health issues. Age was categorized into teenagers (12 - 17 y), young adults (18 - 24 y), adults (25 - 64 y), and older adults (65 - 84 y). Public assistance pay codes referred to the recipients of the Hospital Authority medical fee waiver schemes (Residential Care Service Voucher, Comprehensive Social Security Assistance, and Old Age Living Allowance) [23]. Triage categories 1 to 3 were defined in Table S2 in Multimedia Appendix 1 [18]. Mental health issues included mood disorders, neurosis, alcohol and drug-associated mental health conditions, and other psychotic illnesses and developmental disorders. These diagnoses were identified by *ICD-9-CM* codes, as defined in Table S1 in Multimedia Appendix 1.

Data Analysis

First, a descriptive analysis of the outcome measures was conducted to identify statistical significance before and during

COVID-19. Fisher exact and Mann-Whitney *U* tests were used to compare DSH incidences, sex, age groups, and co-occurring mental health issues.

Next, a quasi-experiment was conducted using an interrupted time series (ITS) design to estimate the impact of the COVID-19 pandemic on DSH-related ED visits. Quasi-experiments estimate causal relationships between different factors and outcomes when random assignment is not possible. The ITS design enabled the estimation of DSH-related ED visits during the pandemic period (2020 to 2022) based on prepandemic data (2016 to 2019). This allowed for a comparison between the observed DSH-related ED visits and the expected DSH-related ED visits in the absence of the pandemic.

Segmented regression is a method commonly used in ITS, typically involves assuming linear trends and identifying significant break points. Given the complexity of COVID-19's impact on self-harm ED visits, a nonsegmented regression model was chosen, fitted with prepandemic data instead of a segmented-regression model fitted with all data. Self-harm–related ED visits were aggregated into monthly counts to ITS analysis for each group. This analysis was done by fitting prepandemic data to the negative binomial models, with a linear term for trend, and optionally with Fourier terms to account for seasonality, using population terms as an offset. The candidate models can be described by:

$$E[y]=\exp(\beta0+\beta trend\cdot xt+\sum k=1n[\beta1,k\cdot\sin(2\pi k\cdot xt12)+\beta2,k\cdot\cos(2\pi k\cdot xt12)])\cdot population$$

Or

$$E[y]=\exp\ (\beta0+\beta trend\cdot xt)\cdot population$$

The dependent variable y represents the monthly count of self-harm incidents for a particular subgroup. Where xt is the number of months from the start of the study period, n is the highest degree of Fourier terms, β0βtrend is the prepandemic slope of incidence rate, β1,kβ2,k are coefficients for seasonality. The inclusion of the Fourier term and its degree n were guided by the Akaike Information Criterion (AIC). The final model is the one with the minimal AIC value, ensuring the best balance between model fit and complexity.

The findings from this study were presented as observed incidence ratio (oIR) and adjusted incidence ratio (aIR). The respective formulas were as follows:

$$oIR=\text{ObservedannualDSHrelatedEDvisitsPrepandemicannualDSHrelatedEDvisits}$$

$$aIR=\text{ObservedannualDSHrelatedEDvisitrateExpectedannualDSHrelatedEDvisitrate}$$

In the formulas, the expected annual DSH-related ED visits were estimated using negative binomial models and accounted for trends, seasonal variations, and population changes. Annual counts were estimated by summing monthly counts, assuming the population remained constant. The aIR compares the observed number of DSH-related ED visits during the pandemic to the expected number based on prepandemic data, providing an estimate of the COVID-19 pandemic's impact on DSH-related ED visits. Time series graphs were then generated to illustrate these trends. All *P* values less than .05 were considered statistically significant. All statistical analyses were performed using RStudio (version 4.1.0; R Foundation for Statistical Computing).

Ethical Considerations

The study protocol received approval from the institutional review board of the Hospital Authority, Hong Kong West Cluster (reference UW 20 - 112; Multimedia Appendix 2). Informed consent was waivered, and all the data were anonymized.

Results

Overview

A total of 31,893 cases of DSH-related ED visits were identified during the study period. Initial descriptive analysis compared demographic characteristics between the prepandemic and pandemic periods (Table 1). The total number of ED visits with DSH increased by 1175 out of 15,359 (7.7%) patients. There were significantly higher DSH-related ED visits in males (*P*<.001), co-occurring mental health issues diagnosis (*P*<.001), triage category 3 (urgent), and 4 (semiurgent) (*P*<.001) during the pandemic. Conversely, there were significantly less deaths within 28 days (*P*<.001) and less patients with public assistance pay code during the pandemic (*P*=.046). These trends were further explored with subgroup analysis using ITS models (Table 2, Figures 1-8). Further subgroup analysis of co-occurring mental health issues diagnoses were explored in Table S2 in Multimedia Appendix 1.

Table . Demographic characteristics in patients with deliberate self-harm who visited the emergency department before and during the pandemic (total N=31,893).

Characteristic		Time period		*P* value
		Prepandemic (n=15,359)	Pandemic (n=16,534)	
Age (years), median (IQR)		38 (25)	39 (30)	.60[a]
Sex, n (%)				<.001[b]
	Female	8184 (53)	8258 (50)	
	Male	7175 (47)	8276 (50)	
Deaths within 28 days, n (%)				<.001[b]
	No	15,102 (98)	16,352 (99)	
	Yes	257 (2)	182 (1)	
Co-occurring mental health issues diagnosis, n (%)				<.001[b]
	No	5991 (39)	4878 (30)	
	Yes	9368 (61)	11,656 (71)	
Public assistance pay code, n (%)				.046[b]
	No	12,060 (79)	13,134 (79)	
	Yes	3299 (22)	3400 (21)	
Triage, n (%)				<.001[a]
	Category 1 (critical)	1885 (12)	1329 (8)	
	Category 2 (emergent)	5944 (39)	5450 (33)	
	Category 3 (urgent)	6429 (42)	8539 (52)	
	Category 4 (semiurgent)	1055 (7)	1169 (7)	
	Category 5 (nonurgent)	29 (0)	30 (0)	

[a] Mann-Whitney *U* test.

[b] Fisher exact test.

Table . The number of average yearly visits, observed incidence ratio, and adjusted incidence ratio in deliberate self-harm–related emergency department visits between 2016 and 2022.

Patient characteristic	Years									
	2016 - 2019	2020			2021			2022		
	Average yearly visits, n (%)	oIR[a]	aIR[b]	*P* value	oIR	aIR	*P* value	oIR	aIR	*P* value
Female teenager (12 - 17 years old)										
Total	251 (100)	1.71	0.78	.98	3.26	1.10	.27	2.90	0.73	.95
All mental health issues	156 (62)	1.89	0.86	.86	3.99	1.34	.06	3.31	0.82	.81
Alcohol and drug associated mental health conditions	14 (6)	2.69	1.50	.11	4.58	2.07	.05	3.78	1.39	.28
Male teenager (12 - 17 years old)										
Total	77 (100)	1.09	0.92	.70	1.71	1.36	.07	1.65	1.25	.21
All mental health issues	37 (48)	1.47	1.07	.35	2.67	1.73	.04	2.50	1.47	.15
Alcohol and drug associated mental health conditions	6 (8)	2.24	2.46	.01	3.04	3.52	.01	4.16	5.13	.02
Female young adult (18 - 24 years old)										
Total	287 (100)	1.02	1.06	.25	1.38	1.48	.002	1.57	1.71	<.001
All mental health issues	149 (52)	1.21	1.24	.04	1.53	1.60	.005	1.83	1.91	.004
Alcohol and drug associated mental health conditions	47 (16)	1.16	1.64	.008	1.61	2.65	.001	1.91	3.58	.001
Male young adult (18 - 24 years old)										
Total	177 (100)	0.93	1.34	.002	1.15	1.94	<.001	1.30	2.53	<.001
All mental health issues	107 (61)	0.96	1.48	.01	1.37	2.53	<.001	1.61	3.51	<.001
Alcohol and drug associated mental health conditions	66 (37)	1.08	2.04	.001	1.68	4.08	<.001	2.08	6.38	<.001
Female adult (25 - 64 years old)										
Total	1316 (100)	0.94	1.13	.01	1.15	1.52	<.001	1.13	1.64	<.001

Patient characteristic		Years									
		2016-2019	2020			2021			2022		
		Average yearly visits, n (%)	oIR[a]	aIR[b]	*P* value	oIR	aIR	*P* value	oIR	aIR	*P* value
	All mental health issues	765 (58)	0.96	1.27	<.001	1.26	1.89	<.001	1.23	1.23	<.001
	Alcohol and drug associated mental health conditions	320 (24)	1.20	1.88	<.001	1.86	3.55	<.001	1.88	4.35	<.001
Male adult (25-64 years old)											
	Total	1289 (100)	1.17	1.58	<.001	1.72	2.64	<.001	1.78	3.13	<.001
	All mental health issues	915 (71)	1.24	1.83	<.001	2.00	3.49	<.001	2.09	4.28	<.001
	Alcohol and drug associated mental health conditions	713 (55)	1.35	2.07	<.001	2.32	4.27	<.001	2.39	5.26	<.001
Female older adult (65-84 years old)											
	Total	128 (100)	1.25	1.04	.33	1.54	1.19	.15	1.81	1.32	.08
	All mental health issues	59 (46)	1.27	1.27	.09	1.64	1.63	.02	2.14	2.13	.002
	Alcohol and drug associated mental health conditions	14 (11)	1.85	1.29	.23	3.78	2.27	.07	3.93	2.05	.16
Male older adult (65-84 years old)											
	Total	177 (100)	1.37	1.53	<.001	2.03	2.37	<.001	2.43	3.01	<.001
	All mental health issues	104 (59)	1.60	1.71	<.001	2.57	2.84	<.001	3.19	3.66	<.001
	Alcohol and drug associated mental health conditions	70 (40)	1.92	2.11	<.001	3.35	3.83	<.001	4.10	4.93	<.001

[a] oIR: observed incidence ratio.

[b] aIR: adjusted incidence ratio.

Figure 1. Subgroup analysis of female teenagers (12 - 17 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.

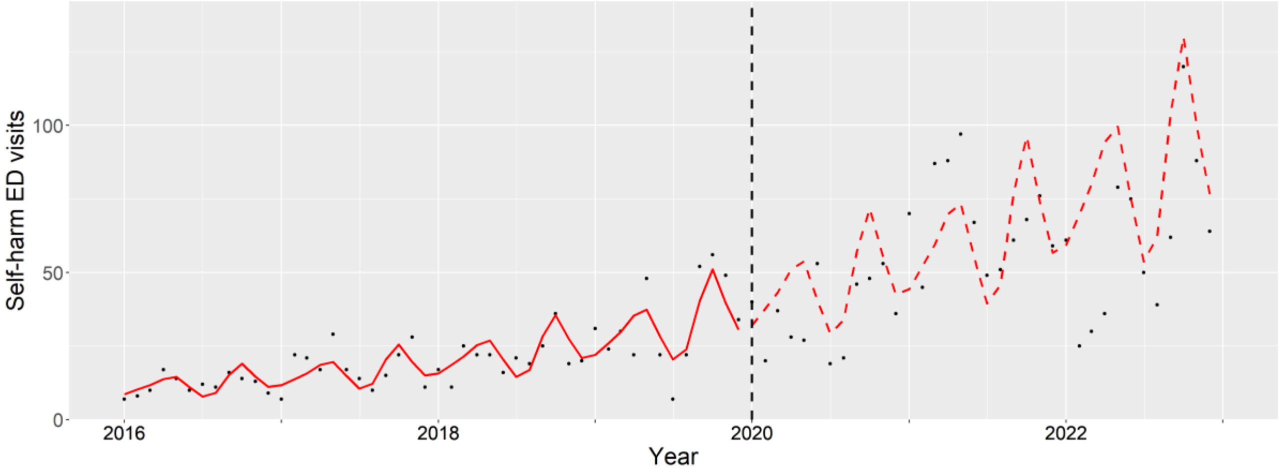


Figure 2. Subgroup analysis of male teenagers (12 - 17 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.

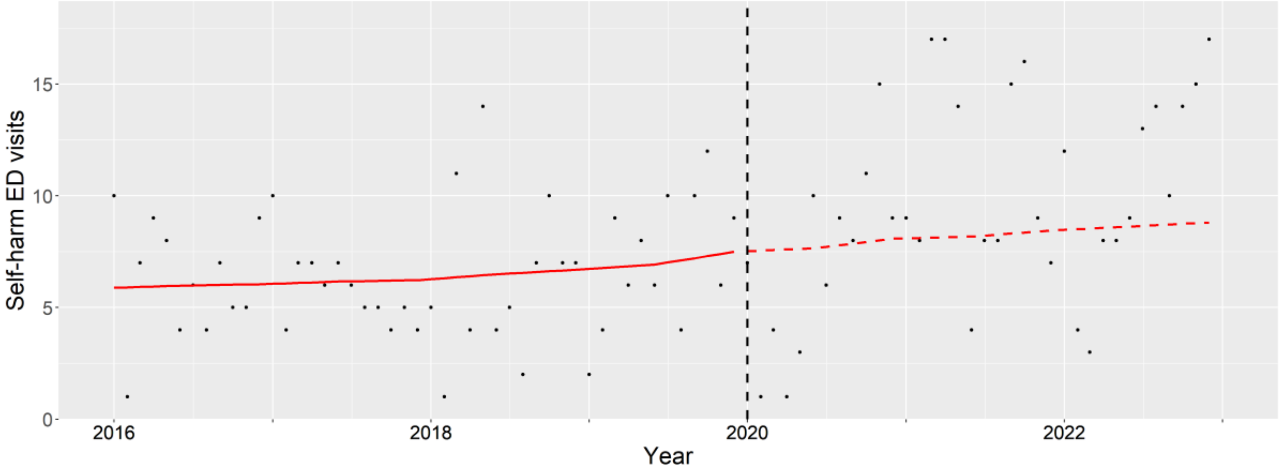


Figure 3. Subgroup analysis of female young adults (18 - 24 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.

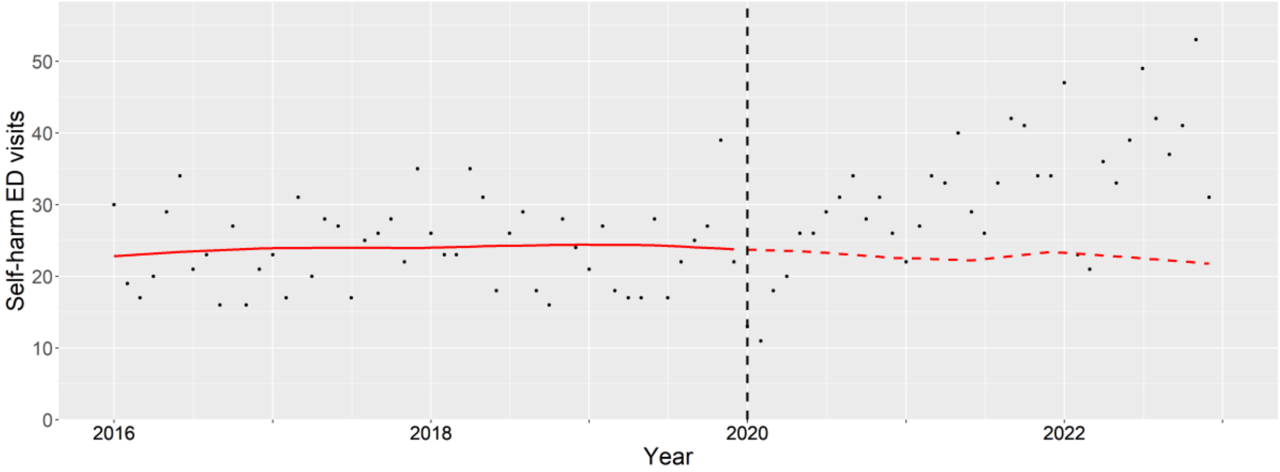


Figure 4. Subgroup analysis of male young adults (18 - 24 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.

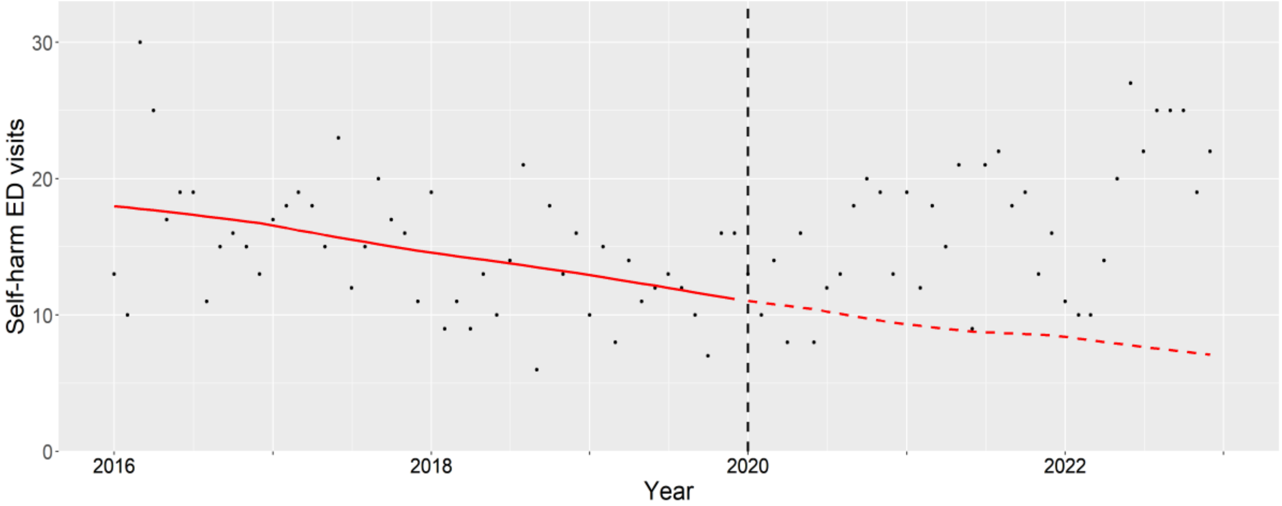


Figure 5. Subgroup analysis of female adults (25 - 64 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.

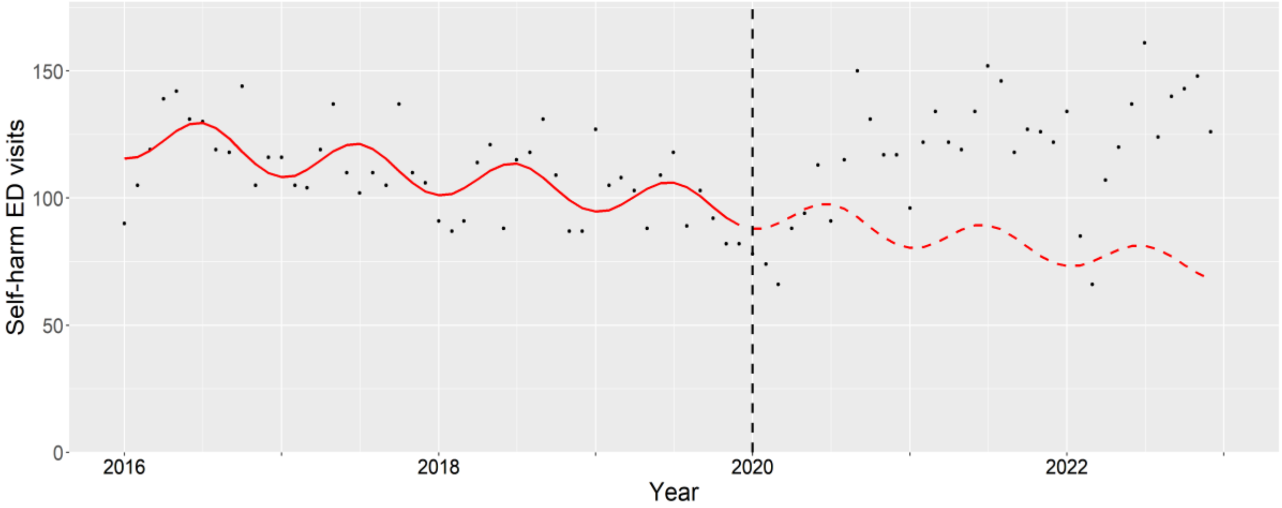


Figure 6. Subgroup analysis of male adults (25 - 64 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020).

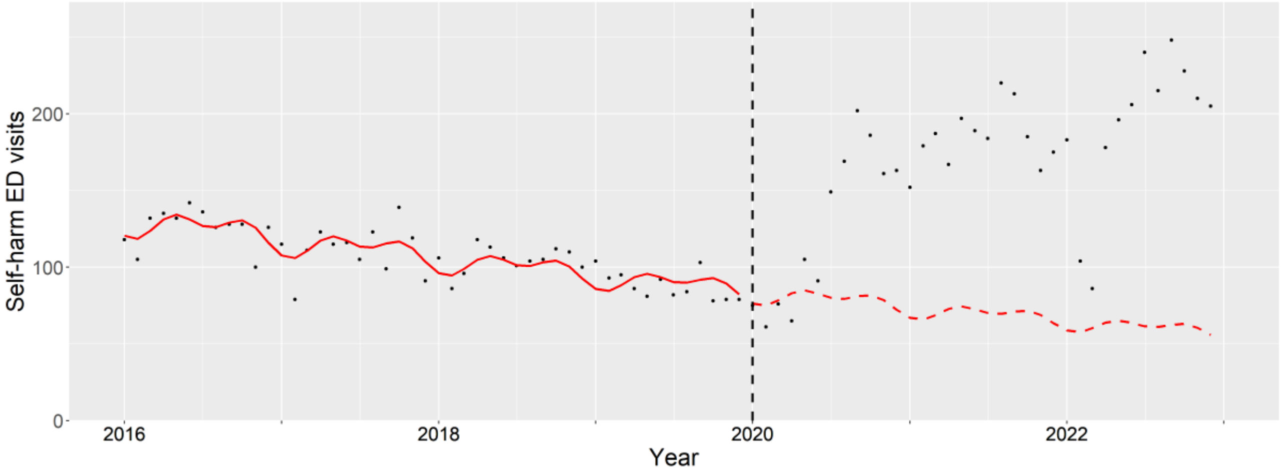


Figure 7. Subgroup analysis of female older adults (65 - 84 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.

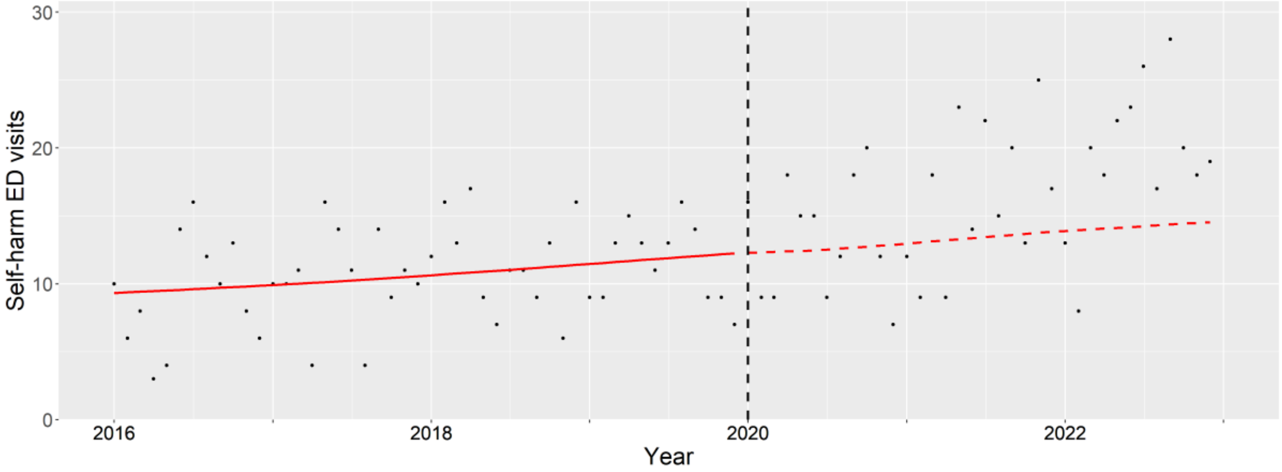
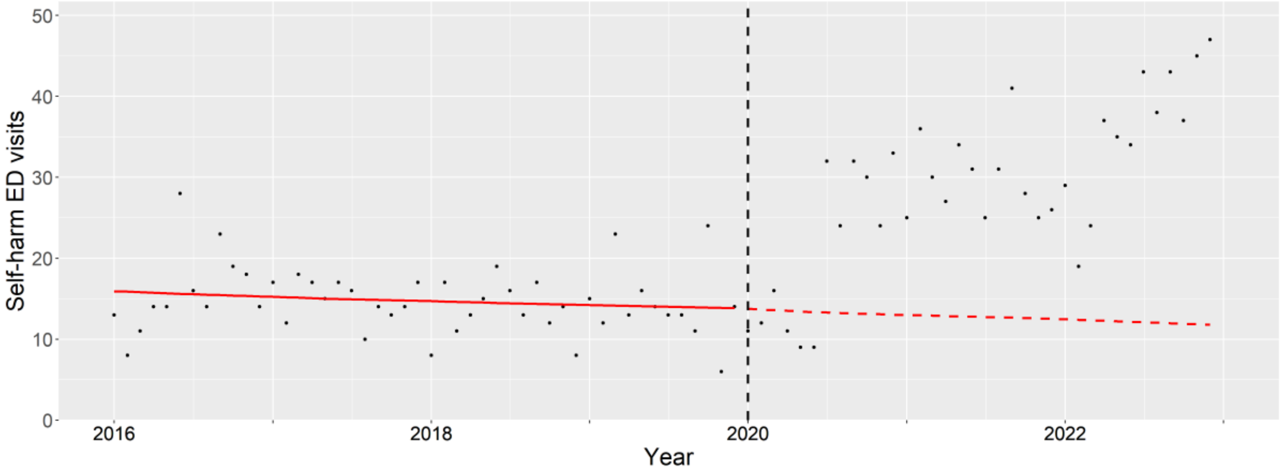


Figure 8. Subgroup analysis of male older adults (65 - 84 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.



Teenagers

For teenage females, the oIRs and aIR showed no difference (2020: *P*=.98, 2021: *P*=.27, and 2022: *P*=.95), as reflected by the overlap of the expected and observed trends (Figure 1).

For teenage males, the aIRs showed no difference (2020: *P*=.70, 2021: *P*=.07, and 2022: *P*=.21), as highlighted by the overlap of the observed and expected trends during the study period (Figure 2). However, the aIRs among those with alcohol and drug-associated mental health conditions were significantly different and increased during the pandemic (2020: *P*=.01, 2021: *P*=.01, and 2022: *P*=.02).

Young Adults

For female young adults, the observed trend was higher than expected during the pandemic (Figure 3). The highest aIR were observed among those with alcohol and drug-associated mental health conditions, with significant difference (2020: *P*=.008, 2021: *P*=.001, and 2022: *P*=.001).

For male young adults, the aIRs between 2020 and 2022 all significantly increased during the pandemic (2020: *P*=.002, 2021: *P*<.001, and 2022: *P*<.001). Further illustrated when the observed trend was higher than the expected trend during the pandemic (Figure 4). Those with alcohol and drug associated mental health conditions had the highest aIRs with statistical significance (2020: *P*=.001, 2021: *P*<.001, and 2022: *P*<.001).

Adults

The aIRs among adult females significantly increased during the pandemic (2020: *P*=.01, 2021: *P*<.001, and 2022: *P*<.001). This was highlighted when the expected trend decreased while the observed trend increased during the pandemic (Figure 5). The highest aIRs were seen in patients with co-occurring alcohol and drug associated mental health conditions with statistical significance (2020: *P*<.001, 2021: *P*<.001, and 2022: *P*<.001).

For adult males, the aIRs significantly increased during the pandemic (2020: *P*<.001, 2021: *P*<.001, and 2022: *P*<.001). As illustrated when the expected trend decreased while the observed trend increased during the pandemic (Figure 6). Significantly highest aIRs were observed among those with co-occurring alcohol and drug associated mental health condition (2020: *P*<.001, 2021: *P*<.001, and 2022: *P*<.001).

Older Adults

During 2021 and 2022, older female adults with mental health issues exhibited significantly higher aIRs (2021: *P*=.02, 2022: *P*=.002). These rates were elevated compared with the expected levels during 2021 and 2022 but declined during 2020 (Figure 7).

For older male adults, the aIRs significantly increased throughout the pandemic period (2020: *P*<.001, 2021: *P*<.001, and 2022: *P*<.001). Further underscored by the divergence between the decreasing expected rates and the increasing observed rates during the pandemic (Figure 8). The most pronounced aIRs with statistical significance were among found in older male adults with co-occurring alcohol and drug associated mental health conditions (2020: *P*<.001, 2021: *P*<.001, and 2022: *P*<.001).

Discussion

Principal Results

The findings from this study revealed differences in DSH incidences before and during the pandemic, with a notable rise among males and individuals with co-occurring alcohol and drug-related mental health issues.

Previous studies conducted before the pandemic also reported higher rates of DSH among males, which contrasts with the trends observed during our study period [12,24]. This discrepancy may be influenced by cultural factors [12]. For instance, past research has shown that while Chinese males in Hong Kong demonstrate health-seeking behavior, they often hesitate to consult doctors, preferring over-the-counter medications instead [25]. Such tendencies could be exacerbated by the mental distress linked to the isolation policies, as noted in another Hong Kong-based study [26]. However, further research is essential to elucidate the impact of COVID-19 policies on the rising incidences of male DSH in Hong Kong.

In addition, the increased incidences of DSH among individuals with alcohol and drug-related mental health issues observed in this study align with local government reports. The Narcotics Division of the Security Bureau in Hong Kong reported a surge in the use of cocaine, cannabis, ketamine, and heroin combined with methamphetamine during the pandemic [27]. In April 2020, the Centre of Health Protection in Hong Kong also noted a 5.5% increase in alcoholic consumption since the outbreak [28]. This increase in substance use may have been fueled by misinformation circulated on social media, suggesting that alcohol or smoking could protect against COVID-19 [29]. Despite the well-established link between self-harm and substance misuse [30], further research is needed to determine the association between these behaviors and COVID-19 policies.

Several additional patterns emerged from this study. First, there was a decline in DSH-related ED visits in early 2020 and 2022 as well as peaking staying times (Multimedia Appendices 3-6), consistent with the findings from both local and international studies [31-33]. This decline may be attributed to the stigma surrounding mental health issues, which can complicate health-seeking behavior, delay necessary care, and result in more severe DSH presentation requiring readmission to the ED [32,34]. Second, a significant decrease in patients requiring social security suggests that patients performing DSH may come from higher socioeconomic backgrounds. However, socioeconomic profiles varied across epidemic waves [35], indicating the need for further research to validate these findings. Finally, patients performing DSH might have been triaged to less urgent categories due to the changes in ED policies aimed at managing overcrowded cases of infection. These additional findings suggest that patients performing DSH had reduced access to health care services during COVID-19, potentially leading to worsening conditions and long-term repercussions postpandemic.

Given the empirical evidence from this study and existing literature, continued surveillance post the COVID-19 pandemic is essential to determine whether a spillover effect occurred. This need is particularly pressing in light of recent changes in the Hong Kong government's policies such as the announcement of a liquor tax cut to promote trade in 2024 [36]. Ongoing research and surveillance will be crucial in developing targeted interventions, guiding policy-making, and enhancing mental health services. By closely monitoring these trends, we can better address the mental health challenges posed by the pandemic and its aftermath, ultimately improving outcomes for vulnerable populations.

Strengths and Limitations

This study has several strengths and limitations. A major strength was the large sample size, which enhances reliability of the findings. In addition, the trend, seasonal, and population factors were adjusted for in the ITS analysis. However, there are notable limitations as well. The absence of data from outpatient and private clinics may reduce the generalizability of our results. Furthermore, the use of *ICD-9-CM* codes to identify DSH cases could compromise the accuracy of our findings, as later versions of the *ICD* (*International Classification of Diseases*) included more specific codes for DSH, but only *ICD-9-CM* was available on CDARS. Also, the data captured patients who presented to the ED with DSH, but neglected those not admitted through the ED. Another limitation is the lack of consideration for other potential confounding variables, such as the changing social conditions in 2019. In addition, the use of linear interpolation to adjust population counts may introduce bias, as population changes during the pandemic could affect the results.

Conclusion

The results indicate increasing adjusted DSH incidence ratios among males and individuals with alcohol and drug-related mental health conditions. This underscores the need for continued surveillance beyond the COVID-19 pandemic to determine if there is a spill-over effect. Importantly, all stakeholders—including policy makers, psychiatrists, and health care professionals—should collaborate to raise awareness and provide support for these vulnerable groups in Hong Kong. Targeted interventions and policies are essential to address the mental health challenges exacerbated by the pandemic and ensure better outcomes for those affected.

Acknowledgments

This work was supported in part by AIR@InnoHK administered by Innovation and Technology Commission of Hong Kong.

Authors' Contributions

CJYC and LYFL contributed to conception and design of the study. CJYC, LYFL, and AKCW performed the study. CJYC, LYFL, SCLC, TFY, CYTL, ETCL, SRBY, and HCP carried out the data analysis. CJYC, LYFL, TFY, JYHW, CKHW, CKT, TY, THR, AKCW, and JWKH contributed to the revision and read and approved the submitted version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional tables.
[DOCX File, 44 KB - publichealth_v11i1e57500_app1.docx]

Multimedia Appendix 2
Ethical approval.
[DOCX File, 128 KB - publichealth_v11i1e57500_app2.docx]

Multimedia Appendix 3
The average observation room staying time between the years 2019 and 2022. The red highlight corresponds to the dip in the monthly deliberate self-harm–related emergency department visits seen in Figure 1.
[PNG File, 45 KB - publichealth_v11i1e57500_app3.png]

Multimedia Appendix 4
The average total staying time between the years 2019 and 2022. The red highlight corresponds to the dip in the monthly deliberate self-harm–related emergency department visits seen in Figure 1.
[PNG File, 41 KB - publichealth_v11i1e57500_app4.png]

Multimedia Appendix 5
The average cubicle time between the years 2019 and 2022. The red highlight corresponds to the dip in the monthly deliberate self-harm–related emergency department visits seen in Figure 1.
[PNG File, 41 KB - publichealth_v11i1e57500_app5.png]

Multimedia Appendix 6
The average waiting time to triage between the years 2019 and 2022. The red highlight corresponds to the dip in the monthly deliberate self-harm–related emergency department visits seen in Figure 1.
[PNG File, 41 KB - publichealth_v11i1e57500_app6.png]

References

1. Zhao SZ, Wong JYH, Luk TT, Wai AKC, Lam TH, Wang MP. Mental health crisis under COVID-19 pandemic in Hong Kong, China. Int J Infect Dis 2020 Nov;100:431-433. [doi: 10.1016/j.ijid.2020.09.030] [Medline: 32947051]
2. Choi EPH, Hui BPH, Wan EYF. Depression and Anxiety in Hong Kong during COVID-19. Int J Environ Res Public Health 2020 May 25;17(10):3740. [doi: 10.3390/ijerph17103740] [Medline: 32466251]
3. Ni MY, Yao XI, Leung KSM, et al. Depression and post-traumatic stress during major social unrest in Hong Kong: a 10-year prospective cohort study. Lancet 2020 Jan;395(10220):273-284. [doi: 10.1016/S0140-6736(19)33160-5]
4. Chair SY, Chien WT, Liu T, et al. Psychological Distress, Fear and Coping Strategies among Hong Kong People During the COVID-19 Pandemic. Curr Psychol 2023;42(3):2538-2557. [doi: 10.1007/s12144-021-02338-7] [Medline: 34690470]
5. Legislative Council of the Hong Kong Special Administrative Region of the People’s Republic of China. Measures to promote mental health during COVID-19. 2021. URL: https://www.legco.gov.hk/research-publications/english/essentials-2021ise16-measures-to-promote-mental-health-during-covid-19.htm#endnote12 [accessed 2025-01-31]
6. The Hong Kong Special Administrative Region of the People’s Republic of China. The Chief Executive’s 2020 Policy Address. 2020. URL: https://www.policyaddress.gov.hk/2020/eng/p144.html [accessed 2025-01-31]

7. Nock MK, Mendes WB. Physiological arousal, distress tolerance, and social problem-solving deficits among adolescent self-injurers. J Consult Clin Psychol 2008 Feb;76(1):28-38. [doi: 10.1037/0022-006X.76.1.28] [Medline: 18229980]
8. Guerry JD, Prinstein MJ. Longitudinal prediction of adolescent nonsuicidal self-injury: examination of a cognitive vulnerability-stress model. J Clin Child Adolesc Psychol 2010;39(1):77-89. [doi: 10.1080/15374410903401195] [Medline: 20390800]
9. Kaess M, Hille M, Parzer P, Maser-Gluth C, Resch F, Brunner R. Alterations in the neuroendocrinological stress response to acute psychosocial stress in adolescents engaging in nonsuicidal self-injury. Psychoneuroendocrinology 2012 Jan;37(1):157-161. [doi: 10.1016/j.psyneuen.2011.05.009] [Medline: 21676550]
10. Liu RT, Frazier EA, Cataldo AM, Simon VA, Spirito A, Prinstein MJ. Negative life events and non-suicidal self-injury in an adolescent inpatient sample. Arch Suicide Res 2014;18(3):251-258. [doi: 10.1080/13811118.2013.824835] [Medline: 24712970]
11. Burke TA, Hamilton JL, Abramson LY, Alloy LB. Non-suicidal self-injury prospectively predicts interpersonal stressful life events and depressive symptoms among adolescent girls. Psychiatry Res 2015 Aug 30;228(3):416-424. [doi: 10.1016/j.psychres.2015.06.021] [Medline: 26165966]
12. Chai Y, Luo H, Yip PSF. Prevalence and risk factors for repetition of non-fatal self-harm in Hong Kong, 2002-2016: A population-based cohort study. Lancet Reg Health West Pac 2020 Sep;2:100027. [doi: 10.1016/j.lanwpc.2020.100027] [Medline: 34327378]
13. Isumi A, Doi S, Yamaoka Y, Takahashi K, Fujiwara T. Do suicide rates in children and adolescents change during school closure in Japan? The acute effect of the first wave of COVID-19 pandemic on child and adolescent mental health. Child Abuse Negl 2020 Dec;110(Pt 2):104680. [doi: 10.1016/j.chiabu.2020.104680] [Medline: 32847679]
14. O’Connor RC, Wetherall K, Cleare S, et al. Mental health and well-being during the COVID-19 pandemic: longitudinal analyses of adults in the UK COVID-19 Mental Health & Wellbeing study. Br J Psychiatry 2021 Jun;218(6):326-333. [doi: 10.1192/bjp.2020.212] [Medline: 33081860]
15. Robillard CL, Turner BJ, Ames ME, Craig SG. Deliberate self-harm in adolescents during COVID-19: The roles of pandemic-related stress, emotion regulation difficulties, and social distancing. Psychiatry Res 2021 Oct;304:114152. [doi: 10.1016/j.psychres.2021.114152] [Medline: 34371298]
16. Hawton K, Lascelles K, Brand F, et al. Self-harm and the COVID-19 pandemic: A study of factors contributing to self-harm during lockdown restrictions. J Psychiatr Res 2021 May;137:437-443. [doi: 10.1016/j.jpsychires.2021.03.028] [Medline: 33774538]
17. Bismark M, Scurrah K, Pascoe A, Willis K, Jain R, Smallwood N. Thoughts of suicide or self-harm among Australian healthcare workers during the COVID-19 pandemic. Aust N Z J Psychiatry 2022 Dec;56(12):1555-1565. [doi: 10.1177/00048674221075540] [Medline: 35128948]
18. Hospital Authority Hong Kong Special Administrative Region of the People's Republic of China. Accident & Emergency (A&E). URL: https://www.ha.org.hk/visitor/ha_serviceguide_details.asp?Content_ID=10051 [accessed 2025-01-31]
19. Hospital Authority Hong Kong Special Administrative Region of the People's Republic of China. Multi-tiered Triage and Treatment Strategy to battle a tsunami of COVID-19 cases. 2022. URL: https://www3.ha.org.hk/ehaslink/issue119/en/cover_story-2.html [accessed 2025-01-31]
20. Cheng MCY, Tong YH, Kwok TC, et al. Development journey of clinical data analysis and reporting system (CDARS) in hospital authority of Hong Kong. MedInfo. 2010.
21. Census and Statistics Department The Government of the Hong Kong Special Adminstrative Region. Table 1B: Population by Sex and Age [database on the Internet]. URL: https://www.censtatd.gov.hk/en/web_table.html?id=1B [accessed 2025-01-31]
22. Self-harm: assessment, management and preventing recurrence. National Institute for Health and Care Excellence. URL: https://www.nice.org.uk/guidance/ng225/chapter/Terms-used-in-this-guideline [accessed 2025-01-31]
23. Hospital Authority Hong Kong Special Administrative Region of the People's Republic of China. Waiving of Medical Charges (For Eligible Persons). URL: https://www.ha.org.hk/visitor/ha_visitor_index.asp?Content_ID=259365&Lang=ENG&Dimension=100 [accessed 2025-01-31]
24. Liu CH, Chen PC, Chen JH, Yeh CC. Changes in self-harm- and violence-related urgent psychiatric consultation in the emergency department during the different stages of the COVID-19 pandemic. BMC Psychiatry 2022 Jun 7;22(1):384. [doi: 10.1186/s12888-022-04029-4] [Medline: 35672738]
25. Rochelle TL. Health and help-seeking behavior among Chinese men in Hong Kong: The influence of culture. Psychol Men Masculinit 2019;20(1):71-81. [doi: 10.1037/men0000146]
26. Fong TCT, Chang K, Ho RTH. Association between quarantine and sleep disturbance in Hong Kong adults: The mediating role of COVID-19 mental impact and distress. Front Psychiatry 2023;14:1127070. [doi: 10.3389/fpsyt.2023.1127070] [Medline: 36926466]
27. Lam RPK. The impact of the COVID-19 pandemic on acute toxicity related to recreational drug abuse presenting to emergency departments. : Narcotics Division, Security Bureau, Hong Kong Special Administrative Region of the People's Republic of China; 2023.

28. The Centre for Health Protection DoH, The Government of Hong Kong Special Administrative Region. Alcohol Drinking, COVID-19 and NCDs. 2020.
29. Luk TT, Zhao S, Weng X, et al. Exposure to health misinformation about COVID-19 and increased tobacco and alcohol use: a population-based survey in Hong Kong. Tob Control 2021 Nov;30(6):696-699. [doi: 10.1136/tobaccocontrol-2020-055960] [Medline: 32855353]
30. Chai Y, Luo H, Wong GHY, et al. Risk of self-harm after the diagnosis of psychiatric disorders in Hong Kong, 2000-10: a nested case-control study. Lancet Psychiatry 2020 Feb;7(2):135-147. [doi: 10.1016/S2215-0366(20)30004-3] [Medline: 31974072]
31. Wai AKC, Wong CKH, Wong JYH, et al. Changes in Emergency Department Visits, Diagnostic Groups, and 28-Day Mortality Associated With the COVID-19 Pandemic: A Territory-Wide, Retrospective, Cohort Study. Ann Emerg Med 2022 Feb;79(2):148-157. [doi: 10.1016/j.annemergmed.2021.09.424] [Medline: 34742591]
32. Lopes S, Soares P, Gama A, et al. Association between perception of COVID-19 risk, confidence in health services and avoidance of emergency department visits: results from a community-based survey in Portugal. BMJ Open 2022 Jul 8;12(7):e058600. [doi: 10.1136/bmjopen-2021-058600] [Medline: 35803630]
33. DelPozo-Banos M, Lee SC, Friedmann Y, et al. Healthcare contacts with self-harm during COVID-19: An e-cohort whole-population-based study using individual-level linked routine electronic health records in Wales, UK, 2016-March 2021. PLoS ONE 2022;17(4):e0266967. [doi: 10.1371/journal.pone.0266967] [Medline: 35476839]
34. Maguire E, Kavalidou K, Bannan N, Doherty AM, Jeffers A. Substance use and self-harm emergency department presentations during COVID19: evidence from a National Clinical Programme for Self-Harm. Ir J Psychol Med 2023 Sep;40(3):424-429. [doi: 10.1017/ipm.2022.36] [Medline: 35915057]
35. Zhang A, Shi W, Tong C, et al. The fine-scale associations between socioeconomic status, density, functionality, and spread of COVID-19 within a high-density city. BMC Infect Dis 2022 Mar 21;22(1):274. [doi: 10.1186/s12879-022-07274-w] [Medline: 35313829]
36. The Hong Kong Special Administrative Region of the People's Republic of China. The Chief Executive’s 2024 Policy Address. 2024. URL: https://www.policyaddress.gov.hk/2024/en/p65.html [accessed 2025-01-31]

Abbreviations

AIC: Akaike Information Criterion
aIR: adjusted incidence ratio
CDARS: Clinical Data Analysis and Reporting System
DSH: deliberate self-harm
ED: emergency department
***ICD*:** *International Classification of Diseases*
***ICD-9-CM*:** *International Classification of Diseases, Ninth Revision, Clinical Modification*
ITS: interrupted time series
oIR: observed incidence ratio

Edited by A Nakhostin-Ansari; submitted 19.02.24; peer-reviewed by A Costanza, AH Hoveidaei, MA Khadembashiri; revised version received 16.12.24; accepted 21.12.24; published 10.02.25.

Please cite as:
Luk LYF, Ching CJY, Yip TF, Chan SCL, Lam CYT, Lam ETC, Yue SRB, Pang HC, Wong JYH, Wong CKH, Tong CK, Yamamoto T, Rainer TH, Wai AKC, Ho JWK
Incidence of Deliberate Self-Harm in Hong Kong Before and During the COVID-19 Pandemic: Population-Wide Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e57500
URL: https://publichealth.jmir.org/2025/1/e57500
doi:10.2196/57500

The Association Between Posting WeChat Moments and the Risk of Depressive Symptoms Among Middle-Aged and Older Chinese Adults: Prospective National Cohort Study

Wei Wang[1*], MMed; Hui Wang[2*], BM; Xinru Hu[3], BSc; Qian Yu[4], MSc; Fangyi Chen[3], BM; Xirui Qiu[5], BM; Xiaoxiao Wang[6], PhD

[1]GCP Center, Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), Nanjing, China

[2]Public Health Service Center of Nanjing Jiangbei New District, Nanjing, China

[3]Department of Public Health, School of Medicine and Holistic Integrative Medicine, Nanjing University of Chinese Medicine, Nanjing, China

[4]Department of Science and Technology, Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), Nanjing, China

[5]Department of Respiratory and Critical Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), Nanjing, China

[6]Clinical Research Institute, Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), 155 Hanzhong Road, Nanjing, China

[*]these authors contributed equally

Corresponding Author:
Xiaoxiao Wang, PhD
Clinical Research Institute, Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), 155 Hanzhong Road, Nanjing, China

Abstract

Background: The association between social media usage and the risk of depressive symptoms has attracted increasing attention. WeChat is a popular social media software in China. The impact of using WeChat and posting WeChat moments on the risk of developing depressive symptoms among community-based middle-aged and older adults in China is unknown.

Objective: The objective was to assess the association between using WeChat and posting WeChat moments and the risk of depressive symptoms among middle-aged and older adults in China.

Methods: A prospective national cohort study was designed based on the data obtained from the fourth and fifth waves of the China Health and Retirement Longitudinal Study (CHARLS). The strength of association between using WeChat and posting WeChat moments and the risk of depressive symptoms was estimated by modified Poisson regressions. Depressive symptoms were determined using the 10-item Center for Epidemiologic Studies Depression Scale. Meanwhile, the heterogeneity of the associations was explored through multiple subgroup analyses. Moreover, multiple sensitivity analyses were performed to verify the robustness of the associations between the exposures and depressive symptoms.

Results: A total of 9670 eligible participants were included in the cohort study, and the incidence rate of depressive symptoms was 19.08% (1845/9670, 95% CI 19.07% - 19.09%) from the fourth to fifth waves of the CHARLS. Using WeChat (adjusted relative risk [aRR] 0.691, 95% CI 0.582 - 0.520) and posting WeChat moments (aRR 0.673, 95% CI 0.552 - 0.821) reduced the risk of depressive symptoms among middle-aged and older Chinese adults. The association between the exposures and depressive symptoms was robust, proved through multiple sensitivity analyses (all *P*<.05). However, the associations were heterogeneous in certain subgroup catagories, such as solitude, duration of sleep at night, nap after lunch, physical activity, and having multiple chronic conditions.

Conclusions: Using WeChat and especially posting WeChat moments can mitigate the risk of depressive symptoms among community-based middle-aged and older Chinese adults. However, there is likely a need for a longer follow-up period to explore the impact of the exposures on the risk of long-term depressive outcomes.

(JMIR Public Health Surveill 2025;11:e62730) doi:10.2196/62730

KEYWORDS

cohort study; depression; depressive symptoms; mental health; middle-aged adults; modified Poisson regression; older adults; WeChat

Introduction

Depression is a common emotional disorder [1-4] resulting in limited psychosocial functioning and a decreased quality of life [5] and is also an independent risk factor of all-cause mortality and other chronic conditions [6-8]. A meta-analysis [3] reported that the prevalence of depressive symptoms among older adults in mainland China was 20% (95% CI 17.5% - 22.8%), and Chen et al [9] reported an unweighted incidence rate of depressive symptoms of 34.77% based on the China Health and Retirement Longitudinal Study (CHARLS). These findings highlight the need for an increased attention to depressive symptoms in China. Therefore, understanding modifiable factors associated with depressive symptoms is essential for developing effective public health strategies to prevent them. Multiple well-known factors associated with depressive symptoms include genetic factors, psychosocial stress, changes in the hypothalamic-pituitary-adrenal axis, inflammation, and others [5,10,11].

In terms of psychosocial factors, there has been a growing body of studies on the association or correlation between social media usage (SMU) and depression in recent years. Studies conducted by Lin et al [12], Primack et al [13] and Perlis et al [14] demonstrated that SMU, particularly on Facebook, TikTok (known as Douyin in China) and Snapchat, was associated with an increased risk of depression among US adults. However, Cotten et al [15] analyzed the data from 6 waves of the National Health and Aging Trends Study and suggested that SMU may not be related to depression in older adults. Based on the above studies, the association between SMU and depression could be attributed to problematic SMU [16,17], while moderate and reasonable use of social media may not cause harm to mental health.

WeChat (known as Weixin in China) is a popular social media software in China, which had 1.04 billion monthly active users according to Tencent [18]. The primary functions of WeChat include texting, voice and video calling, and browsing and posting moments that encompass both positive and negative events happening among friends or oneself, as well as facilitating financial management. From the perspective of these functionalities, WeChat distinguishes itself from platforms like Facebook or TikTok, which focus on interests and preferences that can foster addiction or problematic SMU.

It is hypothesized that using WeChat and posting WeChat moments would not result in an increased risk of depressive symptoms among community-based middle-aged and older Chinese adults, but rather potentially mitigate the risk of developing depressive symptoms. In fact, two cross-sectional studies using data from the fourth wave of the CHARLS and the fifth wave of the China Family Panel Study reported an association between WeChat usage and a decreased risk of depressive symptoms among middle-aged and older Chinese adults [19,20]. However, the two studies were unable to determine a temporal relationship between WeChat usage and depressive symptoms, thus precluding any causal inferences regarding the relationship between the observed exposures and depressive symptoms. Additionally, findings from the study conducted by Qu et al [19] indicated that participants who posted WeChat moments did not exhibit a decreased risk of depressive symptoms, which contradicted the findings of Zhang and Liang [20].

In response to these conflicting findings, a national, prospective cohort study was designed, leveraging data from both the fourth and fifth waves of the CHARLS. The aim of this study was to elucidate the relationship between the use of WeChat—with particular emphasis on posting WeChat moments—and the risk of depressive symptoms.

Methods

Study Design and Study Population

The study was characterized as a national, prospective cohort study, deriving its data from the fourth and fifth waves of the CHARLS. Initiated in 2011, the CHARLS represents a longitudinal national survey aimed at collecting extensive data pertinent to community-based middle-aged and older adults in China. The scope of the CHARLS encompasses high-quality, representative panel data covering a broad spectrum of topics. These topics include demographic variables, socioeconomic status, familial relationships, health status, health care usage, and others. More information about the CHARLS is accessible at their official website [21].

The intersection of the following 3 datasets from the CHARLS were compiled for the raw dataset: demographic backgrounds from the fourth wave, health status and functioning from the fourth wave, and the 10-item Center for Epidemiologic Studies Depression Scale (CESD-10) from the fifth wave. Participants from the raw dataset that (1) were aged <45 years, (2) were missing scores for the CESD-10 in the fourth wave, (3) had a CESD-10 score ≥10 in the fourth wave, or (4) were missing scores for the CESD-10 in the fifth wave were excluded.

Assessment of Exposures

In the fourth wave of the CHARLS, participants were asked, “Do you use WeChat?” with possible responses being yes or no. Participants who answered yes were categorized as WeChat users; otherwise, they were categorized as non-WeChat users. Subsequently, the participants using WeChat were further asked, “Do you post WeChat moments?” with possible responses being yes or no. Participants who answered yes were categorized as WeChat moments users; otherwise, they were categorized as non-WeChat moments users.

Therefore, according to whether or not they used WeChat, the participants were stratified into 2 groups: WeChat users and non-WeChat users. Furthermore, according to whether or not the participant posted WeChat moments, they were stratified into 3 groups: non-WeChat users, non-WeChat moments users and WeChat moments users. In this study, we first focused on the comparison between the incidence of depressive symptoms among the WeChat users and non-WeChat users, then that among non-WeChat users, non-WeChat moments users and WeChat moments users.

Assessment of the Outcome

The outcome of this study was depressive symptoms, which were measured using the CESD-10 scores from the fifth wave of the CHARLS as a dichotomous variable (yes or no). Each item of the CESD-10 was rated on a 4-point Likert scale ranging from 0 to 3, corresponding to the frequency of the symptoms (0=rarely or not at all; 1=some or a little of the time; 2=occasionally or a moderate amount of time; 3=most or all of the time). The aggregate score, which was the sum of the responses to all items, thus ranged from 0 to 30. A cumulative score of 10 or above was used as a threshold for identifying participants with depressive symptoms, whereas scores below this threshold indicated the absence of such symptoms. Additionally, for sensitivity analysis, alternative cutoff points of 11 and 12 on the CESD-10 were also employed to further ascertain the presence of depressive symptoms.

Assessment of Covariates

The following characteristics obtained from the fourth wave of the CHARLS were considered as covariates: age, sex (male or female), race (Han or minority), faith (yes or no), type of community (village or city/town), marital status (married and cohabiting, married and separated, or others), education (illiterate, not finished primary school/home school/primary school, junior high school, or senior high school and above), smoking habits (currently, ever, or never), alcohol consumption (yes or no), duration of sleep at night (<6 hours, ≥6 hours to <8 hours, or ≥8 hours), nap after lunch (yes or no), and physical activity (none, mild, moderate, or vigorous). Additionally, the social activity score was a covariate [22], which was the frequency of 11 kinds of activities rated as never (score=0), not regularly (score=1), almost every week (score=2), or almost daily (score=3). These activities were assembled to a sum score based on the frequency level, and the total scores for social activities could range from 0-33 points.

The basic activities of daily living (BADL) score was also a covariate [23], which included items for dressing, bathing, feeding, moving from a bed to a chair, using the toilet, and maintaining continence. All items had 4 potential answers, including (1) do it without difficulty, (2) do it but with difficulty, (3) do it with difficulty and need help, or (4) cannot do it, which were rated as 0, 1, 2, or 3 points, respectively. BADL scores were the sum of points for all items. Another covariate was life satisfaction. This question had 5 alternative answers: completely satisfied, very satisfied, somewhat satisfied, not very satisfied, or not at all satisfied. The responses were then stratified into 3 categories: completely/very satisfied, somewhat satisfied, or not very/not at all satisfied. Last, the number of chronic conditions (5 categories: no chronic conditions, 1 chronic condition, 2 chronic conditions, 3 chronic conditions, and ≥4 chronic conditions) and the number of parts with body pain were covariates.

Statistical Methods

Descriptive analyses and differential comparisons were conducted on the characteristics of the eligible participants. The quantitative variables with a normal distribution were described using mean and SD, and a 2-tailed *t* test or one-way ANOVA was used for comparing the difference between two groups or among three groups, respectively. Quantitative variables with skewed distributions were described using the median and IQR, and a Wilcoxon rank sum test or Kruskal-Wallis H test was used for the comparison. Disordered multinomial variables were described by the number of cases and percentage, and a χ^2 test or Fisher exact test was used for the comparison. Ordinal multinomial variables were described by the frequency and percentage, and a Wilcoxon rank sum test or Kruskal-Wallis H test was used for comparison.

For all eligible participants, univariable modified Poisson regressions were used to analyze the association between using WeChat and posting moments with the risk of depressive symptoms. Next, variables with $P<.05$ from between-group comparisons, along with adjusted variables in previous studies [19,20], were included as confounders in the multivariable modified Poisson regressions to analyze the association between the exposures and depressive symptoms. Meanwhile, the generalized variance-inflation factor of covariates in the multivariable regressions were calculated to test for collinearity (Table S1 in Multimedia Appendix 1).

For subgroup analyses, all eligible participants were divided into different subgroups according to sex (male or female), age (<60 years old or ≥60 years old), type of community (village or city/town), solitude (living with a heterosexual partner or not), education (primary school and below or junior high school and above), alcohol consumption (yes or no), duration of sleep at night (<6 hours, 6 - 8 hours, or ≥8 hours), nap after lunch (yes or no), physical activity (vigorous, moderate, mild, or none), and chronic conditions (≥2 chronic conditions or not). In the same way, multivariable modified Poisson regressions with the confounders, except the grouping variable, were performed to estimate the strength of association between the exposures and depressive symptoms.

There were missing values in variables for smoking habits, alcohol consumption, duration of sleep at night, nap after lunch, physical activity, social activity score, BADL score, and the number of body parts with pain, of which the proportions of missing values are shown in the Table S2 in Multimedia Appendix 1. Hot deck imputation was used to impute the missing values, and all statistical analyses were based on that data. Smoking habits, social activity score, BADL score, and the number of body parts with pain were not considered as confounders in the main analysis, due to the proportion of missing values being >40%. However, the above variables were adjusted for in the sensitivity analyses.

Multiple sensitivity analyses were conducted to validate the robustness of the association between the exposures and the outcome. First, the outcome of depressive symptoms was redefined as a CEDS-10 score of ≥11 or 12 in the multivariable modified Poisson regressions. Second, eligible participants diagnosed with emotional, nervous, or psychiatric problems in the fourth wave of the CHARLS were excluded. Third, to improve the comparability of the exposed groups with the control group, eligible participants aged >83 years in the non-WeChat users group were excluded since the oldest user in the WeChat users group was 83 years old (Table S3 in

Multimedia Appendix 1). Fourth, additional confounders (smoking habits, BADL score, social activity score, and the number of body parts with pain) were adjusted for in the main analysis. Last, multivariable binary logistic regressions were performed to estimate the association between exposures and depressive symptoms.

All statistical analyses were completed by R 4.3.3 (R Core Team) with R packages *VIM*, *rqlm*, *car* and *forestploter* and the Mengte Cloud version 1.0 statistics platform (Wuxi Mengte Yi Shu Tong Co, Ltd). A 2-sided hypothesis test was used, and the significance level was set at α=.05.

Ethical Considerations

The ethical approval for the CHARLS was approved by the Ethics Review Committee of Peking University (IRB00001052-11015). Prior to participation, all study participants were required to sign an informed consent form, thereby confirming their voluntary engagement in the survey process. The study data were anonymous. China’s Health Commission, Ministry of Science and Technology, and other departments jointly issued the Ethical Review Measures for Life Science and Medical Research involving Human Beings in February 2023, in which Article 32 mentioned 4 situations of exemption from ethical review. The secondary analysis of this public database CHARLS satisfied the first 2 of these: the data were free and publicly available and were available through a web-based application. Importantly, the data obtained from the platform were anonymous, and the identity information of all respondents was unknown. In addition, the study did not use biological samples from the survey subjects. Regarding compensation and informed consent, because this study was a secondary analysis, along with the subjects being anonymous, the respondents did not receive secondary compensation or sign a new informed consent form.

Results

The Characteristics of Eligible Participants

The process for screening participants is shown in Figure 1. A total of 9670 eligible participants were included in the final analysis. The mean age of these participants was 61.01 (SD 9.35) years. Of all eligible participants, 47.55% (n=4598) were female. A majority (n=6733, 69.63%) resided in villages, and 82.99% (n=8025) reported being married or cohabiting. Alcohol consumption was reported by 38.73% (n=3745) of participants, and 62.79% (n=6072) indicated they took a nap after lunch. A total of 92.84% (n=8978) engaged in physical activity, and 74.96% (n=7249) had been diagnosed with chronic conditions. The median scores of the CESD-10 were 4 (IQR 2-7) in the fourth wave and 6 (IQR 3-9) in the fifth wave. The incidence rate of depressive symptoms was 19.08% (95% CI 19.07% - 19.09%) from the fourth to fifth waves of the CHARLS (Table 1). Participants who used WeChat, regardless of whether they posted on WeChat moments, tended to be younger, male, non-religious, urban residents, or have a higher level of education. They were also more likely to be married or cohabiting, be current or former smokers, be alcohol consumers, sleep less than 8 hours at night, take afternoon naps, engage in physical activity, have high social activity scores, have low BADL scores, have fewer body parts in pain, and have no chronic conditions (Table 1 and S4 in Multimedia Appendix 1).

Figure 1. The flowchart for screening eligible participants, data governance, and statistical analysis. CESD-10: 10-item Center for Epidemiologic Studies Depression Scale; CHARLS: China Health and Retirement Longitudinal Study.

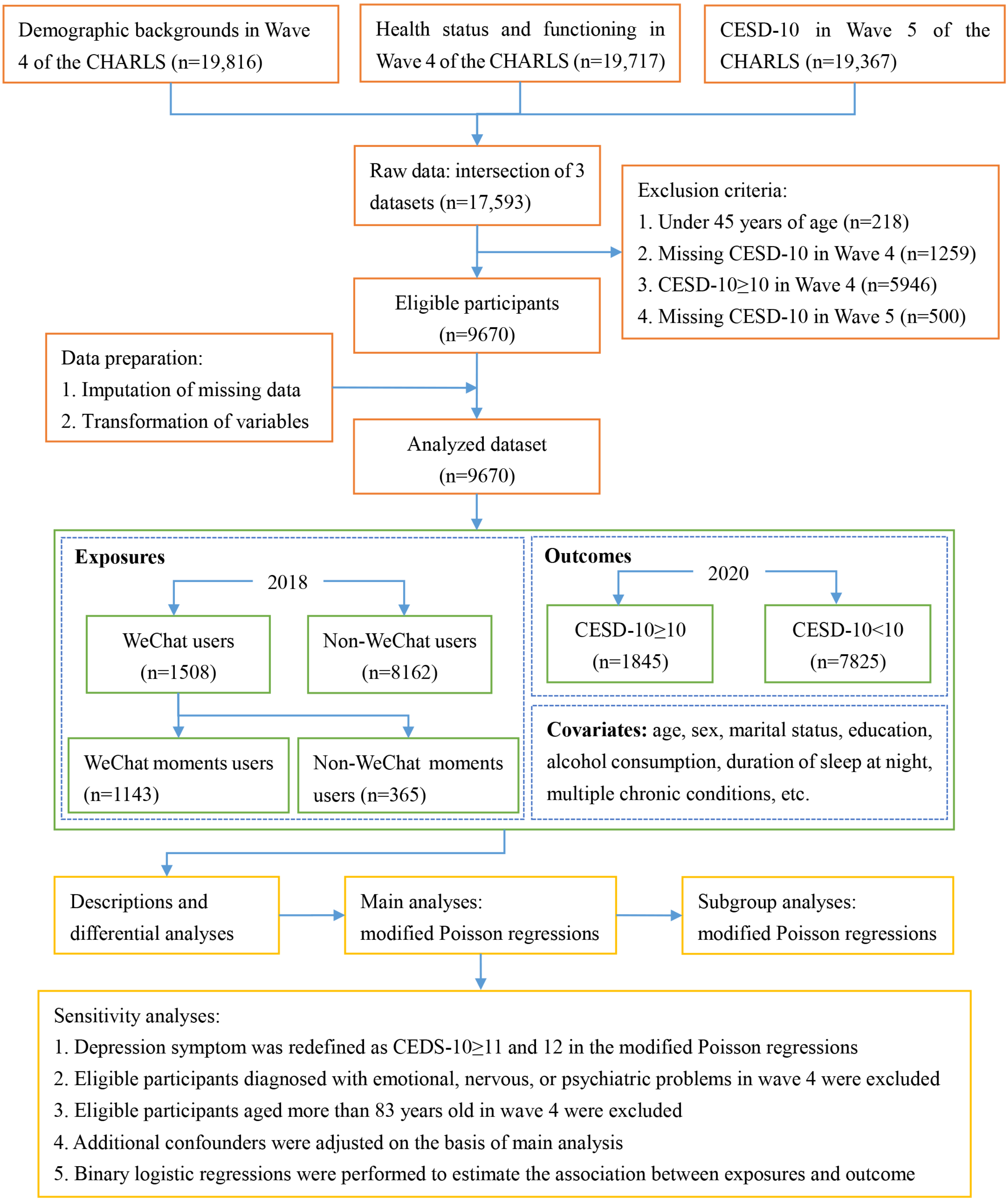


Table . The baseline characteristics and incidence of depressive symptoms of eligible participants in the cohort study based on the China Health and Retirement Longitudinal Study.

Characteristics	Total (n=9670)	Non-WeChat users (n=8162)	WeChat users (n=1508)		*P* value
			Non-WeChat moments users (n=365)	WeChat moments users (n=1143)	
Age (years), mean (SD)	61.01 (9.35)	62.05 (9.32)	55.68 (7.71)	55.28 (7.14)	<.001
Sex, n (%)					<.001
Male	5072 (52.45)	4204 (51.51)	204 (55.89)	664 (58.09)	
Female	4598 (47.55)	3958 (48.49)	161 (44.11)	479 (41.91)	
Race, n (%)					.48
Han	8990 (92.97)	7578 (92.84)	344 (94.25)	1068 (93.44)	
Minority	680 (7.03)	584 (7.16)	21 (5.75)	75 (6.56)	
Faith, n (%)					.02
Yes	946 (9.78)	821 (10.06)	40 (10.96)	85 (7.44)	
No	8724 (90.22)	7341 (89.94)	325 (89.04)	1058 (92.56)	
Type of community, n (%)					<.001
Village	6733 (69.63)	6085 (74.55)	176 (48.22)	472 (41.29)	
City/Town	2937 (30.37)	2077 (25.45)	189 (51.78)	671 (58.71)	
Education, n (%)					<.001
Illiterate	1721 (17.80)	1704 (20.88)	12 (3.29)	5 (0.44)	
Not finish primary school/home school/primary school	4040 (41.78)	3713 (45.49)	100 (27.40)	227 (19.86)	
Junior high school	2453 (25.37)	1903 (23.32)	133 (36.44)	417 (36.48)	
Senior high school and above	1456 (15.06)	842 (10.32)	120 (32.88)	494 (43.22)	
Marital status, n (%)					<.001
Married and cohabiting	8025 (82.99)	6727 (82.42)	307 (84.11)	991 (86.70)	
Married and separated	574 (5.94)	454 (5.56)	32 (8.77)	88 (7.70)	
Others	1071 (11.08)	981 (12.02)	26 (7.12)	64 (5.60)	
Smoking habits, n (%)					.006
Currently	333 (3.44)	271 (3.32)	8 (2.19)	54 (4.72)	
Ever	394 (4.07)	319 (3.91)	13 (3.56)	62 (5.42)	
Never	8943 (92.48)	7572 (92.77)	344 (94.25)	1027 (89.85)	
Alcohol consumption, n (%)					<.001
No	5925 (61.27)	5237 (64.16)	194 (53.15)	494 (43.22)	
Yes	3745 (38.73)	2925 (35.84)	171 (46.85)	649 (56.78)	
Duration of sleep at night, n (%)					<.001
<6 hours	2573 (26.61)	2229 (27.31)	88 (24.11)	256 (22.40)	
≥6 hours to <8 hours	4282 (44.28)	3456 (42.34)	190 (52.05)	636 (55.64)	
≥8 hours	2815 (29.11)	2477 (30.35)	87 (23.84)	251 (21.96)	
Nap after lunch, n (%)					<.001
No	3598 (37.21)	3129 (38.34)	131 (35.89)	338 (29.57)	
Yes	6072 (62.79)	5033 (61.66)	234 (64.11)	805 (70.43)	

Characteristics	Total (n=9670)	Non-WeChat users (n=8162)	WeChat users (n=1508)		*P* value
			Non-WeChat moments users (n=365)	WeChat moments users (n=1143)	
Physical activity, n (%)					<.001
No	692 (7.16)	657 (8.05)	13 (3.56)	22 (1.92)	
Mild	2741 (28.35)	2321 (28.44)	96 (26.30)	324 (28.35)	
Moderate	3072 (31.77)	2462 (30.16)	143 (39.18)	467 (40.86)	
Vigorous	3165 (32.73)	2722 (33.35)	113 (30.96)	330 (28.87)	
Social activity score, median (IQR)	3 (2-5)	3 (2-4)	5 (3-6)	5 (3-8)	<.001
BADL[a] score, mean (SD)	6.32 (1.02)	6.34 (1.04)	6.23 (0.71)	6.25 (0.90)	.006
Life satisfaction, n (%)					<.001
Completely/very	4010 (41.47)	3508 (42.98)	135 (36.99)	367 (32.11)	
Somewhat	5262 (54.42)	4319 (52.92)	211 (57.81)	732 (64.04)	
Not very/not at all	398 (4.12)	335 (4.10)	19 (5.21)	44 (3.85)	
Number of chronic conditions, n (%)					<.001
0	2421 (25.04)	1960 (24.01)	113 (30.96)	348 (30.45)	
1	2574 (26.62)	2179 (26.70)	98 (26.85)	297 (25.98)	
2	1971 (20.38)	1699 (20.82)	79 (21.64)	193 (16.89)	
3	1243 (12.85)	1051 (12.88)	45 (12.33)	147 (12.86)	
≥4	1461 (15.11)	1273 (15.60)	30 (8.22)	158 (13.82)	
Number of body parts with pain, median (IQR)	3 (2-6)	3 (2-6)	3 (1-5)	3 (1-5)	.007
CESD-10[b] in 2020, median (IQR)	6 (3-9)	6 (3-10)	4 (2-7)	4 (1-7)	<.001
Depressive symptoms, n (%)					<.001
No	7825 (80.92)	6457 (79.11)	326 (89.32)	1042 (91.16)	
Yes	1845 (19.08)	1705 (20.89)	39 (10.68)	101 (8.84)	

[a] BADL: basic activities of daily living.

[b] CESD-10: 10-item Center for Epidemiologic Studies Depression Scale.

Association Between Exposures and Depressive Symptoms

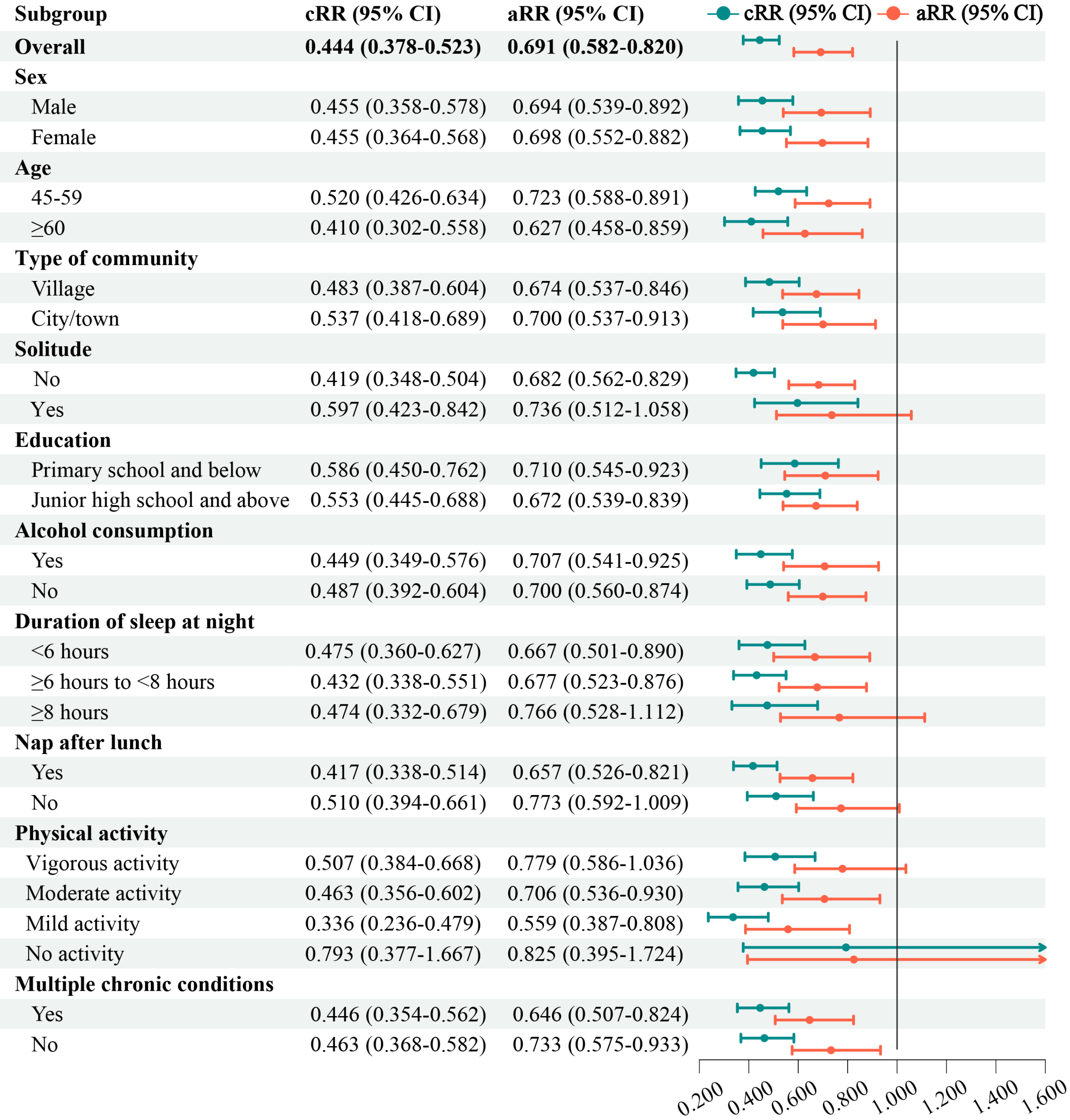
As shown in Figure 2 and Table S5 in Multimedia Appendix 1, among all eligible participants, the risk of depressive symptoms in the WeChat users group was 0.556 times lower than that in the non-WeChat users group, as indicated by the crude relative risk (cRR) of 0.444 (95% CI 0.378 - 0.523; *P*<.001) using a univariable modified Poisson regression. Furthermore, after adjusting for potential confounders in the multivariable modified Poisson regression models, the risk of depressive symptoms in the WeChat users group was 0.309 times lower than that in the non-WeChat users group (adjusted relative risk [aRR] 0.691, 95% CI 0.582 - 0.820; *P*<.001). These results suggested a significant association between WeChat usage and a decreased risk of depressive symptoms, implying that WeChat may serve as a protective factor against depressive symptoms in community-based middle-aged and older Chinese adults.

Subgroup analyses showed that WeChat usage did not reduce the risk of depressive symptoms among those who were solitary, had a sleep duration of ≥8 hours at night, did not have nap after lunch, engaged in vigorous physical activity, or did not engage in any physical activity. However, in other subgroups, the result indicated that WeChat usage was associated with a lower risk of depressive symptoms.

As depicted in Figure 3 and Table S6 in Multimedia Appendix 1, compared to non-WeChat users, both the non-WeChat moments users (aRR 0.739, 95% CI 0.550 - 0.992; *P*=.044) and WeChat moments users (aRR 0.673, 95% CI 0.552 - 0.821; *P*<.001) exhibited a lower risk of depressive symptoms. In the subgroups of females (aRR 0.655, 95% CI 0.432 - 0.996) and those without multiple chronic conditions (aRR 0.589, 95% CI

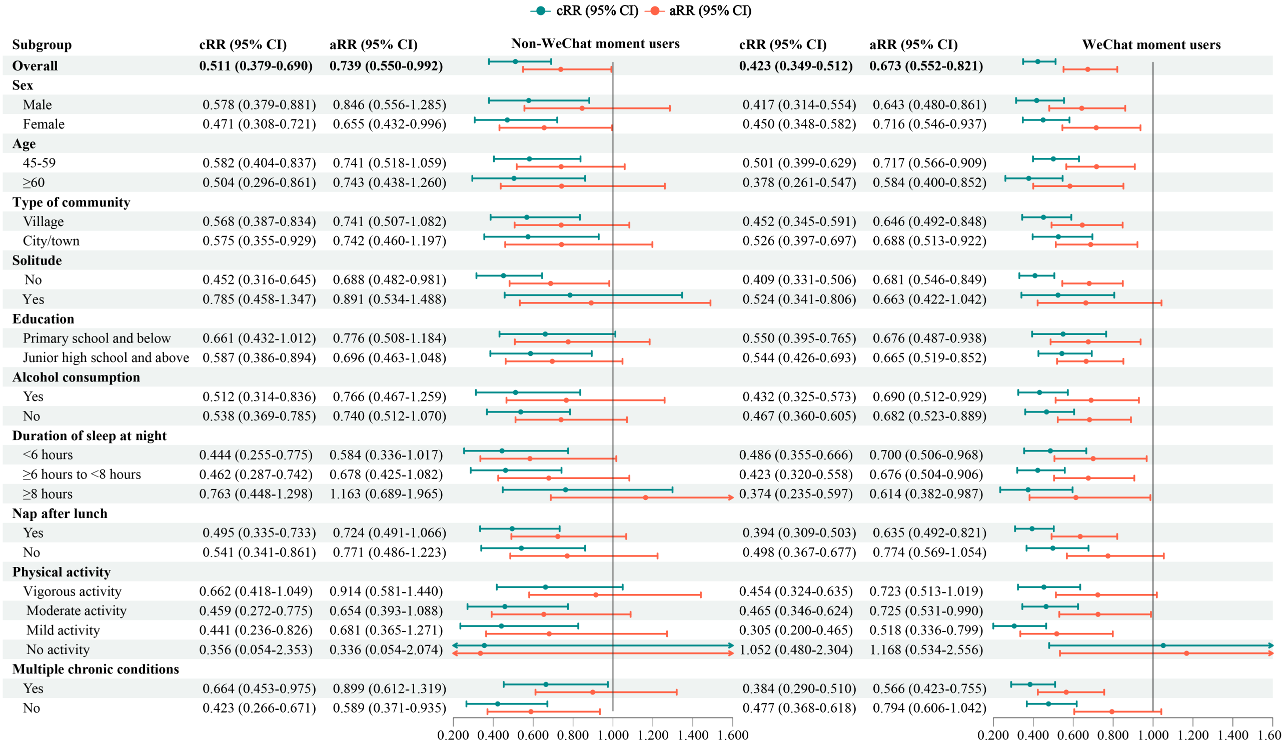
0.371–0.935), non-WeChat moments users had a lower risk of depressive symptoms compared to non-WeChat users. Additionally, with the exception of individuals who did not live with a spouse or opposite sex, did not nap after lunch, engaged in vigorous physical activity, did not engage in any physical activity, or did not have multiple chronic conditions, results from the subgroup analyses suggested that posting WeChat moments was associated with a reduced risk of depressive symptoms.

Figure 2. The association between using WeChat and risk of depressive symptoms among middle-aged and older Chinese adults in the cohort study based on the CHARLS. aRR: adjusted relative risk; cRR: crude relative risk.



Subgroup	cRR (95% CI)	aRR (95% CI)
Overall	**0.444 (0.378-0.523)**	**0.691 (0.582-0.820)**
Sex		
Male	0.455 (0.358-0.578)	0.694 (0.539-0.892)
Female	0.455 (0.364-0.568)	0.698 (0.552-0.882)
Age		
45-59	0.520 (0.426-0.634)	0.723 (0.588-0.891)
≥60	0.410 (0.302-0.558)	0.627 (0.458-0.859)
Type of community		
Village	0.483 (0.387-0.604)	0.674 (0.537-0.846)
City/town	0.537 (0.418-0.689)	0.700 (0.537-0.913)
Solitude		
No	0.419 (0.348-0.504)	0.682 (0.562-0.829)
Yes	0.597 (0.423-0.842)	0.736 (0.512-1.058)
Education		
Primary school and below	0.586 (0.450-0.762)	0.710 (0.545-0.923)
Junior high school and above	0.553 (0.445-0.688)	0.672 (0.539-0.839)
Alcohol consumption		
Yes	0.449 (0.349-0.576)	0.707 (0.541-0.925)
No	0.487 (0.392-0.604)	0.700 (0.560-0.874)
Duration of sleep at night		
<6 hours	0.475 (0.360-0.627)	0.667 (0.501-0.890)
≥6 hours to <8 hours	0.432 (0.338-0.551)	0.677 (0.523-0.876)
≥8 hours	0.474 (0.332-0.679)	0.766 (0.528-1.112)
Nap after lunch		
Yes	0.417 (0.338-0.514)	0.657 (0.526-0.821)
No	0.510 (0.394-0.661)	0.773 (0.592-1.009)
Physical activity		
Vigorous activity	0.507 (0.384-0.668)	0.779 (0.586-1.036)
Moderate activity	0.463 (0.356-0.602)	0.706 (0.536-0.930)
Mild activity	0.336 (0.236-0.479)	0.559 (0.387-0.808)
No activity	0.793 (0.377-1.667)	0.825 (0.395-1.724)
Multiple chronic conditions		
Yes	0.446 (0.354-0.562)	0.646 (0.507-0.824)
No	0.463 (0.368-0.582)	0.733 (0.575-0.933)

Figure 3. The association between using WeChat with or without posting WeChat moments and risk of depressive symptoms among middle-aged and older Chinese adults in the cohort study based on CHARLS. aRR: adjusted relative risk; cRR: crude relative risk.



Subgroup	cRR (95% CI)	aRR (95% CI)	cRR (95% CI)	aRR (95% CI)
Overall	**0.511 (0.379-0.690)**	**0.739 (0.550-0.992)**	**0.423 (0.349-0.512)**	**0.673 (0.552-0.821)**
Sex				
Male	0.578 (0.379-0.881)	0.846 (0.556-1.285)	0.417 (0.314-0.554)	0.643 (0.480-0.861)
Female	0.471 (0.308-0.721)	0.655 (0.432-0.996)	0.450 (0.348-0.582)	0.716 (0.546-0.937)
Age				
45-59	0.582 (0.404-0.837)	0.741 (0.518-1.059)	0.501 (0.399-0.629)	0.717 (0.566-0.909)
≥60	0.504 (0.296-0.861)	0.743 (0.438-1.260)	0.378 (0.261-0.547)	0.584 (0.400-0.852)
Type of community				
Village	0.568 (0.387-0.834)	0.741 (0.507-1.082)	0.452 (0.345-0.591)	0.646 (0.492-0.848)
City/town	0.575 (0.355-0.929)	0.742 (0.460-1.197)	0.526 (0.397-0.697)	0.688 (0.513-0.922)
Solitude				
No	0.452 (0.316-0.645)	0.688 (0.482-0.981)	0.409 (0.331-0.506)	0.681 (0.546-0.849)
Yes	0.785 (0.458-1.347)	0.891 (0.534-1.488)	0.524 (0.341-0.806)	0.663 (0.422-1.042)
Education				
Primary school and below	0.661 (0.432-1.012)	0.776 (0.508-1.184)	0.550 (0.395-0.765)	0.676 (0.487-0.938)
Junior high school and above	0.587 (0.386-0.894)	0.696 (0.463-1.048)	0.544 (0.426-0.693)	0.665 (0.519-0.852)
Alcohol consumption				
Yes	0.512 (0.314-0.836)	0.766 (0.467-1.259)	0.432 (0.325-0.573)	0.690 (0.512-0.929)
No	0.538 (0.369-0.785)	0.740 (0.512-1.070)	0.467 (0.360-0.605)	0.682 (0.523-0.889)
Duration of sleep at night				
<6 hours	0.444 (0.255-0.775)	0.584 (0.336-1.017)	0.486 (0.355-0.666)	0.700 (0.506-0.968)
≥6 hours to <8 hours	0.462 (0.287-0.742)	0.678 (0.425-1.082)	0.423 (0.320-0.558)	0.676 (0.504-0.906)
≥8 hours	0.763 (0.448-1.298)	1.163 (0.689-1.965)	0.374 (0.235-0.597)	0.614 (0.382-0.987)
Nap after lunch				
Yes	0.495 (0.335-0.733)	0.724 (0.491-1.066)	0.394 (0.309-0.503)	0.635 (0.492-0.821)
No	0.541 (0.341-0.861)	0.771 (0.486-1.223)	0.498 (0.367-0.677)	0.774 (0.569-1.054)
Physical activity				
Vigorous activity	0.662 (0.418-1.049)	0.914 (0.581-1.440)	0.454 (0.324-0.635)	0.723 (0.513-1.019)
Moderate activity	0.459 (0.272-0.775)	0.654 (0.393-1.088)	0.465 (0.346-0.624)	0.725 (0.531-0.990)
Mild activity	0.441 (0.236-0.826)	0.681 (0.365-1.271)	0.305 (0.200-0.465)	0.518 (0.336-0.799)
No activity	0.356 (0.054-2.353)	0.336 (0.054-2.074)	1.052 (0.480-2.304)	1.168 (0.534-2.556)
Multiple chronic conditions				
Yes	0.664 (0.453-0.975)	0.899 (0.612-1.319)	0.384 (0.290-0.510)	0.566 (0.423-0.755)
No	0.423 (0.266-0.671)	0.589 (0.371-0.935)	0.477 (0.368-0.618)	0.794 (0.606-1.042)

Sensitivity Analyses

The results of sensitivity analyses are presented in Table 2. (1) When depressive symptoms were redefined as CESD-10 scores of ≥11 or 12, the multivariable Poisson regressions indicated that WeChat usage reduced the risk of depressive symptoms in middle-aged and older individuals, regardless of whether they posted WeChat moments. (2) The multivariable Poisson regression was performed again after excluding participants who had emotional, nervous, or psychiatric problems in the fourth wave of the CHARLS, and results were consistent with the main analysis. (3) Given that none of the WeChat users were over 83 years old, another multivariable Poisson regression was performed after excluding non-WeChat users over 83 years old, which yielded results similar to the main analysis. (4) Additional adjustments were made for smoking habits, social activity score, BADL score, and the number of body parts with pain. The outcome of depressive symptoms remained consistent with those of the main analysis. (5) Odds ratios from multivariable binary logistic regressions were used to estimate the associations, and the results demonstrated the robustness of the main analysis.

Table . Multiple sensitivity analyses to determine the association between exposures and depressive symptoms that were identified by the main analysis.

Sensitivity analysis and variables	β[a] (SE)	aRR[b]/aOR[c] (95% CI)	*P* value
Depressive symptoms was defined as CESD-10[d] of ≥11 in the modified Poisson regressions			
Use WeChat (vs no)			
Yes	–0.370 (0.087)	0.691 (0.582 - 0.820)	<.001
Use WeChat and post moments (vs do not use WeChat)			
Use WeChat and not post moments	–0.303 (0.151)	0.738 (0.549 - 0.992)	.04
Use WeChat and post moments	–0.396 (0.101)	0.673 (0.552 - 0.821)	<.001
Depressive symptoms was defined as CESD-10 of ≥12 in the modified Poisson regressions			
Use WeChat (vs no)			
Yes	–0.402 (0.100)	0.669 (0.550 - 0.813)	<.001
Use WeChat and post moments (do not use WeChat)			
Use WeChat and not post moments	–0.267 (0.168)	0.766 (0.551 - 1.064)	.11
Use WeChat and post moments	–0.457 (0.117)	0.633 (0.504 - 0.796)	<.001
Participants diagnosed with emotional, nervous, or psychiatric problems in Wave 4 were excluded			
Use WeChat (vs no)			
Yes	–0.399 (0.089)	0.671 (0.564 - 0.799)	<.001
Use WeChat and post moments (do not use WeChat)			
Use WeChat and not post moments	–0.325 (0.153)	0.723 (0.536 - 0.975)	.03
Use WeChat and post moments	–0.428 (0.103)	0.652 (0.532 - 0.799)	<.001
Participants aged more than 83 years in Wave 4 were excluded			
Use WeChat (vs no)			
Yes	–0.363 (0.088)	0.696 (0.586 - 0.826)	<.001
Use WeChat and post moments (vs do not use WeChat)			
Use WeChat and not post moments	–0.298 (0.151)	0.742 (0.552 - 0.998)	.048
Use WeChat and post moments	–0.388 (0.101)	0.678 (0.556 - 0.827)	<.001
Smoking habits, social activity score, BADL[e] score, and the number of body parts with pain were adjusted for the main analysis			
Use WeChat (vs no)			
Yes	–0.376 (0.090)	0.687 (0.576 - 0.819)	<.001
Use WeChat and post moments (vs do not use WeChat)			
Use WeChat and not post moments	–0.306 (0.152)	0.736 (0.546 - 0.992)	.045
Use WeChat and post moments	–0.403 (0.103)	0.668 (0.546 - 0.818)	<.001
Multivariable logistic regressions for analyzing the association between exposures and depressive symptoms			
Use WeChat (vs no)			
Yes	–0.431 (0.102)	0.650 (0.532 - 0.794)	<.001
Use WeChat and post moments (vs do not use WeChat)			
Use WeChat and not post moments	–0.361 (0.178)	0.697 (0.492 - 0.988)	.04
Use WeChat and post moments	–0.457 (0.117)	0.633 (0.504 - 0.797)	<.001

[a] β: regression coefficient.

[b] aRR: adjusted relative risk.

[c] aOR: adjusted odds ratio.

[d] CESD-10: 10-item Center for Epidemiologic Studies Depression Scale.

[e] BADL: basic activities of daily living.

Discussion

Principal Findings

This prospective national cohort study confirmed that using WeChat and posting WeChat moments was a protective factor for depressive symptoms among community based middle-aged and older adults in China. In terms of the absolute magnitude of the aRR, participants who post WeChat moments had a lower risk of depressive symptoms than those who do not post WeChat moments. The results of multiple sensitivity analysis also illustrated the robustness of the association between the exposures and depressive symptoms. In addition, the associations were heterogeneous across several subgroups. For example, irrespective of posting WeChat moments, the association between WeChat usage and the risk of depression was not significant among participants whose duration of sleep at night was more than or equal to 8 hours, while the association was statistically significant in the other two subgroups for the duration of sleep. Also, only in the subgroups where participants engaged in mild and moderate physical activity did using WeChat and posting WeChat moments mitigate the risk of depressive symptoms.

First, this study demonstrated that WeChat usage mitigated the risk of future depressive symptoms, which was consistent with the conclusions of 3 existing cross-sectional studies carried out by Qu et al [19], Wang et al [24], and Zhang and Liang [20]. Second, the study further suggested that posting WeChat moments was significantly associated with a decreased risk of depressive symptoms, and that was consistent with the finding of the cross-sectional study conducted by Zhang and Liang [20], but it was inconsistent with the finding of the cross-sectional study conducted by Qu et al [19]. The results from this prospective cohort study are more convincing because it meets a necessary and inarguable criterion for causal inference: the temporality of cause and effect [25]. Last, multiple subgroup analyses were performed to explore the heterogeneity of the associations for the different characteristics of the participants, which were not conducted in previous studies about WeChat usage and depression.

Many studies have shown that loneliness, social isolation, lack of family and social support, and lack of social activity are risk factors for anxiety and depression in adults [26-30]. On the one hand, WeChat is an instant message platform where WeChat users can communicate with family and friends to relieve the feeling of loneliness, especially for older adults whose children or relatives are not around them [31]. Importantly, a video call on WeChat allows users to see each other, which is more convenient for emotional expression and perception and users can obtain emotional support. WeChat users not only can maintain and improve existing social relationships through WeChat, but they can also expand their social circle to make more like-minded friends and engage in more social activity [32].

WeChat moments is another common and primary function of WeChat, which is an emotional space composed of words, pictures, and short videos posted by users. Users of WeChat moments can post what is happening around them or express their feelings by text or video, browse the content posted or reproduced by their WeChat friends, and interact with their friends through the act of liking and commenting on posts and short videos. Similarly, users can obtain emotional support and emotional communication, as well as release negative emotions by posting WeChat moments. Therefore, the results of Nan et al [33] were supported, which found that internet usage, including WeChat usage, could directly or indirectly reduce the risk of depression through improving interpersonal relationships. Since solitude, a longer sleep duration, and no physical activity are all risk factors for depressive symptoms, it is difficult to offset the negative impact of these factors by using WeChat or posting WeChat moments. Therefore, the heterogeneity of the associations among certain subgroups was observed. Importantly, the heterogeneity of the associations shown in the subgroup analysis suggested that a one-size-fits-all model should be used with caution, and recommendations for individualization or stratified treatment are necessary, especially when research is used to inform policy and clinicians' decisions [34-36].

Strengths and Limitations

There are several strengths of this study. First, this study is a prospective national cohort study, of which the sample sizes are large and the participants are representative. Second, modified Poisson regressions were used to precisely estimate the strength of the associations, which is a statistical approach more appropriate for this scenario given the high incidence of depressive symptoms. Third, multiple subgroups analyses were performed to explore the heterogeneity of the associations between the exposures and depressive symptoms among participants with different characteristics. Last, multiple sensitivity analyses were performed to validate the robustness of the associations between the exposures and depressive symptoms based on the main analysis.

Meanwhile, there are several limitations of this study. First, as the questions about WeChat usage and WeChat moments usage were only collected in the 2018 and 2020 surveys in the CHARLS database, the data for only these two surveys were used in this study. The corresponding follow-up period in this study was short, and a longer follow-up is needed to further confirm our conclusion. Second, the use of subgroups and more exposure groups (WeChat users were further divided into WeChat moments users and non-WeChat moments users) reduced the sample size and statistical power of the regression analyses. In particular, fewer participants used WeChat but did not post moments, which may have resulted in wide CIs for the associations and heterogeneity of the associations that are not easily explained. Third, some studies have suggested that excessive SMU or problematic SMU may increase the risk of depression [16,37,38]. However, the CHARLS database lacks relevant information on the frequency or duration of WeChat and WeChat moments usage, so this study cannot analyze the association between the frequency of WeChat use and depressive symptoms. Fourth, limited by the information collected in the database, other factors behind the use of WeChat and WeChat moments that may influence depressive symptoms were not considered in this study, such as the personality of participants (ie, outgoing or introverted). These factors can relate to whether participants post moments and affect the risk of depressive symptoms. Last, WeChat is only one of many social media software, and the participants may also use other social media software in addition to WeChat. Also, limited by the relevant information provided by CHARLS, this study was unable to investigate the effects of other social media and the joint use of them on mental health.

Conclusion

In summary, the prevention of depression in community-based middle-aged and older adults calls for the joint efforts of individuals, families, and society. On the one hand, middle-aged and older adults are recommended to moderately use social media and the internet to learn more about new things or keep in touch with family and friends [39]. At the same time, middle-aged and older adults are recommended to participate in offline social activities and physical activity, if their physical capacity allows. On the other hand, it is suggested that children should actively contact their older family members, both electronically and offline, to provide them with emotional and material support. In addition, especially for older adults, it is recommended that the government and community organizations provide more community support [40], such as group recreational activities, which can also be important for the prevention of depression. In addition, the next waves of the CHARLS data will be obtained to determine the impact of using WeChat and posting WeChat moments on the long-term risk of depressive symptoms.

Acknowledgments

This work was supported by Preventive Medicine Scientific Research Project of Jiangsu Province in 2023 (Ym2023103), Medical Research Program of Health Commission of Jiangsu Province (M2021001), Natural Science Foundation of Nanjing University of Traditional Chinese Medicine (XZR2021005), and Academic Promotion Special Project of Department Director in 2022 (Y2022ZR14). All funding organizations had no role in the study design, implementation, analysis, or interpretation of the data. All authors thank the CHARLS team of Peking University for its excellent contribution and all participants surveyed for their efforts.

Data Availability

The data used in the cohort study can be accessed through the official website of the CHARLS [21].

Authors' Contributions

WW and HW are credited as cofirst authors for their significant contributions. XW is recognized as the corresponding author. WW, HW, and XW jointly worked on the study's concept and design. WW and HW also took on the analysis and interpretation of data. The drafting of the manuscript was a joint effort by WW and HW, with guidance from XW and with additional support from XH, QY, and FC who concentrated on assessing the results, especially focusing on the methods and analytics. All authors participated in critically revising the manuscript for important intellectual content, had full access to all the study's data, reviewed and revised the manuscript drafts, and gave their final approval for the version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional data tables.
[DOC File, 252 KB - publichealth_v11i1e62730_app1.doc]

References

1. James SL, Abate D, Abate KH. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1789-1858. [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
2. Charlson FJ, Baxter AJ, Cheng HG, Shidhaye R, Whiteford HA. The burden of mental, neurological, and substance use disorders in China and India: a systematic analysis of community representative epidemiological studies. Lancet 2016 Jul 23;388(10042):376-389. [doi: 10.1016/S0140-6736(16)30590-6] [Medline: 27209143]
3. Tang T, Jiang J, Tang X. Prevalence of depressive symptoms among older adults in mainland China: a systematic review and meta-analysis. J Affect Disord 2021 Oct 1;293:379-390. [doi: 10.1016/j.jad.2021.06.050] [Medline: 34246000]
4. Gu L, Xie J, Long J, et al. Epidemiology of major depressive disorder in mainland china: a systematic review. PLoS One 2013;8(6):e65356. [doi: 10.1371/journal.pone.0065356] [Medline: 23785419]
5. Malhi GS, Mann JJ. Depression. Lancet 2018 Nov 24;392(10161):2299-2312. [doi: 10.1016/S0140-6736(18)31948-2] [Medline: 30396512]
6. Bennett S, Thomas AJ. Depression and dementia: cause, consequence or coincidence? Maturitas 2014 Oct;79(2):184-190. [doi: 10.1016/j.maturitas.2014.05.009] [Medline: 24931304]

7. Meng R, Yu C, Liu N, et al. Association of depression with all-cause and cardiovascular disease mortality among adults in China. JAMA Netw Open 2020 Feb 5;3(2):e1921043. [doi: 10.1001/jamanetworkopen.2019.21043] [Medline: 32049295]
8. Wang S, Mao S, Xiang D, Fang C. Association between depression and the subsequent risk of Parkinson's disease: a meta-analysis. Prog Neuropsychopharmacol Biol Psychiatry 2018 Aug 30;86:186-192. [doi: 10.1016/j.pnpbp.2018.05.025] [Medline: 29859854]
9. Chen R, Chen Q, Lu G, et al. Sleep duration and depressive symptoms in Chinese middle-aged and older adults: the moderating effects of grip strength. J Affect Disord 2023 Oct 15;339:348-354. [doi: 10.1016/j.jad.2023.07.059] [Medline: 37451435]
10. Hasler G. Pathophysiology of depression: do we have any solid evidence of interest to clinicians? World Psychiatry 2010 Oct;9(3):155-161. [doi: 10.1002/j.2051-5545.2010.tb00298.x] [Medline: 20975857]
11. Mendelson T, Tandon SD. Prevention of depression in childhood and adolescence. Child Adolesc Psychiatr Clin N Am 2016 Apr;25(2):201-218. [doi: 10.1016/j.chc.2015.11.005] [Medline: 26980124]
12. Lin LY, Sidani JE, Shensa A, et al. Association between social media use and depression among U.S. young adults. Depress Anx 2016 Apr;33(4):323-331. [doi: 10.1002/da.22466] [Medline: 26783723]
13. Primack BA, Shensa A, Sidani JE, Escobar-Viera CG, Fine MJ. Temporal associations between social media use and depression. Am J Prev Med 2021 Feb;60(2):179-188. [doi: 10.1016/j.amepre.2020.09.014] [Medline: 33309454]
14. Perlis RH, Green J, Simonson M, et al. Association between social media use and self-reported symptoms of depression in US adults. JAMA Netw Open 2021 Nov 1;4(11):e2136113. [doi: 10.1001/jamanetworkopen.2021.36113] [Medline: 34812844]
15. Cotten SR, Ghaiumy Anaraky R, Schuster AM. Social media use may not be as bad as some suggest: implication for older adults. Innov Aging 2023;7(3):igad022. [doi: 10.1093/geroni/igad022] [Medline: 37179658]
16. Shannon H, Bush K, Villeneuve PJ, Hellemans KG, Guimond S. Problematic social media use in adolescents and young adults: systematic review and meta-analysis. JMIR Ment Health 2022 Apr 14;9(4):e33450. [doi: 10.2196/33450] [Medline: 35436240]
17. Lopes LS, Valentini JP, Monteiro TH, et al. Problematic social media use and its relationship with depression or anxiety: a systematic review. Cyberpsychol Behav Soc Netw 2022 Nov;25(11):691-702. [doi: 10.1089/cyber.2021.0300] [Medline: 36219756]
18. Tencent announces 2018 first quarter results. Tencent. 2018. URL: https://static.www.tencent.com/storage/uploads/2019/11/09/3e68a59c9f2568245fd0532efa0d7f78.pdf [accessed 2018-05-16]
19. Qu X, Houser SH, Zhang J, Wen J, Zhang W. Association between using social media WeChat and depressive symptoms among middle-aged and older people: findings from a national survey. BMC Geriatr 2022 Apr 22;22(1):351. [doi: 10.1186/s12877-022-03054-y] [Medline: 35448984]
20. Zhang C, Liang X. Association between WeChat use and mental health among middle-aged and older adults: a secondary data analysis of the 2020 China Family Panel Studies database. BMJ Open 2023 Oct 16;13(10):e073553. [doi: 10.1136/bmjopen-2023-073553] [Medline: 37844989]
21. China Health and Retirement Longitudinal Study. National Development Institute, Peking University. URL: http://charls.pku.edu.cn/ [accessed 2024-10-28]
22. Li H, Li C, Wang A, et al. Associations between social and intellectual activities with cognitive trajectories in Chinese middle-aged and older adults: a nationally representative cohort study. Alzheimers Res Ther 2020 Sep 25;12(1):115. [doi: 10.1186/s13195-020-00691-6] [Medline: 32977839]
23. Liu H, Ma Y, Lin L, Sun Z, Li Z, Jiang X. Association between activities of daily living and depressive symptoms among older adults in China: evidence from the CHARLS. Front Public Health 2023;11:1249208. [doi: 10.3389/fpubh.2023.1249208] [Medline: 38035294]
24. Wang G, Duan J, Kan Q, Zhou Y, Cheng Z, Tang S. The correlation analysis of WeChat usage and depression among the middle-aged and elderly in China: the mediating role of social participation. BMC Public Health 2023 Mar 10;23(1):462. [doi: 10.1186/s12889-023-15349-9] [Medline: 36899336]
25. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health 2005;95 Suppl 1:S144-S150. [doi: 10.2105/AJPH.2004.059204] [Medline: 16030331]
26. Lee SL, Pearce E, Ajnakina O, et al. The association between loneliness and depressive symptoms among adults aged 50 years and older: a 12-year population-based cohort study. Lancet Psychiatry 2021 Jan;8(1):48-57. [doi: 10.1016/S2215-0366(20)30383-7] [Medline: 33181096]
27. Zhang Y, Kuang J, Xin Z, et al. Loneliness, social isolation, depression and anxiety among the elderly in Shanghai: findings from a longitudinal study. Arch Gerontol Geriatr 2023 Jul;110:104980. [doi: 10.1016/j.archger.2023.104980] [Medline: 36863166]
28. Elmer T, Geschwind N, Peeters F, Wichers M, Bringmann L. Getting stuck in social isolation: solitude inertia and depressive symptoms. J Abnorm Psychol 2020 Oct;129(7):713-723. [doi: 10.1037/abn0000588] [Medline: 32672987]
29. Ibarra-Rovillard MS, Kuiper NA. Social support and social negativity findings in depression: perceived responsiveness to basic psychological needs. Clin Psychol Rev 2011 Apr;31(3):342-352. [doi: 10.1016/j.cpr.2011.01.005] [Medline: 21382539]

30. Roh HW, Hong CH, Lee Y, et al. Participation in physical, social, and religious activity and risk of depression in the elderly: a community-based three-year longitudinal study in Korea. PLoS One 2015;10(7):e0132838. [doi: 10.1371/journal.pone.0132838] [Medline: 26172441]
31. Song L, Ge Y, Zhang X. The relationship between WeChat use by Chinese urban older adults living alone and their subjective well-being: the mediation role of intergenerational support and social activity. Psychol Res Behav Manag 2021;14:1543-1554. [doi: 10.2147/PRBM.S330827] [Medline: 34629911]
32. Guo L. WeChat as a semipublic alternative sphere: exploring the use of WeChat among Chinese older adults. Int J Commun 2017 Jan 27;11:408-428 [FREE Full text]
33. Nan Y, Xie Y, Hu Y. Internet use and depression among Chinese older adults: the mediating effect of interpersonal relationship. Front Public Health 2023;11:1102773. [doi: 10.3389/fpubh.2023.1102773] [Medline: 36935716]
34. Cao Q, Buskens E, Hillege HL, Jaarsma T, Postma M, Postmus D. Stratified treatment recommendation or one-size-fits-all? A health economic insight based on graphical exploration. Eur J Health Econ 2019 Apr;20(3):475-482. [doi: 10.1007/s10198-018-1013-z] [Medline: 30374630]
35. Wang W, Yu Q, Hu XR, Wang XX. Heterogeneous analysis on association between exposure and outcome in cohort studies: based on propensity score matching and subpopulation treatment effect pattern plot. Chin J Epidemiol 2024 May 14;45(5):748-754. [doi: 10.3760/cma.j.cn112338-20231121-00303]
36. Goligher EC, Lawler PR, Jensen TP, et al. Heterogeneous treatment effects of therapeutic-dose heparin in patients hospitalized for COVID-19. JAMA 2023 Apr 4;329(13):1066-1077. [doi: 10.1001/jama.2023.3651] [Medline: 36942550]
37. Valkenburg PM. Social media use and well-being: what we know and what we need to know. Curr Opin Psychol 2022 Jun;45:101294. [doi: 10.1016/j.copsyc.2021.12.006] [Medline: 35016087]
38. Majeed M, Irshad M, Fatima T, Khan J, Hassan MM. Relationship between problematic social media usage and employee depression: a moderated mediation model of mindfulness and fear of COVID-19. Front Psychol 2020;11:557987. [doi: 10.3389/fpsyg.2020.557987] [Medline: 33391075]
39. Chiu CJ, Hu YH, Lin DC, Chang FY, Chang CS, Lai CF. The attitudes, impact, and learning needs of older adults using apps on touchscreen mobile devices: results from a pilot study. Comput Hum Behav 2016 Oct;63:189-197. [doi: 10.1016/j.chb.2016.05.020]
40. Zhang C, Tang D, Wang Y, Jiang S, Liu X. Community support and promoting cognitive function for the elderly. Front Psychol 2022;13:942474. [doi: 10.3389/fpsyg.2022.942474] [Medline: 36148108]

Abbreviations

aRR: adjusted relative risk
BADL: basic activities of daily living
CESD-10: 10-item Center for Epidemiologic Studies Depression Scale
CHARLS: China Health and Retirement Longitudinal Study
cRR: crude relative risk
SMU: social media usage

Edited by A Mavragani; submitted 04.06.24; peer-reviewed by S Sheth, S Munusamy, W Zheng; revised version received 15.11.24; accepted 17.11.24; published 13.01.25.

Please cite as:
Wang W, Wang H, Hu X, Yu Q, Chen F, Qiu X, Wang X
The Association Between Posting WeChat Moments and the Risk of Depressive Symptoms Among Middle-Aged and Older Chinese Adults: Prospective National Cohort Study
JMIR Public Health Surveill 2025;11:e62730
URL: https://publichealth.jmir.org/2025/1/e62730
doi:10.2196/62730

Original Paper

Identifying Community-Built Environment’s Effect on Physical Activity and Depressive Symptoms Trajectories Among Middle-aged and Older Adults: Chinese National Longitudinal Study

Kaili Zhang[1]; Bowen Huang[2]; Prasanna Divigalpitiya[3]

[1]Graduate School of Human-Environment Studies, Kyushu University, Fukuoka, Japan
[2]Zigong Academy of Urban Planning and Design, Zigong, Sichuan Province, China
[3]Faculty of Human-Environment Studies, Kyushu University, Fukuoka, Japan

Corresponding Author:
Bowen Huang
Zigong Academy of Urban Planning and Design
Number 19, Donghuan Road
Bancang Street, High-Tech Zone
Zigong, Sichuan Province, 643031
China
Phone: 86 15882073304
Email: jpby5223248@gmail.com

Abstract

Background: The effects of physical activity (PA) across different domains and intensities on depressive symptoms remain inconclusive. Incorporating the community-built environment (CBE) into longitudinal analyses of PA’s impact on depressive symptoms is crucial.

Objective: This study aims to examine the effects of PA at different intensities—low-intensity PA (eg, walking activities) and moderate-to-vigorous-intensity PA (eg, activities requiring substantial effort and causing faster breathing or shortness of breath)—across leisure-time and occupational domains on depressive symptom trajectories among middle-aged and older adults. Additionally, it investigated how CBEs influence depressive symptoms and PA trajectories.

Methods: This longitudinal study included 6865 middle-aged and older adults from the China Health and Retirement Longitudinal Survey. A CBE variable system was developed using a community questionnaire to assess attributes of the physical built environment. Depressive symptoms were measured using the Center for Epidemiologic Studies Depression Scale. Latent growth curve modeling was applied to analyze 3 waves of the cohort data (2015, 2018, and 2020) to explore the differential effects of PA on depressive symptoms and the role of the CBE.

Results: In the 2015 and 2018 waves, higher low-intensity leisure-time physical activity (LTPA) was associated with lower depressive symptoms (β=–.025, *P*=.01 and β=–.027, *P*=.005, respectively). Across all waves, moderate-to-vigorous-intensity LTPA showed no significant predictive effects (*P*=.21 in 2015, *P*=.57 in 2018, and *P*=.85 in 2020, respectively). However, higher occupational physical activity (OPA), particularly at moderate-to-vigorous intensities, was consistently associated with higher depressive symptoms. Parallel process latent growth curve modeling revealed that the initial level of total LTPA negatively predicted the initial level of depressive symptoms (β=–.076, *P*=.01). OPA exhibited dual effects, positively predicting the initial level of depressive symptoms (β=.108, *P*<.001) but negatively predicting their upward trajectory (β=–.136, *P*=.009). Among CBE variables, better infrastructure conditions (β=–.082, *P*<.001) and greater accessibility to public facilities (β=–.036, *P*=.045) negatively predicted the initial level of depressive symptoms. However, greater accessibility to public facilities positively predicted the upward trajectory of depressive symptoms (β=.083, *P*=.04). Better infrastructure conditions (β=.100, *P*=.002) and greater accessibility to public transport (β=.060, *P*=.01) positively predicted the initial level of total LTPA. Meanwhile, better infrastructure conditions (β=–.281, *P*<.001) and greater accessibility to public facilities (β=–.073, *P*<.001) negatively predicted the initial level of total OPA. Better infrastructure conditions positively predicted the declining trajectory of total OPA (β=.100, *P*=.004).

Conclusions: This study underscores the importance of considering the differential effects of PA across domains and intensities on depressive symptoms in public policies and guidelines. Given the influence of the environment on PA and depressive symptoms, targeted community measures should be implemented.

(JMIR Public Health Surveill 2025;11:e64564) doi:10.2196/64564

KEYWORDS

community-built environment; physical activity; depressive symptom trajectories; middle-aged and older adults; latent growth curve modeling; longitudinal study

Introduction

In communities, the physical built environment and health behaviors, such as physical activity (PA), can significantly influence residents’ mental health [1]. Overall, well-designed community-built environments (CBEs)—including factors such as walkability, appropriate density, and safety—are beneficial in reducing depressive symptoms. By contrast, unfavorable built environments, such as poor-quality housing and insufficient green spaces, may exacerbate depressive symptoms [2]. Exercise and PA have been shown to be as effective as some antidepressant medications in improving depressive symptoms [3].

With the acceleration of the global aging process [4], depression among middle-aged and older adults has become a significant public health issue worldwide. The World Health Organization (WHO) estimates that the prevalence of depression in the global aging population ranges from 10% to 20% [5]. In China, the pooled prevalence of depressive symptoms among seniors aged 60 years and older is approximately 22.7% [6]. Depression poses a serious threat to the mental health of middle-aged and older adults and can also lead to physical illness and an increased risk of death [7].

In all studies examining factors influencing depressive symptoms, PA has become a focal point. PA can be easily incorporated into the daily context of a community, whether in a living or work environment, making it a universally accessible and straightforward intervention. Its practicality and ease of implementation have made PA a key topic in discussions about improving mental health [8]. The beneficial effects of PA on alleviating depressive symptoms are well-documented. Physically inactive individuals are more likely to experience depressive symptoms compared with those who engage in regular PA [9]. Recent studies have shown that this effect is not influenced by age [10]. Middle-aged and older adults are particularly susceptible to depressive disorders due to factors such as the gradual decline in physical function, an increase in chronic illnesses, and the potential for significant life changes, including widowhood, retirement, and living alone after middle age [11]. Older adults who engage in insufficient PA are at a higher risk of depression compared with those who maintain adequate levels of PA [12].

Many studies exploring the relationship between PA and depressive symptoms have primarily focused on exercise-based PA [13]. However, PA can also occur in occupational settings or outside structured exercise programs. PA shows different health outcomes depending on the context, such as work versus leisure settings [14]. Recently, researchers have begun to examine the “PA paradox”—the concept that leisure-time physical activity (LTPA), such as sports and recreational activities, has positive effects on health, while PA in occupational settings may have negative consequences for both physical and mental health [15]. LTPA, regardless of intensity, is associated with a lower risk of depressive symptoms. By contrast, occupational PA (OPA) or transportation-related PA does not show a similar reduction in the risk of depressive symptoms [16]. Workers who engage in work-related physical activities tend to report higher levels of depressive symptoms [17], and individuals with the highest levels of OPA are more likely to experience depressive symptoms [18]. Globally, a significant proportion of people continue to engage in manual labor, particularly in many developing countries. Therefore, it is important to distinguish between different domains or motivations when examining the effect of PA on depressive symptoms, as this distinction may lead to entirely different results. However, only a small number of studies have made such distinctions to examine the effects of various PA domains on depressive symptoms [19].

There are still several limitations and gaps in the existing research on the effects of PA on depressive symptoms. First, while the beneficial effects of exercise on depressive symptoms are well-established, the complex impacts of different domains (such as exercise and work) or PA intensity remain inconclusive [20]. There is still a lack of basic evidence, particularly for middle-aged and older adults, and further research is urgently needed to provide clearer insights. Second, most studies on the effect of PA on depressive symptoms offer only a cross-sectional perspective, with few utilizing cohort data to explore the longitudinal impact. By focusing on individual trajectories and patterns over time, longitudinal studies can provide valuable insights into the developmental continuum, help determine the stability of the association between PA and depressive symptoms, and analyze the dynamics of these interactions [21].

Furthermore, the CBE is an important factor influencing mental health, including depressive symptoms and mental health inequalities [22]. Depressive symptoms are associated with the built environment, and different physical attributes of the community may impact them through various behavioral or social pathways [23]. Many characteristics of the CBE may expose residents to environmental stressors—such as poor housing quality, uncomfortable population densities, low safety, and poor accessibility to amenities—that can negatively affect their mental health [24]. As for built environment measures, the most common ones are related to the “5D” elements: density, diversity, design, destination accessibility, and distance to transit

[25]. These elements are often evaluated individually or together.

Despite the growing body of literature on the effects of CBE on depressive symptoms, there is still limited evidence concerning middle-aged and older adults. Few studies have highlighted that improving community walkability, particularly the accessibility and safety of public transportation, is crucial for middle-aged individuals with depression [2]. Higher levels of residential greenery have been significantly associated with lower odds of depressive symptoms among older adults living in the community [26]. Additionally, high population density and proximity to destinations contribute to better mental health, while walkable environments are linked to reduced depression, particularly among Asian women [27]. Conversely, the availability of local retail stores has been associated with an increased likelihood of depressive symptoms in older men [28]. Thus, inconsistent results may arise in different contexts or regions. Therefore, it is crucial to consider the local community context when assessing the effects of CBE on depressive symptoms and PA. Furthermore, among longitudinal studies exploring the relationship between PA and depressive symptoms, few have taken CBE into account [29]. Therefore, providing insights into how CBE influences the trajectories of depressive symptoms and PA is invaluable.

To address these gaps, this study focused on middle-aged and older adults aged 45 years and older in China, using data from 3 waves of repeated measurements to identify the trajectory of depressive symptoms, various types of PA, and the built environment of communities. The study aimed to answer the following 2 research questions: (1) What are the longitudinal effects of LTPA and OPA on depressive symptoms among middle-aged and older adults? (2) How does the CBE influence the trajectory of depressive symptoms and the 2 domains of PA in this population?

Methods

Data Source and Study Sample

The data for this study were obtained from the latest 3 waves of tracking data (2015, 2018, and 2020) from the China Health and Retirement Longitudinal Survey (CHARLS), a large microsurvey database. CHARLS covers 450 communities across 28 provinces in China, using a multistage stratified probability proportional to size sampling method to conduct the survey. A detailed description can be found in Appendix S1 in Multimedia Appendix 1 (see also [30-32]). As part of the International Survey of Health in Aging, CHARLS has been extensively utilized in interdisciplinary studies focusing on aging and health issues [33].

Figure S1 in Multimedia Appendix 1 illustrates the sample screening process. For each wave of data, this study excluded samples with missing PA data (missing PA data were randomly absent: participants were randomly selected to answer PA-related questions, and the absence of available records indicated that they did not respond to PA-related questions) and missing depressive symptom scores (individuals who left more than 2 questions unanswered on the scale were considered invalid for the depressive symptom measure. Among the respondents excluded from the analysis, approximately 80% had missing data for all 10 items on the scale in each wave). Additionally, samples lost due to attrition or participant death were excluded. After matching data across the 3 waves, participants younger than 45 years were also excluded. This resulted in a final sample of 6865 respondents, with 20,595 observation points. For samples with missing values for sex and education level, data from adjacent waves were used.

Ethics Approval

The original data collection was approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052-11015). All data used in this study were fully anonymized, and participants provided informed consent. The data for this study were registered and obtained through the official CHARLS website [34].

Depressive Symptoms

Depressive symptoms were assessed using the Center for Epidemiologic Studies Depression Scale (CES-D), as provided by CHARLS. Respondents answered 10 questions reflecting feelings and behaviors associated with depression [35]. The total score, which ranged from 0 to 30, was calculated by summing the responses to these 10 questions, with reverse scoring applied to the 2 positive questions. Higher CES-D scores indicated more severe depressive symptoms. The Cronbach α coefficients for the CES-D scale in this study were 0.794, 0.800, and 0.796 in 2015, 2018, and 2020, respectively. In the descriptive statistical analysis, a CES-D score of ≥10 was classified as indicative of depression [36]. Additionally, we calculated the Reliable Change Index (RCI) for the CES-D scores across waves to assess whether changes in participants' depressive symptom scores over multiple waves were statistically significant ($|RCI|>1.96$) [37].

Physical Activity Level

The CHARLS used items similar to those in the International Physical Activity Questionnaire (short form) to assess respondents' levels of PA. It collected data on the duration and weekly frequency of each type of PA: low intensity (walking activities), moderate intensity (activities that cause faster breathing than usual, such as cycling, playing table tennis, and practicing yoga), and high intensity (activities that cause shortness of breath, such as fast running and playing basketball) [38]. This study used the metabolic equivalent (MET) method to calculate PA levels. To determine an individual's weekly MET for each PA intensity, we multiplied the corresponding duration, frequency, and coefficient for each activity. Following previous studies [30,31], we performed data organization and calculations for the 3 intensity levels of PA, including data cleaning, outlier elimination, and data truncation. Details regarding the CHARLS Questionnaire on PA, as well as the data processing and calculation methods used in this study, are provided in Appendix S1 in Multimedia Appendix 1. We scaled all PA data by a factor of 100 to meet the data requirements for statistical analysis before inputting them into the model calculation. As a result, the actual unit of MET reported in the models was 100 MET minutes per week.

The CHARLS database collected information on the purpose of each PA (4 categories: work, exercise, recreation, and others), allowing for the examination of the differential effects of various PA domains on depressive symptoms. Activities labeled as "other" in CHARLS were not explicitly defined and included a mix of multiple purposes, so they were excluded from the calculations for this study. PA data were categorized into 2 domains: LTPA and OPA. Within each domain, PA was further classified by intensity: low intensity (walking activities) and moderate-to-vigorous intensity (activities that require significant effort and lead to faster breathing or shortness of breath). The corresponding MET values were calculated and aggregated.

Construction of the Community-Built Environment Variables System

In this study, a set of variables was developed to assess the physical built environments of communities based on the community questionnaire data set provided by CHARLS. This data set was released in a single issue as part of the national baseline survey, updated in December 2013, with no subsequent updates in later waves. Based on the concepts of 3D, 5D, and other typical indicator systems and their extensions [25,39], and considering data availability and the characteristics of middle-aged and older adults, this study selected 7 core variables to evaluate the CBE attributes of each community: residential density, density of public facilities, diversity of sports venues, diversity of senior care facilities, accessibility of public facilities, accessibility of public transport, and infrastructure conditions. These variables are outlined in Table 1.

Table 1. Community-built environment variable system constructed from the community questionnaire of the China Health and Retirement Longitudinal Survey.

Built environment element	Variable	Calculation method: unit (weight)
Density	• Residential density • Density of public facilities	• Community population/total area of the community: pieces/km2 (N/A[a]) • The total number of basic public facilities in the community/total area of the community: pieces/km2 (N/A)
Diversity	• Diversity of sports venues • Diversity of senior care facilities	• Whether or not the community has any of the 8 sports venues: a "yes" answer will be scored as 1 point for that venue, a "no" answer will be scored as 0 points for that venue, and the total values will be added up. The maximum score is 8: points (N/A) • Whether or not the community has any of the 7 senior care facilities; diversity levels were calculated as described above.
Destination accessibility	• Accessibility of public facilities	• The average distance from the community office to the most commonly used type of each facility (if this facility is located within the community, the answer is 0): km (N/A).
Distance to transit	• Accessibility of public transport	• The actual distance from the community office to the most commonly used bus stop: km (43.86%). • The total number of actual bus lines accessible to the community: pieces (56.14%).
Infrastructure	• Infrastructure conditions	• Pavement quality of the main roads, where pathway/dirt/unpaved road=1 point, sandstone road=2 points, and paved road=3 points: points (13.62%) • The total number of days roads are impassable: days (12.31%) • The proportion of households that use purified tap water: percentage (18.30%) • Availability of sewer system in the community. yes=1 point and no=0 points: points (17.56%) • The level of waste disposal, where moved away by truck=5 points, buried in the village=4 points, burn away=3 points, put into nearby river=2 points, and do not manage=1 point: points (17.13%) • The proportion of households that use electricity: percentage (17.07%) • The condition of community toilets, where inside toilet with water=5, inside toilet without water=4, outside toilet with water=3, outside public toilet without water=2, and open air=1: points (4.00%)

[a] N/A: not applicable.

Among the variables, the accessibility of public transport and infrastructure conditions were assessed by calculating multiple subvariables. The CRITIC weighting method was used to determine the weight of each subvariable. Detailed information on the content of the facilities and the source of the question numbers can be found in Table S1 in Multimedia Appendix 1. Furthermore, due to limitations in the survey data, this study could not include variables such as green spaces and road network connectivity, despite their recognized importance [40,41]. All CBE variables were normalized using the min-max normalization method, resulting in a processed data range of 0-1, before being included in the model calculations. Further details can be found in Appendix S1 in Multimedia Appendix 1.

Covariates

Three types of variables were used as control variables: demographic and socioeconomic status, health behavior and status, and housing characteristics. These variables were selected due to their association with mental health levels or their potential impact on depressive symptoms [42]. Baseline data were chosen as covariates for inclusion in the model calculations. Specifically, demographic and socioeconomic status included age, sex (male or female), education level (categorized into 4 groups: illiteracy; primary and lower; middle, high, and vocational school; and 3-year college/bachelor's degree and higher), marital status (married or other), cohabiting status (with or without a cohabiting partner), and personal annual income. Personal annual income encompasses salaries and transfer income, such as pensions, benefits, and subsidies.

Health behavior and status included smoking (current smoker or nonsmoker), drinking (categorized into 3 groups: drinking more than once a month, drinking less than once a month, and nondrinker), the number of chronic diseases (total of 10, as outlined in Appendix S2 and Table S2 in Multimedia Appendix 1), and disability status (disabled or nondisabled). Housing characteristics were assessed based on the presence of various amenities: elevator, barrier-free facilities, toilet flushing, electricity, running water, bathing facility, gas/natural gas, heating, broadband, air purifier, and tidiness, for a total of 11 items. One point was assigned for each present characteristic or amenity, and the total score was calculated. Higher scores indicate better housing quality.

Statistical Analysis

This study utilized latent growth curve modeling (LGCM) to analyze the trajectories and longitudinal associations among the variables. LGCM, a variant of structural equation modeling, is a statistical method commonly used in longitudinal studies with latent variables [43]. It uses the intercept to represent the initial level and the slope to capture the rate of change. First, an unconditional linear LGCM was applied to each of the 3 measures of PA and depressive symptoms to examine their trajectories in middle-aged and older adults. In the second step, all control variables were incorporated into a linear growth model to construct a conditional linear LGCM for depressive symptoms. PA was included as a time-varying covariate to examine the effect of each wave of PA on depressive symptoms. To test whether these effects remained stable over time (2015-2018-2020), equality constraints were applied to the effects at each wave, and likelihood ratio tests (LRTs) were conducted to assess the statistical equivalence of these effects. In the third step, an LGCM was constructed for both depressive symptoms and PA, with 7 categories of built environment (CBE) as time-invariant covariates. A parallel process LGCM was used to examine the dynamic associations among CBE, PA, and depressive symptoms. The study also incorporated the adjusted individual weights provided by CHARLS for each wave of data in all models to minimize representational bias due to response differences. The acceptable thresholds for model fit indices were as follows: Comparative Fit Index>0.90, Tucker-Lewis Index>0.90, root mean square error of approximation<0.08, and standardized root mean square residual<0.08 [44]. Additionally, the Harman single-factor test was conducted to examine the data for common method bias. The total variance explained by the first common factor was 19.76% (<40%), indicating that the study was not significantly affected by common method bias [45]. Data cleaning was performed using Stata 18.0 (StataCorp), all LGCM analyses were conducted with Mplus 8.10 (Muthén and Muthén), and the Harman single-factor test was executed using IBM SPSS 26.

Results

Characteristics of the Sample at Baseline

A total of 6865 participants were included in the study sample, with the detailed statistical characteristics presented in Table S2 in Multimedia Appendix 1. The male-to-female ratio in the sample was 0.88, indicating a nearly balanced distribution. The respondents' education levels were relatively low: of the 6865 participants, 4652 (67.76%) did not receive education beyond secondary school, reflecting the social context of the middle-aged and older adult generation in China. The marriage rate in the sample was very high (6836/6865, 99.58%), and a majority of respondents (5767/6865, 84%) lived with a cohabiting partner. The mean annual personal income of participants was 15,712.63 Chinese yuan (1 yuan=US $0.14). Among the respondents, the majority were nonsmokers (4986/6865, 72.63%) and nondrinkers (4400/6865, 64.09%). On average, participants reported having 2.10 chronic diseases, and the average score for housing characteristics was 4.06. Additionally, 872 (12.70%) participants reported having a disability.

Developmental Trajectories of Physical Activity and Depressive Symptoms

The statistics for depressive symptoms and various types of PA among middle-aged and older adults across the 3 data waves are presented in Table 2. The proportion of participants with depression (CES-D scores≥10) ranged from approximately 30% to 40% over the study period (2178/6865, 31.73%, in 2015; 2469/6865, 35.97%, in 2018; and 2666/6865, 38.83%, in 2020). Additionally, in each wave, participants consistently reported higher levels of OPA than LTPA. The amount of moderate-to-vigorous-intensity OPA was greater than that of low-intensity OPA, while low-intensity LTPA was more prevalent than moderate-to-vigorous-intensity LTPA.

Unconditional LGCM was used to examine the developmental trends of each variable. The model fit indices, along with the mean intercepts and slopes for the 3 models, are presented in Table 3. Overall, the models demonstrated a good fit. Depressive symptoms and LTPA levels showed a linear upward trend across the 3 measurement waves (with positive and significant slopes), while OPA levels displayed a linear downward trend. Furthermore, based on the RCI statistics for changes in depressive symptoms (Appendix S2 and Table S3 in Multimedia Appendix 1), despite the overall increasing trend in depressive symptoms, most participants did not show significant changes across waves (–1.96≤RCI≤1.96). Specifically, 5506 of 6865 (80.20%) individuals from 2015 to 2018, 5483 of 6865 (79.87%) individuals from 2018 to 2020, and 5391 of 6865 (78.53%)

individuals from 2015 to 2020 showed no significant changes (–1.96≤RCI≤1.96). Over the longer period from 2015 to 2020, 948 of 6865 (13.81%) participants experienced a significant increase (RCI>1.96) in depressive symptoms, while 526 of 6865 (7.66%) participants experienced a significant decrease (RCI<–1.96).

Table 2. Statistical characteristics of depressive symptoms and physical activity among middle-aged and older adults across three waves of data.[a]

Variable	Wave 2015	Wave 2018	Wave 2020
Depressive symptoms, mean (SD)	7.81 (6.27)	8.46 (6.32)	8.85 (6.46)
Number of depression, n/N (%)	2178/6865 (31.73)	2469/6865 (35.97)	2666/6865 (38.83)
Number of nondepression, n/N (%)	4687/6865 (68.27)	4396/6865 (64.03)	4199/6865 (61.17)
Total LTPA[b], mean (SD)	974.38 (2130.84)	1248.28 (2412.47)	1590.66 (2748.37)
Low-intensity LTPA, mean (SD)	593.47 (1198.14)	792.45 (1356.73)	902.70 (1356.9)
Moderate-to-vigorous-intensity LTPA, mean (SD)	380.91 (1514.96)	455.83 (1719.31)	687.96 (2140.36)
Total OPA[c], mean (SD)	5529.50 (6988.45)	4834.27 (6593.79)	3745.39 (5744.90)
Low-intensity OPA, mean (SD)	1234.15 (1907.24)	1190.91 (1925.46)	719.76 (1570.59)
Moderate-to-vigorous-intensity OPA, mean (SD)	4295.35 (5985.89)	3643.36 (5600.78)	3025.63 (5064.63)

[a] A score of the Center for Epidemiologic Studies Depression Scale ≥10 was defined as depression. The unit of physical activity in this table is metabolic equivalent minutes/week.

[b] LTPA: leisure-time physical activity.

[c] OPA: occupational physical activity.

Table 3. Estimates and fit index of unconditional LGCMs.[a]

Estimates	LGCM[b] for depressive symptoms		LGCM for total LTPA[c]		LGCM for total OPA[d]	
	Estimate	SE	Estimate	SE	Estimate	SE
Means						
Intercept	7.855[e]	0.073	9.661[e]	0.246	56.000[e]	0.803
Slope	0.516[e]	0.038	3.048[e]	0.190	–9.087[e]	0.439
Covariance (intercept with slope)	–0.319	0.486	–18.383[f]	8.409	–46.701[e]	5.570
Fit index						
Chi-square (*df*)	4.180 (1)	N/A[g]	1.182 (1)	N/A	3.442 (1)	N/A
Root mean square error of approximation (90% CI)	0.022 (0.004-0.045)	N/A	0.005 (0.000-0.033)	N/A	0.031 (0.013-0.053)	N/A
Comparative Fit Index	0.999	N/A	1.000	N/A	0.998	N/A
Tucker-Lewis Index	0.998	N/A	0.999	N/A	0.994	N/A
Standardized root mean square residual	0.006	N/A	0.004	N/A	0.008	N/A

[a] The unit of physical activity in this table is 100 metabolic equivalent minutes/week.

[b] LGCM: latent growth curve modeling.

[c] LTPA: leisure-time physical activity.

[d] OPA: occupational physical activity.

[e] *P*<.001.

[f] *P*<.05.

[g] N/A: not applicable.

Effects of Leisure-Time and Occupational Physical Activity on Depressive Symptoms

Based on the time-varying covariate LGCM specified in Figure 1, the direct effects of all types of PA on depressive symptoms were calculated for each wave, and equality constraints were tested. The results are presented in Table 4. The model fit for all pathways was satisfactory (Appendix S2 and Table S4 in Multimedia Appendix 1). In both 2015 and 2018, total LTPA (β=–.023, *P*=.02 and β=–.025, *P*=.01, respectively) and low-intensity LTPA (β=–.025, *P*=.01 and β=–.027, *P*=.005, respectively) negatively predicted depressive symptoms.

However, this effect was not statistically significant in 2020 (*P*=.51). No significant differences were observed in the pathways across multiple waves, as the LRT results were not significant (*P*=.31 in 2015, *P*=.65 in 2018, and *P*=.14 in 2020, respectively). Moderate-to-vigorous-intensity LTPA did not show a significant predictive effect on depressive symptoms in any wave (*P*=.21 in 2015, *P*=.57 in 2018, and *P*=.85 in 2020, respectively).

By contrast, in waves 2015, 2018, and 2020, both total OPA (β=.041, *P*<.001; β=.022, *P*=.02; and β=.027, *P*=.03, respectively) and moderate-to-vigorous-intensity OPA (β=.036, *P*<.001; β=.020, *P*=.04; and β=.019, *P*=.049, respectively) positively predicted depressive symptoms. The positive predictive effect of low-intensity OPA on depressive symptoms was significant only in wave 2015 (β=.030, *P*=.003). Furthermore, the LRT results indicated that the effect of low-intensity OPA on depressive symptoms differed significantly between waves 2015 and 2018 (*P*=.048) and between waves 2015 and 2020 (*P*=.02). Across the 3 waves of data, the effects of both total OPA (*P*=.27 in 2015; *P*=.98 in 2018; and *P*=.33 in 2020, respectively) and moderate-to-vigorous-intensity OPA on depressive symptoms remained stable with no significant differences observed (*P*=.27 in 2015, *P*=.87 in 2018, and *P*=.41 in 2020, respectively). Given the representativeness, statistical significance, and stability of the effects across waves, this study proceeded with the inclusion of total LTPA and total OPA for further analysis.

Figure 1. Graphical representation of the time-varying covariate latent growth curve modeling with the depressive symptoms and physical activity as time-varying variables. The 11 covariates as time-invariant covariates were plotted in a box for mapping simplicity. They were calculated concurrently in each latent growth curve model analysis.

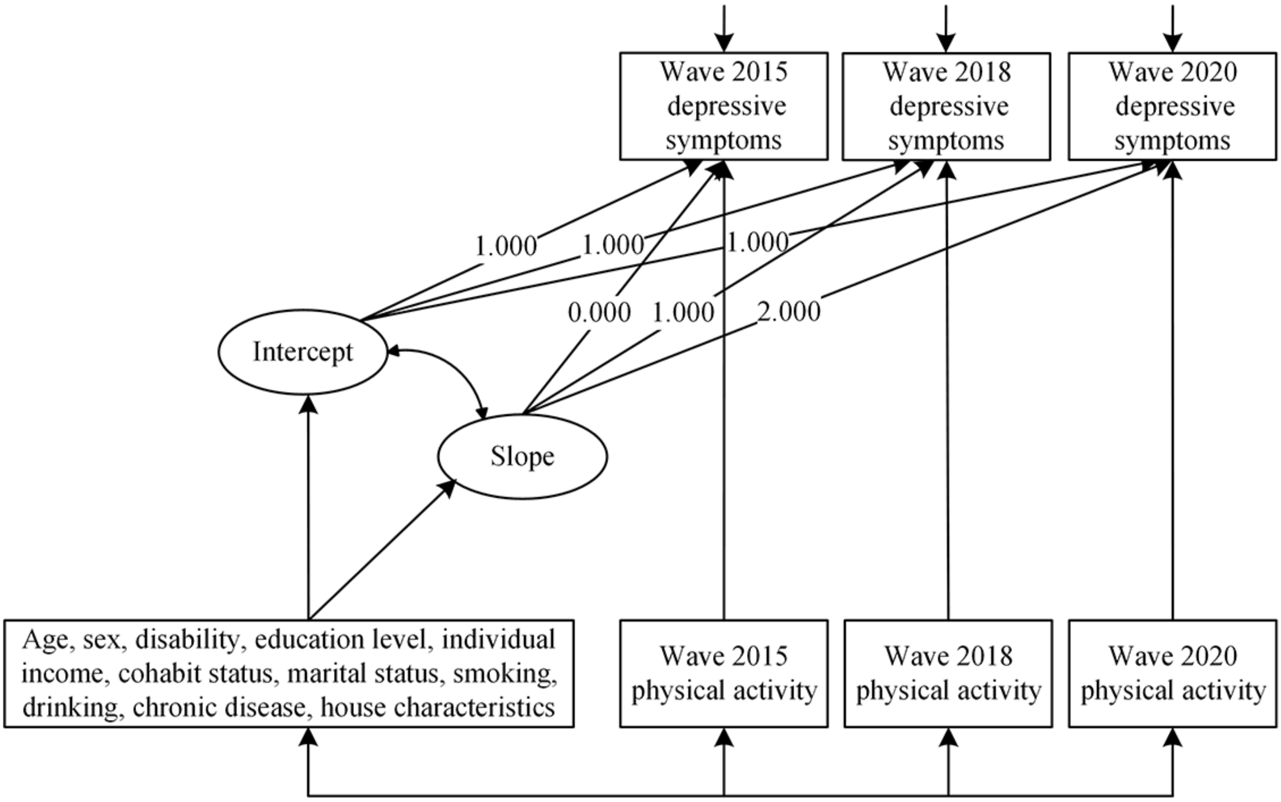


Table 4. Standardized estimate, SE (in parentheses), and significance of the time-varying covariate latent growth curve modeling.[a]

Variables	Wave 2015 depressive symptoms	Wave 2018 depressive symptoms	Wave 2020 depressive symptoms	LRT[b] 2015 vs 2018	LRT 2018 vs 2020	LRT 2015 vs 2020
2015 Total LTPA[c]	*–0.023 (0.010)*[d]	N/A[e]	N/A	*P*=.31	N/A	N/A
2018 Total LTPA	N/A	*–0.025 (0.010)*[d]	N/A	N/A	*P*=.65	N/A
2020 Total LTPA	N/A	N/A	–0.015 (0.022)	N/A	N/A	*P*=.14
2015 Low-intensity LTPA	*–0.025 (0.010)*[d]	N/A	N/A	*P*=.93	N/A	N/A
2018 Low-intensity LTPA	N/A	*–0.027 (0.010)*[f]	N/A	N/A	*P*=.19	N/A
2020 Low-intensity LTPA	N/A	N/A	*–0.010 (0.010)*	N/A	N/A	*P*=.20
2015 Moderate-to-vigorous-intensity LTPA	–0.012 (0.010)	N/A	N/A	*P*=.18	N/A	N/A
2018 Moderate-to-vigorous-intensity LTPA	N/A	0.005 (0.009)	N/A	N/A	*P*=.57	N/A
2020 Moderate-to-vigorous-intensity LTPA	N/A	N/A	–0.002 (0.009)	N/A	N/A	*P*=.35
2015 Total OPA[g]	*0.041 (0.010)*[h]	N/A	N/A	*P*=.27	N/A	N/A
2018 Total OPA	N/A	*0.022 (0.010)*[d]	N/A	N/A	*P*=.98	N/A
2020 Total OPA	N/A	N/A	*0.027 (0.012)*[d]	N/A	N/A	*P*=.33
2015 Low-intensity OPA	*0.030 (0.010)*[f]	N/A	N/A	*P=.048*[d]	N/A	N/A
2018 Low-intensity OPA	N/A	0.012 (0.010)	N/A	N/A	*P*=.22	N/A
2020 Low-intensity OPA	N/A	N/A	–0.005 (0.010)	N/A	N/A	*P=.02*[d]
2015 Moderate-to-vigorous-intensity OPA	*0.036 (0.010)*[h]	N/A	N/A	*P*=.27	N/A	N/A
2018 Moderate-to-vigorous-intensity OPA	N/A	*0.020 (0.010)*[d]	N/A	N/A	*P*=.87	N/A
2020 Moderate-to-vigorous-intensity OPA	N/A	N/A	*0.019 (0.010)*[d]	N/A	N/A	*P*=.41

[a] Italicized values indicate significant results.

[b] LRT: likelihood ratio test.

[c] LTPA: leisure-time physical activity.

[d] *P*<.05.

[e] N/A: not applicable.

[f] *P*<.01.

[g] OPA: occupational physical activity.

[h] *P*<.001.

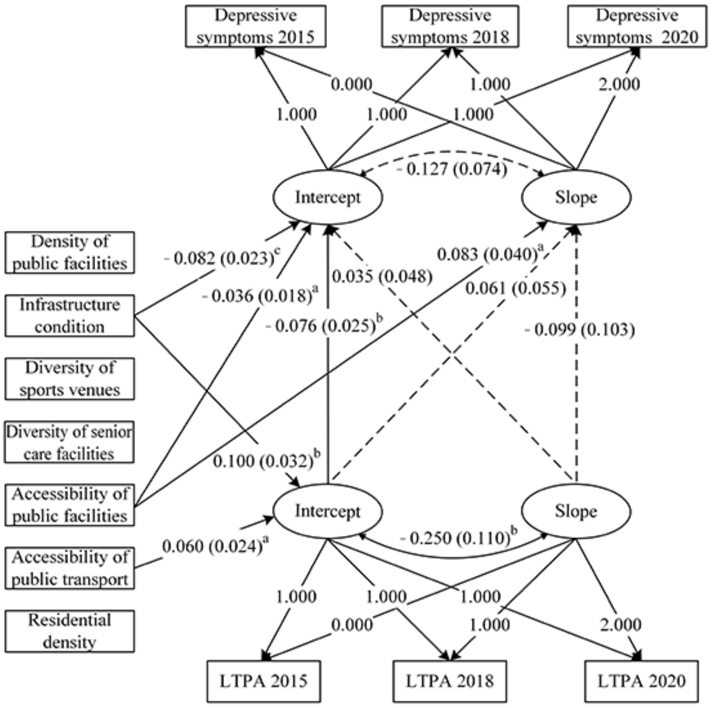
Longitudinal Associations Among Community-Built Environment, Physical Activity, and Depressive Symptoms

This study further used a parallel process LGCM to analyze the dynamic associations between CBE, the 2 domains of PA, and depressive symptoms (Figure 2). The models demonstrate a good fit (Appendix S2 and Table S4 in Multimedia Appendix 1). First, the intercept (initial level) of total LTPA negatively predicted the initial level of depressive symptoms (β=–.076, *P*=.003). The initial level of total OPA positively predicted the initial level of depressive symptoms (β=.108, *P*<.001) and negatively predicted the rate of increase in depressive symptoms (β=–.136, *P*=.009).

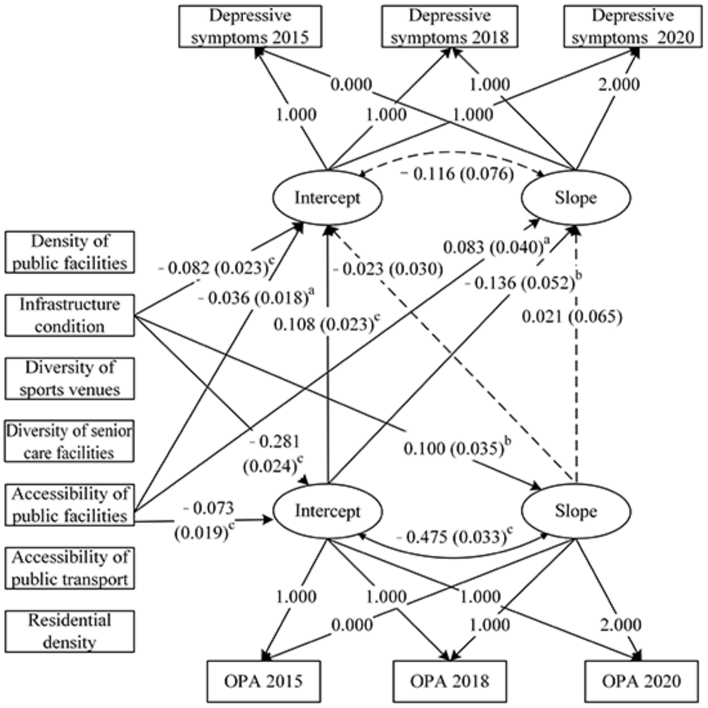
Regarding the effect of CBE on the trajectory of depressive symptoms, infrastructure conditions and accessibility of public facilities negatively predicted the initial level of depressive symptoms (β=–.082, *P*<.001 and β=–.036, *P*=.045, respectively). However, accessibility to public facilities positively predicted the rate of increase in depressive symptoms (β=.083, *P*=.04). Additionally, infrastructure conditions (β=.100, *P*=.002) and accessibility of public transport (β=.060, *P*=.01) positively predicted the intercept of total LTPA. By contrast, infrastructure conditions (β=–.281, *P*<.001) and accessibility of public facilities (β=–.073, *P*<.001) negatively predicted the initial level of total OPA. Infrastructure conditions also positively predicted the declining trend of total OPA over time (β=.100, *P*=.004).

Figure 2. Parallel process latent growth curve modeling to estimate the effect of physical activity on depressive symptoms, with the community-built environment serving as a time-invariant covariate. Dashed lines indicate nonsignificant paths. The nonsignificant paths for the 7 community-built environment variables are not displayed. All control variables are evaluated simultaneously. ^{a}*P*<.05; ^{b}*P*<.01; ^{c}*P*<.001. LTPA: leisure-time physical activity; OPA: occupational physical activity.

(A) Path: Total leisure-time physical activity → Depressive symptoms



(B) Path: Total occupational physical activity → Depressive symptoms



Discussion

Longitudinal Effect of Physical Activity on Depressive Symptoms Among Middle-Aged and Older Adults

First, the analysis based on the linear unconditional LGCM revealed an upward trend in depressive symptoms among middle-aged and older adults from 2015 to 2020. This finding aligns with previous studies using CHARLS data [33], suggesting that depressive symptoms may be influenced by the aging process. However, further analysis using the RCI indicated that while depressive symptoms showed an upward trend, this change appeared to occur primarily at the group level. At the individual level, fluctuations in depressive symptoms were relatively small for most participants, with only a minority experiencing significant increases. For the majority, changes did not reach statistically significant levels (Table S3 in Multimedia Appendix 1). Additionally, the average levels of depressive symptoms across the 3 waves were slightly below the CES-D clinical screening threshold (Table 2) [36], indicating that most participants exhibited mild or subclinical depressive symptoms. Therefore, the core focus of this study is to understand and explore the potential effects of CBE and PA on changes in depressive symptoms. These changes should not be overly simplified as the onset of clinical depression or an increase in its severity.

Existing studies examining the effects of PA on depressive symptoms often focus on a single domain or intensity level. However, given that PA across different intensities and domains may have varying effects on depressive symptoms [46], our study contributes to the literature by further categorizing PA into low- and moderate-to-high-intensity levels within the domains of LTPA and OPA. These results highlight the importance of the differences observed. By analyzing data from 3 waves of repeated measurements, we found that LTPA, particularly low-intensity LTPA, consistently served as a negative predictor of depressive symptoms among middle-aged and older adults in China. This suggests that greater participation in low-intensity LTPA was associated with fewer depressive symptoms in this population (Table 4). Our findings are consistent with previous studies that emphasize the benefits of moderate and light exercise over vigorous exercise for middle-aged and older adults [47]. Engaging in lower-intensity activities, such as walking, may also help prevent depression [48], with lower LTPA levels being associated with fewer depressive symptoms [20].

However, in all waves of the analysis, we did not observe a significant association between moderate-to-vigorous-intensity LTPA and depressive symptoms. This finding is inconsistent with some existing literature. For example, certain cross-sectional studies have suggested that both low- and moderate-to-vigorous-intensity PA positively impact the mental health of middle-aged and older adults [49]. Furthermore, both small and large amounts of LTPA are associated with a lower prevalence of depression, with comparable effects [19]. These discrepancies may arise from differences in how PA intensity is categorized, the participant selection criteria, or regional characteristics across the studies.

This study also clearly observed that after controlling for socioeconomic factors such as income levels, higher moderate-to-vigorous-intensity OPA was associated with higher depressive symptoms. This negative predictive effect remained significant and consistent across all waves of data (Table 4). This result aligns with previous studies. Research from Brazil and South Korea reported that participants with higher levels of OPA were more likely to experience depressive symptoms [20], and that the higher the level of work-related PA, the greater the likelihood of depressive symptoms [50]. Similarly, OPA appears to be associated with an increased risk of depression [51]. However, these results contradict those of other studies. For instance, research from the United States suggested that OPA does not significantly predict depression regardless of gender [52], while a study from Ghana found that depression decreased as work-related PA increased [53]. As our study controlled for socioeconomic characteristics, these differences suggest that the effect of OPA on depressive symptoms may vary depending on the level of economic development or overall social progress in a given region.

Collectively, these findings support the important perspective that the effects of PA on depressive symptoms vary depending on the domain and intensity of the activity. This study adds to the existing research on the PA paradox and PA parameters [54]. Overall, our findings validate and expand upon the PA paradox in the mental health domain, showing that LTPA, particularly low-intensity exercise in LTPA, is associated with lower depressive symptoms, whereas moderate-to-vigorous-intensity OPA is linked to higher depressive symptoms. These 2 distinct forms of PA exert opposite effects on depressive symptoms in middle-aged and older adults [51]. Leisure exercise and occupational labor, as fundamentally different drivers of PA, likely result in varying degrees of subjective willingness and well-being, which may have divergent long-term effects on mental health [55]. Particularly in many developing countries, moderate-to-vigorous-intensity OPA is often linked to limited individual job skills and is usually accompanied by poorer working conditions and lower levels of welfare protection.

When examining the dynamic effects of PA on depressive symptoms, this study further revealed the heterogeneous characteristics of different domains of PA. Using parallel-processed LGCM, we found that the initial level of total LTPA in middle-aged and older adults negatively predicted the initial level of depressive symptoms. This indicated that individuals with higher initial LTPA levels tended to have lower initial levels of depressive symptoms (Figure 2). This finding is consistent with the results of Heesch et al [56], who reported that low LTPA and walking levels were associated with reduced anxiety and depression scores. One possible explanation is that individuals with depressive symptoms are less likely to engage in PA. Additionally, confounding factors may have influenced this association.

This study also identified, for the first time, that OPA is not only significantly positively associated with the initial level of depressive symptoms, but also significantly predicts the trajectory of change in depressive symptoms. Specifically, middle-aged and older adults with higher initial OPA levels tended to exhibit higher initial levels of depressive symptoms. However, higher OPA levels appeared to suppress the upward trajectory of depressive symptoms over time (Figure 2). This finding highlights the complexity and duality of OPA's influence on the depressive symptom trajectory. On the one hand, high-intensity or repetitive OPA may increase physical burdens, contributing to higher initial depressive symptoms. On the other hand, the "steeling effect" resulting from prolonged OPA may enhance individuals' resilience to negative experiences and improve physiological adaptability, thereby mitigating the increase in depressive symptoms over time [57,58]. Moreover, OPA may exert long-term positive effects on mental health through mechanisms such as improving physical metabolism, maintaining a regular daily routine, and reducing sedentary behavior. These benefits may partially offset the negative effects.

Effects of Community-Built Environment on Trajectories of Depressive Symptoms and Physical Activity

To the best of our knowledge, few studies have attempted to develop indicator systems or variable sets for international microsurvey databases to assess a community's physical environment. Jones-Smith and Popkin [59] proposed a scale for assessing community contextual characteristics but did not pay enough attention to physical built environments. Taking the CHARLS community questionnaire as an example, this study developed CBE indicators for international microsurvey databases, which can provide references for other studies. Specifically, incorporating the variables of the diversity of senior care facilities and sports venues directly addresses the unique needs of adults aged 45 years or older. Assessing the level of access to senior care and exercise facilities within the community is particularly important for this age group. The density of public facilities and accessibility of public facilities reflect the richness of community services and the convenience of facility distribution, providing valuable insights into middle-aged and older adults' ease of access to essential daily services. These factors are closely linked to the overall livability and quality of life within the community. Additionally, the infrastructure conditions variable covered several fundamental aspects critical to the community environment, such as road quality, waste disposal systems, sewer availability, and waste management practices. By comprehensively accounting for these infrastructure elements, this system effectively represents the community environmental realities in China.

Our study identified infrastructure conditions and accessibility to public facilities as the 2 most critical CBE factors influencing the trajectory of depressive symptoms among middle-aged and older adults (Figure 2). First, these factors significantly predicted the initial level of depressive symptoms; middle-aged and older adults residing in communities with better infrastructure conditions and higher accessibility to public facilities exhibited significantly lower initial levels of depressive symptoms. This finding aligns with those reported in the existing literature. Fan et al [60] highlighted the crucial role of community factors such as infrastructure and senior centers in mitigating depressive symptoms among middle-aged and older adults. Similarly, Li et al [61] found that a lack of infrastructure in villages was positively associated with the prevalence of depression in older adults in rural China. Moreover, the availability of recreational facilities in urban communities was associated with a lower incidence of depressive symptoms [62]. Perceived proximity to community facilities is an important factor influencing depressive symptoms among low-income older adults. Improving accessibility to such facilities may help alleviate depressive symptoms [63]. From a mechanistic perspective, high-quality infrastructure provides a safer and more comfortable living environment and reduces the sources of psychological stress. Concurrently, convenient public facilities promote social participation and PA, offering environmental support for mental well-being.

Through a long-term analysis of depressive symptom trajectories, our study not only confirmed the effect of accessibility to public facilities on the initial level of depressive symptoms but also revealed its predictive role in the upward trajectory of depressive symptoms. Contrary to expectations, greater accessibility to public facilities was found to significantly accelerate the upward trajectory of depressive symptoms among middle-aged and older adults in China (Figure 2). This can be attributed to 2 reasons. First, higher accessibility may be associated with greater community activity and more complex social interactions, which can introduce additional environmental stressors, such as noise, social conflicts, or other pressures that negatively affect mental health [64]. Second, the inadequate service quality of public facilities and insufficient consideration of the needs of vulnerable populations could also play a key role. Middle-aged and older adults may have higher expectations of accessible public facilities; however, if these facilities fail to meet their needs—due to issues such as overcrowding, usage restrictions, or poor design—it may lead to disappointment or frustration [65]. In developing countries, public facilities are often designed with a focus on quantity and coverage, whereas service capacity and the needs of disadvantaged groups are frequently overlooked. This limitation may exacerbate mental health challenges in middle-aged and older adults by introducing potential risk factors.

Our study further demonstrates the significant effect of CBE on the initial level of total LTPA, emphasizing the importance of infrastructure conditions and accessibility to public transport. The results indicated that middle-aged and older adults living in communities with better infrastructure or higher public transport accessibility had significantly higher initial LTPA levels. These findings provide new insights into how the built environment influences health behaviors among middle-aged and older adults. Additionally, middle-aged and older adults in communities with better infrastructure and higher public facility accessibility engaged in significantly lower initial OPA. Furthermore, infrastructure conditions positively predicted a declining trajectory in OPA levels over time (Figure 2). As residential communities are not direct venues for occupational activity, a plausible explanation is that infrastructure conditions and public facility accessibility reflect the overall quality, location advantages, and amenities provided within the community. In China, residents of higher-tier communities are typically part of the nonlabor force population, while manual laborers tend to reside in communities with poorer infrastructure and fewer amenities [66]. This community-level differentiation significantly impacts the intensity and type of occupational activity, reflecting the observed community segregation and spatial stratification within Chinese society [67]. However, due to the lack of data on community-level economic variables, this study could not fully account for the potential influence of these factors on the results. Further research is needed to validate these associations and explore the underlying mechanisms.

Implications

Our findings could offer valuable insights for public health management practices and community environmental planning, with relevance to other developing countries or regions. When formulating community PA guidelines, it is crucial to consider the distinct effects of different types of PA. LTPA should be integrated into community mental health intervention programs, with a focus on encouraging middle-aged and older adults to engage in activities such as walking. However, relevant authorities should also guide middle-aged and older adults in reducing the intensity of OPA, optimizing workplace conditions and workflows, and providing mental health support. This recommendation is particularly critical for developing countries, where physical labor constitutes a significant portion of the workforce.

From the perspective of CBE, improving community infrastructure and enhancing the accessibility of public facilities are key strategies for preventing and addressing mental health issues in middle-aged and older adults, while also promoting PA. This includes, but is not limited to, upgrading road quality, improving water and sewage systems, and enhancing electricity services. These measures can lead to better depressive symptom outcomes and encourage LTPA in middle-aged and older populations. Furthermore, improving public transportation accessibility has the potential to promote LTPA in this demographic. However, we recommend that governments carefully consider the potential environmental stressors associated with enhanced public facilities and focus on improving their service capacity and quality. Special attention should be given to addressing the needs of vulnerable populations to mitigate negative effects and more effectively promote community mental health.

Limitations and Future Research

First, we excluded participants with missing PA or depressive symptom scores. While this approach is common, it may introduce bias, as it is challenging to determine whether the missing values are associated with participants who have lower PA levels or higher depressive symptom scores. Second, although the findings offer critical empirical support for the PA paradox, the study is constrained by its design and data availability, with limited consideration of factors such as workplace environment, working conditions, and job autonomy. These factors may mediate or confound the relationship between PA and mental health. Future research should incorporate these variables to comprehensively elucidate the complex mechanisms linking OPA and mental health. Third, the wave 2020 data used in this study were collected during the COVID-19 pandemic, which may have influenced the results. Fourth, due to space constraints and the study's objectives, insufficient attention was given to subgroup analyses or population heterogeneity, which future research could explore in greater detail. Finally, the CBE data in the CHARLS were published for only 1 wave without subsequent follow-up, limiting this study's ability to examine changes in the role of CBE in the relationship between PA and depressive symptom trajectories. Future research should incorporate dynamic geographical information to capture the effects of CBE changes on mental health or PA trajectories.

Conclusions

Communities play an integral role in the daily lives of residents. This study offers valuable insights from developing countries by highlighting the effects of PA on depressive symptom trajectories and the role of CBE factors as physical determinants. The findings reveal that the impact of PA on depressive symptoms varies by domain and intensity. Overall, higher levels of LTPA were associated with lower depressive symptom levels among middle-aged and older adults, whereas OPA, particularly at moderate-to-vigorous intensities, had negative effects on depressive symptoms. Additionally, OPA exhibits a dual effect: while a higher initial level of total OPA was associated with a higher initial level of depressive symptoms, it also contributed to slowing the upward trajectory of depressive symptoms over time. Among the CBE factors, infrastructure conditions and accessibility to public facilities emerged as the most critical determinants influencing depressive symptom trajectories. Better infrastructure and greater accessibility to public facilities were linked to lower initial levels of depressive symptoms; however, greater accessibility to public facilities might exacerbate the upward trajectory of depressive symptoms. Furthermore, infrastructure conditions and accessibility to public transport played positive roles in promoting LTPA.

These findings provide novel perspectives on the PA paradox in mental health research, offering valuable guidance for developing more precise public health strategies. Such strategies should aim to promote PA types that effectively reduce depressive symptoms while minimizing potential adverse effects. By identifying the specific impacts of CBE factors, this study underscores that optimizing the built environment is not merely a physical enhancement but also a vital public health approach to improving residents' mental health. We advocate for enhancing community environments to encourage beneficial PA and alleviate depressive symptoms among middle-aged and older adults.

Acknowledgments

This work was supported by JST SPRING (grant JPMJSP2136). We sincerely thank the handling editor and all the reviewers for their time and effort, as their feedback and comments greatly contributed to improving this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional analyses.
[DOCX File , 258 KB - publichealth_v11i1e64564_app1.docx]

References

1. Sokale IO, Conway SH, Douphrate DI. Built environment and its association with depression among older adults: a systematic review. TOPHJ 2022 Mar 18;15(1):e187494452202030 [FREE Full text] [doi: 10.2174/18749445-v15-e2202030]
2. Koohsari M, Yasunaga A, McCormack G, Shibata A, Ishii K, Nakaya T, et al. Depression among middle-aged adults in Japan: the role of the built environment design. Landscape and Urban Planning 2023 Mar;231:104651. [doi: 10.1016/j.landurbplan.2022.104651]
3. Dinas PC, Koutedakis Y, Flouris AD. Effects of exercise and physical activity on depression. Ir J Med Sci 2011 Jun;180(2):319-325. [doi: 10.1007/s11845-010-0633-9] [Medline: 21076975]
4. Dogra S, Dunstan DW, Sugiyama T, Stathi A, Gardiner PA, Owen N. Active aging and public health: evidence, implications, and opportunities. Annu Rev Public Health 2022 Apr 05;43:439-459 [FREE Full text] [doi: 10.1146/annurev-publhealth-052620-091107] [Medline: 34910580]
5. Brundtland GH. From the World Health Organization. Mental health: new understanding, new hope. JAMA 2001 Nov 21;286(19):2391. [doi: 10.1001/jama.286.19.2391] [Medline: 11712923]
6. Zhang L, Xu Y, Nie H, Zhang Y, Wu Y. The prevalence of depressive symptoms among the older in China: a meta-analysis. Int J Geriatr Psychiatry 2012 Sep;27(9):900-906. [doi: 10.1002/gps.2821] [Medline: 22252938]
7. van den Bosch M, Meyer-Lindenberg A. Environmental exposures and depression: biological mechanisms and epidemiological evidence. Annu Rev Public Health 2019 Apr 01;40:239-259. [doi: 10.1146/annurev-publhealth-040218-044106] [Medline: 30633709]
8. Soini E, Rosenström T, Määttänen I, Jokela M. Physical activity and specific symptoms of depression: a pooled analysis of six cohort studies. J Affect Disord 2024 Mar 01;348:44-53 [FREE Full text] [doi: 10.1016/j.jad.2023.12.039] [Medline: 38128736]
9. De Mello MT, Lemos VDA, Antunes HKM, Bittencourt L, Santos-Silva R, Tufik S. Relationship between physical activity and depression and anxiety symptoms: a population study. J Affect Disord 2013 Jul;149(1-3):241-246. [doi: 10.1016/j.jad.2013.01.035] [Medline: 23489405]
10. Wassink-Vossen S, Collard RM, Penninx BW, Hiles SA, Oude Voshaar RC, Naarding P. The reciprocal relationship between physical activity and depression: does age matter? Eur Psychiatry 2018 Jun;51:9-15. [doi: 10.1016/j.eurpsy.2017.12.029] [Medline: 29510298]
11. Abdoli N, Salari N, Darvishi N, Jafarpour S, Solaymani M, Mohammadi M, et al. The global prevalence of major depressive disorder (MDD) among the elderly: a systematic review and meta-analysis. Neurosci Biobehav Rev 2022 Jan;132:1067-1073. [doi: 10.1016/j.neubiorev.2021.10.041] [Medline: 34742925]
12. Kong JY, Hong H, Kang H. Relationship between physical activity and depressive symptoms in older Korean adults: moderation analysis of muscular strength. BMC Geriatr 2022 Nov 21;22(1):884 [FREE Full text] [doi: 10.1186/s12877-022-03610-6] [Medline: 36411413]
13. Kim JH. Regular physical exercise and its association with depression: a population-based study short title: exercise and depression. Psychiatry Res 2022 Mar;309:114406. [doi: 10.1016/j.psychres.2022.114406] [Medline: 35074644]
14. Clays E, Lidegaard M, De Bacquer D, Van Herck K, De Backer G, Kittel F, et al. The combined relationship of occupational and leisure-time physical activity with all-cause mortality among men, accounting for physical fitness. Am J Epidemiol 2014 Mar 01;179(5):559-566. [doi: 10.1093/aje/kwt294] [Medline: 24305575]
15. Temporelli PL. Is physical activity always good for you? The physical activity paradox. Eur Heart J Suppl 2021 Oct;23(Suppl E):E168-E171 [FREE Full text] [doi: 10.1093/eurheartj/suab115] [Medline: 34650379]
16. He F, Li Y, Hu Z, Zhang H. Association of domain-specific physical activity with depressive symptoms: a population-based study. Eur Psychiatry 2022 Dec 12;66(1):e5 [FREE Full text] [doi: 10.1192/j.eurpsy.2022.2350] [Medline: 36503700]
17. Joo MJ, Jang YS, Jang YS, Park EC. Association between work-related physical activity and depressive symptoms in Korean workers: data from the Korea national health and nutrition examination survey 2014, 2016, 2018, and 2020. BMC Public Health 2023 Sep 08;23(1):1752 [FREE Full text] [doi: 10.1186/s12889-023-16631-6] [Medline: 37684616]
18. Oliveira AJ, Pancoti BM, Oliveira BRR, de Figueiredo JA, Paravidino V. Associations of occupational physical activity and active transport with depressive symptoms in a middle-income country. Public Health 2024 Jul;232:195-200. [doi: 10.1016/j.puhe.2024.04.033] [Medline: 38805868]
19. Schuch FB, Werneck AO, Vancampfort D, Stubbs B, Teychene M, Lotufo PA, et al. Cross-sectional associations of leisure and transport related physical activity with depression and anxiety. J Psychiatr Res 2021 Aug;140:228-234. [doi: 10.1016/j.jpsychires.2021.05.053] [Medline: 34118641]
20. Lopes MVV, Matias TS, da Costa BGG, Schuch FB, Chaput JP, Samara Silva K. The relationship between physical activity and depressive symptoms is domain-specific, age-dependent, and non-linear: an analysis of the Brazilian national health survey. J Psychiatr Res 2023 Mar;159:205-212. [doi: 10.1016/j.jpsychires.2023.01.041] [Medline: 36739848]
21. Lu S, Liu Y, Guo Y, Ho HC, Song Y, Cheng W, et al. Neighbourhood physical environment, intrinsic capacity, and 4-year late-life functional ability trajectories of low-income Chinese older population: A longitudinal study with the parallel process

of latent growth curve modelling. EClinicalMedicine 2021 Jun;36:100927 [FREE Full text] [doi: 10.1016/j.eclinm.2021.100927] [Medline: 34189445]

22. Nie P, Clark AE, D'Ambrosio C, Ding L. Income-related health inequality in urban China (1991-2015): the role of homeownership and housing conditions. Health Place 2022 Jan;73:102743. [doi: 10.1016/j.healthplace.2022.102743] [Medline: 35045352]
23. Xiao J, Zhao J, Luo Z, Liu F, Greenwood D. The impact of built environment on mental health: a COVID-19 lockdown perspective. Health Place 2022 Sep;77:102889 [FREE Full text] [doi: 10.1016/j.healthplace.2022.102889] [Medline: 36027740]
24. Lederbogen F, Kirsch P, Haddad L, Streit F, Tost H, Schuch P, et al. City living and urban upbringing affect neural social stress processing in humans. Nature 2011 Jun 22;474(7352):498-501. [doi: 10.1038/nature10190] [Medline: 21697947]
25. Ewing R, Cervero R. Travel and the built environment: a synthesis. Transportation Research Record: Journal of the Transportation Research Board 2001 Jan 01;1780(1):87-114. [doi: 10.3141/1780-10]
26. Wang P, Wang M, Shan J, Liu X, Jing Y, Zhu H, et al. Association between residential greenness and depression symptoms in Chinese community-dwelling older adults. Environ Res 2024 Feb 15;243:117869 [FREE Full text] [doi: 10.1016/j.envres.2023.117869] [Medline: 38070849]
27. Koohsari M, McCormack G, Nakaya T, Shibata A, Ishii K, Yasunaga A, et al. Urban design and Japanese older adults' depressive symptoms. Cities 2019 Apr;87:166-173. [doi: 10.1016/j.cities.2018.09.020]
28. Saarloos D, Alfonso H, Giles-Corti B, Middleton N, Almeida OP. The built environment and depression in later life: the health in men study. Am J Geriatr Psychiatry 2011 May;19(5):461-470. [doi: 10.1097/JGP.0b013e3181e9b9bf] [Medline: 20808136]
29. Gubbels JS, Kremers SPJ, Droomers M, Hoefnagels C, Stronks K, Hosman C, et al. The impact of greenery on physical activity and mental health of adolescent and adult residents of deprived neighborhoods: a longitudinal study. Health Place 2016 Jul;40:153-160. [doi: 10.1016/j.healthplace.2016.06.002] [Medline: 27322564]
30. Li X, Zhang W, Zhang W, Tao K, Ni W, Wang K, et al. Level of physical activity among middle-aged and older Chinese people: evidence from the China Health and Retirement Longitudinal Study. BMC Public Health 2020 Nov 10;20(1):1682 [FREE Full text] [doi: 10.1186/s12889-020-09671-9] [Medline: 33172439]
31. Fan M, Lyu J, He P. Chinese guidelines for data processing and analysis concerning the International Physical Activity Questionnaire. Zhonghua Liu Xing Bing Xue Za Zhi 2014 Aug;35(8):961-964. [Medline: 25376692]
32. Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ)-short and long forms. IPAQ. 2005. URL: http://www.ipaq.ki.se/scoring [accessed 2024-12-23]
33. Gao Y, Jia Z, Zhao L, Han S. The effect of activity participation in middle-aged and older people on the trajectory of depression in later life: national cohort study. JMIR Public Health Surveill 2023 Mar 23;9:e44682 [FREE Full text] [doi: 10.2196/44682] [Medline: 36951932]
34. China Health and Retirement Longitudinal Study (CHARLS) website. China Health and Retirement Longitudinal Study (CHARLS). URL: https://charls.pku.edu.cn/ [accessed 2024-12-12]
35. Park SH, Yu HY. How useful is the center for epidemiologic studies depression scale in screening for depression in adults? An updated systematic review and meta-analysis. Psychiatry Res 2021 Aug;302:114037. [doi: 10.1016/j.psychres.2021.114037] [Medline: 34098160]
36. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). Am J Prev Med 1994;10(2):77-84. [Medline: 8037935]
37. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991 Feb;59(1):12-19. [doi: 10.1037//0022-006x.59.1.12] [Medline: 2002127]
38. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
39. Cervero R, Kockelman K. Travel demand and the 3Ds: density, diversity, and design. Transportation Research Part D: Transport and Environment 1997 Sep;2(3):199-219. [doi: 10.1016/S1361-9209(97)00009-6]
40. Cardinali M, Beenackers MA, van Timmeren A, Pottgiesser U. The relation between proximity to and characteristics of green spaces to physical activity and health: a multi-dimensional sensitivity analysis in four European cities. Environ Res 2024 Jan 15;241:117605 [FREE Full text] [doi: 10.1016/j.envres.2023.117605] [Medline: 37956752]
41. Cui X, Wen J, Yang H, Helbich M, Dijst M, Roberts H, et al. Pathways from street network design to symptoms of depression among emerging adults in China. Travel Behaviour and Society 2023 Oct;33:100644. [doi: 10.1016/j.tbs.2023.100644]
42. Zhang W, Wang T, Wang A. Impact of physical activity intensity on longitudinal trajectories of cognitive function and depressive symptoms in middle-aged and older Chinese adults: eight-year prospective study. J Affect Disord 2022 Oct 15;315:64-69. [doi: 10.1016/j.jad.2022.07.012] [Medline: 35839942]
43. Verbeke G, Fieuws S, Molenberghs G, Davidian M. The analysis of multivariate longitudinal data: a review. Stat Methods Med Res 2014 Feb;23(1):42-59. [doi: 10.1177/0962280212445834] [Medline: 22523185]

44. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
45. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003 Oct;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
46. Meng Y, Luo Y, Qin S, Xu C, Yue J, Nie M, et al. The effects of leisure time physical activity on depression among older women depend on intensity and frequency. J Affect Disord 2021 Dec 01;295:822-830. [doi: 10.1016/j.jad.2021.08.142] [Medline: 34706452]
47. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462 [FREE Full text] [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
48. Mammen G, Faulkner G. Physical activity and the prevention of depression: a systematic review of prospective studies. Am J Prev Med 2013 Nov;45(5):649-657. [doi: 10.1016/j.amepre.2013.08.001] [Medline: 24139780]
49. Loprinzi PD. Objectively measured light and moderate-to-vigorous physical activity is associated with lower depression levels among older US adults. Aging Ment Health 2013;17(7):801-805. [doi: 10.1080/13607863.2013.801066] [Medline: 23731057]
50. Ryu J, Jeong A, Min JH, Lee DH, Lee J, Song IH, et al. The relationship between domain-specific physical activity and depressive symptoms in Korean adults: analysis of the Korea National Health and Nutrition Examination Survey. J Affect Disord 2022 Apr 01;302:428-434. [doi: 10.1016/j.jad.2022.01.097] [Medline: 35090946]
51. Werneck AO, Stubbs B, Szwarcwald CL, Silva DR. Independent relationships between different domains of physical activity and depressive symptoms among 60,202 Brazilian adults. Gen Hosp Psychiatry 2020;64:26-32. [doi: 10.1016/j.genhosppsych.2020.01.007] [Medline: 32086172]
52. Lin L, Halgin RP, Well AD, Ockene I. The relationship between depression and occupational, household, and leisure-time physical activity. Journal of Clinical Sport Psychology 2008;2(2):95-107. [doi: 10.1123/jcsp.2.2.95]
53. Nestor A, Henry Kofi M. The association between work-related physical activity and depression. Journal of Physical Activity Research 2017;2(1):1-6 [FREE Full text] [doi: 10.12691/jpar-2-1-1]
54. Coenen P, Huysmans MA, Holtermann A, Krause N, van Mechelen W, Straker LM, et al. Towards a better understanding of the 'physical activity paradox': the need for a research agenda. Br J Sports Med 2020 Sep;54(17):1055-1057. [doi: 10.1136/bjsports-2019-101343] [Medline: 32265218]
55. Luo M, Gupta N, Holtermann A, Stamatakis E, Ding D. Revisiting the 'physical activity paradox' in a Chinese context: occupational physical activity and mortality in 142,302 urban working adults from the China Kadoorie Biobank study. Lancet Reg Health West Pac 2022 Jun;23:100457. [doi: 10.1016/j.lanwpc.2022.100457] [Medline: 35602414]
56. Heesch KC, Burton NW, Brown WJ. Concurrent and prospective associations between physical activity, walking and mental health in older women. J Epidemiol Community Health 2011 Sep;65(9):807-813. [doi: 10.1136/jech.2009.103077] [Medline: 20515892]
57. Kok AAL, Twisk JWR, Blom F, Beekman ATF, Huisman M. Steeling or sensitizing? A longitudinal examination of how ongoing accumulation of negative life events affects depressive symptoms in older adults. J Gerontol B Psychol Sci Soc Sci 2021 Nov 15;76(10):2041-2053 [FREE Full text] [doi: 10.1093/geronb/gbab114] [Medline: 34171092]
58. Rutter M. Resilience in the face of adversity. Protective factors and resistance to psychiatric disorder. Br J Psychiatry 1985 Dec;147:598-611. [doi: 10.1192/bjp.147.6.598] [Medline: 3830321]
59. Jones-Smith JC, Popkin BM. Understanding community context and adult health changes in China: development of an urbanicity scale. Soc Sci Med 2010 Oct;71(8):1436-1446 [FREE Full text] [doi: 10.1016/j.socscimed.2010.07.027] [Medline: 20810197]
60. Fan VS, Mahadevan R, Leung J. Effect of income inequality, community infrastructure and individual stressors on adult depression. Health Promot Int 2021 Mar 12;36(1):46-57. [doi: 10.1093/heapro/daaa036] [Medline: 32277828]
61. Li LW, Liu J, Zhang Z, Xu H. Late-life depression in rural China: do village infrastructure and availability of community resources matter? Int J Geriatr Psychiatry 2015 Jul;30(7):729-736. [doi: 10.1002/gps.4217] [Medline: 25333218]
62. Deng Y, Paul DR. The relationships between depressive symptoms, functional health status, physical activity, and the availability of recreational facilities: a rural-urban comparison in middle-aged and older Chinese adults. Int J Behav Med 2018 Jun;25(3):322-330. [doi: 10.1007/s12529-018-9714-3] [Medline: 29498014]
63. Chen YY, Wong GHY, Lum TY, Lou VWQ, Ho AHY, Luo H, et al. Neighborhood support network, perceived proximity to community facilities and depressive symptoms among low socioeconomic status Chinese elders. Aging Ment Health 2016;20(4):423-431. [doi: 10.1080/13607863.2015.1018867] [Medline: 25775108]
64. Pineda VS. What is inclusive and accessible public space? jps 2022 Jun 26;7(2):5-8. [doi: 10.32891/jps.v7i2.1607]
65. Huang W, Lin G. The relationship between urban green space and social health of individuals: a scoping review. Urban Forestry & Urban Greening 2023 Jul;85:127969. [doi: 10.1016/j.ufug.2023.127969]
66. Simon A, John K, Lina S, Qingjie X. Contrasting paradigms: segmentation and competitiveness in the formation of the Chinese labour market. In: Unemployment, Inequality and Poverty in Urban China. London, UK: Routledge; Mar 10, 2006:228-251.

67. Zhou S, Deng L, Kwan M, Yan R. Social and spatial differentiation of high and low income groups' out-of-home activities in Guangzhou, China. Cities 2015 Jun;45:81-90. [doi: 10.1016/j.cities.2015.03.002]

Abbreviations

CBE: community-built environment
CES D: Center for Epidemiologic Studies Depression Scale
CHARLS: China Health and Retirement Longitudinal Survey
LGCM: latent growth curve modeling
LRT: likelihood ratio test
LTPA: leisure-time physical activity
MET: metabolic equivalent
OPA: occupational physical activity
PA: physical activity
RCI: Reliable Change Index
WHO: World Health Organization

Edited by A Mavragani; submitted 20.07.24; peer-reviewed by R Liang, Y Liu, S Fujihara, E Luo; comments to author 08.11.24; revised version received 08.12.24; accepted 11.12.24; published 13.01.25.

Please cite as:
Zhang K, Huang B, Divigalpitiya P
Identifying Community-Built Environment's Effect on Physical Activity and Depressive Symptoms Trajectories Among Middle-aged and Older Adults: Chinese National Longitudinal Study
JMIR Public Health Surveill 2025;11:e64564
URL: https://publichealth.jmir.org/2025/1/e64564
doi:10.2196/64564
PMID:39804686

Original Paper

Data Parameters From Participatory Surveillance Systems in Human, Animal, and Environmental Health From Around the Globe: Descriptive Analysis

Carrie McNeil[1], MPH, DVM; Nomita Divi[1], BSc, MSPH; Charles Thomas Bargeron IV[2], MS; Andrea Capobianco Dondona[3], MSc, DVM; Kacey C Ernst[4], MPH, PhD; Angela S Gupta[5], MA; Olukayode Fasominu[6], MPH, MD; Lucy Keatts[7], BVetMed, MVSc, MRCVS; Terra Kelly[8], DVM, Dipl. ACZM, PhD; Onicio B Leal Neto[9], PhD; May O Lwin[10], MBA, PhD; Mvuyo Makhasi[11], BSc, MSc; Eric Beda Mutagahywa[12], MSc; Diego Montecino-Latorre[7], BVSc, MPVM, PhD; Sarah Olson[7], PhD; Pranav S Pandit[13], BVSc & AH, MPVM, PhD; Daniela Paolotti[14], PhD; Matt C Parker[15], MPH; Muhammad Haiman Samad[16], MPH; Kara Sewalk[17], MPH; Anita Sheldenkar[16], MS; Lertrak Srikitjakarn[18], DrMedVet; Channé Suy Lan[19], MCA; Michael Wilkes[20], MD, PhD; Terdsak Yano[21], MS, DVM; Mark Smolinski[1], MPH, MD

[1]Ending Pandemics, San Francisco, CA, United States
[2]Center for Invasive Species & Ecosystem Health, University of Georgia, Tifton, GA, United States
[3]Food and Agriculture Organization of United Nations, Rome, Italy
[4]Department of Epidemiology and Biostatistics, College of Public Health, University of Arizona, Tucson, AZ, United States
[5]College of Extension, University of Minnesota, Rochester, MN, United States
[6]Volte Health Systems Limited, Abuja, Nigeria
[7]Health Program, Wildlife Conservation Society, Bronx, NY, United States
[8]EpiEcos LLC, Flagstaff, AZ, United States
[9]College of Public Health, Department of Epidemiology and Biostatistics, University of Arizona, Tuscon, AZ, United States
[10]Wee Kim Wee School of Communication & Information, Nanyang Technological University, Singapore, Singapore
[11]Centre for Respiratory Diseases and Meningitis, National Health Laboratory Service, National Institute for Communicable Diseases, Johannesburg, South Africa
[12]SACIDS Foundation for One Health, Morogoro, United Republic of Tanzania
[13]Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, Davis, CA, United States
[14]Institute for Scientific Interchange, Turin, Italy
[15]Open Dream, Bangkok, Thailand
[16]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore
[17]Boston Children's Hospital, Boston, MA, United States
[18]PODD Centre, Chiang Mai University, Chiang Mai, Thailand
[19]Kawsang Co, LTD, Phnom Penh, Cambodia
[20]School of Medicine, University of California, Davis, Davis, CA, United States
[21]Faculty of Veterinary Medicine, Chiang Mai University, Chiang Mai University, Thailand

Corresponding Author:
Carrie McNeil, MPH, DVM
Ending Pandemics
870 Market Street, Suite 528
San Francisco, CA, 94102
United States
Phone: 1 9169522185
Email: carrie.s.mcneil@gmail.com

Abstract

Background: Emerging pathogens and zoonotic spillover highlight the need for One Health surveillance to detect outbreaks as early as possible. Participatory surveillance empowers communities to collect data at the source on the health of animals, people, and the environment. Technological advances increase the use and scope of these systems. This initiative sought to collate information from active participatory surveillance systems to better understand parameters collected across the One Health spectrum.

Objective: This study aims to develop a compendium of One Health data parameters by examining participatory surveillance systems active in 2023. The expected outcomes of the compendium were to pinpoint specific parameters related to human, animal, and environmental health collected globally by participatory surveillance systems and to detail how each parameter is collected. The compendium was designed to help understand which parameters are currently collected and serve as a reference for future systems and for data standardization initiatives.

Methods: Contacts associated with the 60 systems identified through the One Health Participatory Surveillance System Map were invited by email to provide specific data parameters, methodologies used for data collection, and parameter-specific considerations. Information was received from 38 (63%) active systems. Data were compiled into a searchable spreadsheet-based compendium organized into 5 sections: general, livestock, wildlife, environmental, and human parameters. An advisory group comprising experts in One Health participatory surveillance reviewed the collected parameters, refined the compendium structure, and contributed to the descriptive analysis.

Results: A comprehensive compendium of data parameters from a diverse array of single-sector and multisector participatory surveillance systems was collated and reviewed. The compendium includes parameters from 38 systems used in Africa (n=3, 8%), Asia (n=9, 24%), Europe (n=12, 32%), Australia (n=3, 8%), and the Americas (n=12, 32%). Almost one-third of the systems (n=11, 29%) collect data across multiple sectors. Many (n=17, 45%) focus solely on human health. Variations in data collection techniques were observed for commonly used parameters, such as demographics and clinical signs or symptoms. Most human health systems collected parameters from a cohort of users tracking their own health over time, whereas many wildlife and environmental systems incorporated event-based parameters.

Conclusions: Several participatory surveillance systems have already adopted a One Health approach, enhancing traditional surveillance by identifying shared health threats among animals, people, and the environment. The compendium reveals substantial variation in how parameters are collected, underscoring the need for further work in system interoperability and data standards to allow for timely data sharing across systems during outbreaks. Parameters collated from across the One Health spectrum represent a valuable resource for informing the development of future systems and identifying opportunities to expand existing systems for multisector surveillance.

(JMIR Public Health Surveill 2025;11:e55356) doi:10.2196/55356

KEYWORDS

participatory surveillance; One Health; citizen science; community-based surveillance; digital disease detection; environmental health; wildlife health; livestock health; human health; data standards

Introduction

Background

In a world where anthropogenic landscape change drives zoonotic spillover, climate change exacerbates vulnerabilities in food security, and diseases further threaten biodiversity, effective and timely One Health surveillance mechanisms are imperative. The One Health United Nations Joint Action Plan (2022-2026) [1], Africa's Centers for Disease Control Event-Based Surveillance (EBS) guidance [2], and World Health Organization's "Defining Collaborative Surveillance" [3] underscore the critical need to integrate data across the animal, human, plant, and environmental health sectors. This paper examines the range of data parameters collected across the One Health spectrum through participatory surveillance methodologies.

One Health Surveillance

The One Health approach recognizes the interdependencies of human, animal, and ecosystem health, aiming "to sustainably balance and optimize the health of people, animals, and ecosystems" [4]. Numerous national and subnational systems have undertaken initiatives for data sharing to develop One Health strategies, with support from organizations such as the United Nations, World Bank, and United States Agency for International Development (USAID) [5,6]. However, challenges remain in the implementation of One Health initiatives due to years of siloed professional training, budgets, data system development, data structures, and policies [7]. Rabies and influenza viruses have served as key drivers of multisectoral, collaborative surveillance and offer valuable models for advancing One Health approaches to address multiple pathogens. The need for the inclusion of plant health has more recently come to the forefront due to its critical role in food safety, food security, and ecosystem health [8].

Historically, traditional disease surveillance systems have largely been led by government entities and are often siloed, with a primary focus on human and livestock health. There is growing awareness of the need to harness more wildlife and environmental health event information to provide early warning of broader health threats facing both humans and animals. The Africa's Centers for Disease Control EBS guidance emphasizes the integration of both formal and informal data sources, including hotlines and community-based EBS, to enhance outbreak detection [2]. To strengthen participatory surveillance systems, greater efforts are needed to capture data on putative drivers of pathogen spillover and disease emergence, including landscape change, environmental and climatic variables, and activities at the human-animal-environmental interface that bring animals and humans into contact with each other [9]. While previous literature has examined One Health tools and individual surveillance systems, no study to date has comprehensively documented the range of parameters collected in One Health participatory surveillance systems [7,10-12].

Participatory Surveillance

Participatory surveillance involves a bidirectional process of receiving and transmitting data for actionable outcomes through the direct engagement of the target population [13]. This approach is increasingly used across multiple sectors to enhance early detection of and rapid response to emerging infections within a One Health framework. By engaging a wide variety of users including the general public, health care workers, rangers, farmers, and outdoor enthusiasts, participatory surveillance systems can provide early warning of potential outbreaks through crowdsourcing self-reported data in near real-time [14]. This active surveillance strategy empowers communities to act as the "eyes and ears" for detecting health threats at the interfaces where animals, humans, and environmental factors converge, creating opportunities for disease emergence.

Participatory surveillance systems play a crucial role in providing evidence-based feedback to users, enabling them to take timely actions to mitigate potential threats and facilitating connections with local resources for prevention and control. In the context of One Health, participatory surveillance extends beyond a human-centered lens to encompass plants, wildlife, livestock, and a multitude of environmental factors—habitats for disease vectors and water and air quality. The organizational structure of system users can vary widely, ranging from the general public to members of sector-specific associations or community alliances. For more structured community-liaison models, feedback is directed to the user group, empowering them to take action on behalf of the communities they serve.

Participatory surveillance has emerged as a powerful tool for enhancing traditional epidemiologic surveillance while fostering community engagement [15]. From reporting disease incidence to promoting mitigation behaviors, participatory surveillance has showcased its versatility across various infectious disease contexts, including influenza, cholera, COVID-19, and Zika [16-22]. One of the key benefits of participatory surveillance is its capacity to address and surmount many of the limitations inherent to traditional surveillance systems. In resource-constrained settings, conventional surveillance efforts often rely on overburdened health care and diagnostic systems and struggle to access difficult-to-reach communities, resulting in reporting delays or gaps. By empowering communities, participatory surveillance enables a proactive role in health monitoring, facilitating both prompt and comprehensive reporting of disease cases [23,24].

Ending Pandemics, a US nonprofit organization working to detect, verify, and contain outbreaks faster (and the supporter of the study presented here) developed the International Workshops on Participatory Surveillance (IWOPS), four of which have been held since 2012 in the United States, the Netherlands, Australia, and Cambodia [25]. IWOPS convened the creators, implementors, and key advocates to advance participatory surveillance approaches and strengthen partnerships among these systems. In 2016, the IWOPS community established a minimum dataset among systems collecting self-reported information to monitor the annual risk of influenza so that a common set of data standards could be developed. As a result, deidentified and disaggregated data are shared on a public platform, Global Flu View, for the purpose of improving situational awareness [26].

While Global Flu View represents one example of what is possible when participatory surveillance systems can "speak" to each other, data standards across the entire spectrum of parameters encompassing One Health are lacking. The One Health Participatory Surveillance Data Parameters Compendium (Multimedia Appendix 1) is a snapshot into the world of participatory surveillance for human, animal, and environmental health. The compendium approach was used as it not only illustrates the breadth and depth of what is being used in currently operating participatory surveillance systems across the globe, but it can serve to inform the needs of those contemplating the development of a participatory surveillance system or expansion of an existing system. As the compendium provides information on how data parameters are collected, it is poised to guide efforts to incorporate additional health sectors, adapt data parameters to align with other systems, and foster collaborations and data sharing. In the future, it could play a pivotal role in advancing data sharing, through informing the development of international data standards.

Methods

Development of Advisory Group

A multisector advisory group was assembled to guide the development of the compendium, support data collection, and provide cross-sectoral reviews encompassing animal, environmental, and public health sector perspectives. Selection criteria for advisory group members included expertise in participatory surveillance systems and integrating data from participatory systems into traditional surveillance systems. Subteams of experts in the areas of human, wildlife, livestock, and environmental health (including plants and vectors) were established. The experts represented a diverse range of public and government organizations, universities or research centers, nongovernmental organizations, and private sector entities. Geographically, they were affiliated with organizations in Asia (n=7), Africa (n=3), Europe (n=2), North America (n=10), and South America (n=2).

Ethical Considerations

The study was reviewed by the University of Arizona Institutional Review Board and deemed not to involve human participants as defined by US Department of Health and Human Services and Food and Drug Administration regulations research (IRB protocol# 00002653). Within the body of the email, individuals were provided a disclosure statement. All participants were notified that they would receive a copy of the compendium, and their system would be listed in the acknowledgments. Data included in the compendium included terms and categories of parameters collected by systems and did not include any identifiable information. Participants were not compensated, and participation was voluntary.

Data Collection

An invitation letter was emailed to a closed sample of representatives of 60 systems, using contact information from the One Health Participatory Surveillance map and 2022 landscape analysis [13]. Follow-up emails were sent twice, as needed, by members of the compendium advisory group to encourage responses from system representatives who did not reply initially. Emails described the goal for developing the One Health participatory surveillance compendium and requested respondents provide data parameters from their systems, methods for collection, and any special considerations for specific parameters using an attached template. The template had been used for parameter collection for IWOPS IV [27]. System representatives were offered the option to provide the requested information in other formats, such as a link to web-based systems allowing authors to independently collect relevant questions and parameters.

Data collected were in English. Parameters had already been collected for IWOPS IV in 2022 from: AfyaData, Cambodia 115 Hotline, EDDMapS or Wild Spotter, Forest First Pest Detector, Kidenga, MoBuzz, Outbreaks Near Me, One Health App Philippines, Participatory One Health Digital Disease Detection (PODD), SickSense/Open Dream, SURPRISE, World Animal Health Information System, SMART for Health, and WildAlert (formally known as the Wildlife Morbidity and Mortality Event Alert System) [27]. For these systems, emails were sent to confirm the parameters collected and to allow for revisions.

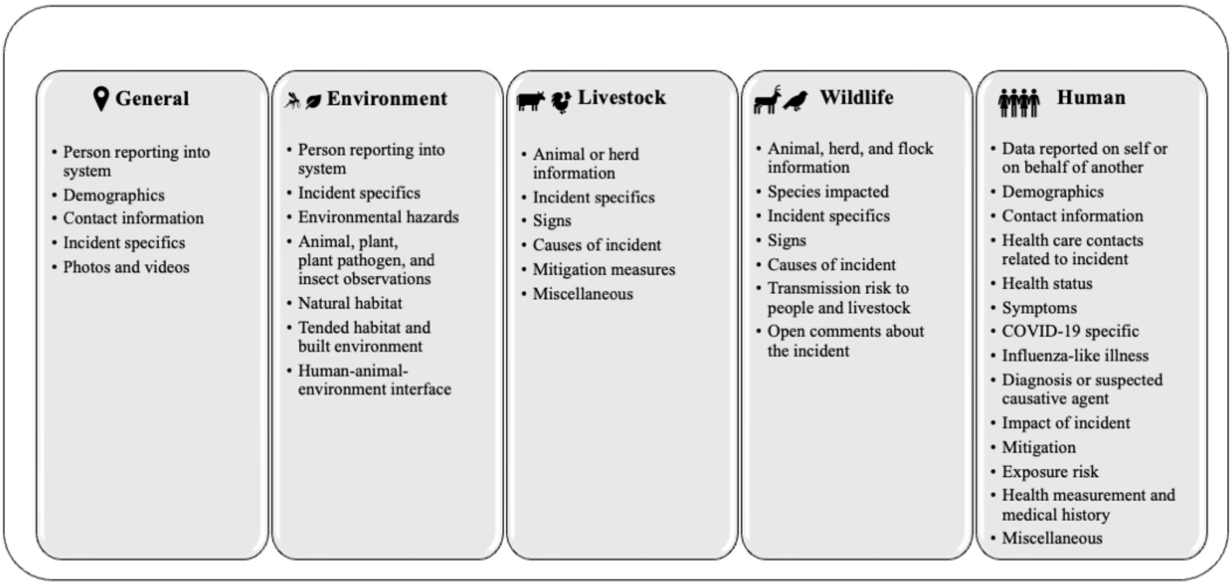
Out of the 60 systems contacted, 41 systems (68%) responded, and 3 systems (5%) responded they were no longer active and were not included in this analysis. Completed responses were received from 36/60 (60%) active systems; data were collected directly from publicly accessible apps or websites from 2/60 (3%) systems. 38 (63%) responses include the 9 country-level programs of InfluenzaNet. Seven of the 9 InfluenzaNet programs submitted 1 set of parameters and 2 (GrippeWeb and Hälsorapport) provided parameters individually.

Data Analysis

The collected parameters were organized into 5 sections in the compendium: general, livestock, wildlife, environmental (including plants and vectors), and human health. Subcategories within each section were reviewed and refined through group discussions with the advisory group subteams specializing in livestock, wildlife, human, and environmental health. The focus of these discussions was to ensure usability, clarity, and accuracy. The human health advisory group subteam determined that parameters would be limited to those they determined directly relevant to infectious disease surveillance. The compendium focuses on parameters and does not describe structure, such as decision trees.

Advisory group representatives reviewed the parameters and overall organization of the compendium and shared perspectives on key aspects to emphasize in the discussion and conclusion of this paper. Descriptive analyses were performed to examine the categories of parameters collected, the number and types of systems represented across sectors, and advisory group feedback on the parameters. Figure 1 outlines the compendium's format, and the categories of data parameters collected.

Figure 1. Categories of data parameters included in the 2023 One Health Participatory Surveillance Compendium.



General
- Person reporting into system
- Demographics
- Contact information
- Incident specifics
- Photos and videos

Environment
- Person reporting into system
- Incident specifics
- Environmental hazards
- Animal, plant, plant pathogen, and insect observations
- Natural habitat
- Tended habitat and built environment
- Human-animal-environment interface

Livestock
- Animal or herd information
- Incident specifics
- Signs
- Causes of incident
- Mitigation measures
- Miscellaneous

Wildlife
- Animal, herd, and flock information
- Species impacted
- Incident specifics
- Signs
- Causes of incident
- Transmission risk to people and livestock
- Open comments about the incident

Human
- Data reported on self or on behalf of another
- Demographics
- Contact information
- Health care contacts related to incident
- Health status
- Symptoms
- COVID-19 specific
- Influenza-like illness
- Diagnosis or suspected causative agent
- Impact of incident
- Mitigation
- Exposure risk
- Health measurement and medical history
- Miscellaneous

Results

Participatory Surveillance Systems Represented in the Compendium

Data parameters were collected in English from across 38 participatory surveillance systems in Africa (n=3, 8%), Asia (n=9, 24%), Europe (n=12, 32%), Australia (n=3, 8%), and the Americas (n=12, 32%); only SMART for Health data were collected in several continents (Table 1). Approximately one-third of the systems (n=11, 29%) collect data from more than 1 sector, while the majority focus on a single sector. Many programs (n=17, 45%) are dedicated solely to human health, with fewer focusing exclusively on environmental or animal health (n=10, 26%). Most users contributing data to these systems are members of the general public (n=30, 79%), though other user groups include trained volunteers (n=6, 16%), wildlife rangers (n=2, 5%), farmers (n=2, 5%), wildlife rehabilitators and biologists (n=2, 5%), informal health care workers (n=1, 3%), and students (n=1, 3%). For example, SURPRISE is a system in which health care workers report on their own health, (n=1, 3%). Systems requesting weekly data reporting (n=16, 42%) primarily focus on human health, though some offer users the flexibility to determine the reporting frequency. Most of the remaining systems (n=20, 53%) allow users to set their reporting frequency.

Table 1. Description of global One Health participatory surveillance systems identified as active in 2023 including sectors included, location, user type, reporting frequency, and technology.

Systems	Sector	Countries	Type of user	Frequency of reporting	User technology
AfyaData	Animal, human, and environment	Tanzania, Mozambique	General public, trained volunteer, farmer, ranger, researchers, and students	Determined by user	SMS and mobile app
Arizona Department of Fish and Game Hotline	Animal and environment	United States	General public	Determined by user	Hotline and email or website
Cambodia 115 Hotline	Animal and human	Cambodia	General public and health officer	Weekly, determined by user	Hotline
California West Nile Bird Reporting	Animal	United States	General public	Determined by user	Hotline and email or website
Colab	Animal, human, and environment	Brazil	General public	Determined by user	Mobile app, hotline, and email or website
CoughWatch SA	Human	South Africa	General public	Weekly	Email or website
DisApp	Environment	India	General public	Determined by user	Mobile app
EDDMapS/Wild Spotter	Animal and environment	United States and Canada	General public and trained volunteer	Determined by user	Mobile app and email or website
eWHIS[a]	Animal and environment	Australia	Trained volunteer	Monthly, determined by user	Hotline and email or website
FluTracking Australia	Human	Australia	General public	Weekly	Email or website
FluTracking HongKong	Human	Hong Kong	General public	Weekly	Email or website
Forest Pest First Detector	Environment	United States	General public	Determined by user	Mobile app and email or website
Guardians of Health	Human	Brazil	General public	Daily	Mobile app
Garden Wildlife Health Project	Animal	Great Britain	General public	Determined by user	Email or website
iMammalia	Animal	Europe	General public	Daily	Mobile app and email or website
InfluenzaNet programs					
FluSurvey.net	Human	United Kingdom	General public	Weekly	Email or website
Grippenet.ch	Human	Switzerland	General public	Weekly	Email or website
Grippenet.fr	Human	France	General public	Weekly	Email or website
GrippeWeb	Human	Germany	General public	Weekly	Email or website
Hälsorapport, Health Report Sweden	Human	Sweden	General public	Weekly	Email or website
Infectieradar.be	Human	Belgium	General public	Weekly	Email or website
Infectieradar.nl	Human	Netherlands	General public	Weekly	Email or website
Influmeter.dk	Human	Denmark	General public	Weekly	Email or website
InfluWeb	Human	Italy	General public	Weekly	Email or website
I-TICK	Environment	United States	General public	Determined by user	Mobile app
Kidenga	Environment and human	United States and Puerto Rico	General public	Weekly	Mobile app
MoBuzz+	Environment and human	Sri Lanka	General public, health authorities	Determined by user	Mobile app and email or website
MosApp	Environment	India	Formal, informal health care workers	Determined by user	Mobile app
Outbreaks Near Me	Human	United States, Canada, Mexico	General public	Weekly, Determined by User	SMS and email or website

Systems	Sector	Countries	Type of user	Frequency of reporting	User technology
PODD[b]	Animal, human, and environment	Thailand	General public, trained volunteer, farmer, and local government	Fortnightly	Mobile app
SMART for Health	Animal and environment	Laos, Cambodia, Vietnam, and Guatemala	Rangers	Determined by user	Mobile app
SickSense or Sabaidee	Human	Thailand	General public	Daily, Weekly, Determined by user	Mobile app
SURPRISE+	Human	Switzerland	Health care workers	Adaptive according to epidemiology	SMS and email or website
Tamil Nadu Population Health Registry	Human	India	Trained volunteer	Daily	Mobile app and paper
Tick Tag Go	Environment	United States	General public	Determined by user	Mobile app and email or website
Weedspotters Australia	Environment	Australia	Trained volunteer	Determined by user	Mobile app
Wild Health Net Ebola	Animal and environment	Republic of Congo	Wildlife biologists	Determined by user	SMS, mobile app, and email or website
WildAlert	Animal	United States	Wildlife rehabilitators, state wildlife agencies, and university researchers	Weekly, determined by user	Email or website

[a] eWHIS: electronic wildlife health information system.

[b] PODD: participatory one health digital disease detection.

General Parameters

General parameters include those not specific to any sector—such as user demographics and contact information, incident location, and options for photo or video uploads. Systems with regular users often collect user demographics and contact information during the initial login, eliminating the need for repeated entry on subsequent logins. Substantial variation exists on how demographic data are collected—particularly as related to gender, ethnicity, education, and profession. In contrast, contact information, such as email and phone numbers, are collected uniformly across most systems. Incident location is commonly collected using automated geolocation. Additionally, some systems include mechanisms for photo or video upload to provide supplementary information.

Environmental Health Parameters

Overview

The range of parameters collected related to environmental health is as broad as the field itself—encompassing activities, such as tracking logging, monitoring water contamination, and detecting invasive species and disease vectors. Over half of the systems (8/15, 53%) collecting environmental health data also include a focus on animal and human health. Kidenga collects data on environmental factors including the risk of mosquito exposure alongside human health symptoms. SMART for Health captures key environmental and epidemiological features as well as information on wildlife and livestock. Participatory surveillance systems vary in scale, with some designed for local, relatively small geographic areas and others operating on global platforms, such as iNaturalist. Several vector-focused systems reported that weather information, such as rainfall and temperature, can be added retrospectively by linking it to date, time, and location. Natural habitat data in these systems includes landscape descriptions, while details of tended habitats include descriptions of yards, gardens, and ponds. The built environment is also considered, with systems collecting data on housing structures, water and sanitation, and energy sources, such as fuel and electricity.

Plants

Three systems collect data on plants, including the Early Detection and Distribution Mapping System, which also gathers data on animals. Collected parameters include date or season, high-resolution digital images, information on the reporter, and detailed location information. Date, time, and GPS coordinates are often extracted directly from digital image metadata and autofilled in the report. Other details include ecosystem type, whether the area is tended or untended, seasonality, and additional species, such as host or insect pests.

Species identification is commonly performed by the primary reporter, with photos allowing for additional verification. Species identification is done in 1 reporting field or across multiple fields, capturing details or images of features such as flowers, seeds, leaves, leaf arrangements, and the entire plant. All the plant-focused systems record data on invasive plants and some databases also include information on legal, noxious weed status. High-quality, high-resolution digital images and tightly calibrated geographic locations are noted to be particularly important in these systems.

Vectors

Almost half (7/15, 46.7%) of the environmental systems are explicitly designed to monitor vectors. Notably, the Early Detection and Distribution Mapping System amalgamates vector, animal, and plant surveillance. Among the 7 vector surveillance systems that primarily focus on the environmental sector, 3 exhibit a multisectoral approach. Kidenga and MoBuzz+ bridge the environmental and human domains.

Vector exposure is collected in several ways across systems including incident-specific data with details on the date and precise location of the vector bite incident, specific vector bite information, and the detection of vectors in the vicinity. Often, observations include the perceived number of vectors and their morphology.

Data encompasses information on any formal diagnoses of vector-borne diseases and community prevalence of disease. Vector identification data incorporates visual documentation using photos uploaded by the user. Photos enable experts receiving the data to visually identify and assess vector size, color, and, where possible, species. Vector habitat data includes breeding sites and the detection of larvae, which use the number of measurement containers at a given location and whether they contain larvae. A number of systems integrate the history of recent outdoor activities to better understand the potential exposure of vectors. Vector mitigation data collects information on the use of pesticides, as well as any treatments for controlling or countering vector exposure, and appropriate clothing.

Animal Health Parameters

Wildlife Health

Nine of the participatory surveillance systems collect data on wild animals. SMART for Health includes parameters to record the health of free-ranging livestock that may be present in protected natural habitats. Most of these systems were developed with the goal of syndromic wildlife disease surveillance except for 2 systems— one that was designed to collect data on geographic distribution and abundance of wild mammalian species and another that focuses on early detection of invasive wildlife species [13]. Among these 9 systems, 4 (44%) also collect information on a range of environmental variables. Many variables focus on risk factors or environmental drivers of adverse health events/disease outbreaks in animals as well as humans such as environmental disturbance, or interactions at the wildlife-livestock-human interface. In addition, WildHealthNet—Ebola integrates data on both wildlife and human health parameters specifically for Ebola Virus Disease surveillance.

Common across most of these systems are parameters specific to wildlife incidents, including species or taxa groups of wild animals involved in the event, date, and numbers and location of affected animals. For location information, a few of the systems have the capability to capture specific GPS coordinates as the user digitally enters data or to generate GPS coordinates from an address through geocoding. Typically, systems also include data on the condition of the animal(s) involved in the incident, including abnormal findings in sick, injured, or dead wild animals and a presenting syndrome or clinical classification of these animals. Depending on the system, data are reported at the level of the individual animal or species.

Wildlife health parameters are collected in a variety of formats including dropdown menus, check boxes, and fields with free text. Categorizing free text data for surveillance purposes through natural language processing models has been applied by WildAlert to categorize data for surveillance purposes. Collection of data related to risk of exposure to humans—such as handling of sick wildlife or bats roosting near a home was also included.

Two of the systems also collect data on definitive diagnosis. For systems that integrate clinical wildlife data, this data, if available, is included for both individual animals as well as events. Other parameters among these systems include circumstances surrounding the incident and the potential causes of illness, injury, or death at both individual and species levels. Furthermore, one system also allows the user to enter data on the estimated population size and number of animals at risk for wildlife incidents enabling estimations of disease occurrence among a population.

Livestock Health

Only 4 surveillance systems collect data on livestock, and all of them are multisectoral. These systems together cover a wide variety of domestic species including livestock, poultry, dogs, and cats. These systems collect data using a range of technologies including calls and mobile apps. They focus on common parameters such as species, age, sex, and location of affected animals, including the magnitude of morbidity and mortality at individual and group (herd or flock) levels. Clinical signs of illness were either reported by body systems (checkboxes) or as free text. AfyaData collects health information on the owner as well as on the animals. Only 2 systems include parameters related to companion animals.

Human Health Parameters

Data parameters on human health are collected by 23/38 (61%) systems, with the majority (17/23, 74%) collecting only human health data. Of the 6 collecting multisector data, 3 systems collect data parameters across all One Health sectors, that is, human, animal, and environmental health.

The data parameters for human health cover virtually every organ system within the human body. Across all the systems monitoring human health, fever-related terms are the most common set of data parameters for reporting symptoms of illness. Additional symptoms indicative of influenza-like illness (ILI), including headache, sore throat, and cough are also among the most common symptom parameters collected. Gastrointestinal symptoms are the next most common set of data parameters across the systems collecting human health data.

In 2020, COVID-19-specific symptoms were added to several of the systems, many of which already collected ILI symptoms to track influenza, but now added “loss of taste or smell,” for example. COVID-19 and influenza diagnostic testing and vaccination history are common among the systems monitoring human health only.

Outbreaks Near Me asks the user to report being "healthy." Only when the user reports not being "healthy" do they receive queries on symptoms, events, or other options available within a specific system. Most systems focus on symptoms of illness and the onset of the timing of the symptoms. Several systems ask about health-seeking behavior and the type of health care system visited. Kidenga asks about the health status of others in the household and others inquire about the number of people in the household.

Absenteeism from work or school or changes in daily routine due to illness is captured by many of the systems. Common parameters also include various risk factors, such as exposure to other people with similar symptoms, exposure to pets, work exposure, and recent travel history. Furthermore, a few of the systems collect data on preexisting medical conditions as well as lifestyle parameters such as exercise, smoking, and alcohol consumption.

Discussion

Overview

The One Health Participatory Surveillance Data Parameters Compendium showcases the breadth and depth of how systems are using community engagement to collect multisector data to better inform shared threats. This resource highlights opportunities to expand One Health surveillance and identifies parameters currently collected across multisector systems. It encompasses a broad array of data modes and mechanisms used for data collection, addressing challenges with optimizing user time for data entry and improving accuracy. Additionally, the compendium reveals sector-specific differences in how and when data are collected.

Principal Findings

Opportunities and Challenges for System Augmentation or Development

Only 3 systems—PODD, AfyaData, and Colab—include data from all 3 sectors. For existing single-sector surveillance systems, the compendium offers a framework to identify key parameters for expanding into this type of multisector, One Health surveillance system. For example, wildlife-focused systems could incorporate data on the surrounding environment or on the health of humans interacting with wildlife. Similarly, human health–focused systems in urban areas could expand to include data parameters on the built environment, proximal wildlife, and the health of companion animals. In rural areas, human health systems could include parameters on the health of livestock and wildlife within the shared environment. Given the increasing incidence of emerging infections, data on drivers of disease emergence—such as environmental and weather changes—are valuable contributions to participatory surveillance systems [7]. With the interconnectedness of human, animal, and environmental health, the expansion of environmental systems to include these components is a logical progression.

Most participatory systems are community-driven and developed; thus, parameters are added to a given system to address specific community needs and concerns. Therefore, piloting terms and questions within a target population is essential to ensure wording and translation are culturally appropriate and that the requested parameters are accurately captured. Ensuring integration with governmental surveillance efforts across multiple ministries and departments is also critical when expanding beyond one sector to support more sustainable and successfully coordinated One Health approaches.

Expanding a surveillance system to collect parameters across sectors requires consideration of the time users are willing to invest. Users can only reasonably be expected to spend a limited amount of time answering questions or selecting from menu options. Many systems address this by incorporating automatic data inputs, such as timestamps or geolocation, to save time and improve accuracy. Decision-tree frameworks are also used to streamline data collection and gather targeted information; for example, a "yes" response to a question about "cough" could lead to a follow-up inquiry about "bloody cough." Wildlife and vector programs frequently rely on images to minimize the number of questions and to ensure accurate species identification.

Advancements in technology allow for automated data collection, such as integrating ambient temperature measurements, which can enhance user experience, save time, and improve data quality. Variations in education and literacy levels within communities must be considered. Working closely with the community during system development can help address these challenges and foster user engagement.

Existing siloes among One Health sectors can limit the effective use of data and hinder cross-sectoral collaboration. Existing system infrastructure as well as sensitivities around data privacy can create barriers that make it difficult to collect and share parameters across sectors. Current systems also operate at varying spatial scales. For example, human health data is often protected by privacy regulations, making it unethical to include precise geocoordinates. On the other hand, some environmental parameters, such as mosquito habitats or locations of dead wildlife, require specific geocoded data to enable required actions. Data privacy is critical and can be addressed through user-specific access layers. Integrated dashboards and similar tools can enable health authorities from multiple sectors to visually identify disease hotspots, pinpoint areas with elevated disease transmission risk, and prioritize mitigation efforts. For instance, in Mo-Buzz+, public health inspectors can review reports on communal dengue cases, identify mosquito breeding sites, and detect community practices that contribute to mosquito breeding. To protect privacy, this type of information could be aggregated to a neighborhood level before being shared with public users, preserving anonymity while providing actional insights.

For livestock and plant health, additional sensitivities regarding the location of incidents may arise due to potential economic impacts. Systems address these concerns in various ways, such as obscuring publicly available information and anonymizing reporter details. Any system handling sensitive information must proactively address these concerns to ensure confidentiality and minimize potential risks.

System Structure and Timing Impact Types of Information Gathered

The types of information gathered by surveillance systems vary based on their structure. Systems designed to collect longitudinal data often include more “general” parameters during the initial registration process, reducing the need to recollect this information during subsequent logins. This approach saves the user time by focusing additional questions only on cases where the user reports “unhealthy” conditions for that week. Collecting longitudinal data from a known population allows these systems to calculate both numerators and denominators for analyzing trends and spatial-temporal patterns. Visualizing reports over time and space strengthens disease tracking and anomaly detection. This approach is most common for human health systems, as well as Garden Wildlife Health and WildAlert. In SMART for Health, users (wildlife rangers) track animal populations over time, allowing for the identification of both population and disease trends.

Some systems that collect data on wildlife and environmental factors rely on users to report opportunistically when they encounter a specific animal or setting. As these encounters may be the only opportunity to collect data on the health event, extensive data may be needed. To make the process user-friendly, these systems often incorporate photographs and automated downloads of data such as geolocation and limit the collection of extensive user and animal demographic details. While one-time reporting for plant or environmental health incidents is highly valuable for enabling prompt action, it rarely supports longitudinal surveillance or trend detection. Identification of ways to expand follow-up reporting by users may be appropriate, particularly in the context of suspected outbreaks. Systems collecting data on human-animal interactions, such as bat exposures, are especially valuable for surveillance initiatives aiming to improve understanding of the underlying factors driving pathogen spillover and disease emergence.

Dropdown menus offer the advantage of standardizing information, which facilitates data recording in the field, promotes consistent data analysis across units, and supports data integration across systems. In contrast, free text data fields capture a wide array of information from the user or when the user may not have the technical background or expertise to properly categorize a case.

Language and cultural barriers within communities can result in disparities in data capture. Addressing these challenges requires strategies such as raising community awareness and education to enhance disease literacy and improve access to health care. Using tailored engagement approaches for specific population segments and fostering collaboration with local stakeholders is essential for improving surveillance and response outcomes.

Data Standards

Improving interoperability and data sharing among participatory surveillance systems can allow for early incident detection. Improving interoperability (the ability of different systems to exchange, interpret, and use data accurately) and data sharing on a regional or global scale will require the development of data standards. Data standards facilitate international data exchange and better collaboration. A data standard is an agreed-upon approach, to allow for consistent measurement, qualification or exchange of an object, process, or unit of information. Data standards include agreements on several topic areas including representation, format, definition, structure, use, documentation, and management of data [28,29]. They enable transparency and understanding, and the use of standards promotes common, clear meanings for data [29]. Several systems in this study collect similar categories of parameters but use different questions or methods to gather the information. This variability is often necessary to account for community-specific differences, such as how questions about ethnic groups are phrased. Ideally, while user interfaces may differ to accommodate these variations, the backend systems should be designed to integrate the collected information.

Sector-Specific Findings

Human Health

It is not surprising to see fever as the most prevalent symptom among systems designed to detect early infection in humans. Many have noted that fever monitored on an ongoing basis with temporal and geospatial visualization could be an early indicator of a communicable disease threat [30,31]. Since respiratory diseases have the highest probability of causing local outbreaks, regional epidemics, and potentially a pandemic, many systems monitor for ILI and this likely accounts for the frequency of ILI-terms among the systems. Several systems collect parameters to complement symptom data reported—such as health care–seeking behavior to help track the severity of illness or travel to help understand exposure risks.

Environmental Health

The inclusion of environmental parameters in multisector systems allows systems to understand the complete context in which an incident occurs. Identification of potential vector-breeding locations, water contamination, and habitat change may allow for intervention before the spread of disease into local animal and human populations. Environmental parameters also capture how animals and people may be impacting the health of the environment.

Like wildlife health, environmental health programs often rely on event-based reporting. For example, a person enters information about vector exposure, seeing a contaminated location, or spotting a presumed invasive species. Automated features, such as automated collection of climate parameters, geolocation, and image uploads, allow the system to accurately capture a situation without solely relying on an observer’s subjective reporting. To help standardize data, systems use tools such as iNaturalist, which may include predefined choices, and have internal experts reviewing photographs to correctly identify vector, animal, and plant species. The collection of incident-specific details is foundational for understanding the spatiotemporal distribution of vectors and their potential role in disease transmission.

Despite the comprehensive parameters used, the underreporting of vector-borne disease cases in humans remains a persistent concern. This phenomenon is especially prevalent in regions with limited access to health care services or where vector-borne diseases are poorly understood [30,31].

Livestock and Wildlife Health

Relatively fewer participatory systems exist for livestock compared to systems dedicated to other sectors. Participatory epidemiology is often used to address specific outbreaks in livestock health [32]; ongoing participatory surveillance systems may not be as common due to heavy reliance on governmental surveillance programs for outbreak detection. In the compendium, livestock parameters are from systems also collecting data on wildlife, people, or the environment, reflecting the interconnectedness of livestock among each of the other sectors. Given the daily proximity of pets to owners and the potential for bidirectional transmission, the expansion of participatory surveillance to inform companion animal outbreak detection is needed.

Similar to environmental health systems, wildlife health systems are often event-based—increasing situational awareness of disease occurrence and informing on common disease threats in real time. Information stemming from these systems are commonly underused as resource for monitoring long-term trends in population health and mostly are untapped as resource by sectors outside of wildlife health.

Across wildlife participatory surveillance systems, data accuracy remains a significant challenge, as users often face difficulties accurately identifying wildlife species and describing abnormalities in sick, injured, or dead wild animals. Additionally, the integration of disparate data sources and achieving interoperability between systems is limited due to the diverse types and formats of data collected. The lack of harmonized data in wildlife health surveillance has been a long-standing issue; however, systems designed for large, multilingual, and strategically diverse user bases hold promise for reversing this trend.

The increasing use of machine learning methods, such as natural language processing, offers new opportunities for categorizing data, and enabling comparisons across multiple surveillance streams and systems [33,34]. However, these approaches require compatible language inputs or effective translation mechanisms. The compendium highlights the value of collecting and analyzing prediagnostic data during incidents, which facilitates the rapid detection of clusters of animals presenting with similar abnormalities that might signify a disease outbreak or other adverse health event in wildlife.

Comparison With Prior Work

This work builds upon the landscape analysis of participatory surveillance programs and IWOPS workshops reviewing data parameters and working toward data standards [13,25]. Prior publications have highlighted frameworks and steps for building One Health surveillance systems, including methods for better cross-sectoral stakeholder engagement and ways to overcome implementation barriers [7,10,11]. Identified barriers to the implementation of One Health surveillance systems include incompatible vocabularies, professional divisions, isolated datasets, and the lack of data standards. When developing new One Health participatory surveillance systems, this compendium provides a concrete starting point with real-world examples to facilitate dialogue on data elements and standards for sharing across siloed sectors. A critical analysis by Behravesh et al [12] of landscaping globally available One Health tools uncovered 50 tools that advance One Health. The results highlight that most of these tools addressed assessment and few addressed planning and prioritization. This compendium provides a practical One Health tool to facilitate and support planning, prioritization discussions, and One Health participatory surveillance system design. Prior studies have noted the value of participatory surveillance in war-affected camps and the use of One Health surveillance to capture the impact of the environment on refugee health [35,36]. Interactions between humans and animals are often ignored with regard to the health of refugees, but animal populations are often closely tied to refugee communities. This compendium can be applied to improve One Health disease surveillance for displaced populations in high-risk regions.

Limitations

Data parameters included in the compendium reflect a point in time and are not fully representative of the entire field of participatory surveillance. The parameters were collected in English, through an English-language survey, which may have limited participation from non-English systems. Furthermore, selection bias could be present, as systems were identified based on the participatory surveillance map. Additional validation of data parameters was not conducted because many systems are not publicly available; the parameters included in the compendium are based on the information directly provided by the systems.

Most systems use a decision-tree framework to structure their questions. The compendium does not include this format, focusing on parameters over the structural design. To partially address this limitation, the “special considerations” column in the compendium provides comments on how specific parameters relate to others. Future iterations of the compendium should explore methods to illustrate system structures, providing context for how parameters are used within their frameworks.

Given the compendium’s focus on outbreak detection, parameters related to chronic disease were not included. Future exploration on how mental health, chronic disease, toxic exposures, and other non-outbreak–related parameters are collected would be informative for participatory surveillance systems aiming to capture a broader set of health parameters.

Conclusions

Participatory surveillance enhances traditional means of health surveillance by empowering communities to identify threats at their source. The compendium highlights how existing participatory surveillance systems are already incorporating a One Health approach. Expanding opportunities for multisector data collection through these participatory systems aligns with recent guidance from the United Nations Quadripartite [3,37]. Additional research is required to inform international data standards for One Health participatory surveillance. Pilot studies could provide useful mechanisms to capture various multisectoral data collection and sharing techniques. Furthermore, ethical issues of data privacy are always essential to consider when capturing personal information. Participatory surveillance provides critical surveillance data that can be deidentified and collated at a specific geographic level, which facilitates anonymity of the responses. The compendium offers a reference to strengthen community capabilities for developing multisector surveillance systems to inform early detection of health incidents threatening themselves, their families, and the environment. Decision makers and practitioners can use the compendium as a tool to build a new participatory surveillance system or augment their existing systems to include a One Health approach. The compendium highlights the need to work toward system interoperability and data standards to enable community-level surveillance to have national, regional, or global implications in outbreak detection and response.

Acknowledgments

The authors would like to acknowledge support from Marlo Libel for his assistance on this project, and Elizabeth Aaron for copyediting the compendium and assisting in data collection. The University of Arizona and Dr Kacey Ernst for providing human participants' review of this project. The authors appreciate the following systems sharing data parameters to ensure that they had a comprehensive compendium as possible: AfyaData, AZDGF, Cambodia 115 Hotline, California West Nile Bird Reporting, Colab, CoughWatch SA, DisApp, EDDMapS/Wild Spotter, eWHIS, FluTracking Australia, FluTracking HongKong, Forest Pest First Detector, GrippeWeb, Guardians of Health, Garden Wildlife Health Project, Health Report (Hälsorapport) Sweden, iMammalia, InfluenzaNet (including individual country-level programs in Italy, France, Germany, United Kingdom, the Netherlands, Belgium, Switzerland, Sweden, and Denmark), I-Tick, Kidenga, MoBuzz+, Outbreaks Near Me, and PODD. Generative artificial intelligence was not used.

Data Availability

All data generated or analyzed during this study are included in this published article [and its supplementary information files—Multimedia Appendix 1].

Conflicts of Interest

None declared.

Multimedia Appendix 1
One Health Participatory Surveillance Compendium.
[XLSX File (Microsoft Excel File), 672 KB - publichealth_v11i1e55356_app1.xlsx]

References

1. Food and Agriculture Organization of the United Nations, United Nations Environment Programme, World Health Organization, World Organisation for Animal Health. The One Health Joint Plan of Action (2022–2026). Working together for the health of humans, animals, plants and the environment. World Organisation for Animal Health. 2022. URL: https://bulletin.woah.org/?panorama=02-1-4-2023-1_oh-jpa [accessed 2024-12-21]
2. Mercy K, Balajee A, Numbere T, Ngere P, Simwaba D, Kebede Y. Africa CDC's blueprint to enhance early warning surveillance: accelerating implementation of event-based surveillance in Africa. J Public Health Afr 2023;14(8):2827 [FREE Full text] [doi: 10.4081/jphia.2023.2827] [Medline: 37753431]
3. Defining collaborative surveillance: a core concept for strengthening the global architecture for health emergency preparedness, response, and resiliency (HEPR). World Health Organization. URL: https://www.who.int/publications/i/item/9789240074064 [accessed 2024-12-06]
4. One Health High-Level Expert Panel (OHHLEP), Adisasmito WB, Almuhairi S, Behravesh CB, Bilivogui P, Bukachi SA, et al. One Health: a new definition for a sustainable and healthy future. PLoS Pathog 2022;18(6):e1010537 [FREE Full text] [doi: 10.1371/journal.ppat.1010537] [Medline: 35737670]
5. Bordier M, Uea-Anuwong T, Binot A, Hendrikx P, Goutard FL. Characteristics of One Health surveillance systems: a systematic literature review. Prev Vet Med 2020;181:104560. [doi: 10.1016/j.prevetmed.2018.10.005] [Medline: 30528937]
6. Kelly TR, Machalaba C, Karesh WB, Crook PZ, Gilardi K, Nziza J, PREDICT Consortium, et al. Implementing One Health approaches to confront emerging and re-emerging zoonotic disease threats: lessons from PREDICT. One Health Outlook 2020;2:1 [FREE Full text] [doi: 10.1186/s42522-019-0007-9] [Medline: 33824944]
7. One Health High-Level Expert Panel (OHHLEP), Hayman DTS, Adisasmito WB, Almuhairi S, Behravesh CB, Bilivogui P, et al. Developing One Health surveillance systems. One Health 2023;17:100617 [FREE Full text] [doi: 10.1016/j.onehlt.2023.100617] [Medline: 38024258]
8. Rizzo DM, Lichtveld M, Mazet JAK, Togami E, Miller SA. Plant health and its effects on food safety and security in a One Health framework: four case studies. One Health Outlook 2021;3:6 [FREE Full text] [doi: 10.1186/s42522-021-00038-7] [Medline: 33829143]

9. Gruetzmacher K, Karesh WB, Amuasi JH, Arshad A, Farlow A, Gabrysch S, et al. The Berlin principles on One Health—bridging global health and conservation. Sci Total Environ 2021;764:142919 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.142919] [Medline: 33097250]
10. Bordier M, Goutard FL, Antoine-Moussiaux N, Pham-Duc P, Lailler R, Binot A. Engaging stakeholders in the design of one health surveillance systems: a participatory approach. Front Vet Sci 2021;8:646458 [FREE Full text] [doi: 10.3389/fvets.2021.646458] [Medline: 34109232]
11. Uchtmann N, Herrmann JA, Hahn EC, Beasley VR. Barriers to, efforts in, and optimization of integrated One Health surveillance: a review and synthesis. Ecohealth 2015;12(2):368-384. [doi: 10.1007/s10393-015-1022-7] [Medline: 25894955]
12. Behravesh CB, Charron DF, Liew A, Becerra NC, Machalaba C, Hayman DT, et al. An integrated inventory of One Health tools: mapping and analysis of globally available tools to advance One Health. CABI One Health 2024;3:1. [doi: 10.1079/cabionehealth.2024.0017]
13. McNeil C, Verlander S, Divi N, Smolinski M. The landscape of participatory surveillance systems across the One Health spectrum: systematic review. JMIR Public Health Surveill 2022;8(8):e38551 [FREE Full text] [doi: 10.2196/38551] [Medline: 35930345]
14. Pruvot M, Denstedt E, Latinne A, Porco A, Montecino-Latorre D, Khammavong K, et al. WildHealthNet: Supporting the development of sustainable wildlife health surveillance networks in Southeast Asia. Sci Total Environ 2023;863:160748 [FREE Full text] [doi: 10.1016/j.scitotenv.2022.160748] [Medline: 36513230]
15. Wójcik OP, Brownstein JS, Chunara R, Johansson MA. Public health for the people: participatory infectious disease surveillance in the digital age. Emerg Themes Epidemiol 2014;11:7 [FREE Full text] [doi: 10.1186/1742-7622-11-7] [Medline: 24991229]
16. Suy Lan C, Sok S, Chheang K, Lan DM, Soung V, Divi N, et al. Cambodia national health hotline—participatory surveillance for early detection and response to disease outbreaks. Lancet Reg Health West Pac 2022;29:100584 [FREE Full text] [doi: 10.1016/j.lanwpc.2022.100584] [Medline: 36605884]
17. Leal-Neto O, Egger T, Schlegel M, Flury D, Sumer J, Albrich W, et al. Digital SARS-CoV-2 detection among hospital employees: participatory surveillance study. JMIR Public Health Surveill 2021;7(11):e33576 [FREE Full text] [doi: 10.2196/33576] [Medline: 34727046]
18. Pagliari C, Vijaykumar S. Digital participatory surveillance and the Zika crisis: opportunities and caveats. PLoS Negl Trop Dis 2016;10(6):e0004795 [FREE Full text] [doi: 10.1371/journal.pntd.0004795] [Medline: 27294787]
19. Leal Neto O, Haenni S, Phuka J, Ozella L, Paolotti D, Cattuto C, et al. Combining wearable devices and mobile surveys to study child and youth development in Malawi: implementation study of a multimodal approach. JMIR Public Health Surveill 2021;7(3):e23154 [FREE Full text] [doi: 10.2196/23154] [Medline: 33536159]
20. Mahmud AS, Chowdhury S, Sojib KH, Chowdhury A, Quader MT, Paul S, et al. Participatory syndromic surveillance as a tool for tracking COVID-19 in Bangladesh. Epidemics 2021;35:100462 [FREE Full text] [doi: 10.1016/j.epidem.2021.100462] [Medline: 33887643]
21. Guenin MJ, de Nys HM, Peyre M, Loire E, Thongyuan S, Diallo A, et al. A participatory epidemiological and One Health approach to explore the community's capacity to detect emerging zoonoses and surveillance network opportunities in the forest region of Guinea. PLoS Negl Trop Dis 2022;16(7):e0010462 [FREE Full text] [doi: 10.1371/journal.pntd.0010462] [Medline: 35816491]
22. Garg S, Bhatnagar N, Gangadharan N. A case for participatory disease surveillance of the COVID-19 pandemic in India. JMIR Public Health Surveill 2020;6(2):e18795 [FREE Full text] [doi: 10.2196/18795] [Medline: 32287038]
23. Worsley-Tonks KEL, Bender JB, Deem SL, Ferguson AW, Fèvre EM, Martins DJ, et al. Strengthening global health security by improving disease surveillance in remote rural areas of low-income and middle-income countries. Lancet Glob Health 2022;10(4):e579-e584 [FREE Full text] [doi: 10.1016/S2214-109X(22)00031-6] [Medline: 35303467]
24. Wittwer S, Paolotti D, Lichand G, Leal Neto O. Participatory surveillance for COVID-19 trend detection in Brazil: cross-sectional study. JMIR Public Health Surveill 2023;9:e44517 [FREE Full text] [doi: 10.2196/44517] [Medline: 36888908]
25. Smolinski MS, Crawley AW, Olsen JM, Jayaraman T, Libel M. Participatory disease surveillance: engaging communities directly in reporting, monitoring, and responding to health threats. JMIR Public Health Surveill 2017;3(4):e62 [FREE Full text] [doi: 10.2196/publichealth.7540] [Medline: 29021131]
26. Leal Neto O, Paolotti D, Dalton C, Carlson S, Susumpow P, Parker M, et al. Enabling multicentric participatory disease surveillance for global health enhancement: viewpoint on global flu view. JMIR Public Health Surveill 2023;9:e46644 [FREE Full text] [doi: 10.2196/46644] [Medline: 37490846]
27. IWOPS program. Ending Pandemics. 2022. URL: https://endingpandemics.org/iwops-program [accessed 2024-12-06]
28. Data standards. National Library of Medicine. URL: https://www.nnlm.gov/guides/data-glossary/data-standards [accessed 2024-12-06]
29. Learn about data standards. United States Environmental Protection Agency. URL: https://www.epa.gov/data/learn-about-data-standards [accessed 2024-12-06]
30. Multisectoral approach to the prevention and control of vector-borne diseases: a conceptual framework. World Health Organization. 2020. URL: https://iris.who.int/bitstream/handle/10665/331861/9789240004788-eng.pdf [accessed 2024-12-06]

31. Mubayi A, Castillo-Chavez C, Chowell G, Kribs-Zaleta C, Ali Siddiqui N, Kumar N, et al. Transmission dynamics and underreporting of Kala-azar in the Indian state of Bihar. J Theor Biol 2010;262(1):177-185. [doi: 10.1016/j.jtbi.2009.09.012] [Medline: 19769990]
32. Jost CC, Mariner JC, Roeder PL, Sawitri E, Macgregor-Skinner GJ. Participatory epidemiology in disease surveillance and research. Rev Sci Tech 2007;26(3):537-549 [FREE Full text] [Medline: 18293603]
33. Kelly TR, Pandit PS, Carion N, Dombrowski DF, Rogers KH, McMillin SC, et al. Early detection of wildlife morbidity and mortality through an event-based surveillance system. Proc Biol Sci 2021;288(1954):20210974 [FREE Full text] [doi: 10.1098/rspb.2021.0974] [Medline: 34256001]
34. Mackey T, Purushothaman V, Li J, Shah N, Nali M, Bardier C, et al. Machine learning to detect self-reporting of symptoms, testing access, and recovery associated with COVID-19 on Twitter: retrospective big data infoveillance study. JMIR Public Health Surveill 2020;6(2):e19509 [FREE Full text] [doi: 10.2196/19509] [Medline: 32490846]
35. Sutradhar I, Zaman MH. One Health approaches to improve refugee health. Lancet Glob Health 2021;9(12):e1646-e1647 [FREE Full text] [doi: 10.1016/S2214-109X(21)00416-2] [Medline: 34798016]
36. Wickramage K, Nellapalli P. Community participatory methods in disease surveillance and public health in war-affected camps, and its potential contribution to peace building. Int Electron J Health Educ 2008:95-108 [FREE Full text]
37. Quadripartite One Health intelligence scoping study: actions to develop an effective global One Health intelligence system. Food and Agriculture Organization of the United Nations.: World Health Organization; 2022. URL: https://www.fao.org/3/cc1533en/cc1533en.pdf [accessed 2024-12-06]

Abbreviations

EBS: event-based surveillance
ILI: influenza-like illness
IWOPS: International Workshops on Participatory Surveillance
PODD: Participatory One Health Digital Disease Detection
USAID: United States Agency for International Development

Edited by A Mavragani; submitted 11.12.23; peer-reviewed by D Balate, H Kinzer, C Argyropoulos, F Fathmawati; comments to author 23.09.24; revised version received 07.12.24; accepted 18.12.24; published 26.03.25.

Please cite as:
McNeil C, Divi N, Bargeron IV CT, Capobianco Dondona A, Ernst KC, Gupta AS, Fasominu O, Keatts L, Kelly T, Leal Neto OB, Lwin MO, Makhasi M, Mutagahywa EB, Montecino-Latorre D, Olson S, Pandit PS, Paolotti D, Parker MC, Samad MH, Sewalk K, Sheldenkar A, Srikitjakarn L, Suy Lan C, Wilkes M, Yano T, Smolinski M
Data Parameters From Participatory Surveillance Systems in Human, Animal, and Environmental Health From Around the Globe: Descriptive Analysis
JMIR Public Health Surveill 2025;11:e55356
URL: https://publichealth.jmir.org/2025/1/e55356
doi:10.2196/55356
PMID:

Population Size Estimation of Men Who Have Sex With Men in Low- and Middle-Income Countries: Google Trends Analysis

Carly M Malburg[1,2*], MPH; Steve Gutreuter[1*], PhD; Horacio Ruiseñor-Escudero[1*], MD, PhD; Abu Abdul-Quader[1*], PhD; Wolfgang Hladik[1*], MD, PhD

[1]Division of Global HIV and TB, Centers for Disease Control and Prevention, 1600 Clifton Rd, Atlanta, GA, United States
[2]CDC Global Health Fellowship Program, Public Health Institute, Oakland, CA, United States
[*]all authors contributed equally

Corresponding Author:
Carly M Malburg, MPH
Division of Global HIV and TB, Centers for Disease Control and Prevention, 1600 Clifton Rd, Atlanta, GA, United States

Abstract

Background: Population size estimation (PSE) for key populations is needed to inform HIV programming and policy.

Objective: This study aimed to examine the utility of applying a recently proposed method using Google Trend (GT) internet search data to generate PSE (Google Trends Population Size Estimate [GTPSE]) for men who have sex with men (MSM) in 54 countries in Africa, Asia, the Americas, and Europe.

Methods: We examined GT relative search volumes (representing the relative internet search frequency of specific search terms) for “porn” and, as a comparator term, “gay porn” for the year 2020. We assumed “porn” represents “men” (denominator) while “gay porn” represents a subset of “MSM” (numerator) in each county, resulting in a proportional size estimate for MSM. We multiplied the proportional GTPSE values with the countries’ male adult population (15 - 49 years) to obtain absolute size estimates. Separately, we produced subnational MSM PSE limited to countries’ (commercial) capitals. Using linear regression analysis, we examined the effect of countries’ levels of urbanization, internet penetration, criminalization of homosexuality, and stigma on national GTPSE results. We conducted a sensitivity analysis in a subset of countries (n=14) examining the effect of alternative English search terms, different language search terms (Spanish, French, and Swahili), and alternative search years (2019 and 2021).

Results: One country was excluded from our analysis as no GT data could be obtained. Of the remaining 53 countries, all national GTPSE values exceeded the World Health Organization’s recommended minimum PSE threshold of 1% (range 1.2% - 7.5%). For 44 out of 49 (89.8%) of the countries, GTPSE results were higher than Joint United Nations Programme on HIV/AIDS (UNAIDS) Key Population Atlas values but largely consistent with the regional UNAIDS Global AIDS Monitoring results. Substantial heterogeneity across same-region countries was evident in GTPSE although smaller than those based on Key Population Atlas data. Subnational GTPSE values were obtained in 51 out of 53 (96%) countries; all subnational GTPSE values exceeded 1% but often did not match or exceed the corresponding countries’ national estimates. None of the covariates examined had a substantial effect on the GTPSE values (R^2 values 0.01 - 0.28). Alternative (English) search terms in 12 out of 14 (85%) countries produced GTPSE>1%. Using non-English language terms often produced markedly lower same-country GTPSE values compared with English with 10 out of 14 (71%) countries showing national GTPSE exceeding 1%. GTPSE used search data from 2019 and 2021, yielding results similar to those of the reference year 2020. Due to a lack of absolute search volume data, credibility intervals could not be computed. The validity of key assumptions, especially who (males and females) searches for porn and gay porn, could not be assessed.

Conclusions: GTPSE for MSM provides a simple, fast, essentially cost-free method. Limitations that impact the certainty of our estimates include a lack of validation of key assumptions and an inability to assign credibility intervals. GTPSE for MSM may provide an additional data source, especially for estimating national-level PSE.

(JMIR Public Health Surveill 2025;11:e58630) doi:10.2196/58630

KEYWORDS

population size estimation; men who have sex with men; MSM; PSE; google trends; HIV; AIDS; programming and policy; internet; porn; gay porn; male adult; geriatric; linear regression; homosexuality; sensitivity analysis; World Health Organization; WHO; epidemiology

Introduction

The Joint United Nations Programme on HIV/AIDS (UNAIDS) estimated that in 2022, about 39 million people were living with HIV worldwide [1]. HIV burden is higher among men who have sex with men (MSM), people who inject drugs, sex workers, and transgender persons, which together are often described as key populations (KP) [1]. KPs and their paying or nonpaying sexual partners may account for 70% of new HIV infections worldwide, with an estimated 80% of new HIV infections outside sub-Saharan Africa (SSA) and 55% of all new infections within SSA [1,2].

Key population size estimation (PSE) is needed to estimate the number of individuals belonging to a KP in a given geographical area [3,4]. PSEs provide the denominator values to inform KP programming and policy [5]. However, PSE is a difficult field and its methods often lack rigor in design or implementation, and the many methods available reflect the lack of an acceptable gold standard [3,6]. Challenges to PSE include lack of sampling frames, mobility, and nondisclosure of KP-defining behaviors [3,4]. Further, most PSE methods produce local estimates whereas national PSE estimates are often obtained through “expert opinion,” simple projection, or modeling and less often through national-level empirical data such as direct survey questions or the network scale-up method, both used in general population-based surveys [6,7]. Direct survey questions about KP-defining traits experience reporting bias and require a major effort unless they can be added to an already planned general population survey. The frequent lack of reliable national-level PSE constitutes an even larger challenge compared with the availability of local PSE and complicates national, regional, and global HIV estimation work [3,8-10].

The rise of the internet facilitates web-based activities to improve public health, including in the field of digital epidemiology and infoveillance [11]. Recently, a new PSE method using Google Trends (GT) internet search data was proposed in a proof of concept paper by Card et al [12] GT is a free cloud-based app that displays the relative frequency of user-specified Google search terms as trends across time and user-selected geographical areas [12-14]. Card et al [12] used GT and Canadian census data to estimate the local PSE of MSM in urban and rural locations throughout Canada. Card et al [12] related search terms presumed to be representative of MSM (“gay porn”) to that presumed to be representative of the general (male) population (“porn”). By relating these 2 sets of values, Card et al [12] estimated the relative size of MSM in these Canadian towns. To date, no other published PSE exists using this method.

The literature on pornography consumption by sex and sexual orientation is limited and often the MSM population is not represented. However, a major porn website reported that about a third of its visitors globally in 2021 were reportedly women [15,16]. Further, women, regardless of sexual orientation, may also watch gay porn, possibly in substantial numbers [17]. Beyond this, we found no meaningful gray literature or peer-reviewed articles about internet pornography consumption in low- and middle-income countries (LMICs) or pornography consumption by MSM in LMICs. We are also not aware of (gray) literature about the proportion of heterosexual and homosexual men searching Google for (gay) porn in LMICs.

We expanded the literature search to include high-income settings. A study conducted in the United States reported that more men than women consume pornography (92%:68%, respectively) over the span of a year [18]. The study did not report the type of pornography consumed or disaggregate male respondents by sexual orientation or practice [18]. A separate study from Norway with a sample of some 2300 male and female participants suggested that more men than women consume some pornography (94% of men and 68% of women) [19]. However, only 5% (n=106) of participants identified as gay/lesbian/bisexual, no breakdown of sexual orientation by sex was given, and no information on the type of pornography consumed by participants was available [19].

The aim of this study was to examine the potential utility of using GT data to obtain MSM PSE in selected LMICs.

Methods

Preliminary Literature Search

A nonsystematic literature search was conducted to better understand the behavior of pornography consumption of the general population and sexual minorities, by sex, as well as the relative frequencies with which these populations search for (gay) porn in general (via Google) or by directly accessing specific porn sites.

Selection of Countries

We analyzed GT data for a selected set of 54 countries that receive support from the US President’s Emergency Plan for AIDS Relief, the US Government’s initiative to support global HIV responses, for which information on MSM PSE has been sought [2,20]. These countries are located in SSA (n=29), Asia (n=13), the Americas (n=11), and Ukraine (Tables 1 and 2).

Table . National men who have sex with men (MSM) population size estimation (PSE) for US President’s Emergency Plan for AIDS Relief supported countries (n=53) using Google Trends (GT) data for the year 2020[a,b].

Region and country		GT (number of MSM), n	GT, %	UNAIDS[c] GAM[d] regional %, median (IQR)[e]	UNAIDS KP[f] Atlas, %
East Africa[g]				1.67	
	Burundi	48,500	1.77		0.34
	Ethiopia	365,000	1.28		—[h]
	Kenya	276,000	1.99		0.24
	Rwanda	51,300	1.54		0.15
	Tanzania	243,000	1.73		0.35
	Uganda	154,000	1.47		0.23
Southern Africa[i]				1.67	
	Angola	106,000	1.44		—
	Botswana	13,000	2.12		0.43
	Eswatini	4500	1.57		1.38
	Lesotho	10,000	1.71		1.05
	Malawi	52,500	1.16		0.94
	Mozambique	134,000	1.87		0.22
	Namibia	16,500	2.60		—
	South Africa	393,000	2.46		1.94
	Zambia	51,800	1.18		0.15
	Zimbabwe	53,000	1.64		0.71
West Central Africa[j]				1.28 (IQR 0.45 - 1.50)	
	Benin	34,000	1.18		0.20
	Burkina Faso	88,000	1.80		0.07
	Cameroon	148,000	2.29		0.11
	Cote d’Ivoire	166,000	2.68		0.90
	DRC	33,000	1.66		0.98
	Ghana	112,000	1.40		0.69
	Liberia	22,600	1.83		6.04
	Mali	70,500	1.55		0.09
	Nigeria	614,000	1.26		0.49
	Senegal	73,600	1.94		1.38
	Sierra Leone	27,000	1.36		0.16
	Togo	53,100	2.65		0.30
Asia[k]				1.63 (IQR 0.26 - 3.10)	
	Burma	664,000	4.53		1.72
	Cambodia	258,000	5.67		1.93
	India	6,460,000	1.18		0.06
	Indonesia	1,180,000	1.61		1.03
	Kazakhstan	137,000	2.99		1.35
	Kyrgyz Rep.	53,000	3.10		0.99
	Lao PDR	53,000	2.73		2.96

Region and country		GT (number of MSM), n	GT, %	UNAIDS[c] GAM[d] regional %, median (IQR)[e]	UNAIDS KP[f] Atlas, %
	Nepal	83,000	1.19		0.86
	PNG	31,000	1.30		1.58
	Tajikistan	52,000	2.14		0
	Thailand	215,000	1.25		3.08
	Philippines	1,260,000	4.27		2.33
	Viet Nam	1,953,000	7.46		0.98
Europe					
	Ukraine	366,000	3.48	2.11 (IQR 1.75 - 2.49)	1.71
Caribbean[l]				2.71	
	Dominican Rep.	124,000	4.26		4.90
	Guyana	8200	3.60		1.45
	Haiti	108,000	3.60		1.03
	Jamaica	24,000	2.91		5.15
	Trinidad and Tobago	11,000	3.04		—
Central and South America[m]				3.37	
	Brazil	2,960,000	5.18		3.50
	El Salvador	85,000	5.20		3.31
	Guatemala	245,000	5.09		2.42
	Honduras	147,000	5.32		1.48
	Nicaragua	114,000	6.32		1.97
	Panama	81,000	7.23		2.65

[a] These estimates are for descriptive purposes only, to examine issues related to the potential utility of the method proposed by Card et al [12]. They represent the MSM population national population size estimates (percentage of MSM) for the year 2020. The percentage of MSM was calculated by taking the average relative search volume score produced by Google Trends for “gay porn” and dividing it by the average relative search volume score produced by Google Trends for “porn.” MSM population size estimate (number of MSM) was calculated by taking the percentage of MSM population size estimate and dividing it by the total male population (ages 15 - 49 years). Key populations (KPs) Atlas percentage of MSM population size estimate was calculated by dividing the absolute MSM population size estimate taken from the United Nations Programme on HIV/AIDS (UNAIDS) KPs Atlas dashboard by the total adult male population (ages 15 - 49 years), and then multiplying by 100. The absolute value difference was calculated by subtracting the GT absolute MSM population size estimate value from the KPs Atlas MSM population size estimate absolute value. All absolute values under 10,000 are rounded to the nearest 100. All other absolute values are rounded to the nearest 1000. UNAIDS Global AIDS Monitoring system (GAM) values are regional values transcribed from the UNAIDS open-source Spectrum 6 guide. The countries used to create these regions and respective values may not be in full alignment with the countries included in the population size estimate analysis, therefore direct 1:1 comparisons should not be made. Max:Min ratio: The ratio based on the largest and smallest PSE % value in each region.

[b] Absolute values are not provided as Google Trends does not provide absolute search frequency values.

[c] UNAIDS: United Nations Programme on HIV/AIDS.

[d] GAM: Global AIDS Monitoring system.

[e] IQR values were included for available regions. Regions without an IQR listed did not have one available.

[f] KP: key population.

[g] Max:Min ratio: 1.6 (GT) and 2.3 (UNAIDS KP).

[h] Not available (data missing for the country).

[i] Max:Min ratio: 2.2 (GT) and 12.9 (UNAIDS KP).

[j] Max:Min ratio: 2.3 (GT) and 86.3 (UNAIDS KP).

[k] Max:Min ratio: 6.3 (GT) and 51.3 (UNAIDS KP).

[l] Max:Min ratio: 1.5 (GT) and 5 (UNAIDS KP).

[m] Max:Min ratio: 1.2 (GT) and 2.4 (UNAIDS KP).

Table . Regional median Google Trends Population Size Estimate, United Nations Programme on HIV/AIDS (UNAIDS) Global AIDS Monitoring system (GAM), and key populations (KP) Atlas for men who have sex with men (MSM) populations for the year 2020.[a]

Region	Median regional percentage MSM population size estimation[b]		
	GT[c], %	UNAIDS GAM, %	UNAIDS KP Atlas, %
Eastern Africa	1.64	1.67	0.24
Southern Africa	1.68	1.67	0.83
West Central Africa	1.73	1.28	0.40
Asia	2.86	1.63	—[d]
Europe	2.86	2.11	1.47
Caribbean	3.60	2.71	3.17
Central & South America	5.26	3.37	2.54

[a] Absolute values are not provided as Google Trends does not provide absolute search frequency values.

[b] Google Trends (GT) and KP Atlas regional estimates only include estimates from included countries with available data (Table 1). UNAIDS GAM data separate regions differently and include countries that vary from our GT or the KP Atlas regional data: UNAIDS GAM includes eastern and southern Africa in 1 estimate and separates Asia and Europe into 2 estimates (1.63% for Asia and the Pacific, 2.11% for Eastern Europe and Central Asia). Region names were not adjusted in the above table to align with GAM data.

[c] GT: Google Trends.

[d] Not available.

Ethical Considerations

No ethics or review board approval or informed consent was obtained or applicable for this work. All data used in this paper are anonymous, aggregate, and publicly available and sourced.

GT-Based Population Size Estimation

GT provides results based on exact search terms, unlike the “topical” search results that Google’s main search engine provides. GT does not provide absolute search frequency values; instead, GT offers relative search volume (RSV) values across time (eg, 52 wk) in a specified space (eg, Kenya), ie, it normalizes search frequencies for specific search terms (eg, porn) to a range from 0 to 100, where a search term’s maximum frequency (for the specified geographic area and during the specified time frame) is set at 100 and 0 reflects no search for that term [11,13,14]. Importantly, GT allows users to add “comparator” terms (eg, gay porn) next to the main term (eg, porn); the RSV values for such comparator terms are normalized against the main term’s RSV values [13,21]. For the purpose of PSE calculation, the main term “porn” may represent all men whereas the comparator term “gay porn” may be viewed as a subset of men who represent the subpopulation of gay men or MSM. To generate an MSM PSE from the RSV values we divide the comparator RSV value (gay porn) by the larger same-time, same-place RSV value (porn).

National Size Estimates

PSE data collection was carried out through GT’s application [13]. We applied this analytic approach for the year 2020 using “porn” and “gay porn” as the main and comparator search terms for each of the 54 countries. The time period for data collection was set as the year 2020, the most recent year for which we could obtain all necessary data for this analysis. Weekly RSV values for “porn” and “gay porn” for the year 2020 were exported, summed, and proportional size estimates obtained. For example, for Botswana, the average of the weekly RSV values for “porn” was 78.3, the corresponding average for “gay porn” was 1.66 and the proportional PSE was therefore calculated as 1.66/78.3=2.1%. This was repeated for all countries. We then calculated the absolute Google Trends Population Size Estimate (GTPSE) by multiplying the proportional GTPSE by the total male population aged 15‐49 years in each country, the most used age range for KPs. The sizes for countries’ 15‐49 year-old male general population in 2020 were obtained through Spectrum (version 6.1, Avenir Health).

Local Size Estimates

GT data can be restricted to subnational areas. Separately from national estimates, for each country, we also attempted to obtain local GTPSE for the political (or, if different, commercial) capital city. Where data were unavailable for the political or commercial capital city, we used data from the district that contained the capital city. The calculation to obtain relative GTPSE was then the same as for the national level. We did not produce absolute subnational GTPSE.

Consistency of GTPSE Results With WHO-Recommended Minimum Estimate

We assessed whether the GTPSE results met the World Health Organization (WHO) and UNAIDS recommendation that national MSM PSE should represent at least 1% of the general adult male population [22,23].

Comparability

We compared the country-level GTPSE against 2 reference data sources used by UNAIDS: the KP Atlas database and the Global AIDS Monitoring system (GAM) [22,24,25]. The KP Atlas database stores countries’ self-reported absolute MSM size estimates using a wide range of PSE methods, often projected up to national scale from local estimates, with primary data collected over different periods of time. Proportional KP Atlas

PSE values were computed by dividing the absolute MSM PSE values from the KP Atlas over the male general population (15 ‐ 49 years). UNAIDS’ GAM is a global data warehousing system that informs policy and facilitates monitoring, including KP size estimates. Using GAM data, UNAIDS curated a table with regional relative MSM PSE (median and IQR) deemed reasonable.

Covariates Potentially Affecting GTPSE

Overview

We examined the potential effect of select covariates on the relative GTPSE values by performing regression analysis for each covariate. The country-specific covariates we examined included internet penetration, urbanization, stigma, and criminalization of homosexuality. The covariate data are provided in Table S1 in Multimedia Appendix 1; these data were not used to adjust GTPSE values.

Internet Penetration and Urbanization

Internet penetration data were extracted from the World Development Indicators database through the World Bank and the Internet World Statistics database, indicating the percentage of each country’s total population with access to the internet. Urbanization data were obtained from the World Development Indicators database through the World Bank, indicating the percent of the total population in each country considered urban [26,27].

Stigma

Country-level stigma values were extracted from the Global Acceptance Index [28]. This index was developed using computer modeling informed by responses to questions that measure attitudes toward lesbian, gay, bisexual, transgender, or intersex people from 11 different global surveys to create a stigma score in 175 countries toward lesbian, gay, bisexual, transgender, or intersex persons. The system scores countries on a scale of 1 to 10; higher scores indicate less stigma [28].

Criminalization

The State-Sponsored Homophobia International Lesbian, Gay, Bisexual, Trans, and Intersex Association report was used to evaluate the effects of criminalization of homosexual orientation or behavior on GTPSE [28]. The report classifies countries based on their level of legal protection or criminalization of sexual orientation and same-sex sexual acts. These classifications, ranging from most severe to most protected, include the death penalty, up to lifelong imprisonment, up to 8 years imprisonment, de facto criminalization, no criminalization or legal protections, limited protections, employment protections, broad protections, and constitutional protections. We converted these classifications into a quantitative ranking ranging from +4 to −4. The most severe classification (death penalty) was assigned the rank value “+4” and descended to the least severe/most protective classification (constitutional protection) with a rank value “−4.”

Sensitivity Analysis

Using a subset (n=14) of the 53 countries we performed 3 sensitivity analyses at the national level. The 14 countries were randomly selected among countries with prominent languages being French, Spanish, or Swahili. The first sensitivity analysis probed the effect of select non-English search languages. The 14 countries comprised 4 using Swahili (Kenya, Tanzania, Uganda, Democratic Republic of Congo [DRC]), 5 using French (Cote d’Ivoire, Senegal, Cameroon, Mali, Haiti), and 5 using Spanish (Dominican Republic, Panama, El Salvador, Nicaragua, Honduras) as their national/dominant language. We generated GTPSE using search terms in Swahili (“ngono” and “ngono za mashoga”), French (“porno” and “porno gay”), and Spanish (“porno” and “porno gay”) and compared them to the original relative GTPSE values. Using the same 14 countries, the second sensitivity analysis probed the effect of different search terms in English on GTPSE, that is, “sex,” ”gay sex” as well as “sex,” ”anal sex” and compared them to the original GTPSE (porn and gay porn). The third sensitivity analysis probed the effect of using different calendar years, ie, (2019 [pre-COVID] and 2021) and compared them to the original 2020 GTPSE values, using the original English language search terms.

Results

GTPSE and Comparability

Of the 54 countries examined, 1 (South Sudan), was omitted for lack of RSV values. All remaining 53 countries had GTPSE exceeding 1% (Table 1), similar to GAM values (all exceeding 1% as well) and compared with KP Atlas values where 24 out of 53 (45%) countries showed values above 1%. GTPSE ranged from 1.16% to 7.46% (median 1.99%, IQR 1.54% ‐ 3.48%), compared with 0.06% to 6.04% (median 0.99%, IQR 0.34 ‐ 1.93%) in the KP Atlas, and 1.38% to 2.82% in GAM regions. In 48 out of 53 (91%) countries, relative GTPSE exceeded estimates in the KP Atlas values; KP Atlas values were larger in 5 countries (DRC, Liberia, Lao People’s Democratic Republic [PDR], Thailand, and Jamaica). Absolute differences between GTPSE and KP Atlas ranged from −312,900 (Thailand) to 6,221,800 (India). Table 2 displays regional median GTPSE, ranging from 1.64% (East Africa) to 5.26% (Central/South America), larger in all regions than the corresponding KP Atlas values and largely similar to GAM values in most regions. Table 1 also displays the ratios between the largest and smallest country-level %PSE for each region, separately for GT and KP Atlas values. While substantial variability is seen in all regions and for both data sources (GT and KP Atlas), in all regions the observed heterogeneity was consistently higher for KP Atlas values compared with GT values.

Local GTPSE pertaining to political or commercial capitals or the larger sub-national areas encompassing these are displayed in Table 3. We could obtain local estimates for 51 out of 53 (96%) countries’ capital cities; GT did not provide data for Nairobi (Kenya) and Kathmandu (Nepal). Among the 51 cities with estimates, the GTPSE ranged from 0% to 13% (median 2.2%); most cities’ estimates (44/51, 86%) exceeded 1%. Five cities yielded noncredible GTPSE values of 0%, including Bujumbura (Burundi), Dodoma (encompassing Dar es Salaam, Tanzania), Ouagadougou (Burkina Faso), Monrovia (Liberia), and Vientiane (Laos PDR). Of the 44 subnational GTPSE with values >1%, 18 (41%) were below the same-country national GTPSE.

Table . Reported local men who have sex with men (MSM) Google Trends Population Size Estimate (GTPSE) (n=53) in the year 2020.[a]

Region and country		Local area[b]	Relative national GTPSE, %	Relative local GTPSE, %	Absolute percentage difference national and local GTPSE, %
East Africa					
	Burundi	Bujumbura	1.77	0	−1.77
	Ethiopia	Addis Ababa	1.28	1.30	0.02
	Kenya	Nairobi	1.99	—[c]	—
	Rwanda	Kigali	1.54	1.70	0.16
	Tanzania	Dodoma	1.73	0	−1.73
	Uganda	Kampala	1.47	1.60	0.13
Southern Africa					
	Angola	Luanda	1.44	2.04	0.60
	Botswana	Gaborone	2.12	0	−2.12
	Eswatini	Mbabane	1.57	1.48	−0.09
	Lesotho	Maseru	1.71	1.01	−0.70
	Malawi	Lilongwe	1.16	2.24	1.08
	Mozambique	Maputo	1.87	2.04	0.17
	Namibia	Windhoek	2.60	2.53	−0.07
	South Africa	Johannesburg (Gauteng)	2.46	0.99	−1.47
	Zambia	Lusaka	1.18	1.56	0.38
	Zimbabwe	Harare	1.64	1.56	−0.08
West Central Africa					
	Benin	Littoral (Cotonou)	1.18	4.11	2.93
	Burkina Faso	Centre (Ouagadougou)	1.80	0	−1.80
	Cameroon	Littoral (Douala)	2.29	2.47	0.18
	Cote d'Ivoire	Abidjan	2.68	1.01	−1.67
	DRC	Kinshasa	1.66	2.04	0.38
	Ghana	Accra	1.40	1.42	0.02
	Liberia	Monrovia	1.83	0	−1.83
	Mali	Bamako	1.55	2.93	1.38
	Nigeria	Abuja (Federal Capital Terriorty)	1.26	1.44	0.18
	Senegal	Dakar	1.94	2.85	0.91
	Sierra Leone	Freetown	1.36	1.01	−0.35
	Togo	Lome	2.65	2.04	−0.61
Asia					
	Burma	Yangon (Yangon Region)	4.53	4.79	0.26
	Cambodia	Phnom Penh	5.67	5.43	−0.24
	India	New Delhi (Uttar Pradesh)	1.18	1.15	−0.03
	Indonesia	Jakarta	1.61	2.20	0.59
	Kazakhstan	Almaty (Almaty Region)	2.99	5.52	2.53

Region and country		Local area[b]	Relative national GTPSE, %	Relative local GTPSE, %	Absolute percentage difference national and local GTPSE, %
	Kyrgyz Rep.	Bishkek	3.10	3.09	−0.01
	Lao PDR	Vientiane	2.73	0	−2.73
	Nepal	Katmandu/Kantipur	1.19	—	—
	PNG	Port Moresby	1.30	1.01	−0.29
	Tajikistan	Dushanbe	2.14	1.01	−1.13
	Thailand	Bangkok	1.25	3.24	1.99
	Philippines	Manila	4.27	5.51	1.24
	Viet Nam	Hanoi	7.46	4.56	−2.90
Europe					
	Ukraine	Kyiv	3.48	4.14	0.66
Caribbean					
	Dominican Rep.	Santo Domingo	4.26	3.99	−0.27
	Guyana	Georgetown	3.60	3.09	−0.51
	Haiti	Port-au-Prince	3.60	3.09	−0.51
	Jamaica	Kingston (St. Andrew Parish)	2.91	3.38	0.47
	Trinidad and Tobago	Port of Spain	3.04	13	9.96
Central and South America					
	Brazil	São Paulo (State of São Paulo)	5.18	5.92	0.74
	El Salvador	San Salvador	5.20	5.73	0.53
	Guatemala	Guatemala City (Guatemala Department)	5.09	4.91	−0.18
	Honduras	Tegucigalpa (Comayagua)	5.32	8.13	2.81
	Nicaragua	Managua	6.32	5.31	−1.01
	Panama	Panama City	7.23	7.44	0.21

[a] Absolute values are not provided as Google Trends does not provide absolute search frequency values.

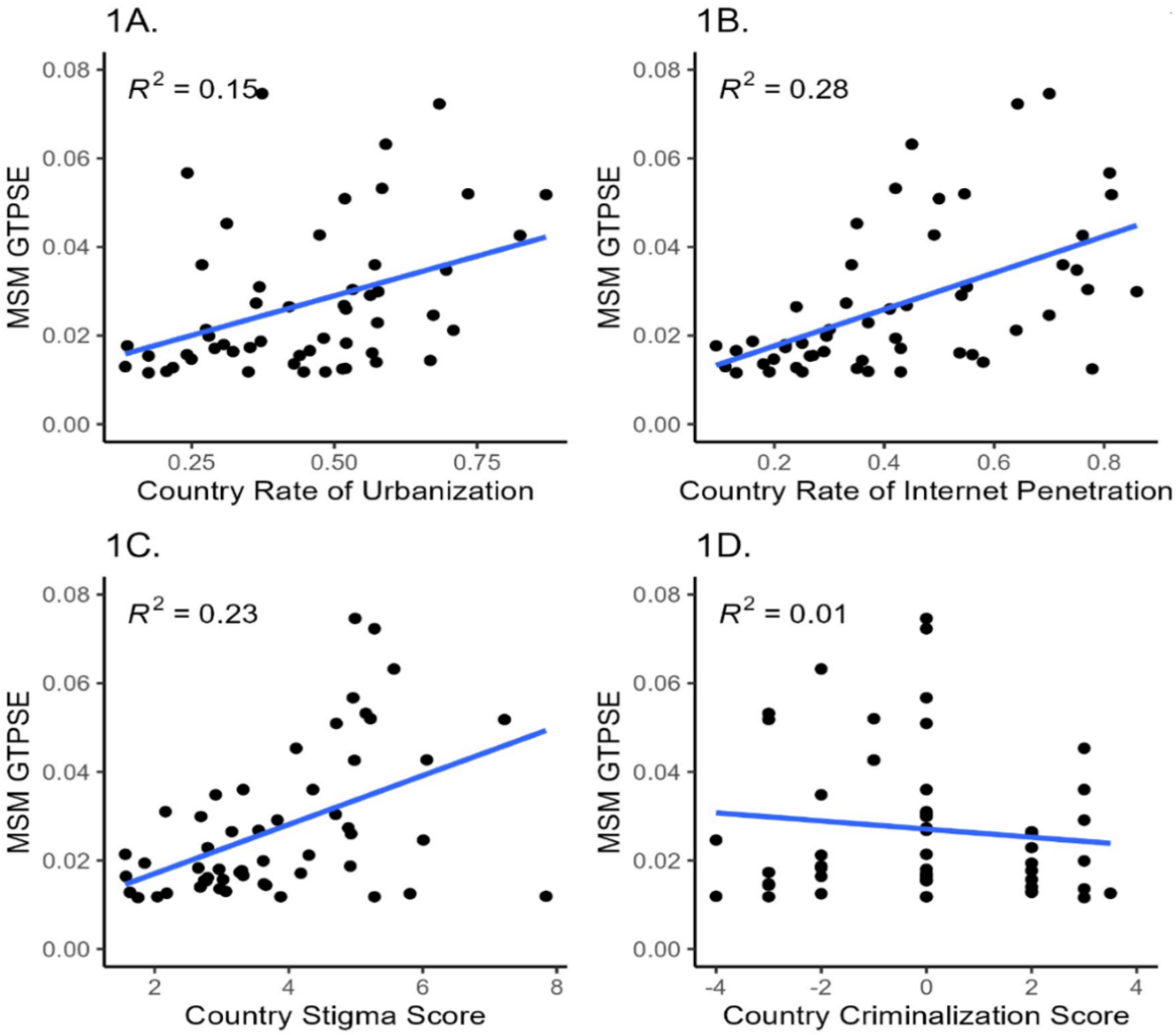
[b] Local MSM GTPSE for 53 countries for the year 2020 was calculated by restricting the geographic entity to the desired capital city or commercial hub. Where Google Trends (GT) did not provide data for a given city, we substituted the place name with the largest city by population or by district that had data available in GT. This is noted by listing what was available in GT in parenthesis next to the capital city. Kenya and Nepal were excluded from this analysis due to insufficient regional data available in GT.

[c] Not available (data missing for that country).

Effect of Covariates

Figure 1A-D displays the correlations between national-level GTPSE and urbanization, internet penetration, stigma, and criminalization. Coefficients ranged from 0.01 (criminalization) to 0.28 (internet penetration).

Figure 1. The linear relationship between the Google Trends national population size estimates and the rate of urbanization in each country (n=53). (A) The linear relationship between the Google Trends national population size estimates and the rate of urbanization in each country (n=53). (B) The linear relationship between the Google Trends national population size estimates and the rate of internet penetration in each country (n=53). (C) The linear relationship between the Google Trends national population size estimates and the level of stigma against LGBTQ+ persons in each country (n=53). (D) The linear relationship between the Google Trends national population size estimates and the degree of criminalization against men who have sex with men population in each country (n=53). LGBTQ+: lesbian, gay, bisexual, transgender, queer, and other identities; MSM: men who have sex with men; GTPSE: Google Trends Population Size Estimate.



Sensitivity Analysis

Table 4 displays how the GTPSE generated from the alternative search terms compares to the original search term GTPSE. In most countries “Porn/Gay Porn” produced higher PSE values compared with “sex/anal sex” (13/14, 93%) as well as compared with “sex/gay sex” (12/14, 86%). For “sex/gay sex,” all 14 countries produced estimates exceeding 1%. For “sex/anal sex”, 3 out of 14 (21%) countries did not produce estimates reaching the 1% threshold, including Mali for which zero search results were reported for “anal sex.”

Table . Sensitivity analysis using alternative search terms in Google Trends to calculate national population size estimations (PSEs) for select US President's Emergency Plan for AIDS Relief countries (n=53) in 2020.[a]

	Original GTPSE[b]	SA alternate search term GTPSE[c]			
	Porn/gay porn PSE	Sex/gay sex PSE	Absolute percentage difference	Sex/anal sex PSE	Absolute percentage difference
Country, %					
Kenya	1.99	1.37	0.62	1.37	0.62
Tanzania	1.73	1.46	0.27	3.54	−1.81
Uganda	1.47	1.38	0.09	1.26	0.21
DRC	1.66	1.55	0.11	1.15	0.51
Cameroon	2.29	1.28	1.01	0.90	1.39
Mali	1.55	1.65	−0.10	0	1.55
Cote d'Ivoire	2.68	1.90	0.78	1.74	0.94
Senegal	1.94	1.50	0.44	0.88	1.06
Haiti	3.60	2.60	1	2.83	0.77
Dominican Rep.	4.26	3.36	0.90	1.83	2.43
Panama	7.23	5.17	2.06	3.71	3.52
El Salvador	5.20	5.34	−0.14	4.19	1.01
Nicaragua	6.32	7.10	−0.78	4.82	1.50
Honduras	5.32	4.85	0.47	2.96	2.36
Median (IQR)	2.49 (1.78-4.97)	1.78 (1.47-4.48)	0.45 (0.10-0.87)	1.79 (1.18-3.40)	1.03 (0.66-1.54)

[a] Absolute values are not provided as Google Trends does not provide absolute search frequency values.

[b] GTPSE: Google Trends Population Size Estimate.

[c] Alternative search terms were chosen based on words that represented the general male population and men who have sex with men subset population in each country (n=53) in the year 2020.

Table 5 shows how GTPSE was generated using alternative language terms compared with the original GT search terms. For Swahili, only 1 country yielded a PSE in that language. All countries using French (n=5), or Spanish (n=5) search terms yielded estimates, all exceeding 1%. All alternative language estimates were lower than the original "porn/gay porn" PSE values.

Table . Sensitivity analysis using alternate national language searches in Google Trends to calculate national population size estimation for select US President's Emergency Plan for AIDS Relief countries (n=14) in 2020.[a]

Language and country		Original GTPSE[b] (English), %	Alternate language term GTPSE, %[c]	Absolute percentage difference, %
Swahili				
	Kenya	1.99	0	1.99
	Tanzania	1.73	0.52	1.21
	Uganda	1.47	0	1.47
	DRC	1.66	0	1.66
French				
	Cameroon	2.29	1.36	0.93
	Mali	1.55	1.07	0.48
	Cote d'Ivoire	2.68	1.35	1.33
	Senegal	1.94	1.28	0.66
	Haiti	3.60	2.23	1.37
Spanish				
	Dominican Rep.	4.26	2.56	1.70
	Panama	7.23	5.14	2.09
	El Salvador	5.20	4.36	0.84
	Nicaragua	6.32	4.13	2.19
	Honduras	5.32	4.07	1.25

[a] Absolute values are not provided as Google Trends does not provide absolute search frequency values.

[b] GTPSE: Google Trends Population Size Estimate.

[c] Alternative language search terms included "ngono/ngono za mashoga" (Swahili), "porno/porno gay" (French), "porno/porno gay" (Spanish).

Table 6 displays how GTPSE generated for alternative years (2019 and 2021) compared with the original 2020 GT searches. All 14 countries in both years produced estimates exceeding 1%. No large discrepancies in PSE between the years were observed; 13 out of 14 in 2019 values were larger than the 2020 values whereas the 2021 values were largely similar to the 2020 values.

Table . Sensitivity analysis for men who have sex with men population size estimates for select US President's Emergency Plan for AIDS Relief supported countries (n=14) using Google Trends data in years 2019 and 2021 compared with the year 2020.[a, b]

	2019 PSE[c], %	2020 PSE, %	2021 PSE, %
Kenya	2.37	1.99	1.99
Tanzania	1.96	1.73	1.85
Uganda	1.73	1.47	1.69
DRC	1.95	1.66	1.60
Cameroon	2.70	2.29	2.25
Mali	2.30	1.55	2.17
Cote d'Ivoire	2.52	2.68	2.23
Senegal	2.54	1.94	1.90
Haiti	4.33	3.60	2.92
Dominican Republic	4.91	4.26	4.34
Panama	9.36	7.23	6.74
El Salvador	6.19	5.20	4.77
Nicaragua	7.31	6.32	4.93
Honduras	6.79	5.32	5.51

[a] 2019 and 2021 values were computed in the same way as the reference 2020 estimates.

[b] Absolute values are not provided as Google Trends does not provide absolute search frequency values.

[c] PSE: population size estimation.

Discussion

Principal Findings

Our analysis suggests that national-level MSM GTPSE is feasible in almost all countries. Importantly, all estimates appeared plausible, that is, they exceeded the WHO/UNAIDS suggested minimum threshold of 1%. Heterogeneity of GTPSE across same-region countries was pronounced within all regions yet smaller than the ratios based on the UNAIDS KP Atlas values which contained numerous PSE values well below the 1% threshold.

Our analysis draws on several strengths. We successfully applied the GTPSE method to many low and middle-income countries, suggesting that GTPSE appears to have wide geographic applicability. We compared the values against 2 PSE data sources at UNAIDS, assessed the potential effect of various covariates on GTPSE values, and conducted a sensitivity analysis with varying English search terms, non-English search languages, and different calendar years. Google is the dominant search engine in all countries covered in this analysis, with a market share ranging between 84% and 99% (data shown in Table S1 in Multimedia Appendix 1) [27]. Although no absolute search volume data were available to us, searches for "porn" globally were among the top 20 search terms in 2023 with about 65 million searches globally each month according to one source [29] although this is still well behind the largest porn site-specific searches. GTPSE may emerge as another example of digital public health and epidemiology that includes real-time surveillance of disease outbreaks [30], assessing the impact of global public health days [31], informing health and health policy research [32], or understanding spatiotemporal patterns of dry eye disease [33].

While most local estimates were plausible (>1%), 14% (n=7) did not reach the WHO/UNAIDS minimum threshold, and 2 more locations did not produce a GTPSE value at all due to lack of GT data and how GT organized the subnational data despite some of the affected cities' large population sizes. This is not an uncommon finding, as other PSE methods in active use typically do not meet the WHO/UNAIDs minimum threshold. For a few other country or commercial capital cities with no direct GT data available, such as Johannesburg (South Africa), we could obtain a subnational estimate using the larger district or province within which the city (eg, Johannesburg and Pretoria) are located. This may limit the utility and comparability of such local estimates. About one-third of the local (relative) estimates did not reach or exceed the same country national level estimates, somewhat contrary to our expectation that rural-to-urban migration among MSM may be more pronounced than that of other men and so yielding higher GTPSE values [9]. In Card et al's [12] study on Canadian towns and cities the estimates ranged from 2% to 4% compared with 0% to 13% among the local estimates, whereas the Canadian national estimate was 2.8% compared with 1.2% - 7.5% across all countries we examined. While not a limitation, it is worth noting that weekly RSV data varied widely (data not shown), confirming the recommendation to use GT data for size estimation only over longer time periods, such as a full calendar year.

Limitations

Like most PSE methods, GTPSE has limitations. In particular, the assumptions underlying the GTPSE method deserve close scrutiny: straight men only search for porn, MSM only search for gay porn, MSM and straight men search for (gay) porn in equal proportions, and women do not search for (gay) porn at all or do not affect the generated GTPSE for MSM. Violations of these assumptions will result in bias if they affect RSV for porn and gay porn to differing extents, hence altering the proportion of porn searches that are directed at gay porn. While the literature from LMIC settings on this topic is very sparse, reports and literature from high-income settings suggest that gay porn is also consumed by heterosexual men and women, suggesting that some bias may be present. Complicating speculations about the magnitude and direction of bias is the fact that specific porn websites' user statistics may not accurately reflect searches for (gay) porn on Google. Women's search behavior on Google regarding gay porn may increase or decrease the GTPSE estimates depending on the frequency relative to searches for just porn.

Regrettably, Google does not provide access to its algorithm generating the RSV data nor can users filter GT searches by age or gender. An inherent limitation in using GT data includes the lack of deduplication in the search data (although repeated searches by the same user within a short time period are not counted multiple times by Google) and the lack of absolute search volume data. Not having access to the absolute search volume data impedes the computation of uncertainty intervals (which in most national settings may be expected to be small due to the large search volumes involved). However, absolute search volume information may eventually be made available by Google and is already offered to some extent by select third-party companies. Absolute search volume data may also inform the choice of search language and even search terms and may facilitate composite GTPSE metrics by incorporating multiple GTPSE metrics stemming from different language search terms. Restricting GTPSE-relevant data to male users may further refine GTPSE values by excluding female users, a limitation our analysis could not overcome. VPN (virtual private network) also has the potential to introduce errors if users select a country other than their place of residence. The adoption of VPN may vary considerably across time and by country, and, among US President's Emergency Plan for AIDS Relief countries. According to one industry website in 2020, VPN was highest in Ukraine (7.9%) and lowest in Kenya (0.5%) [34]. Taken together, these limitations constitute a major source of uncertainty about the bias and precision of GTPSE. For that reason, GPTSE should be regarded as an approximate reference value. Clearly, they do not attain the rigor or transparency of statistically principled estimation from accurately measured data, which the currently best available PSE methods do offer. Additionally, GTPSE may not be feasible for a few countries, perhaps due to poor or little data availability on search terms and frequencies.

GTPSE seems infeasible for size estimation among transgender persons, sex workers, or people who inject drugs. Unlike (gay) porn, where the search is about a web-based product (visual depictions of porn), searches for sex work or clients, transgenderism, or injecting drug use are not directly tied to the internet, and may exhibit a more variable search terminology, and may lack fitting "denominator" search terms (analogous to "porn").

Overall, the GTPSEs often were substantially higher than the KP Atlas estimates but were more closely aligned with the reported GAM regional estimates. The KP Atlas estimates are based on a broad range of PSE methods typically generating local PSE that may or may not be projected to national scale, or summed or averaged across multiple localities, and may refer to various time points (calendar years) and various age ranges. Many KP Atlas based MSM PSE were implausibly low (<1%), suggesting that substantial differences to GTPSE may often be due to KP Atlas underestimates. The regional GAM estimates are based on a more curated database of PSE after excluding estimates with subpar quality and hence of perhaps more trustworthy quality [22]. However, GAM regions do not exactly overlap with the regions we used for GTPSE and the KP Atlas estimates.

The national MSM GTPSE values were robust against varying levels of urbanization, internet penetration, stigma, and criminalization or protection of homosexuality, negating the need for adjustment and increasing comparability across different settings. The largest influence was seen with internet penetration which can be expected to increase over time. In the sensitivity analysis, the largest differences to the original GTPSE values were seen using alternate English language search terms. Among the 14 examined countries, almost half (43%) of the alternate estimates were below 1% and hence considered implausibly low. This indicates that search term selection is important, especially for comparison across time and space. Further exploration may be warranted to evaluate if country or region-specific English or non-English slang terms may produce plausible estimates; however, the limited sensitivity analysis suggests that "Porn/Gay Porn" may be dependable and consistently produces plausible values. The use of similar search terms in French, Spanish, and Swahili yielded universally lower results; Swahili, not a nationally dominant language in most countries, appears particularly unsuitable as it frequently produced 0% PSE values. As most countries display prominent non-English language use, countries may want to consider using the predominant language (used for web searches) when applying this method while considering any language's geographic scope in-country. The results also appeared robust across time (two years affected by the COVID pandemic plus 1-year pre-COVID) as the 2 adjacent years produced plausible and (same country) consistent results. The lack of uncertainty intervals however impeded a more meaningful interpretation of the results from the sensitivity analyses.

Conclusions

Generating national-level PSEs for KPs is challenging for many countries. GTPSE is a simple method with the potential to address this problem efficiently without the need of additional resources. However, the lack of validation of key assumptions and the inability to generate credibility intervals suggest important uncertainty regarding the accuracy and precision of the estimates. Additional research, such as expanding or building on our sensitivity and covariate analysis, to address or better understand these limitations may further improve the quality and utility of GTPSE for MSM in LMICs.

Acknowledgments

We wish to acknowledge all authors who contributed to this paper. WH and AAQ contributed to the study concept and design. All authors substantially contributed to the study's conduct, or to data analysis and interpretation. All authors contributed to writing, reviewing, or editing the manuscript. We also wish to acknowledge all colleagues within the Center for Disease Control and Prevention who offered insights and alternative perspectives during the analysis and writing of this paper.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the funding agencies. This research publication has been supported by the President's Emergency Plan for AIDS Relief (PEPFAR) through the Centers for Disease Control and Prevention (CDC) under the terms of project number 0900f3eb820619c0. This publication was also supported by Cooperative Agreement number NU2GGH002093 between the Centers for Disease Control and Prevention and the Public Health Institute.

Data Availability

All data generated or analyzed during this study are publicly accessible data (with the exception that only aggregate data can be obtained through Google Trends). Links to where the data can be accessed are included in this published article and its supplementary information files. Further, all relative search volume data points can be reproduced through Google Trends.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary Table 1.
[XLSX File, 18 KB - publichealth_v11i1e58630_app1.xlsx]

References

1. Global HIV & AIDS statistics — fact sheet. UNAIDS. URL: https://www.unaids.org/en/resources/fact-sheet [accessed 2024-12-31]
2. Indicators and questions for monitoring progress on the 2021 Political Declaration on HIV and AIDS — Global AIDS Monitoring 2025. UNAIDS. 2024. URL: https://www.unaids.org/en/resources/documents/2024/global-aids-monitoring-guidelines [accessed 2024-12-31]
3. Neal JJ, Prybylski D, Sanchez T, Hladik W. Population size estimation methods: searching for the holy grail. JMIR Public Health Surveill 2020 Dec 3;6(4):e25076. [doi: 10.2196/25076] [Medline: 33270035]
4. Stone J, Mukandavire C, Boily MC, et al. Estimating the contribution of key populations towards HIV transmission in South Africa. J Int AIDS Soc 2021 Jan;24(1):e25650. [doi: 10.1002/jia2.25650] [Medline: 33533115]
5. Viswasam N, Lyons CE, MacAllister J, et al. The uptake of population size estimation studies for key populations in guiding HIV responses on the African continent. PLoS One 2020;15(2):e0228634. [doi: 10.1371/journal.pone.0228634] [Medline: 32101551]
6. Xu C, Jing F, Lu Y, et al. Summarizing methods for estimating population size for key populations: a global scoping review for human immunodeficiency virus research. AIDS Res Ther 2022 Feb 19;19(1):9. [doi: 10.1186/s12981-022-00434-7] [Medline: 35183203]
7. Teo AKJ, Prem K, Chen MIC, et al. Estimating the size of key populations for HIV in Singapore using the network scale-up method. Sex Transm Infect 2019 Dec;95(8):602-607. [doi: 10.1136/sextrans-2018-053747] [Medline: 31073092]
8. Sabin K, Zhao J, Garcia Calleja JM, et al. Availability and quality of size estimations of female sex workers, men who have sex with men, people who inject drugs and transgender women in low- and middle-income countries. PLOS ONE 2016;11(5):e0155150. [doi: 10.1371/journal.pone.0155150] [Medline: 27163256]
9. Edwards JK, Hileman S, Donastorg Y, et al. Estimating sizes of key populations at the national level: considerations for study design and analysis. Epidemiology 2018 Nov;29(6):795-803. [doi: 10.1097/EDE.0000000000000906] [Medline: 30119057]
10. Bozicevic I, Sharifi H, Haghdoost A, Sabry A, Hermez J. Availability of HIV surveillance data in key populations in the countries of the World Health Organization Eastern Mediterranean Region. Int J Infect Dis 2022 Aug;121:211-216. [doi: 10.1016/j.ijid.2022.05.003] [Medline: 35533830]
11. Mavragani A, Ochoa G. Google Trends in infodemiology and infoveillance: methodology framework. JMIR Public Health Surveill 2019 May 29;5(2):e13439. [doi: 10.2196/13439] [Medline: 31144671]

12. Card KG, Lachowsky NJ, Hogg RS. Using Google Trends to inform the population size estimation and spatial distribution of gay, bisexual, and other men who have sex with men: proof-of-concept study. JMIR Public Health Surveill 2021 Nov 29;7(11):e27385. [doi: 10.2196/27385] [Medline: 34618679]
13. Porn, gay porn - explore. Google Trends. URL: https://trends.google.com/trends/explore?date=2020-01-01%202020-12-31&geo=DZ&q=Porn,Gay%20Porn [accessed 2024-12-31]
14. FAQ about google trends data - trends help. Google. URL: https://support.google.com/trends/answer/4365533?hl=en [accessed 2024-12-31]
15. How many women watch porn? Fight the New Drug. URL: https://fightthenewdrug.org/how-do-men-and-womens-porn-site-searches-differ [accessed 2024-12-31]
16. Pornhub stats reveal women search for more hardcore genres than you might expect. Fight the New Drug. URL: https://fightthenewdrug.org/data-reveals-women-are-searching-hardcore-genres [accessed 2024-12-31]
17. When no one is looking, many women are watching gay porn. NBC News. URL: https://www.nbcnews.com/feature/nbc-out/when-no-one-looking-many-women-are-watching-gay-porn-n894266 [accessed 2024-12-31]
18. Solano I, Eaton NR, O'Leary KD. Pornography consumption, modality and function in a large internet sample. J Sex Res 2020 Jan 2;57(1):92-103. [doi: 10.1080/00224499.2018.1532488]
19. Træen B, Daneback K. The use of pornography and sexual behaviour among Norwegian men and women of differing sexual orientation. Sexologies 2013 Apr;22(2):e41-e48. [doi: 10.1016/j.sexol.2012.03.001]
20. PEPFAR 2022 country and regional operational plan (COP/ROP) guidance for all PEPFAR-supported countries. US Department of State.: PEPFAR; 2022. URL: https://www.state.gov/wp-content/uploads/2022/02/COP22-Guidance-Final_508-Compliant-3.pdf [accessed 2025-01-02]
21. Dangerous inequalities: world AIDS day report 2022. UNAIDS. 2022. URL: https://www.unaids.org/sites/default/files/media_asset/dangerous-inequalities_en.pdf [accessed 2025-01-02]
22. Quick start guide for Spectrum. UNAIDS. 2020. URL: https://www.unaids.org/sites/default/files/media_asset/QuickStartGuide_Spectrum_en.pdf [accessed 2025-01-02]
23. World Health Organization. Recommended population size estimates of men who have sex with men. UNAIDS. 2020. URL: https://www.unaids.org/sites/default/files/media_asset/2020-recommended-population-size-estimates-of-men-who-have-sex-with-men_en.pdf [accessed 2025-01-02]
24. The Key Populations Atlas. UNAIDS. 2022. URL: https://kpatlas.unaids.org/dashboard [accessed 2024-12-31]
25. Key Populations Atlas data sources. UNAIDS - Key Populations Atlas. 2017. URL: https://kpatlas.unaids.org/document/kp_data_sources.pdf [accessed 2025-01-02]
26. Internet world stats - usage and population statistics. ResearchGate. URL: https://www.researchgate.net/publication/258847361_Internet_World_Stats_-_Usage_and_Population_Statistics [accessed 2025-01-02]
27. World development indicators. World Bank Group. URL: https://databank.worldbank.org/source/world-development-indicators [accessed 2024-12-31]
28. Mendos LR, Botha K, Lelis RC, de la Peña EL, Savelev I, Tan D. State-Sponsored Homophobia report. ILGA World. 2020. URL: https://ilga.org/state-sponsored-homophobia-report/ [accessed 2025-01-02]
29. Most searched things on Google in 2024. SimilarWeb. URL: https://www.similarweb.com/blog/marketing/seo/top-keywords/ [accessed 2024-12-31]
30. Carneiro HA, Mylonakis E. Google trends: a web-based tool for real-time surveillance of disease outbreaks. Clin Infect Dis 2009 Nov 15;49(10):1557-1564. [doi: 10.1086/630200] [Medline: 19845471]
31. Havelka EM, Mallen CD, Shepherd TA. Using Google Trends to assess the impact of global public health days on online health information seeking behaviour in Central and South America. J Glob Health 2020 Jun;10(1):010403. [doi: 10.7189/jogh.10.010403] [Medline: 32373327]
32. Arora VS, McKee M, Stuckler D. Google Trends: opportunities and limitations in health and health policy research. Health Policy 2019 Mar;123(3):338-341. [doi: 10.1016/j.healthpol.2019.01.001] [Medline: 30660346]
33. Azzam DB, Nag N, Tran J, et al. A novel epidemiological approach to geographically mapping population dry eye disease in the United States through Google Trends. Cornea 2021 Mar 1;40(3):282-291. [doi: 10.1097/ICO.0000000000002579] [Medline: 33177410]
34. NordVPN survey reveals: users still trust free vpns. NordVPN. URL: https://nordvpn.com/blog/nordvpn-usage-survey/ [accessed 2025-01-02]

Abbreviations

GAM: Global AIDS Monitoring system
GT: Google Trend
GTPSE: Google Trend Population Size Estimate
KP: key population
LMIC: low- and middle-income country
MSM: men who have sex with men

PSE:	Population size estimation
PSE:	population size estimation
RSV:	relative search volume
SSA:	sub-Saharan Africa
UNAIDS:	United Nations Programme on HIV/AIDS
VPN:	virtual private network
WHO:	World Health Organization

Edited by A Mavragani; submitted 20.03.24; peer-reviewed by E Tuyishime, K Card; revised version received 28.05.24; accepted 28.05.24; published 09.01.25.

Please cite as:
Malburg CM, Gutreuter S, Ruiseñor-Escudero H, Abdul-Quader A, Hladik W
Population Size Estimation of Men Who Have Sex With Men in Low- and Middle-Income Countries: Google Trends Analysis
JMIR Public Health Surveill 2025;11:e58630
URL: https://publichealth.jmir.org/2025/1/e58630
doi:10.2196/58630

Prevalence and Economic Impact of Acute Respiratory Failure in the Prehospital Emergency Medical Service of the Madrid Community: Retrospective Cohort Study

Ana María Cintora-Sanz[1], MRes, RN; Cristina Horrillo-García[2], MD, PhD; Víctor Quesada-Cubo[3], MD; Ana María Pérez-Alonso[4], MD; Alicia Gutiérrez-Misis[5], MD, PhD

[1]Servicio de Urgencias Médicas de Madrid (SUMMA112), Universidad Autónoma de Madrid, Madrid, Spain
[2]Servicio de Urgencias Médicas de Madrid (SUMMA112), Ventilation Commission, Calle Antracita S/N, Madrid, Spain
[3]Technical Health Research, Madrid, Spain
[4]Servicio de Urgencias Médicas de Madrid (SUMMA112), Research Commission, Madrid, Spain
[5]Department of Medicine, Division of Family Medicine and Primary Care, Clinical Simulation Laboratory, Faculty of Medicine, Universidad Autónoma de Madrid, Madrid, Spain

Corresponding Author:
Cristina Horrillo-García, MD, PhD
Servicio de Urgencias Médicas de Madrid (SUMMA112), Ventilation Commission, Calle Antracita S/N, Madrid, Spain

Abstract

Background: Chronic obstructive pulmonary disease (COPD), congestive heart failure (CHF), and acute pulmonary edema (APE) are serious illnesses that often require acute care from prehospital emergency medical services (EMSs). These respiratory diseases that cause acute respiratory failure (ARF) are one of the main reasons for hospitalization and death, generating high health care costs. The prevalence of the main respiratory diseases treated in a prehospital environment in the prepandemic period and during the COVID-19 pandemic in Spain is unknown. The Madrid Community EMS is a public service that serves all types of populations and represents an epidemiological reference for supporting a population of 6.4 million inhabitants. The high volume of patients treated by Madrid's medical advanced life supports (ALSs) allows us to analyze this little-studied problem.

Objectives: Our goal was to lay the groundwork for comprehensive data collection and surveillance of respiratory failure, with an emphasis on the most prevalent diseases that cause it, an aspect that has been largely overlooked in previous initiatives. By achieving these objectives, we hope to inform efforts to address respiratory failure and establish a standardized methodology and framework that can facilitate expansion to a continuous community-wide registry in Madrid, driving advances in emergency care and care practices in these pathologies. The aim of this retrospective observational study was to determine the pathologies that have mainly caused respiratory failure in patients and required medicalized ALS and to evaluate the cost of care for these pathologies collected through this pilot registry.

Methods: A multicenter descriptive study was carried out in the Madrid Community EMS. The anonymized medical records of patients treated with medical ALS, who received any of the following medical diagnoses, were extracted: ARF not related to chronic respiratory disease, ARF in chronic respiratory failure, exacerbations of COPD, APE, CHF, and bronchospasm (not from asthma or COPD). The prevalence of each pathology, its evolution from 2014 to 2020, and the economic impact of the Medical ALSs were calculated.

Results: The study included 96,221 patients. The most common pathology was exacerbation of COPD, with a prevalence of 0.07% in 2014; it decreased to 0.03% in 2020. CHF followed at 0.06% in 2014 and 0.03% in 2020. APE had a prevalence of 0.01% in 2014, decreasing to 0.005% in 2020 with the pandemic. The greatest economic impact was on exacerbation of COPD in 2015, with an annual cost of €2,726,893 (which equals to US $2,864,628).

Conclusions: COPD exacerbations had the higher prevalence in the Madrid region among the respiratory diseases studied. With the COVID-19 pandemic, the prevalence and costs of almost all these diseases decreased, except for ARF not related to chronic disease. The cost of these pathologies over 5 years was €58,791,031 (US $61,832,879).

(JMIR Public Health Surveill 2025;11:e66179) doi:10.2196/66179

KEYWORDS

acute respiratory failure; COVID-19; chronic obstructive respiratory insufficiency; congestive heart failure; bronchospasm; emergency medical services costs; ambulances; SARS-CoV-2; coronavirus; respiratory; pulmonary; pandemic; economic impact;

observational; Madrid; community; medical records; health records; medical advanced life support; ALS; acute pulmonary edema; chronic obstructive pulmonary disease; COPD; prevalence

Introduction

Chronic obstructive pulmonary disease (COPD) is a public health epidemic. The current state of the art displays a high prevalence and high burden [1]. COPD is the fourth leading cause of death worldwide, accounting for 3.5 million deaths in 2021, approximately 5% of all global deaths [2]. Although systematic reviews on the prevalence of COPD at the hospital level have been published previously [3,4], the literature on current prevalence estimates of COPD exacerbations, treated in prehospital emergency medical services (EMSs), is sparse. In particular, there are few published data on differences in COPD prevalence between rural and urban areas and between countries [5-7]. In this context, relevant and timely information on COPD prevalence in this context is essential to inform, develop, and implement context-appropriate policies and programs for the prevention and optimal control of COPD exacerbations.

Congestive heart failure (CHF) is another highly prevalent disease with significant morbidity and mortality worldwide [8]. There are geographical variations in the epidemiology of CHF [9]. There is a substantial lack of data from the out-of-hospital setting as well, where CHF has different evolution depending on the acute treatment administered [10,11]. Acute pulmonary edema is associated with high mortality. It requires emergency management and usually admission to hospital [7,12,13]. Prospective studies have shown that early treatment applied by prehospital EMS improves short time outcomes and lower hospital readmissions, but the population susceptible to early treatment is unknown [14,15].

COPD, CHF, and acute pulmonary edema (APE) are serious illnesses that often require acute care from EMS [5,6,16]. Respiratory diseases that cause acute respiratory failure (ARF) are one of the main reasons for hospitalization [7,17], generating high health care costs [18,19]. These diseases include bronchospasm as well [20]. Studies on the prevalence and expenditures of prehospital EMSs are scarce [5,21-23]. Consequently, more precise epidemiological data are needed to address these issues and obtain a more accurate picture of the amount of care and the costs generated by these pathologies [24,25] in the prehospital EMS setting.

The main objective of the Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA) project is to determine the prevalence of the most frequent respiratory diseases assisted by the Autonomous Community of Madrid's (CAM) advanced life supports (ALSs). The secondary objective is to measure the EMS economic cost of health care for patients. We present the following article in accordance with the STROBE (Strengthening the Reporting of Observational studies in Epidemiology) reporting checklist (Multimedia Appendix 1).

Methods

Study Design

The SUMMIRA study is an observational, quantitative, multicenter, and cross-sectional study of patients attended by the EMS of the CAM for 6 years.

Sample Size

Between 2014 and 2020, 583,984 persons of the population of the Madrid Community were clinically treated by the CAM ALSs, after a triage in the 112 health emergency telephone number. Based on these medical examinations, the subjects were included in the study. A cohort of 22,085 subjects was identified as having decompensated COPD; 22,085 with a main diagnosis of CHF; 4676 with APE; 17,843 presented ARF not related to chronic respiratory failure (CRF); 6921 were diagnosticated with bronchospasm; and 17,851 were related with ARF in CRF. This study involved a total of 96,221 patients included from the electronic medical records (EMRs).

Study Setting

The CAM has a population of 6.4 million inhabitants from 2014, both urban and rural areas. The CAM has 47 ALSs [26]. The ALSs are responsible for (1) emergency care, namely, patient care at home or on public roads in the Madrid Community and (2) interhospital transfers. The CAM is also in charge of the Helicopter Emergency Medical Service.

Each emergency care is provided by a health professional team formed by an emergency physician, an emergency nurse, and 1 or 2 emergency technicians; these teams assist at home and other places and take care of the patient in the mobile intensive critical unit (ALS).

The EMRs analyzed are those performed by the ALS physicians who attended the pathologies under study. These are the patients attended by the ALS of the Community of Madrid after the patients or their families called 112 (both in the city and in the outskirts).

The EMRs are located in a program named Integrated Management System of the Coordinating Center [27]. Integrated Management System of the Coordinating Center is a highly secure system of clinical data that has received urgent health care from the ALSs of the CAM.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki. This study was reviewed and approved by the Madrid Regional Ethics Committee on Medication Research (dated: June 6, 2022/N° V8.2022). All data used in this study have been deidentified or anonymized to protect the privacy of participants. No information is available that would allow for the direct or indirect identification of individuals. The Madrid Regional Ethics Committee approval covers secondary analysis without additional consent.

Since the study involved the secondary analysis of previously collected data, the ethics committee determined that the project was exempt from further review under the category of anonymized, preexisting data analysis. The study complied fully with international ethical standards and the protection of participants' rights.

All EMRs in this study have been anonymized to protect the privacy of participants. Rigorous measures were taken to ensure the confidentiality of the data and the integrity of the information collected complying, therefore, with the principles regarding data processing referred to in Article 5 of European Union Regulation 2016/679 of the European Parliament and Council of April 27, 2016 (General Data Protection Regulation).

Sampling and Recruitment Procedure

Inclusion Criteria

Patients with EMRs were treated by the CAM ALSs who, according to usual clinical practice, had been diagnosed with ARF not related to CRF, exacerbation of chronic obstructive pulmonary disease (ECOPD), CHF, bronchospasm, or APE according to the codes of the *International Classification of Diseases, Ninth Revision* [9] between 2014 and 2020. Medical diagnostic criteria based on clinical and historical data were followed. The population residing in the CAM per year was taken as the reference population, according to the Spanish National Institute of Statistics [28].

The diagnoses of APE and ECOPD were defined as a "clinical picture" in the prehospital and hospital emergency manuals [29-31], with consensus documents [32,33] and systematic review such as by Bello et al [34] . The diagnostic criteria for COPD and APE are described in Multimedia Appendix 2 and Table 1. The other pathologies were diagnosed following the diagnostic criteria of the 12 Octubre Physician Manual [35] in accordance with another prehospital emergency manual [36] and the Servicio de Urgencias Medicas de Madrid (SUMMA112; Prehospital Emergency Medical Service of Autonomous Community of Madrid) Physician Clinical Guide for Acute Heart Insufficiency (Multimedia Appendix 3).

Table . Prevalence of congestive heart failure, exacerbation of chronic obstructive pulmonary disease (ECOPD), acute pulmonary edema, and bronchospasm acute respiratory insufficiency not related to chronic respiratory failure, of Madrid region population, based to the Madrid Community census [37], and the prehospital emergency medical service´s electronic medical record of Madrid from 2014 to 2020 (n=583,984).

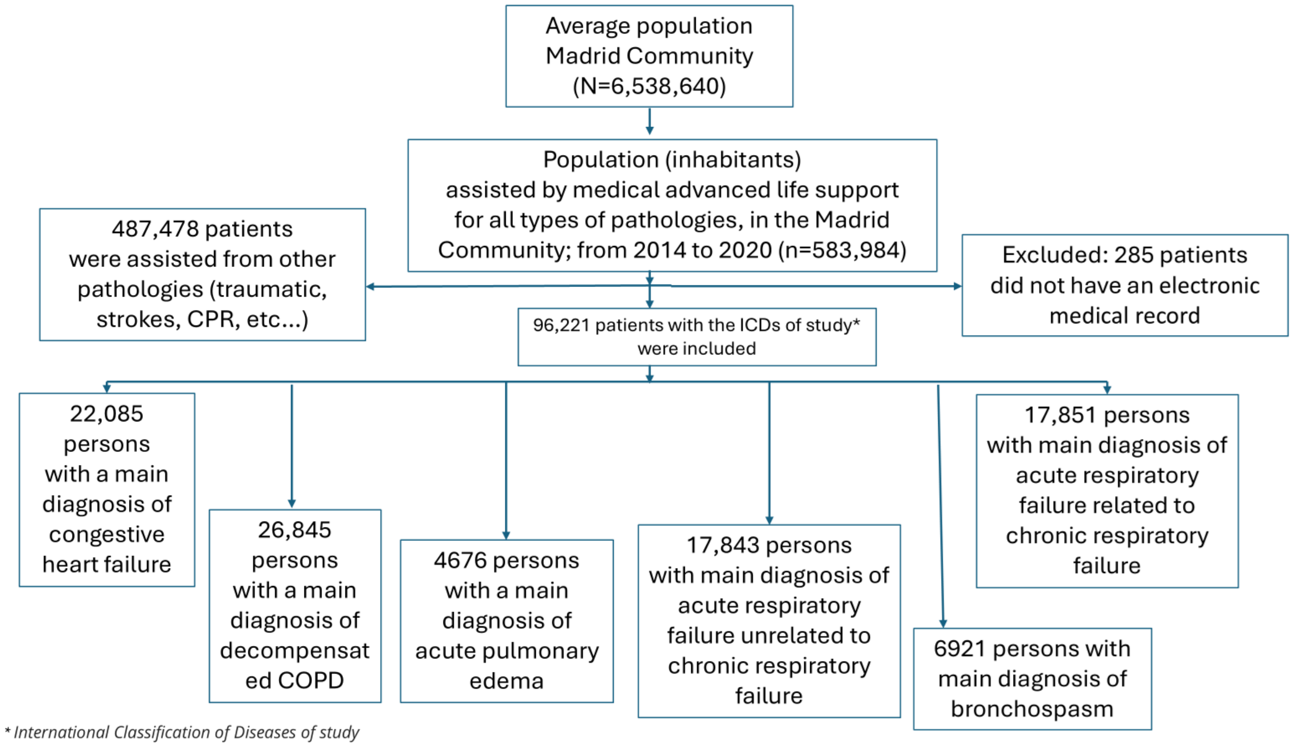
Diseases[a]	2014	2015	2016	2017	2018	2019	2020
Congestive heart failure, n (%)	3593 (0.056)	3696 (0.057)	3651 (0.056)	3484 (0.054)	2895 (0.044)	2828 (0.042)	1938 (0.029)
ECOPD, n (%)	4415 (0.068)	4463 (0.069)	4112 (0.064)	4284 (0.066)	3935 (0.060)	3616 (0.054)	2020 (0.030)
Acute pulmonary edema, n (%)	819 (0.013)	800 (0.012)	725 (0.011)	711 (0.011)	652 (0.010)	616 (0.009)	353 (0.005)
Acute respiratory insufficiency in chronic respiratory failure, n (%)	2872 (0.044)	3034 (0.047)	2834 (0.044)	2797 (0.043)	2399 (0.036)	2362 (0.035)	1553 (0.023)
Bronchospasm, n (%)	1000 (0.015)	1104 (0.017)	1157 (0.018)	1157 (0.018)	1062 (0.016)	979 (0.015)	462 (0.007)
Acute respiratory insufficiency (not related to chronic respiratory failure), n (%)	2648 (0.041)	2736 (0.043)	2717 (0.042)	2573 (0.040)	2349 (0.036)	2289 (0.034)	2531 (0.038)

[a] Results from the retrospective observational research Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA).

Exclusion Criteria

Patients without a health identification number or without an adequate record in the basic minimum dataset were excluded. All patients diagnosed with asthma were excluded since they could be included among the patients with ARF related to chronic respiratory disease (Figure 1).

Figure 1. Cohort selection of patients assisted by the Madrid Community emergency medical service from January 1, 2014, to December 31, 2020. Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA) Study. COPD: chronic obstructive pulmonary disease.



Economic Analysis

The prevalence for each pathology was calculated by dividing the total number of individuals who presented the corresponding medical diagnosis in a given year by the population of Madrid at that time, according to the Spanish National Institute of Statistics.

The direct economic cost of CAM EMS care for the pathologies treated was calculated from the expenditure according to public references [38]. The health care cost data were extracted from the public prices for the provision of public health services and activities of the ALSs published in the Official Gazette of the CAM on August 21, 2017 (price list with an application date until December 2023). The reference cost for this study was the average cost of €611 (US $64,261), which is the arithmetic mean of mobilizing and assisting a patient by a terrestrial ALS unit in the public EMS of the Madrid Community [39]. The ALS provides advanced care, as it contains all the electromedical material and personnel that are required to form a mobile intensive care unit. The cost of ALS care (emergency medical services costs) includes physician care, emergency nursing assistance, and all required ventilatory support techniques, depending on the severity of the patient's condition, with pharmacological treatment, oxygen therapy, invasive and noninvasive mechanical ventilation, and transfer to the hospital.

The direct economic impact was obtained by multiplying the cost of care for each pathology treated in the CAM's ALSs by the number of times care was required. All persons within the CAM were able to receive care, as it is included in the universal emergency health care coverage of the Madrid Health Regional Government (public administration). The economic analysis was done in Microsoft Excel 365 and SPSS software (version 26; IBM Corp), with a copyright license from the Madrid Health Regional Government.

Results

Table 1 shows a total of 96,221 patients who were included between 2014 and 2020. The most prevalent pathology among all those studied was COPD, with an initial prevalence of 0.068% (4415/6,454,440) in 2014, which decreased to 0.030% (2020/6,779,888) in 2020. CHF had a prevalence of 0.056% (3593/6,454,440) in 2014, decreasing to 0.029% (1938/6,779,888) in 2020. The prevalence of ARF in CRF was 0.044% (2872/ 6,454,440) in 2014 but decreased to 0.023% (1553/6,779,888) in 2020. The prevalence of bronchospasm (not from asthma or COPD) decreased from 0.015% (1000/6,454,440) in 2014 to 0.007% (462/6,779,888) in 2020. APE's prevalence decreased from 0.013% (819/6,454,440) in 2014 to 0.005% (353/6,779,888) in 2020. The prevalence of ARF in patients not related to CRF decreased from 0.041% (2648/6,454,440) in 2014 to 0.034% (2289/6,663,394) in 2019 and increased to 0.038% (2531/6,779,888) in 2020, reversing this trend of ARF observed, especially in CHF and ECOPD.

The prevalence of APE shows a decreasing pattern from the onset. The other pathologies start with an upward trend, although in the second year, they reverse the trend and gradually decrease. The prevalence of bronchospasm maintains an upward trend until 2017 and from then on it starts to decrease progressively. In 2020, a sharper peak of decrease is assessed in all pathologies, except ARF not related to CRF, the latter being the only one to increase.

A decrease in most respiratory diseases (Table 2) was observed in 2020, both for patients with ECOPD, APE, or bronchospasm (nonasthmatic, non-COPD) and for patients with CHF (Figures 2A-F and 3), except for ARF not associated with CRF, which corresponded in time and form with the COVID-19 pandemic period [40]. From 2017 on, a decrease in all types of ARF was observed, especially CHF and COPD.

Table 2 shows the economic cost of each pathology and its economic impact in each year. The greatest economic impact was observed in 2015 for ECOPD, with an annual cost of €2,726,893 (US $2,867,982). The estimated cost of all pathologies over 5 years was €58,791,031 (US $61,832,879).

Table . Direct economic impact of the congestive heart failure, exacerbation of chronic obstructive pulmonary disease (ECOPD), acute pulmonary edema (APE), bronchospasm, and acute respiratory failure not related to chronic respiratory failure of the Madrid region population; based on the prehospital Madrid emergency medical service's electronic medical records[a].

ICD-CM[b] type	2014	2015	2016	2017	2018	2019	2020
Congestive heart failure							
Cases	3593	3696	3651	3484	2895	2828	1938
Cost (€)	2,195,323	2,258,256	2,230,761	2,128,724	1,768,845	1,727,908	1,184,118
ECOPD							
Cases	4415	4463	4112	4284	3935	3616	2020
Cost (€)	2,697,565	2,726,893	2,512,432	2,617,524	2,404,285	2,209,376	1,234,220
APE							
Cases	819	800	725	711	652	616	353
Cost (€)	500,409	488,800	442,975	434,421	398,372	376,376	215,683
Acute respiratory failure in chronic respiratory failure							
Cases	2872	3034	2834	2797	2399	2362	1553
Cost (€)	1,754,792	1,853,774	1,731,574	1,708,967	1,465,789	1,443,182	948,883
Bronchospasm							
Cases	1000	1104	1157	1157	1062	979	462
Cost (€)	611,000	674,544	706,927	706,927	648,882	598,169	282,282
Acute respiratory failure (not related to chronic failure)							
Cases	2648	2736	2717	2573	2349	2289	2531
Cost (€)	1,617,928	1,671,696	1,660,087	1,572,103	1,435,239	1,398,579	1,546,441
Total cost (€58,791,031) (€)	9,377,017	9,673,963	9,284,756	9,168,666	8,121,412	7,753,590	5,411,627
Total cost (US $61,832,879) (US $)	9,862,184	10,174,494	9,765,149	9,643,053	8,541,614	8,154,761	5,691,625

[a] Database from 2014 to 2020 included. Results from the retrospective observational research Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA).

[b] *ICD-CM*: *International Classification of Diseases, Clinical Modification*.

Figure 2. Changes in the prevalence of (A) congestive heart failure, (B) exacerbation of COPD, (C) acute pulmonary edema, (D) acute respiratory failure in chronic respiratory failure, (E) bronchospasm, and (F) acute respiratory insufficiency not related to chronic respiratory failure based on electronic medical record of the prehospital Madrid Community emergency medical service (2014-2020), assisted by prehospital emergency medical services. Results from the retrospective observational research Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA). COPD: chronic obstructive pulmonary disease; CRF: chronic respiratory failure.

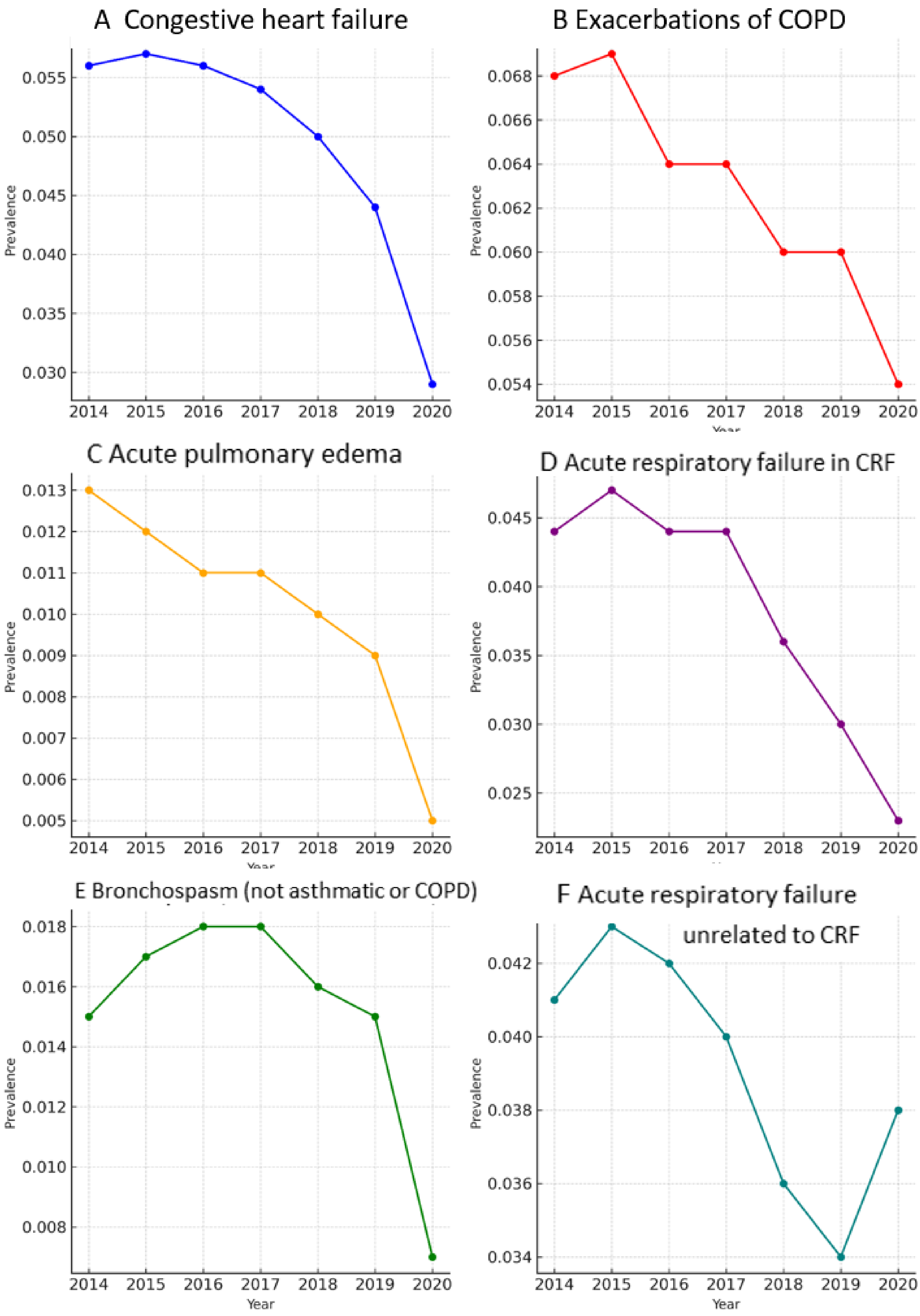
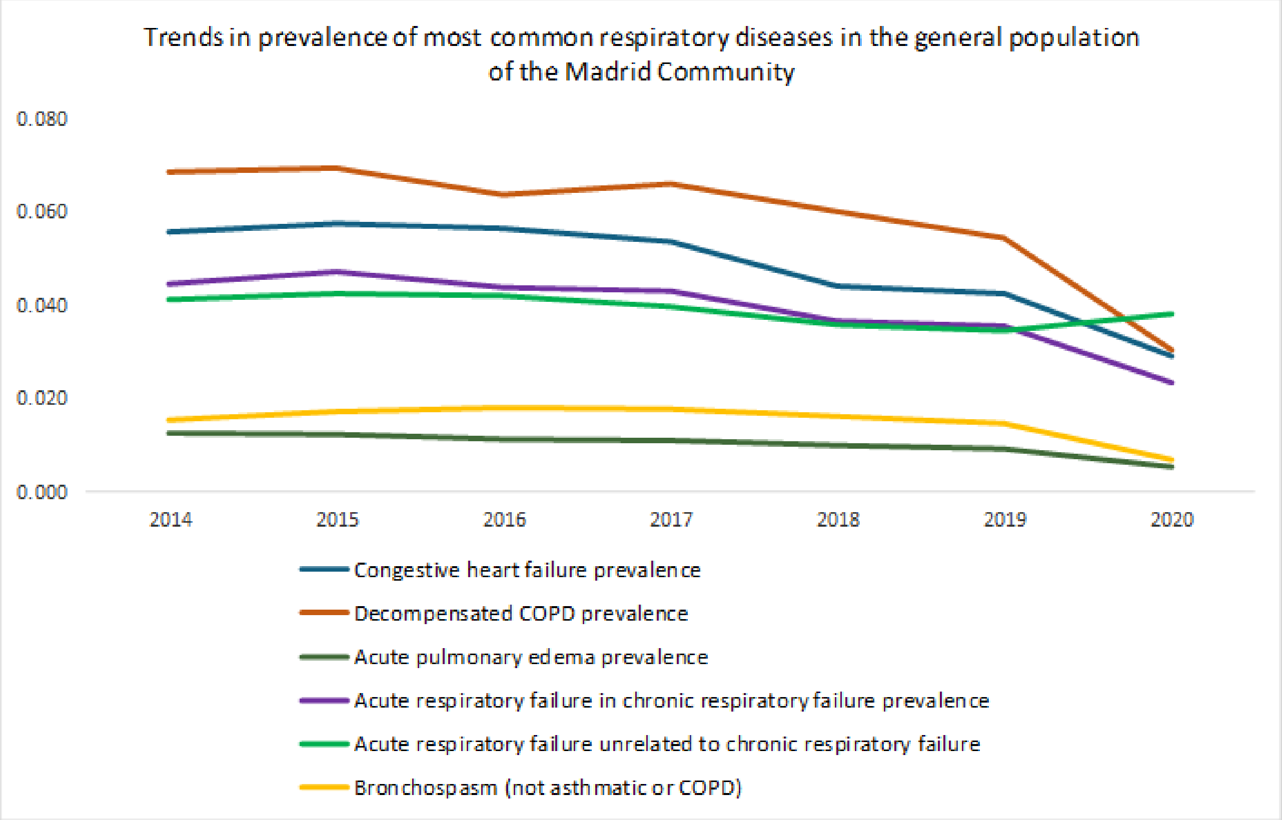

A Congestive heart failure
Prevalence
0.055
0.050
0.045
0.040
0.035
0.030
2014
2015
2016
2017
2018
2019
2020
B Exacerbations of COPD
Prevalence
0.068
0.066
0.064
0.062
0.060
0.058
0.056
0.054
2014
2015
2016
2017
2018
2019
2020
C Acute pulmonary edema
Prevalence
0.013
0.012
0.011
0.010
0.009
0.008
0.007
0.006
0.005
2014
2015
2016
2017
2018
2019
2020
D Acute respiratory failure in CRF
Prevalence
0.045
0.040
0.035
0.030
0.025
2014
2015
2016
2017
2018
2019
2020
E Bronchospasm (not asthmatic or COPD)
Prevalence
0.018
0.016
0.014
0.012
0.010
0.008
2014
2015
2016
2017
2018
2019
2020
Year
F Acute respiratory failure unrelated to CRF
Prevalence
0.042
0.040
0.038
0.036
0.034
2014
2015
2016
2017
2018
2019
2020
Year

Figure 3. Evolution of congestive heart failure, exacerbation of chronic obstructive pulmonary disease, acute pulmonary edema, and acute respiratory failure in chronic respiratory failure prevalences in the Madrid region, based on the general population electronic medical records, assisted by prehospital emergency medical services, from 2014 to 2020. Results from the retrospective observational research Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA). COPD: chronic obstructive pulmonary disease.



Discussion

Principal Findings

Among the findings, we can mention that to our knowledge this is the only study in the prehospital EMS that has analyzed the prevalence of these pathologies in Spain and in Europe, so it can serve as a reference for future European studies on this topic. It was measured by the costs categorized by diagnosis. Finally, it is worth mentioning the large sample size, which provides greater reliability to the study. Few studies have quantified the direct costs incurred by ALSs in these pathologies [25,41-45]. A study carried out in Sweden and Belgium quantified a type of direct cost, such as transport for ECOPD, at 3% of the total direct costs [42]. In France, the cost of medical transport per patient was estimated to be €192 - 273 (US $202-287) in 1996 [46]. In this study, the cost of air ambulance transportation, but not that of emergency medical care, was estimated.

In the ALS of the CAM, the demand by patients with dyspnea as the main diagnosis among patients with ECOPD was known, although there were no data to differentiate among ECOPD, bronchospasm, and other processes. Therefore, we needed to accurately calculate the prevalence of ARF in ALS to more precisely estimate the economic impact of these pathologies and their complete clinical picture, which includes emergent prehospital care [47].

The results of the SUMMIRA project have made it possible to estimate the global quantification for prehospital emergency care due to SARS-CoV-2 in the CAM, served by the ALSs of the CAM. The pathologies studied showed a continuously decreasing trend throughout the period analyzed, coinciding with the mortality data related to ECOPD published in the CAM [22]. The only different trend was the peak produced by ARF not associated with chronic failure, which, despite showing a decreasing trend from 2014 to 2019, increased the number of cases in 2020 at the same time as the COVID-19 pandemic in Spain [48].

These other declines previous to 2020 may be the result of many factors, such as regional success of tobacco control measures, improved information and prevention of allergens, better treatment of respiratory and cardiac comorbidities, and a reduction of universally high rates of underdiagnosis of chronic respiratory diseases [25]. The decreasing trend in the number of cases of most of the chronic respiratory diseases studied, such as COPD, CHF, and APE, before 2020, led us to estimate an improvement in disease control in chronic patients to date and a subsequent influence of the cases due to the COVID-19 pandemic [49].

The economic impact of chronic diseases on patients has not been quantified in the most recent COPD studies [7,21,50], as only primary care and hospital expenditure have been taken into account [22,45,51]. A descriptive North American study lasting 5 years with a large sample (166,908 patients) in which patients with respiratory distress transferred to the hospital was analyzed [6], but the population prevalences were not calculated.

Furthermore, there is a comparative dilemma between the different activities of the EMS's ALSs in different countries. Some of them in Europe, such as the paramedical system, cannot discharge patients since transfer to hospital is part of the work protocol because some of them do not have doctors within health care teams who can discharge patients [52-54]. In this sense, the Spanish system of EMS [55] helps prevent care overload in hospitals since the physicians of ALSs allow patients to remain at home and can be treated there, improving adequately after receiving treatment and thus achieving benefits for both patients and the health system. Considering the number of cases of the diseases found, it would be advisable to evaluate the prevalence of these diseases in other similar European and national emergency health systems to have an updated global vision.

The set of decreasing trends in the prevalence of ECOPD is in line with the trend in industrialized countries [56]. Plausible causes include the greater control of adverse risk factors, as well as the increase in public health measures that protect against respiratory diseases (tobacco, asbestos, protective measures in occupational risks, etc) [18,25,57,58].

Limitations

This work had a shortcoming in the population sample. This cohort did not include the care provided to the patients who go directly to the hospital emergency department; it is a prehospital cohort. The cohort would be bigger if it would include the patients of the Madrid Municipal Emergency Assistance and Rescue Service (SAMUR), though the number of patients treated by them is minimal because their assistance is limited to the public thoroughfare of Madrid capital, and by regular form this type of patient is treated at home and primary care that is covered by SUMMA112.

We have separated the average cost of the assistance provided by the advanced terrestrial life support services of the CAM to estimate the economic cost. Advanced air life support has also treated some patients, which is valued in Table 3.

Table . Case study of congestive heart failure, exacerbation of chronic obstructive pulmonary disease (ECOPD), acute pulmonary edema, bronchospasm, and acute respiratory insufficiency not related to chronic respiratory failure of the Madrid region population per year treated by Servicio de Urgencias Medicas de Madrid Helicopter Emergency Medical Service (SUMMA112 HEMS), from 2014 to 2020[a].

Condition	2015	2016	2017	2018	2019	2020
ECOPD	1 ECOPD	1 ECOPD	2 ECOPD	2 ECOPD	3 ECOPD	1 ECOPD
Acute respiratory failure related to chronic respiratory failure	1 Acute respiratory failure in chronic respiratory failure	2 Acute respiratory failure in chronic respiratory failure	2 Acute respiratory failure related to chronic respiratory failure	1 Acute respiratory failure related to chronic respiratory failure	2 Acute respiratory failure related to chronic respiratory failure	1 Acute respiratory failure related to chronic respiratory failure
Acute pulmonary edema	1 Acute pulmonary edema	1 Acute pulmonary edema	2 Acute pulmonary edema	1 Acute pulmonary edema	3 Acute pulmonary edema	N/A[b]
Congestive heart failure	1 Congestive heart failure	4 Congestive heart failure	2 Congestive heart failure	N/A	N/A	3 Congestive heart failure
Bronchospasm	N/A	1 Bronchospasm	N/A	1 Bronchospasm	2 Bronchospasm	N/A
Acute respiratory failure not related to chronic respiratory failure	N/A	2 Acute respiratory failure not related.to chronic respiratory failure	2 Acute respiratory failure not related to chronic respiratory failure	1 Acute respiratory failure not related to chronic respiratory failure	3 Acute respiratory failure not related to chronic respiratory failure	N/A

[a] As can be valued in the Official Gazette of the Madrid Council [38], the cost per medical helicopter intervention in this period time is €5746 (US $6043).

[b] N/A: not applicable.

Our study shows a reduction in the costs of CHF over the years in EMS. The decrease may be explained by factors such as reduced severity and better treatment of acute coronary syndrome [21]. We have not found any similar bibliography in the prehospital setting. The 2019 Heart Failure Association Atlas reported a heart failure prevalence ranging from ≤12 in Spain [51]. At the European level, there is considerable heterogeneity of data between European countries, ranging from an overall prevalence of 14 per 1000 in the Netherlands [24] to 39 in Slovenia, Lithuania, and Germany but with a common point: EMS use in heart failure is an independent predictor of 30-day mortality [50]. In exacerbations of COPD, the patients have a significant comorbidity burden and experience high rates of hospitalization and mortality [59]. It can lead to several conclusions in the area of public health: (1) the measures taken by primary and specialized care have been effective in control of exacerbations of this type of pathology, and furthermore (2) despite reducing costs, in seasonal outbreaks a greater investment [51] of costs would be needed to reinforce the resources that deal with the ECOPD in the prehospital emergency services.

This study opens the possibility of including these values in the database of the economic impact of chronic respiratory diseases such as COPD, CHF, and APE, which are currently not considered [6-8]. In addition to the registration of their prevalence in the different months of the year, an analysis of the needs required in public health can be made to make investment plans in the times of the year when there is a higher prevalence. Moreover, it could be taken into account when assessing the costs associated with SARS-CoV-2–related bronchospasm and ARF in prehospital critical care for future respiratory transmission pandemics.

Conclusions

The results found in this study are in line with the general hospitalization downward trend in high-income countries of a decrease in the number of cases from 2014 to 2020 for the most common respiratory diseases treated in hospital emergencies [21,51,56]. The prevalence of ECOPD in the Madrid EMS was higher than the other respiratory diseases studied. In 2020, with the COVID-19 pandemic, a decrease in the prevalence and costs of the pathologies studied was observed, except for the prevalence of ARFs not related to chronic failure. The greatest economic impact was for ECOPD in 2015, with an annual cost of €2,726,893 (US $2,867,982). The estimated cost of these pathologies in 5 years was €58,791,031 (US $61,832,879).

Acknowledgments

The authors thank the support of the medical director and the medical deputy director of the Servicio of Urgencias Médicas of Madrid (SUMMA112); data exploration team of the IT Department of SUMMA112; Mr Juan Pajuelo Ayuso; Mr José Gómez Martínez, and Mrs Nuria Rodríguez Rodil; Candelas Gil, Head of the Medical Library Service of the Autonomous University of Madrid, Spain; and the Primary Care Biomedical Research and Innovation Foundation (FIIBAP). Funding for open-access publishing was provided by the FIIBAP (grant 883345).

This work has been presented as oral communication at the XXXIX Annual Meeting of the Spanish Epidemiology Society (SEE) and at the XIX Congress of the Spanish Society of Public Health and Health Administration (SESPAS). This article sheds light on the changing prevalence of the most frequent respiratory diseases treated by emergency medical services, the influence of the COVID-19 pandemic on them, and the costs involved in their care in the out-of-hospital emergency setting.

Data Availability

The datasets generated analyzed during this study are not publicly available due to privacy requirements and ethical restrictions. Reasonable requests for data access are available by contacting the corresponding author and submitting application documents. Anonymized data from this study are available from the corresponding author upon reasonable request. The descriptive multicenter study that is presented is based on a health care casuistry by filtering the codes of the respiratory pathologies registered in the Electronic Medical Record on the Integrated Management System of the Coordinating Center platform.

Integrated Management System of the Coordinating Center is a computer application that operates under the supervision of the Information Systems Security Office of the Department of Health of the Community of Madrid. It is included in the Integrated Management System of the Prehospital Emergency Medical Service of Autonomous Community of Madrid (SUMMA112) Coordination Center. These data are protected by security systems and guarded by the technical managers of the Information Systems Security Office of the Department of Health of the Community of Madrid (OSSI). There is a high level of security in the treatment and storage of medical records from the Medical Emergency Service of the Madrid Region, which depends on the General Directorate of Information Systems and Health Equipment of the Madrid Health Service.

Authors' Contributions

AMCS, CHG, VQC, AMPA, and AGM were involved in the conception and design of the work. AMCS, CHG, and VQC contributed to the screening process, data extraction, and quality assessment of papers. AMCS, CHG, and VQC interpreted the data and drafted the manuscript. The manuscript was critically reviewed and enriched by CHG, AMPA, and AGM. All authors read and approved the final manuscript and agreed to be accountable for all aspects of this work.

Conflicts of Interest

None declared. All authors have confirmed the maintenance of confidentiality and respect for the rights of the patient in the document of responsibilities of the author, the publication agreement, and the transfer of rights to the journal.

Multimedia Appendix 1
STROBE (Strengthening the Reporting of Observational studies in Epidemiology) reporting checklist.
[DOCX File, 72 KB - publichealth_v11i1e66179_app1.docx]

Multimedia Appendix 2
Exacerbation of chronic obstructive pulmonary disease and acute pulmonary edema diagnosis.
[DOCX File, 67 KB - publichealth_v11i1e66179_app2.docx]

Multimedia Appendix 3

Congestive heart failure diagnosis.
[DOCX File, 60 KB - publichealth_v11i1e66179_app3.docx]

References

1. Ananth S, Hurst JR. ERJ advances: state of the art in definitions and diagnosis of COPD. Eur Respir J 2023 Apr;61(4):2202318. [doi: 10.1183/13993003.02318-2022] [Medline: 36796835]
2. Chronic obstructive pulmonary disease (COPD). World Health Organization. 2024 Nov 6. URL: https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd) [accessed 2025-01-03]
3. Adeloye D, Song P, Zhu Y, et al. Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis. Lancet Respir Med 2022 May;10(5):447-458. [doi: 10.1016/S2213-2600(21)00511-7] [Medline: 35279265]
4. Zou W, Ou J, Wu F, et al. Association of mild chronic obstructive pulmonary disease with all-cause mortality: a systematic review and meta-analysis. Pulmonology 2023 Oct 27:S2531-0437(23)00165-4. [doi: 10.1016/j.pulmoe.2023.09.002] [Medline: 38093693]
5. Fitzpatrick D, Duncan EAS, Moore M, et al. Epidemiology of emergency ambulance service calls related to COVID-19 in Scotland: a national record linkage study. Scand J Trauma Resusc Emerg Med 2022 Jan 28;30(1):9. [doi: 10.1186/s13049-022-00995-6] [Medline: 35090527]
6. Prekker ME, Feemster LC, Hough CL, et al. The epidemiology and outcome of prehospital respiratory distress. Acad Emerg Med 2014 May;21(5):543-550. [doi: 10.1111/acem.12380] [Medline: 24842506]
7. Zhou J, Nehme E, Dawson L, et al. Epidemiology, outcomes and predictors of mortality in patients transported by ambulance for dyspnoea: a population-based cohort study. Emerg Med Australas 2023 Feb;35(1):48-55. [doi: 10.1111/1742-6723.14053] [Medline: 35918062]
8. Cardiovascular Research. Global burden of heart failure: a comprehensive and updated review of epidemiology. Oxford Academic. 2022 Feb 12. URL: https://academic.oup.com/cardiovascres/article/118/17/3272/6527627?login=false [accessed 2025-01-03]
9. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats AJS. Global burden of heart failure: a comprehensive and updated review of epidemiology. Cardiovasc Res 2023 Jan 18;118(17):3272-3287. [doi: 10.1093/cvr/cvac013]
10. Miró Ò, Hazlitt M, Escalada X, et al. Effects of the intensity of prehospital treatment on short-term outcomes in patients with acute heart failure: the SEMICA-2 study. Clin Res Cardiol 2018 Apr;107(4):347-361. [doi: 10.1007/s00392-017-1190-2]
11. Miró Ò, Martínez G, Masip J, et al. Effects on short term outcome of non-invasive ventilation use in the emergency department to treat patients with acute heart failure: a propensity score-based analysis of the EAHFE Registry. Eur J Intern Med 2018 Jul;53:45-51. [doi: 10.1016/j.ejim.2018.03.008] [Medline: 29572091]
12. Purvey M, Allen G. Managing acute pulmonary oedema. Aust Prescr 2017 Apr;40(2):59-63. [doi: 10.18773/austprescr.2017.013]
13. Wang J, Duan J, Zhou L. Incidence of noninvasive ventilation failure and mortality in patients with acute respiratory distress syndrome: a systematic review and proportion meta-analysis. BMC Pulm Med 2024 Jan 22;24(1):48. [doi: 10.1186/s12890-024-02839-8] [Medline: 38254064]
14. Quintana MGM, García CH, Misis AG, et al. Factors associated with in-hospital mortality and readmission in a cohort of patients treated with noninvasive ventilation during emergency department or out-of-hospital emergency care: the VentilaMadrid study. Emerg Rev Soc Esp Med Emerg 2022 Feb;34(1):7-14.
15. Muñoz SS, Canela À, Chao JRC, Doblado I, Fàbrega XJ, Rodríguez JJ. Impact of pre-hospital application of non-invasive mechanical ventilation in acute pulmonary oedema. NIVEMS cohort study. Rev Esp Urgenc Emerg Span J Acute Emerg Care 2024;3(4) [FREE Full text]
16. Sneath E, Bunting D, Hazell W, Tippett V, Yang IA. Pre-hospital and emergency department pathways of care for exacerbations of chronic obstructive pulmonary disease (COPD). J Thorac Dis 2019 Oct;11(Suppl 17):S2221-S2229. [doi: 10.21037/jtd.2019.10.37] [Medline: 31737349]
17. Boers E, Barrett M, Su JG, et al. Global burden of chronic obstructive pulmonary disease through 2050. JAMA Netw Open 2023 Dec 1;6(12):e2346598. [doi: 10.1001/jamanetworkopen.2023.46598] [Medline: 38060225]
18. Duan KI, Birger M, Au DH, Spece LJ, Feemster LC, Dieleman JL. Health care spending on respiratory diseases in the United States, 1996-2016. Am J Respir Crit Care Med 2023 Jan 15;207(2):183-192. [doi: 10.1164/rccm.202202-0294OC] [Medline: 35997678]
19. Boucher PE, Taplin J, Clement F. The cost of ARDS: a systematic review. Chest 2022 Mar;161(3):684-696. [doi: 10.1016/j.chest.2021.08.057] [Medline: 34478719]
20. Peters GA, Cash RE, Goldberg SA, Kolb LM, Ordoobadi AJ, Camargo CA Jr. Emergency medical services management of bronchospasm in the United States: a cross-sectional analysis and nationwide quality assessment. Prehosp Emerg Care 2024;28(2):231-242. [doi: 10.1080/10903127.2023.2220021] [Medline: 37276151]
21. Rosano GMC, Seferovic P, Savarese G, et al. Impact analysis of heart failure across European countries: an ESC-HFA position paper. ESC Heart Fail 2022 Oct;9(5):2767-2778. [doi: 10.1002/ehf2.14076] [Medline: 35869679]

22. White paper on the socio-economic burden of COPD [Article in Spanish]. Weber. 2015. URL: https://weber.org.es/publicacion/libro-blanco-sobre-la-carga-socio-economica-de-la-epoc/ [accessed 2025-01-03]
23. Brown R, Goldstein J, Jensen JL, Travers AH, Carter A. Evaluating the impact of a novel mobile care team on the prevalence of ambulatory care sensitive conditions presenting to emergency medical services in Nova Scotia. Cureus 2023 Apr;15(4):e37280. [doi: 10.7759/cureus.37280] [Medline: 37168216]
24. Sicras-Mainar A, Sicras-Navarro A, Palacios B, Varela L, Delgado JF. Epidemiology and treatment of heart failure in Spain: the HF-PATHWAYS study. Rev Esp Cardiol (Engl Ed) 2022 Jan;75(1):31-38. [doi: 10.1016/j.rec.2020.09.033] [Medline: 33380382]
25. Soriano JB, Kendrick PJ, Paulson KR, et al. Prevalence and attributable health burden of chronic respiratory diseases, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Respir Med 2020 Jun;8(6):585-596. [doi: 10.1016/S2213-2600(20)30105-3]
26. Sacristán FR. El SUMMA 112 uno de los principales baluartes de la sanidad madrileña [Article in Spanish]. Gaceta Médica. 2024. URL: https://gacetamedica.com/politica/el-summa-112-uno-de-los-principales-baluartes-de-la-sanidad-madrilena/ [accessed 2025-01-03]
27. Sistemas y Tecnologías de la Información [Article in Spanish]. URL: https://www.comunidad.madrid/hospital/summa112/profesionales/sistemas-tecnologias-informacion [accessed 2025-01-03]
28. Instituto nacional de estadística [Article in Spanish]. INE. URL: https://www.ine.es/ [accessed 2025-01-02]
29. Manual de ventilación mecánica del SUMMA112 comunidad de madrid [Article in Spanish]. Comunidad de Madrid. URL: https://www.comunidad.madrid/publicacion/ref/50596 [accessed 2025-01-03]
30. Manual de Procedimientos · SAMUR Protección Civil [Article in Spanish]. Madrid. URL: https://www.madrid.es/ficheros/SAMUR/ [accessed 2025-01-03]
31. Bibiano C, Sanz M, Esparza C, Caussade F, Montero M, Puig R. Emergency manual 2014. URL: https://www.urgenciasyemergen.com/nuevo-manual-urgencias-3a-ed-bibiano/ [accessed 2024-12-20]
32. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. American Heart Association Journals. 2022 May 3. URL: https://www.ahajournals.org/doi/epub/10.1161/CIR.0000000000001063 [accessed 2025-01-03]
33. Luján M, Peñuelas Ó, Gómez CC, et al. Sumario de las recomendaciones y puntos clave del Consenso de las Sociedades Científicas Españolas (SEPAR, SEMICYUC, SEMES; SECIP, SENeo, SEDAR, SENP) para la utilización de la ventilación no invasiva y terapia de alto flujo con cánulas nasales en el paciente adulto, pediátrico y neonatal con insuficiencia respiratoria aguda grave [Article in Spanish]. Med Intensiva 2021 Jun;45(5):298-312. [doi: 10.1016/j.medin.2020.08.016]
34. Bello G, De Santis P, Antonelli M. Non-invasive ventilation in cardiogenic pulmonary edema. Ann Transl Med 2018 Sep;6(18):355. [doi: 10.21037/atm.2018.04.39]
35. Manual 12 de Octubre [Article in Spanish]. MSD Profesionales Sanitarios. URL: https://profesionales.msd.es/manual-12-de-octubre/ [accessed 2025-01-03]
36. Manual de actuación clínica en las unidades medicalizadas de emergencia [Article in Spanish]. Portal de Salud de la Junta de Castilla y León. URL: https://www.saludcastillayleon.es/institucion/es/biblioteca/materiales-consejeria-sanidad/buscador/manual-actuacion-clinica-unidades-medicalizadas-emergencia [accessed 2025-01-03]
37. Población de la Comunidad de Madrid (según Censo), por sexo y edad (simple y quinquenal). Series desde el año 1900 | Data bank [Population of the community of Madrid (according to the census), by sex and age (simple and five-year). Series since 1900] [Article in Spanish]. Instituto de Estadística. URL: https://iestadis.edatos.io/statistical-visualizer/visualizer/data.html?resourceType=dataset&agencyId=IECM&resourceId=006_000008&version=~latest#visualization/table [accessed 2025-01-03]
38. Precios públicos por la prestación de los servicios y actividades de naturaleza sanitaria. tasas y precios sanitarios de la comunidad de madrid. BOCM A) disposiciones generales—BOCM https://www.BOCM.es. anexo II asistencia sanitaria en atención primaria, actividad no asistencial, transporte sanitario y asistencia sanitaria móvil en urgencias y emergencias [Article in Spanish]. Comunidad de Madrid. URL: https://sede.comunidad.madrid/medias/precios-publicos-anexos-iiiiii-iv/download [accessed 2025-01-03]
39. UVI Móvil [Article in Spanish]. SUMMA 112. URL: https://www.comunidad.madrid/hospital/summa112/ciudadanos/medios/uvi-movil [accessed 2025-01-03]
40. Impacto socioeconómico del coronavirus (COVID-19) y la cuarentena en España [Article in Spanish]. Statista. URL: https://es.statista.com/estudio/73004/el-coronavirus-covid-19-y-su-impacto-economico-y-social-en-espana/ [accessed 2025-01-03]
41. Briseño DM, Plata MRF, Figueroa M, Padilla RP. La carga económica de la EPOC. Análisis de los costos a nivel internacional [Article in Spanish]. Neumol Cir Torax 2011;70 [FREE Full text]
42. Jansson SA, Backman H, Stenling A, Lindberg A, Rönmark E, Lundbäck B. Health economic costs of COPD in Sweden by disease severity—has it changed during a ten years period? Respir Med 2013 Dec;107(12):1931-1938. [doi: 10.1016/j.rmed.2013.07.012] [Medline: 23910072]
43. Gutiérrez Villegas C, Paz-Zulueta M, Herrero-Montes M, Parás-Bravo P, Madrazo Pérez M. Cost analysis of chronic obstructive pulmonary disease (COPD): a systematic review. Health Econ Rev 2021 Dec;11(1):31. [doi: 10.1186/s13561-021-00329-9]

44. Pincombe A, Schultz TJ, Hofmann D, Karnon J. Economic evaluation of a medical ambulatory care service using a single group interrupted time-series design. J Eval Clin Pract 2023 Mar;29(2):329-340. [doi: 10.1111/jep.13771] [Medline: 36156337]
45. Rehman AU, Hassali MAA, Muhammad SA, Harun SN, Shah S, Abbas S. The economic burden of chronic obstructive pulmonary disease (COPD) in Europe: results from a systematic review of the literature. Eur J Health Econ 2020 Mar;21(2):181-194. [doi: 10.1007/s10198-019-01119-1] [Medline: 31564007]
46. Laurendeau C, Chouaid C, Roche N, Terrioux P, Gourmelen J, Detournay B. Costs of COPD in France: a national database analysis. Value Health 2013 Nov;16(7):A370. [doi: 10.1016/j.jval.2013.08.273]
47. Memorias SUMMA 112 [Article in Spanish]. SUMMA 112. URL: https://www.comunidad.madrid/hospital/summa112/nosotros/memorias-summa-112 [accessed 2025-01-03]
48. Garcia-Carretero R, Vazquez-Gomez O, Gil-Prieto R, Gil-de-Miguel A. Hospitalization burden and epidemiology of the COVID-19 pandemic in Spain (2020-2021). BMC Infect Dis 2023 Jul 18;23(1):476. [doi: 10.1186/s12879-023-08454-y] [Medline: 37464303]
49. Emami A, Javanmardi F, Pirbonyeh N, Akbari A. Prevalence of underlying diseases in hospitalized patients with COVID-19: a systematic review and meta-analysis. Arch Acad Emerg Med 2020;8(1):e35. [Medline: 32232218]
50. Harjola P, Tarvasmäki T, Barletta C, et al. The emergency department arrival mode and its relations to ED management and 30-day mortality in acute heart failure: an ancillary analysis from the EURODEM study. BMC Emerg Med 2022 Feb 14;22(1):27. [doi: 10.1186/s12873-022-00574-z] [Medline: 35164693]
51. Becher PM, Lund LH, Coats AJS, Savarese G. An update on global epidemiology in heart failure. Eur Heart J 2022 Aug 21;43(32):3005-3007. [doi: 10.1093/eurheartj/ehac248] [Medline: 35578978]
52. Chenaitia H, Massa H, Noel C, et al. Paramedics in prehospital emergency medical systems across Europe. Int Paramed Pract 2011 Sep 5;1(1):33-39. [doi: 10.12968/ippr.2011.1.1.33]
53. Sanitäts- und ambulanzdienste [Article in German]. Johanniter Österreich. URL: https://www.johanniter.at/hilfe-angebote/sanitaets-und-ambulanzdienste [accessed 2025-01-03]
54. Dúason S, Ericsson C, Jónsdóttir HL, Andersen JV, Andersen TL. European paramedic curriculum-a call for unity in paramedic education on a European level. Scand J Trauma Resusc Emerg Med 2021 May 31;29(1):72. [doi: 10.1186/s13049-021-00889-z] [Medline: 34059084]
55. Barroeta-Urquiza J, Boada-Bravo N. Los Servicios de Emergencia y Urgencias Médicas Extrahospitalarias En España [Emergency Services and Out-of-Hospital Medical Emergencies in Spain] [Book in Spanish]: Mensor; 2011. URL: https://www.madrid.es/UnidadesDescentralizadas/Emergencias/Samur-PCivil/Samur/ApartadosSecciones/10_DescargasPublicacionesWebsInteres/Ficheros/Los_SEM_en_Espana.pdf [accessed 2024-03-15]
56. 2023 GOLD report. Global Initiative for Chronic Obstructive Lung Disease. URL: https://goldcopd.org/2023-gold-report-2/ [accessed 2025-01-03]
57. Ancochea J, Soriano JB. La EPOC en España al inicio de una nueva década [Article in Spanish]. Arch Bronconeumol 2021 Jan;57(1):1-2. [doi: 10.1016/j.arbres.2020.01.025]
58. Bueno H, Bernal JL, Jiménez-Jiménez V, et al. The clinical outcomes, healthcare resource utilization, and related costs (COHERENT) model. Application in heart failure patients. Rev Esp Cardiol (Engl Ed) 2022 Jul;75(7):585-594. [doi: 10.1016/j.rec.2021.08.009] [Medline: 34688580]
59. Sneath E, Tippett V, Bowman RV, et al. The clinical journey of patients with a severe exacerbation of chronic obstructive pulmonary disease (COPD): from the ambulance to the emergency department to the hospital ward. J Thorac Dis 2022 Dec;14(12):4601-4613. [doi: 10.21037/jtd-22-328] [Medline: 36647501]

Abbreviations

ALS: advanced life support
APE: acute pulmonary edema
ARF: acute respiratory failure
CAM: Autonomous Community of Madrid
CHF: congestive heart failure
COPD: chronic obstructive pulmonary disease
CRF: chronic respiratory failure
ECOPD: exacerbation of chronic obstructive pulmonary disease
EMR: electronic medical record
EMS: emergency medical service
STROBE: Strengthening the Reporting of Observational studies in Epidemiology
SUMMA112: Servicio de Urgencias Medicas de Madrid (Prehospital Emergency Medical Service of Autonomous Community of Madrid)
SUMMIRA: Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid

Edited by A Mavragani, T Sanchez; submitted 05.09.24; peer-reviewed by R Brown, S Moitra; revised version received 15.11.24; accepted 17.11.24; published 16.01.25.

Please cite as:
Cintora-Sanz AM, Horrillo-García C, Quesada-Cubo V, Pérez-Alonso AM, Gutiérrez-Misis A
Prevalence and Economic Impact of Acute Respiratory Failure in the Prehospital Emergency Medical Service of the Madrid Community: Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e66179
URL: https://publichealth.jmir.org/2025/1/e66179
doi:10.2196/66179

Original Paper

Trends in Early-Onset Colorectal Cancer in Singapore: Epidemiological Study of a Multiethnic Population

Hui Lionel Raphael Chen[1,2], MBBS, MRCS, FRCS; Qingqing Dawn Chong[2,3], MBBS, MRCP, MPH; Brenda Tay[3], BSc; Siqin Zhou[4], BSc, MSc; Evelyn Yi Ting Wong[2,3], MBBS, MRCP, MMed, FAMS; Isaac Seow-En[1,2], MBBS, MRCS, FRCS; Ker Kan Tan[5,6], MBBS, MMED, FRCS, PhD; Yi Wang[6], PhD; Adeline Seow[6], MBBS, MMED, FFPH; Kwong-Wei Emile Tan[1,2], BSc, MBBS, PhD, FRCS; Bee Huat Iain Tan[2,3], MBBS, MRCP, PhD; Sze Huey Tan[4], PhD

[1]Department of Colorectal Surgery, Singapore General Hospital, Singapore, Singapore
[2]Duke-NUS Medical School, Singapore, Singapore
[3]Division of Medical Oncology, National Cancer Centre, Singapore, Singapore
[4]Division of Clinical Trials and Epidemiological Sciences, National Cancer Centre, Singapore, Singapore
[5]Division of Colorectal Surgery, Department of Surgery, National University Hospital, Singapore, Singapore
[6]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

Corresponding Author:
Hui Lionel Raphael Chen, MBBS, MRCS, FRCS
Department of Colorectal Surgery
Singapore General Hospital
20 College Road
Singapore, 169856
Singapore
Phone: 65 63214677
Email: lionel.raphael.chen.h@singhealth.com.sg

Abstract

Background: Colorectal cancer (CRC) incidence and mortality in those aged 50 years and above have decreased over the past 2 decades. However, there is a rising incidence of CRC among individuals under 50 years of age, termed early-onset colorectal cancer (EOCRC). Patients with EOCRC are diagnosed at an advanced stage and may be in more psychosocial, emotional, and financial distress.

Objective: Our study examined the epidemiological shifts in CRC in Singapore, a multiethnic country.

Methods: CRCs diagnosed at age 20 years and above were identified from the Singapore Cancer Registry (SCR) from 1968 to 2019. Patient characteristics included gender, ethnicity, and age of CRC diagnosis. Population information was obtained from the Department of Statistics Singapore (SingStat). Age-specific incidence rates (ASRs) and age-standardized incidence rates (ASIRs) were calculated. The cohort was divided into 3 age groups: 20-49, 50-64, and ≥65 years. Temporal trends in incidence rates were modeled with joinpoint regression. Birth cohort models were fitted using the National Cancer Institute (NCI) age-period-cohort analysis tool. Cancer-specific survival analysis was performed with the Cox proportional hazards model.

Results: In total, 53,044 CRCs were included, and 6183 (11.7%) adults aged 20-49 years were diagnosed with EOCRC. The ASR of EOCRC rose from 5 per 100,000 population in 1968 to 9 per 100,000 population in 1996 at 2.1% annually and rose to 10 per 100,000 population in 2019 at 0.64% annually. The ASR for CRC among adults aged 50-64 years rose at 3% annually from 1968 to 1987 and plateaued from 1987, while the ASR for adults aged 65 years and above rose at 4.1% annually from 1968 to 1989 and 1.3% annually from 1989 to 2003 but decreased from 2003 onwards at 1% annually. The ASR of early-onset rectal cancer increased significantly at 1.5% annually. There was a continued rise in the ASR of EOCRC among males (annual percentage change [APC] 1.5%) compared to females (APC 0.41%). Compared to the 1950-1954 reference birth cohort, the 1970-1984 birth cohort had a significantly higher incidence rate ratio (IRR) of 1.17-1.36 for rectal cancer, while there was no significant change for colon cancer in later cohorts. There were differences in CRC trends across the 3 ethnic groups: Malays had a rapid and persistent rise in the ASR of CRC across all age groups (APC 1.4%-3%), while among young Chinese, only the ASR of rectal cancer was increasing (APC 1.5%). Patients with EOCRC had better survival compared to patients diagnosed at 65 years and above (hazard ratio [HR] 0.73, 95% CI 0.67-0.79, *P*<.001) after adjusting for covariates.

Conclusions: The rise in the incidence of rectal cancer among young adults, especially among Chinese and Malays, in Singapore highlights the need for further research to diagnose CRC earlier and reduce cancer-related morbidity and mortality.

(JMIR Public Health Surveill 2025;11:e62835) doi:10.2196/62835

KEYWORDS

early-onset colorectal cancer; epidemiology; Singapore; joinpoint regression; age-period-cohort; public health; health disparity

Introduction

Colorectal cancer (CRC) presents a significant health care problem as it is the third-most common cancer diagnosed and the second-leading cause of death worldwide. In 2020, CRC accounted for 10% of the worldwide cancer incidence and 9% of cancer deaths. Currently, 2 million new cases are diagnosed annually worldwide, which is projected to rise to 3.2 million per annum by 2040 [1,2].

With the implementation of population-wide, average risk CRC screening for all adults aged 50 years or above, CRC incidence and mortality in this age group have progressively decreased over the past 2 decades. However, there is emerging evidence that the incidence of CRC in people under 50 years of age is rising, especially in high-income countries. Based on the World Health Organization–International Agency of Research on Cancer Incidence in Five Continents Plus database, the incidence of early-onset colorectal cancer (EOCRC) rose in 27 (54%) of 50 countries, with the highest incidence rates of 14-16 per 100,000 person-years seen in the United States, Australia, New Zealand, and South Korea [3]. In United States, since the 1990s, colon and rectal cancer incidence has been increasing at 1.3% and 2.3% per year, respectively, for adults aged 40-49 years [4]. This is in contrast to the declining incidence of CRC among adults aged 55 years or older. Similar trends in colon and rectal cancer incidence were observed in Australia, with an increase of 1.7% and 0.9% per year, respectively, for adults aged 40-49 years [5].

EOCRC may be associated with hereditary syndromes, such as Lynch syndrome and familial adenomatous polyposis, but these only account for 1%-3% of the total number of new CRC cases. The large majority of EOCRC cases are still sporadic and contribute 10%-12% of new CRC cases. EOCRC is more likely to occur in the left colon, with higher rates of mucinous and poorly differentiated cancers, which may be associated with poor prognosis [6]. Patients with EOCRC tend to possess an underappreciation of symptoms; more reluctance to seek medical assistance, leading to delayed diagnosis; and more advanced stage at diagnosis [7]. They present with unique challenges—as younger patients with young children, heavy financial commitments, and fertility preservation [8]. The rise in the incidence of EOCRC may be a consequence of increased endoscopic capacity, which may lower the threshold for performing colonoscopy in younger patients. However, another important factor is the strong birth cohort effect observed worldwide, which may be the result of changes in exposure to risk factors. Although most studies have found EOCRC prevalence to be independent of gender, racial disparities in EOCRC incidence have been described in the United States, with a higher incidence seen in African Americans [9,10].

Singapore has a multiethnic resident population of 4 million, of which 74% are Chinese, 13% Malays, and 9% Indians [11]. In Singapore, CRC has the highest incidence among all cancers for the past several years, which, therefore, imposes a substantial burden in terms of complications, mortality, health care resources, and medical costs [12]. According to the *Singapore Cancer Registry Annual Report 2021*, there were a total of 12,239 new cases of CRC diagnosed from 2017 to 2021. There was a notable rise in the age-standardized incidence rate (ASIR) of CRC from 19.4 per 100,000 population in 1968-1972 to 37.5 per 100,000 population in 2017-2021 among men. A similar trend among females was observed, at 15.4 per 100,000 population in 1968-1972 to 26.9 per 100,000 population in 2017-2021. Among those aged 40-49 years, CRC was the most common cause of cancer among men and the fourth-most common cause of cancer among women. There were differing trends in the ASIRs of cancer among the 3 main ethnic groups. Although the Chinese had the highest ASIR of cancer throughout the years, there is evidence of a closing ethnic gap in cancer incidence between the Chinese and Malays [13]. Although there has been an overall increase in survival rates for CRC over the past decades, the age-specific patterns and effects of gender and ethnicity on the incidence trends in CRC in Singapore are not well studied. Therefore, the aim of this study was to perform a comprehensive analysis of temporal trends in colon and rectal cancer incidence in Singapore.

Methods

Study Design and Data Source

The Singapore Cancer Registry (SCR) is a population-based registry that was established in 1968 and is part of the National Registry of Disease Office. The SCR collects notifications from medical practitioners about all patients with cancer diagnosed in Singapore, as mandated by the National Registry of Diseases Act 2007. The data are internationally comparable and valid, with completeness of about 97.5%. Annual audits of the data are performed to ensure data accuracy of at least 95% as the data are critical to guide cancer prevention policies and cancer research [14].

Patients aged 20 years and older diagnosed with colon cancer (*International Classification of Diseases* [ICD]-9 code 153, ICD-10 code C18) and rectal cancer (ICD-9 codes 154.0, 154.1; ICD-10 codes C19.9, C20.9) from 1968 to 2019 from the SCR were included in this study. Recurrent and metachronous CRCs, premalignant tumors, carcinoma in situ, extranodal colorectal lymphoma, and sarcomas were excluded. Information about the Singapore population was obtained from the Department of Statistics Singapore (SingStat), which provides publicly available data that were used to define the denominator for incidence rates [11]. Patient characteristics included gender, ethnicity, and age at diagnosis. The tumor, node, and metastasis (TNM) stage at diagnosis was available for patients diagnosed from 2003 onward.

Colon cancers were separated from cancers of the rectosigmoid and rectum, which were grouped together as rectal cancer for analysis. Age-specific incidence rates (ASRs) were calculated as the number of new CRC cases diagnosed each year divided by the population at risk for that age group. The population at risk was defined as the group of individuals susceptible to CRC during the period of interest. ASIRs were calculated as the sum of weighted ASRs using the direct method and the Segi-Doll World Standard as the reference population.

Joinpoint Analysis

The cohort was initially divided into 3 age groups: 20-49, 50-64, and ≥65 years. EOCRC was defined as CRC diagnosed at age 20-49 years. Annual incidence rates from 1968 to 2019 were estimated for the 3 age groups. Joinpoint regression analysis was used to describe the temporal trends in cancer incidence, identify points of change in direction, and estimate the magnitude of the change in rates. Joinpoint analysis estimates the annual percentage change (APC) by assessing observed data and fitting a series of joined straight lines to determine the APC over each period. The changes are tested for statistical significance [15,16]. It assumes that the regression function is piecewise linear, and the segments are continuously connected at unknown change points [17]. Compared to the traditional Monte Carlo permutation method, the data-driven weighted Bayes information criterion (BIC) was used for model selection as it was computationally more efficient yet produced similar results [18].

Age-Period-Cohort Modeling

Birth cohort models were fitted using the National Cancer Institute (NCI) Age Period Cohort web tool [19]. Age-period-cohort modeling provides estimates of parameters that describe relationships between observed incidence rates and age, calendar period, and birth cohort based on age groups and time periods of equal length. There was no information available for those aged ≥70 years for several time points from SingStat population estimates. Therefore, age-period-cohort analysis was limited to those aged 20-69 years. Input data were CRC cases and population counts for 10 five-year time periods (1970-1974, 1975-1979, …, 2015-2019), 10 five-year age groups (20-24, 25-29, …, 65-69 years), and 19 birth cohorts generated starting from the 1905-1909 to the 1995-1999 birth cohort. Cohort effects were presented as incidence rate ratios (IRRs) for a given birth cohort, with the 1950-1954 as the reference birth cohort. Additionally, local drifts estimated the age-specific net APC in incidence rates across the different age groups.

Survival Analysis

Cancer-specific survival was defined as the period from the date of CRC diagnosis to the date of death due to CRC. Patients who died of other causes were censored at their date of death; similarly, patients who were still alive were censored at the date of last follow-up. Clinical characteristics, such as surgical, radiation, and chemotherapy treatments received, were available. Multivariable survival analysis was performed with the Cox proportional hazards model.

Ethical Considerations

The study was approved by the SingHealth Institutional Review Board (IRB 2022/2415) to access and analyze data from the SCR. Informed consent was waived (waiver of consent was approved by the SingHealth IRB) as study data were deidentified.

Results

Characteristics of the Study Population

In total, 53,044 CRCs were included, with 32,880 (62%) colon and 20,164 (38%) rectal cancers. Figure 1 describes the flowchart of identifying and extracting patients from the SCR. Furthermore, 6183 (11.7%) EOCRCs were diagnosed, with 3585 (58%) colon and 2598 (42%) rectal cancers. Clinical characteristics of the study population are described in Table 1. There was a higher proportion of stage 3 (37.4% vs 32.4%) and stage 4 (26.1% vs 24%) CRCs and a lower proportion of stage 2 (21.2% vs 28.1%) CRCs among EOCRCs compared to patients diagnosed at an age of 65 years and above (*P*<.001).

Figure 1. Flowchart of the patients identified and extracted from the SCR. ICD: *International Classification of Diseases*; SCR: Singapore Cancer Registry.

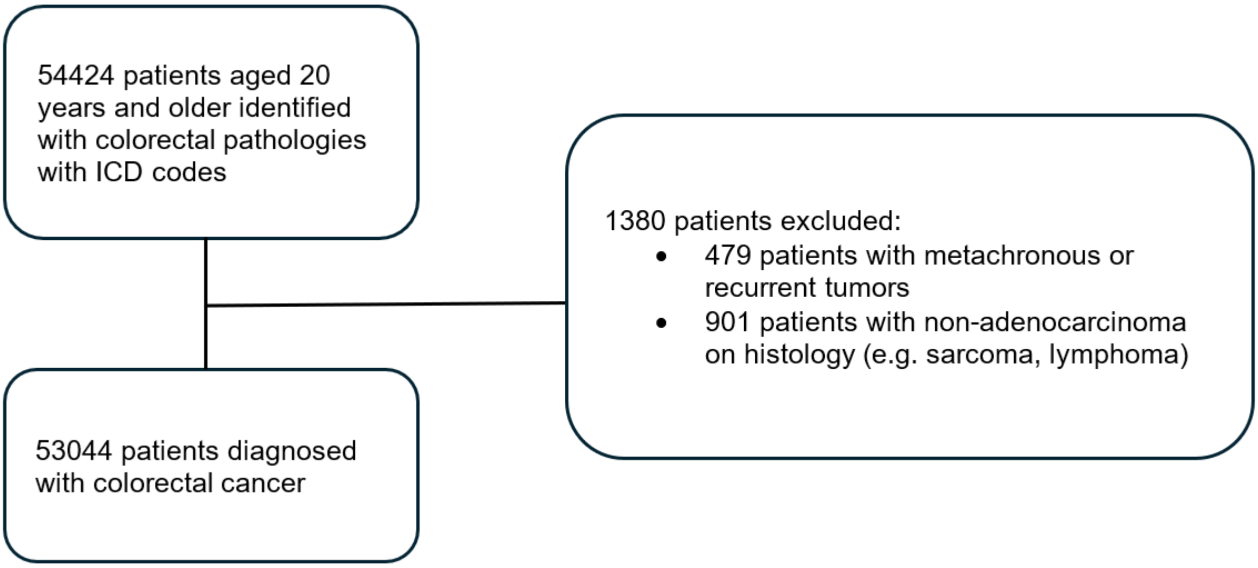


Table 1. Characteristics of CRCs[a] from the SCR[b].

Characteristics	Total cases (N=53,044), n (%)	Group 1 (20-49 years; n=6183), n (%)	Group 2 (50-64 years; n=17,862), n (%)	Group 3 (≥65 years; n=28,999), n (%)
Gender; *P*<.001				
Male	28,429 (53.6)	3235 (52.3)	10,245 (57.4)	14,949 (51.6)
Female	24,615 (46.4)	2948 (47.7)	7617 (42.6)	14,050 (48.4)
Ethnicity; *P*<.001				
Chinese	46,831 (88.3)	5167 (83.6)	15,495 (86.7)	26,169 (90.2)
Indian	1767 (3.3)	305 (4.9)	671 (3.8)	791 (2.7)
Malay	3723 (7.0)	583 (9.4)	1448 (8.1)	1692 (5.8)
Others	723 (1.4)	128 (2.1)	248 (1.4)	347 (1.2)
Tumor site; *P*<.001				
Colon	32,880 (62)	3585 (58.0)	10,332 (57.8)	18,963 (65.4)
Rectum	20,164 (38)	2598 (42.0)	7530 (42.2)	10,036 (34.6)
TNM[c] stage (2003 and beyond); *P*<.001				
1	4436 (15.9)	397 (15.3)	1643 (16.6)	2396 (15.6)
2	7178 (25.7)	551 (21.2)	2299 (23.2)	4328 (28.1)
3	9495 (34)	969 (37.4)	3534 (35.6)	4992 (32.4)
4	6830 (24.4)	677 (26.1)	2448 (24.7)	3705 (24.0)

[a] CRC: colorectal cancer.

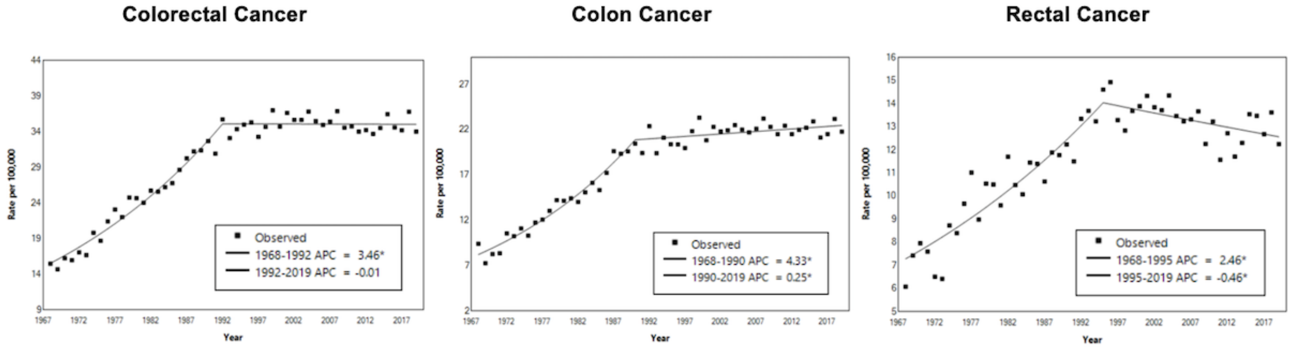
[b] SCR: Singapore Cancer Registry.

[c] TNM: tumor, node, and metastasis.

Joinpoint Regression Analysis of Incidence Rates of Colorectal, Colon, and Rectal Cancers

The ASIR of CRC rose from 15 in 1968 to 36 per 100,000 population in 1992 at an APC of 3.5% and plateaued (Figure 2). The ASIR of colon cancer rose from 8 in 1968 to 20 per 100,000 population in 1990 at an APC of 4.3% and plateaued. The ASIR of rectal cancer rose from 7 in 1968 to 14 per 100,000 population in 1990 at an APC of 2.5% and decreased to 12 per 100,000 population in 2019 at an APC of –0.5%.

Figure 2. Joinpoint regression of the ASIR of CRC, colon cancer, and rectal cancer. APC: annual percentage change; ASIR: age-standardized incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



The results of joinpoint regression are shown in Table 2 and Figure 3 for CRC, colon cancer, and rectal cancer stratified by age group. For EOCRC, the ASR rose from 5 in 1968 to 9 in 1997 at an APC of 2% and increased to 10 per 100,000 population in 2019 at an APC of 0.6%. For those aged 50-64 years, trends were similar to the ASIR of CRC for the overall Singapore population. For those aged 65 years and above, the ASR of CRC rose from 113 in 1968 to 299 in 2002 at an APC of 4.11% and decreased to 242 per 100,000 population in 2019 at an APC of –1.03%.

The ASR of early-onset colon cancer rose from 3 in 1968 to 5 per 100,000 population in 1995 at an APC of 2.2% and thereafter had a nonsignificant increase at an APC of 0.4%. For those aged 50-64 years, colon cancer trends were similar to the ASIR of CRC for the overall Singapore population. For those aged 65 years and above, the ASR of colon cancer rose from 57 in 1968 to 193 in 2005 at an APC of 4.95% and decreased to 170 per 100,000 population in 2019 at an APC of –0.92%.

The ASR of early-onset rectal cancer rose from 2 in 1968 to 5 per 100,000 population in 2019 at an APC of 1.5% per year. For those aged 50-64 years and 65 years and above diagnosed with rectal cancer, trends were similar to the ASIR of rectal cancer for the overall Singapore population (Table 2 and Figure 3).

Table 2. APC[a] of CRC[b], colon cancer, and rectal cancer ASRs[c] by age group and tumor location. See Multimedia Appendix 1 for a high-resolution version.

Type of cancer and age groups	Patients, n (%)	Trend 1		Trend 2		Trend 3	
		Years	APC, %	Years	APC, %	Years	APC, %
CRC							
20-49 years	6183 (11.6)	1968-1996	2.07[d]	1996-2019	0.64[d]	—[e]	—
50-64 years	17,862 (33.7)	1968-1987	2.97[d]	1987-2019	0.09	—	—
≥65 years	28,999 (54.7)	1968-1989	4.11[d]	1989-2003	1.30[d]	2003-2019	–1.03[f]
Colon cancer							
20-49 years	3585 (10.9)	1968-1995	2.16[d]	1995-2019	0.37	—	—
50-64 years	10,332 (31.4)	1968-1986	3.45[d]	1986-2019	0.29[d]	—	—
≥65 years	18,963 (57.7)	1968-1989	4.95[d]	1989-2005	1.30[d]	2005-2019	–0.92[f]
Rectal cancer							
20-49 years	2598 (12.9)	1968-2019	1.49[d]	—	—	—	—
50-64 years	7530 (37.3)	1968-1992	1.97[d]	1992-2019	–0.34[f]	—	—
≥65 years	10,036 (49.8)	1968-1996	2.67[d]	1996-2019	–1.15[f]	—	—

[a] APC: annual percentage change.

[b] CRC: colorectal cancer.

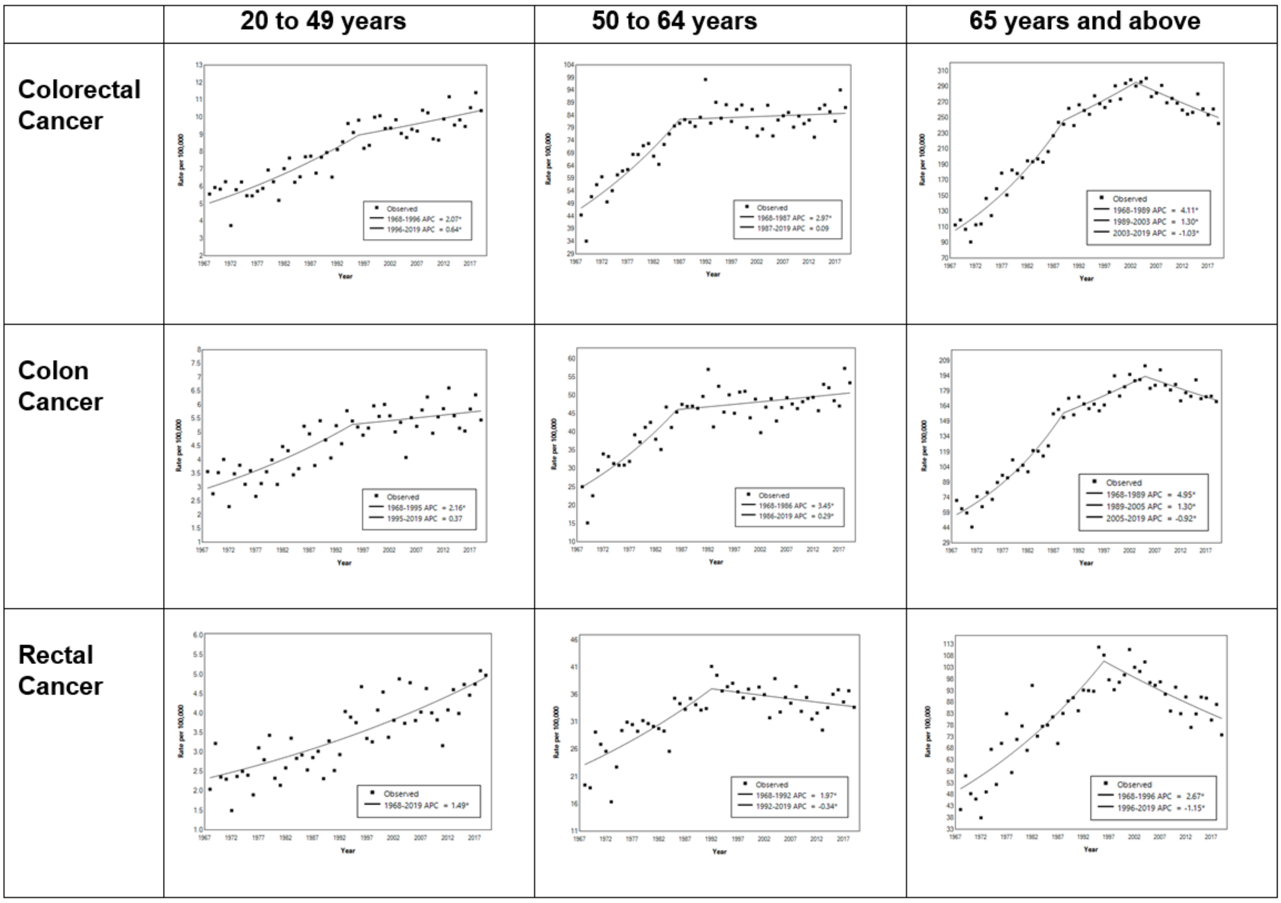
[c] ASR: age-specific incidence rate.

[d] Statistically significant positive APCs.

[e] Not applicable.

[f] Statistically significant negative APCs.

Figure 3. Joinpoint regression of the ASRs of CRC, colon cancer, and rectal cancer by age group and tumor location. APC: annual percentage change; ASR: age-specific incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



Effect of Gender

The male population had higher incidence rates for colon and rectal cancers. Similar to the ASIR, for both genders aged 50-64 years and 65 years and above, the incidence of CRC, colon cancer, and rectal cancer rose before it plateaued or decreased in the recent 2 decades (Table 3 and Figures 4 and 5). Among those aged 50-64 years, there was a rapid initial rise in colon cancer among women compared to men (APC 5% vs 3.8%).

For males aged 20-49 years, there was a rising incidence in CRC (APC 1.5%), mainly rectal cancer (APC 1.9%), with a smaller rise in colon cancer (APC 1.1%). For females aged 20-49 years, similar to older females aged 50-64 years, there was a rapid initial rise, particularly in colon cancer (APC 3.8%); however, in recent years, for colon cancer, rectal cancer, and CRC, the APC has been <1% (Table 3 and Figure 5).

Table 3. APC[a] of CRC[b], colon cancer, and rectal cancer ASRs[c] by age group, tumor location, and gender.

Type of cancer and age groups	Patients, n (%)	Trend 1		Trend 2		Trend 3	
		Years	APC, %	Years	APC, %	Years	APC, %
CRC (20-49 years)							
Male	3235 (52.3)	1968-2019	1.48[d]	—[e]	—	—	—
Female	2948 (47.7)	1968-1994	2.10[d]	1994-2019	0.41	—	—
CRC (50-64 years)							
Male	10,245 (57.4)	1968-1989	3.34[d]	1989-2019	0.11	—	—
Female	7617 (42.6)	1968-1980	4.20[d]	1980-2019	0.04	—	—
CRC (≥65 years)							
Male	14,949 (51.6)	1968-1989	3.76[d]	1989-2003	1.79	2003-2019	–0.95
Female	14,050 (48.4)	1968-1981	6.18[d]	1981-1999	2.17[d]	1999-2019	–1.01[f]
Colon cancer (20-49 years)							
Male	1715 (47.8)	1968-2019	1.10[d]	—	—	—	—
Female	1870 (52.2)	1968-1986	3.81[d]	1986-2019	0.55[d]	—	—
Colon cancer (50-64 years)							
Male	5471 (53.0)	1968-1988	3.76[d]	1988-2019	0.34	—	—
Female	4861 (47.0)	1968-1980	4.99[d]	1980-2019	0.19	—	—
Colon cancer (≥65 years)							
Male	9289 (49.0)	1968-1988	4.92[d]	1988-2005	1.64	2005-2019	–0.77
Female	9674 (51.0)	1968-1989	5.04[d]	1989-2004	1.18[d]	2004-2019	–1.04[f]
Rectal cancer (20-49 years)							
Male	1520 (58.5)	1968-2019	1.91[d]	—	—	—	—
Female	1078 (41.5)	1968-2019	0.93[d]	—	—	—	—
Rectal cancer (50-64 years)							
Male	4774 (63.4)	1968-1989	2.61[d]	1992-2019	–0.27	—	—
Female	2756 (36.6)	1968-2019	0	—	—	—	—
Rectal cancer (≥65 years)							
Male	5660 (56.4)	1968-2002	2.18[d]	2002-2019	–1.59[f]	—	—
Female	4376 (43.6)	1968-1995	2.69[d]	1995-2004	–1.55[f]	—	—

[a] APC: annual percentage change.

[b] CRC: colorectal cancer.

[c] ASR: age-specific incidence rate.

[d] Statistically significant positive APCs.

[e] Not applicable.

[f] Statistically significant negative APCs.

Figure 4. Joinpoint regression of ASRs of CRC, colon cancer, and rectal cancer by gender (male: red; female: blue). APC: annual percentage change; ASR: age-specific incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.

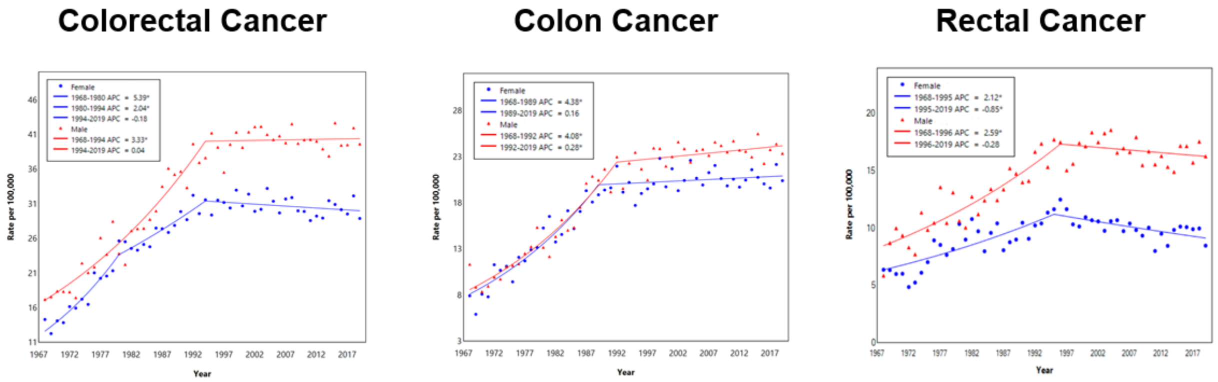
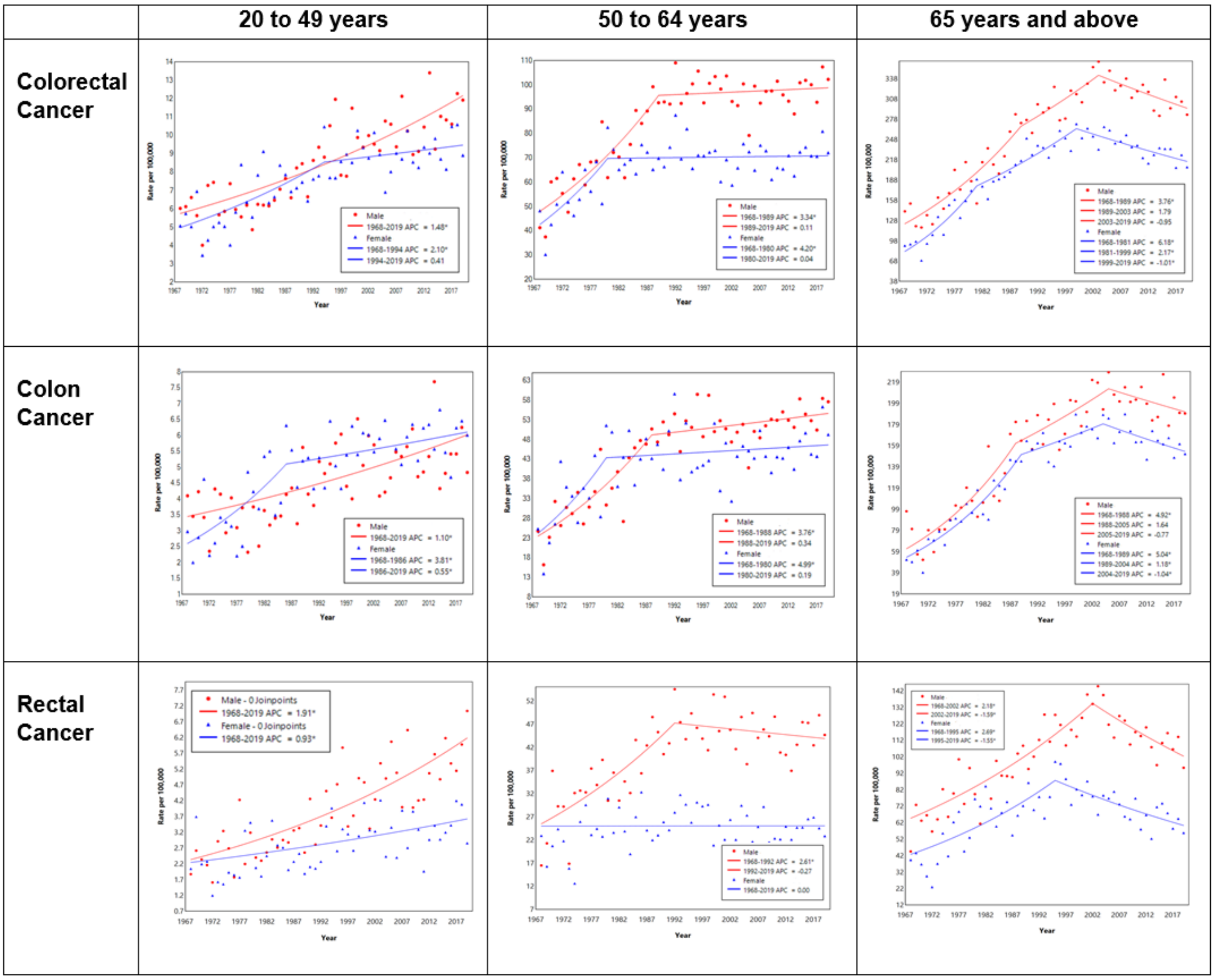


Figure 5. Joinpoint regression of ASRs of CRC, colon cancer, and rectal cancer by age group and gender (male: red; female: blue). APC: annual percentage change; ASR: age-specific incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



Effect of Ethnicity

Similar to the ASIR, for Chinese patients aged 50-64 years and 65 years and above, the ASR of CRC, colon cancer, and rectal cancer showed an initial rise but plateaued or decreased in the recent 2 decades. In Malays across all 3 age groups, there was a persistent rise in the ASR of CRC, colon cancer, and rectal cancer (APC 1.4%-3.2%) without evidence of it plateauing. In Indians aged 50-64 years and 65 years and above, there was a smaller persistent rise in the ASR of CRC, colon cancer, and rectal cancer (APC 1%-1.4%) without evidence of it plateauing (Tables 4-6 and Figures 6 and 7).

Table 4. APC[a] of CRC[b], colon cancer, and rectal cancer ASRs[c] for patients diagnosed at age 20-49 years by tumor location and ethnicity.

Tumor location and ethnicity	Patients, n (%)	Trend 1		Trend 2		Trend 3	
		Years	APC, %	Years	APC, %	Years	APC, %
CRC							
Chinese	5167 (85.3)	1968-1999	1.96[d]	1999-2019	0.23	—[e]	—
Indian	305 (5.1)	1968-2019	0.47	—	—	—	—
Malay	583 (9.6)	1968-2019	2.90[d]	—	—	—	—
Colon cancer							
Chinese	2992 (85.4)	1968-1999	1.75[d]	1999-2019	–0.09	—	—
Indian	164 (4.7)	1968-2019	0.03	—	—	—	—
Malay	346 (9.9)	1968-2019	1.96[d]	—	—	—	—
Rectal cancer							
Chinese	2175 (85.2)	1968-2019	1.48[d]	—	—	—	—
Indian	141 (5.5)	1968-2019	0.45	—	—	—	—
Malay	237 (9.3)	1968-2019	1.96[d]	—	—	—	—

[a] APC: annual percentage change.

[b] CRC: colorectal cancer.

[c] ASR: age-specific incidence rate.

[d] Statistically significant positive APCs.

[e] Not applicable.

Table 5. APC[a] of CRC[b], colon cancer, and rectal cancer ASRs[c] for patients diagnosed at age 50-64 years by tumor location and ethnicity.

Tumor location and ethnicity	Patients, n (%)	Trend 1		Trend 2		Trend 3	
		Years	APC, %	Years	APC, %	Years	APC, %
CRC							
Chinese	15,495 (88.0)	1968-1985	3.45[d]	1985-2019	–0.12	—[e]	—
Indian	671 (3.8)	1968-2019	0.95[d]	—	—	—	—
Malay	1448 (8.2)	1968-2019	2.30[d]	—	—	—	—
Colon cancer							
Chinese	9012 (88.4)	1968-1981	5.57[d]	1981-2019	0.13	—	—
Indian	356 (3.5)	1968-2019	1.22[d]	—	—	—	—
Malay	823 (8.1)	1968-2019	2.69[d]	—	—	—	—
Rectal cancer							
Chinese	6483 (87.3)	1968-1988	2.31[d]	1988-2019	–0.42[f]	—	—
Indian	315 (4.3)	1968-2019	0.45	—	—	—	—
Malay	625 (8.4)	1968-2019	1.42[d]	—	—	—	—

[a] APC: annual percentage change.

[b] CRC: colorectal cancer.

[c] ASR: age-specific incidence rate.

[d] Statistically significant positive APCs.

[e] Not applicable.

[f] Statistically significant negative APCs.

Table 6. APC[a] of CRC[b], colon cancer, and rectal cancer ASRs[c] for patients diagnosed at age ≥65 years by tumor location and ethnicity.

Tumor location and ethnicity	Patients, n (%)	Trend 1		Trend 2		Trend 3	
		Years	APC, %	Years	APC, %	Years	APC, %
CRC							
Chinese	26,169 (91.3)	1968-1990	4.32[d]	1990-2002	1.35[d]	2002-2019	–1.47[e]
Indian	791 (2.8)	1968-2019	1.32[d]	—[f]	—	—	—
Malay	1692 (5.9)	1968-2019	2.97[d]	—	—	—	—
Colon cancer							
Chinese	17,254 (92.0)	1968-1990	5.21[d]	1990-2003	1.35[d]	2003-2019	–1.26[e]
Indian	483 (2.6)	1968-2019	1.38[d]	—	—	—	—
Malay	1006 (5.4)	1968-2019	3.24[d]	—	—	—	—
Rectal cancer							
Chinese	8915 (90.0)	1968-1996	2.93[d]	1996-2019	–1.60[e]	—	—
Indian	308 (3.1)	1968-2019	0.99[d]	—	—	—	—
Malay	686 (6.9)	1968-2019	1.88[d]	—	—	—	—

[a] APC: annual percentage change.

[b] CRC: colorectal cancer.

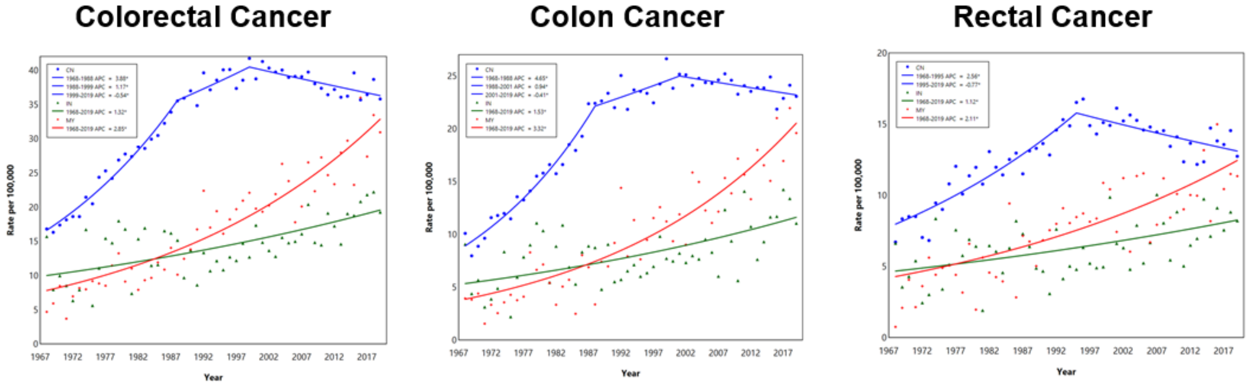
[c] ASR: age-specific incidence rate.

[d] Statistically significant positive APCs.

[e] Statistically significant negative APCs.

[f] Not applicable.

Figure 6. Joinpoint regression of ASIRs of CRC, colon cancer, and rectal cancer by ethnic group (Chinese/CN: blue; Indian/IN: green; Malay/MY: red). APC: annual percentage change; ASIR: age-standardized incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



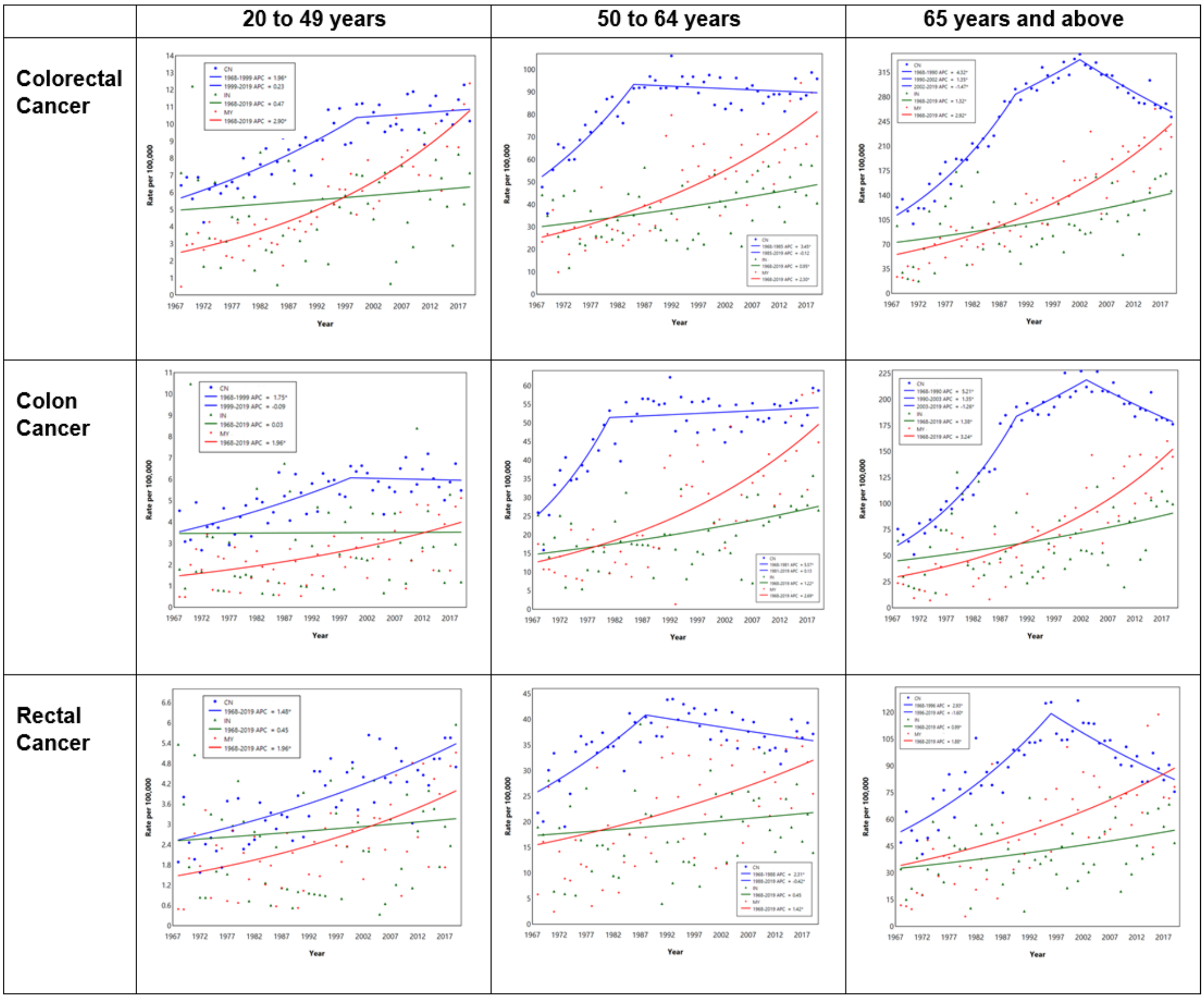
Colorectal Cancer

Colon Cancer

Rectal Cancer

Malay patients aged 20-49 years had the highest rise in the ASR of CRC (APC 2.9%), colon cancer (APC 2%), and rectal cancer (APC 2%). Among Chinese patients aged 20-49 years, the ASR of rectal cancer persistently rose (APC 1.5%), while the ASR of CRC and colon cancer plateaued from 1999. In Indians aged 20-49 years, there was a small and nonsignificant rise in the ASR of CRC (APC 0.5%) and rectal cancer (APC 0.5%), as shown in Tables 4-6 and Figure 7.

Figure 7. Joinpoint regression of ASRs of CEC, colon cancer, and rectal cancer by age group, tumor location, and ethnic group (Chinese/CN: blue; Indian/IN: green; Malay/MY: red). APC: annual percentage change; ASR: age-specific incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



Age-Period-Cohort Model of IRRs and APCs for CRC, Colon Cancer, and Rectal Cancer

Table 7 and Figure 8 show the IRRs by birth cohort for CRC, colon cancer, and rectal cancer. Compared to the reference birth cohort of 1950-1954, there was a trend toward higher IRRs for colon cancer among the 1970-1990 birth cohorts; however, these changes were not statistically significant. In addition, compared to the reference birth cohort of 1950-1954, the IRRs for rectal cancer among the birth cohorts from 1905 onward were marginally lower. However, there was significant increase in the IRR for rectal cancer among the birth cohorts from 1970 onward, and this peaked in the 1980-1984 birth cohort (IRR 1.35, 95% CI 1.043-1.748). Subsequent birth cohorts from 1985 onward had similar IRRs as the reference birth cohort (Figure 7 and Table 7).

Figure 9 shows the net age-specific APC or local drift for CRC, colon cancer, and rectal cancer incidence rates. There was an increase in the incidence rates for CRC among those aged 25-69 years, with an average APC of 0.61% (95% CI 0.44-0.79) for all age groups over the entire period. There was an increase in incidence rates for colon cancer among those aged 40-69 years, while there was an increase in incidence rates for rectal cancer among those aged 35-69 years.

Table 7. IRRs[a] for CRC[b], colon cancer, and rectal cancer by birth cohort.

Birth cohort	CRC IRR (95% CI)	Colon cancer IRR (95% CI)	Rectal cancer IRR (95% CI)
1905-1909	0.516 (0.435-0.613)	0.434 (0.345-0.547)	0.660 (0.503-0.867)
1910-1914	0.697 (0.633-0.768)	0.648 (0.570-0.736)	0.781 (0.668-0.912)
1915-1919	0.763 (0.706-0.825)	0.696 (0.626-0.771)	0.867 (0.765-0.983)
1920-1924	0.761 (0.708-0.817)	0.745 (0.679-0.817)	0.784 (0.696-0.883)
1925-1929	0.856 (0.802-0.912)	0.846 (0.779-0.920)	0.873 (0.784-0.972)
1930-1934	0.889 (0.837-0.943)	0.863 (0.799-0.933)	0.929 (0.842-1.025)
1935-1939	0.939 (0.889-0.993)	0.916 (0.852-0.984)	0.977 (0.892-1.070)
1940-1944	0.948 (0.901-0.997)	0.937 (0.878-1.000)	0.966 (0.889-1.050)
1945-1949	0.970 (0.926-1.015)	0.956 (0.901-1.015)	0.988 (0.915-1.067)
1950-1954	1.000 (1.000)	1.000 (1.000)	1.000 (1.000)
1955-1959	1.030 (0.986-1.075	1.047 (0.991-1.107)	1.010 (0.940-1.086)
1960-1964	1.006 (0.956-1.059))	0.989 (0.925-1.057)	1.030 (0.947-1.120)
1965-1969	1.063 (0.998-1.133)	1.047 (0.964-1.137)	1.086 (0.978-1.205)
1970-1974	1.091 (1.002-1.187)	1.015 (0.907-1.136)	1.199 (1.045-1.375)
1975-1979	1.120 (0.999-1.255)	1.016 (0.873-1.182)	1.28 (1.067-1.541)
1980-1984	1.172 (1.001-1.372)	1.057 (0.858-1.302)	1.35 (1.043-1.748)
1985-1989	1.184 (0.940-1.491)	1.251 (0.947-1.652)	0.971 (0.623-1.512)
1990-1994	1.266 (0.893-1.795)	1.269 (0.841-1.914)	1.031 (0.511-2.080)
1995-1999	0.967 (0.517-1.807)	0.951 (0.443-2.040)	0.884 (0.280-2.792)

[a] IRR: incidence rate ratio.

[b] CRC: colorectal cancer.

Figure 8. IRRs by birth cohort for CRC, colon cancer, and rectal cancer. CRC: colorectal cancer; IRR: incidence rate ratio. See Multimedia Appendix 1 for a high-resolution version.

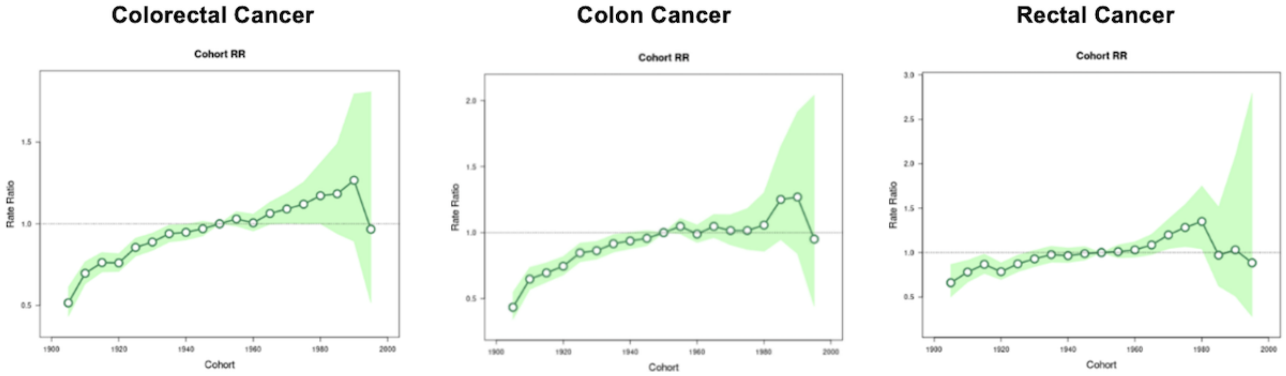
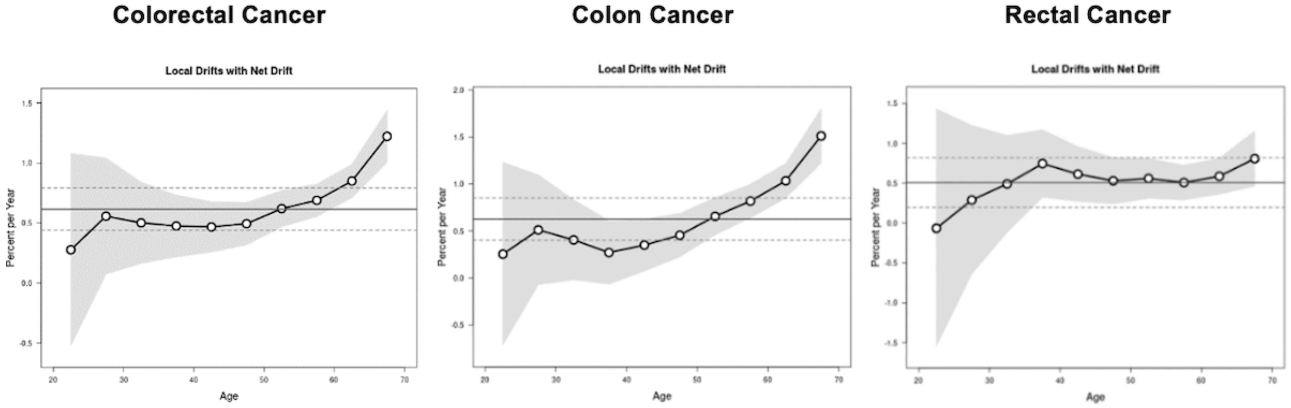


Figure 9. Age-specific APC (local drift) for CRC, colon cancer, and rectal cancer. APC: annual percentage change; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



Survival Outcomes

Patients with EOCRC had better cancer-specific survival compared to patients diagnosed at 65 years and above (hazard ratio [HR] 0.73, 95% CI 0.67-0.79, *P*<.001) after adjusting for the TNM stage, gender, tumor site, surgery, radiotherapy, and chemotherapy (Table 8).

Table 8. Multivariable cancer-specific survival analysis of patients with CRC[a] using the Cox proportional hazards model.

Variables and groups	HR[b] (95% CI)	*P* value
Age (years)		
20-49 vs ≥65	0.73 (0.67-0.79)	<.001
50-64 vs ≥65	0.74 (0.70-0.78)	<.001
TMN[c] stage		
2 vs 1	3.84 (3.23-4.56)	<.001
3 vs 1	11.14 (9.43-13.16)	<.001
4 vs 1	58.70 (49.66-69.38)	<.001
Gender: male vs female	1.06 (1.02-1.11)	.006
Tumor site: rectum vs colon	0.55 (0.52-0.58)	<.001
Treatment		
Surgery vs no surgery	0.43 (0.41-0.46)	<.001
Radiotherapy vs no radiotherapy	1.10 (1.01-1.19)	<.026
Chemotherapy vs no chemotherapy	0.54 (0.51-0.57)	<.001

[a] CRC: colorectal cancer.

[b] HR: hazard ratio.

[c] TNM: tumor, node, and metastasis.

Discussion

Principal Findings

Historically, CRC has been more commonly associated with older age groups, but in recent years, many countries have reported a concerning rise in cases among younger individuals. Although the overall incidence of CRC has been declining in older people, there has been a notable increase in the incidence of EOCRC, particularly in individuals younger than 50 years. In the United States, the overall annual, CRC ASIR has decreased by 46% from 66.2 per 100,000 population in 1985 to 35.7 per 100,000 population in 2019. Notably, CRC is now the most common cause of cancer deaths among individuals less than 50 years old [20]. However, from 2011 through 2019, the CRC incidence rate increased by 1.9% per year for individuals younger than 50 years. Rectal cancer has driven the EOCRC epidemiological trend with an increase of 2% per year. There was a steep decline in incidence among adults aged 50 years and older in the United States at 3%-5% annually in the late 2000s, and this is likely attributed to changing patterns in the risk factors and uptake of CRC screening. Incidence rates have been increasing in adults aged 20-39 years since the mid-1980s and in those aged 40-54 years since the mid-1990s [21]. Similar trends are seen in Canada, the United Kingdom, Australia, New Zealand, Europe, and Asia [5,22-25]. The rising incidence of EOCRC in East Asia may be due to an increase in the prevalence of risk factors, such as smoking, alcoholism, and obesity among young individuals [26]. A population-based study in Taiwan, Korea, Japan, and Hong Kong suggests that from 1995 to 2014, the pooled incidence of early-onset colon and rectal cancers increased by 2.1% and 3.8%, respectively, among men. Similarly, the pooled incidence of early-onset colon and rectal cancers increased by 2.2% and 4%, respectively, among women [27]. In China, the incidence rate of EOCRC rose from 3.6 to 12.1 per 100,000 population from 1990 to 2019 at a rate of 4.6% per year [25].

Our analysis of incidence trends over time suggests that the rising incidence of EOCRC in Singapore is nuanced. The rising trend for people under 50 years of age is particularly for rectal cancer, especially in males. This trend also affects, in particular, the Malays but also the Chinese, with the Indians less affected. Our study supports the increasing literature that the increase in EOCRC is not confined to the West and may be a worldwide phenomenon. The cause of these observations is still unclear, but it can be explained by the cohort effects. The IRR for rectal cancer has been increasing among the 1970-1984 birth cohorts and suggests that it is the changes in exposure to these birth cohorts that influence the risk of carcinogenesis. High consumption of processed meat, alcohol, a sedentary lifestyle, a lack of dietary fiber consumption, smoking, and obesity are known lifestyle factors associated with CRC. However, these modifiable risk factors are present in all age groups and not specific to young individuals and cannot explain fully the rising incidence among young people [28,29]. It has also been reported that there are similar rates of obesity between EOCRC and older patients with CRC [30]. The increasing prevalence of the use of antibiotics and alterations in the gut microbiome may play a role in the development of CRC through effects of bacterial-derived metabolites and virulence factors. Results of a recent study show an increase in *Fusobacterium* spp. in the microbiome of EOCRC [31]. Understanding the generational shifts in early life exposures may provide further insights into the etiologies of EOCRC [7].

Most tumors in EOCRC are in the sigmoid colon and rectum [32]. The reasons regarding the steep rise in the incidence of rectal compared to colon cancer among young adults are unclear. Understanding the differences in the molecular profile between rectal and colon cancers may be key toward developing preventative strategies. EOCRC in young adults has been associated with synchronous and metachronous tumors, a more poorly differentiated histology, and mucinous and signet ring histology compared to older patients [7,33]. The rates of high microsatellite instability (MSI-H) are higher in EOCRC due to the higher incidence of Lynch syndrome. However, among the microsatellite-stable (MSS) cohort of metastatic CRC, patients with EOCRC have lower rates of *KRAS* and *BRAF* V600 mutations [31]. These observations suggest that the underlying biological process and basis of these clinicopathological conditions in EOCRC are different from those found in people above 50 years old.

The male gender is associated with a higher incidence of CRC compared to females. The incidence of rectal cancer is higher among young males compared to females. In 2019, the incidence rate of early-onset rectal cancer was 6.2 per 100,000 population among men, while it was 3.6 per 100,000 population among women. The higher proportion of men diagnosed with rectal cancer suggests a differential distribution of the tumor location between genders. Although the etiology between the gender differences is uncertain, it may be attributed to the varied exposure in dietary and lifestyle risk factors between both gender groups [34]. Some studies have explored the protective effect of estrogens and progestins against CRC in hormone replacement therapy in postmenopausal women [35,36].

The Chinese population in Singapore had the highest ASIR of CRC when compared to Indians and Malays. Previously, the Asia Pacific Consensus identified Chinese as an ethnic group that is more susceptible to CRC, and this was supported by cohort studies in Singapore and Malaysia [37-40]. In this study, although the incidence initially increased among the Chinese, the incidence rates have plateaued and declined over the past decade, except for early-onset rectal cancer, which continues to rise at 1.5% per year among young Chinese adults. However, the steep increase in the incidence of both colon and rectal cancers among the Malays across all age groups is concerning. This observation is supported by another study of age-period-cohort analyses in the SCR, stratified by ethnicity, that suggests a rising risk of CRC across all periods among the Malay population [41]. The rising incidence of EOCRC predominantly among Malays was also observed in Malaysia [42]. The varied epidemiologic trends across the 3 ethnic groups may be attributed to the differences in health-seeking behavior and lifestyle exposures between the groups in Singapore [41]. Modifiable risk factors, such as the rising rates and high prevalence of obesity among the Malays and the increased consumption of sweetened food, may play a role [43]. Malays were also less likely to participate in cancer screening compared to the Chinese. There is a need for tailored cancer-screening promotion campaigns to narrow the knowledge-behavior gap among the Malays [44]. Socioeconomic indicators, such as educational attainment, homeownership rates, and income levels, have shown improvement across all racial groups in Singapore [45]. Although there is a relatively similar socioeconomic environment in Singapore, the varying rates of the rise of EOCRC among the 3 races suggest that cultural, dietary, or genetic factors may potentially contribute to these differences. Further studies should be conducted that compare the lifestyle factors and social practices among the 3 ethnic groups and evaluate the impact on CRC incidence.

Our study suggests that patients with EOCRC have better cancer-specific survival compared to patients with CRC aged 65 years and older. However, there are conflicting data with regard to the survival outcomes of EOCRC compared to patients with late-onset CRC after adjustment for stage. A survival analysis of 35,713 patients with stage 3 colon cancer from the ACCENT database revealed that patients with EOCRC have improved overall survival, disease-free survival, and survival after recurrence. After adjusting for molecular markers (MMR, KRAS, BRAF), the prognostic impact of age of onset was lost [46]. However, there was a poorer 3-year relapse-free survival among patients with stage 3 EOCRC based on the pooled analyses of 16,349 patients from the International Duration Evaluation of Adjuvant Chemotherapy (IDEA) database. This was despite better treatment adherence and higher administered treatment intensity among the younger patients, hence suggestive of a more aggressive disease biology [47].

The current screening guidelines for CRC in Singapore start at 50 years of age for asymptomatic, average-risk individuals and have no impact on the incidence of early-onset CRC. Recently, the American Cancer Society recommended initiating CRC screening at age 45 years instead of 50 years for average-risk individuals to address the rising CRC incidence among young adults [48]. This recommendation was based on modeling data that suggest that this will prevent 29,400 CRC cases and 11,000 CRC deaths in the United States over the next 5 years [49]. Compared to the rising incidence of colon and rectal cancers in the United States, Europe, and Australia, our study suggests that rectal cancer among young adults is rising. Therefore, it is important for physicians to investigate young adults who present with anorectal symptoms, such as rectal bleeding, tenesmus, anal pain, and changes in bowel habits. Although these findings may suggest a need to re-evaluate the recommended age range for screening and consider extending screening to those 45-49 years of age, the incidence of EOCRC in Singapore remains low, and this strategy may not be cost effective. Despite the launch of the National Colorectal Screening Programme in 2011, the uptake of CRC screening remains low nationally at 41.7% [50,51]. However, flexible sigmoidoscopy among young adults is a less costly alternative as compared to colonoscopy and may capture the majority of EOCRCs since the tumor location of EOCRC is usually in the sigmoid colon or the rectum.

Strengths and Limitations

Although this study provides a comprehensive overview of CRC trends in Singapore, further research into the biological basis of carcinogenesis among young adults and differences in modifiable risk factors among ethnic groups is warranted. Age-period-cohort analysis was limited due to missing population data of those aged 70 years and above. Additionally, the study was limited due to the lack of information about the proportion of hereditary cancer syndromes in the cohort. The prevalence of hereditary cancer syndromes among EOCRCs ranges from 5% to 35% compared to 2%-5% for CRCs overall [52-55]. Although genetic predisposition may be associated with EOCRC, the majority of EOCRCs are still sporadic. Additionally, screening of families with hereditary cancer syndromes may only capture a small proportion of EOCRCs [52]. Future research should focus on identification of risk factors and predictors of EOCRC to determine high-risk groups that should be targeted for screening before 50 years of age. It may be a more viable strategy to develop risk stratification tools that combine family history and blood-based tests to identify high-risk neoplasia among young adults.

Conclusion

The incidence of rectal cancer among people under 50 years of age in Singapore appears to be increasing, especially among the Chinese and Malays. The differences in the epidemiologic trends among the 3 main ethnic groups may be due to varied lifestyle factors and social practices. Although it is possible that the observed increase may be attributable to lifestyle changes in the population over time, further studies are required to determine the underlying etiology. In contrast, CRC rates are decreasing among those above 65 years of age, and this may be due to the impact of the national screening program. These findings suggest the need for further research to diagnose CRC earlier among young adults and reduce the associated cancer-related morbidity and mortality.

Acknowledgments

The authors would like to acknowledge the National Registry of Diseases Office, Health Promotion Board, Singapore, for providing access to colorectal cancer data from the Singapore Cancer Registry. CHLR is supported by the Research Training Fellowship (MOH-001300-00) from the National Medical Research Council Singapore. The work was also supported by the Goh Hak Su Professorship in Colorectal Surgery.

Data Availability

Data collected from the Singapore Cancer Registry are not available publicly.

Authors' Contributions

The authors attest that there was no use of generative artificial intelligence in the generation of text, figures, or other content of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
High-resolution versions of Figures 2-9.
[PDF File (Adobe PDF File), 1739 KB - publichealth_v11i1e62835_app1.pdf]

References

1. Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. Transl Oncol 2021 Oct;14(10):101174 [FREE Full text] [doi: 10.1016/j.tranon.2021.101174] [Medline: 34243011]

2. Rabeneck L, Chiu H, Senore C. International perspective on the burden of colorectal cancer and public health effects. Gastroenterology 2020 Jan;158(2):447-452. [doi: 10.1053/j.gastro.2019.10.007] [Medline: 31622620]
3. Sung H, Siegel RL, Laversanne M, Jiang C, Morgan E, Zahwe M, et al. Colorectal cancer incidence trends in younger versus older adults: an analysis of population-based cancer registry data. Lancet Oncol 2024 Dec 11:S1470-2045 [FREE Full text] [doi: 10.1016/S1470-2045(24)00600-4] [Medline: 39674189]
4. Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2023. CA Cancer J Clin 2023;73(3):233-254 [FREE Full text] [doi: 10.3322/caac.21772] [Medline: 36856579]
5. Feletto E, Yu XQ, Lew J, St John DJB, Jenkins MA, Macrae FA, et al. Trends in colon and rectal cancer incidence in Australia from 1982 to 2014: analysis of data on over 375,000 cases. Cancer Epidemiol Biomarkers Prev 2019 Jan;28(1):83-90. [doi: 10.1158/1055-9965.EPI-18-0523] [Medline: 30530848]
6. Patel SG, Karlitz JJ, Yen T, Lieu CH, Boland CR. The rising tide of early-onset colorectal cancer: a comprehensive review of epidemiology, clinical features, biology, risk factors, prevention, and early detection. Lancet Gastroenterol Hepatol 2022 Mar;7(3):262-274. [doi: 10.1016/s2468-1253(21)00426-x]
7. Akimoto N, Ugai T, Zhong R, Hamada T, Fujiyoshi K, Giannakis M, et al. Rising incidence of early-onset colorectal cancer - a call to action. Nat Rev Clin Oncol 2021 Apr;18(4):230-243 [FREE Full text] [doi: 10.1038/s41571-020-00445-1] [Medline: 33219329]
8. Lieu CH, Golemis EA, Serebriiskii IG, Newberg J, Hemmerich A, Connelly C, et al. Comprehensive genomic landscapes in early and later onset colorectal cancer. Clin Cancer Res 2019 Oct 01;25(19):5852-5858 [FREE Full text] [doi: 10.1158/1078-0432.CCR-19-0899] [Medline: 31243121]
9. Saraiva MR, Rosa I, Claro I. Early-onset colorectal cancer: a review of current knowledge. World J Gastroenterol 2023 Feb 28;29(8):1289-1303 [FREE Full text] [doi: 10.3748/wjg.v29.i8.1289] [Medline: 36925459]
10. Pan H, Zhao Z, Deng Y, Zheng Z, Huang Y, Huang S, et al. The global, regional, and national early-onset colorectal cancer burden and trends from 1990 to 2019: results from the Global Burden of Disease Study 2019. BMC Public Health 2022 Oct 12;22(1):1896 [FREE Full text] [doi: 10.1186/s12889-022-14274-7] [Medline: 36221047]
11. Resident population by ethnic group, age group and sex dashboard. Department of Statistics Singapore. URL: https://www.singstat.gov.sg/find-data/search-by-theme/population/population-and-population-structure/visualising-data/resident-population-by-ethnic-group-age-group-and-sex-dashboard [accessed 2024-01-02]
12. Lim GH, Chow KY, Lee HP. Singapore cancer trends in the last decade. Singapore Med J 2012 Jan;53(1):3-9 [FREE Full text] [Medline: 22252175]
13. Singapore Cancer Registry Annual Report 2021. National Registry of Diseases Office. 2023 Aug. URL: https://nrdo.gov.sg/publications/cancer/ [accessed 2024-01-02]
14. Fung JWM, Lim SBL, Zheng H, Ho WYT, Lee BG, Chow KY, et al. Data quality at the Singapore Cancer Registry: an overview of comparability, completeness, validity and timeliness. Cancer Epidemiol 2016 Aug;43:76-86. [doi: 10.1016/j.canep.2016.06.006] [Medline: 27399312]
15. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2001 Feb 15;19(3):335-351. [doi: 10.1002/(sici)1097-0258(20000215)19:3<335::aid-sim336>3.0.co;2-z] [Medline: 10649300]
16. Clegg LX, Hankey BF, Tiwari R, Feuer EJ, Edwards BK. Estimating average annual per cent change in trend analysis. Stat Med 2009 Dec 20;28(29):3670-3682 [FREE Full text] [doi: 10.1002/sim.3733] [Medline: 19856324]
17. Kim H, Luo J, Kim J, Chen H, Feuer EJ. Clustering of trend data using joinpoint regression models. Stat Med 2014 Oct 15;33(23):4087-4103 [FREE Full text] [doi: 10.1002/sim.6221] [Medline: 24895073]
18. Kim H, Chen H, Midthune D, Wheeler B, Buckman DW, Green D, et al. Data-driven choice of a model selection method in joinpoint regression. J Appl Stat 2023;50(9):1992-2013 [FREE Full text] [doi: 10.1080/02664763.2022.2063265] [Medline: 37378270]
19. Rosenberg PS, Check DP, Anderson WF. A web tool for age-period-cohort analysis of cancer incidence and mortality rates. Cancer Epidemiol Biomarkers Prev 2014 Nov;23(11):2296-2302 [FREE Full text] [doi: 10.1158/1055-9965.EPI-14-0300] [Medline: 25146089]
20. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024;74(1):12-49 [FREE Full text] [doi: 10.3322/caac.21820] [Medline: 38230766]
21. Siegel RL, Fedewa SA, Anderson WF, Miller KD, Ma J, Rosenberg PS, et al. Colorectal cancer incidence patterns in the United States, 1974-2013. J Natl Cancer Inst 2017 Aug 01;109(8):djw322 [FREE Full text] [doi: 10.1093/jnci/djw322] [Medline: 28376186]
22. Siegel RL, Torre LA, Soerjomataram I, Hayes RB, Bray F, Weber TK, et al. Global patterns and trends in colorectal cancer incidence in young adults. Gut 2019 Dec;68(12):2179-2185. [doi: 10.1136/gutjnl-2019-319511] [Medline: 31488504]
23. Chambers AC, Dixon SW, White P, Williams AC, Thomas MG, Messenger DE. Demographic trends in the incidence of young-onset colorectal cancer: a population-based study. Br J Surg 2020 Apr;107(5):595-605 [FREE Full text] [doi: 10.1002/bjs.11486] [Medline: 32149386]

24. Brenner DR, Ruan Y, Shaw E, De P, Heitman SJ, Hilsden RJ. Increasing colorectal cancer incidence trends among younger adults in Canada. Prev Med 2017 Dec;105:345-349 [FREE Full text] [doi: 10.1016/j.ypmed.2017.10.007] [Medline: 28987338]
25. Li Q, Yu M, Lv H, Zhang L, Deng Y, Yu H. Burden of early-onset colorectal cancer along with attributable risk factors from 1990 to 2019: a comparative study between China and other G20 countries. BMC Public Health 2023 Jul 31;23(1):1463 [FREE Full text] [doi: 10.1186/s12889-023-16407-y] [Medline: 37525147]
26. Huang J, Lucero-Prisno DE, Zhang L, Xu W, Wong SH, Ng SC, et al. Updated epidemiology of gastrointestinal cancers in East Asia. Nat Rev Gastroenterol Hepatol 2023 May;20(5):271-287. [doi: 10.1038/s41575-022-00726-3] [Medline: 36631716]
27. Sung JJY, Chiu H, Jung K, Jun JK, Sekiguchi M, Matsuda T, et al. Increasing trend in young-onset colorectal cancer in Asia: more cancers in men and more rectal cancers. Am J Gastroenterol 2019 Feb;114(2):322-329. [doi: 10.14309/ajg.0000000000000133] [Medline: 30694865]
28. AlZaabi A, AlHarrasi A, AlMusalami A, AlMahyijari N, Al Hinai K, ALAdawi H, et al. Early onset colorectal cancer: challenges across the cancer care continuum. Ann Med Surg (Lond) 2022 Oct;82:104453 [FREE Full text] [doi: 10.1016/j.amsu.2022.104453] [Medline: 36268309]
29. Lingas EC. Early-onset colon cancer: a narrative review of its pathogenesis, clinical presentation, treatment, and prognosis. Cureus 2023 Sep;15(9):e45404 [FREE Full text] [doi: 10.7759/cureus.45404] [Medline: 37854763]
30. Chen FW, Sundaram V, Chew TA, Ladabaum U. Advanced-stage colorectal cancer in persons younger than 50 years not associated with longer duration of symptoms or time to diagnosis. Clin Gastroenterol Hepatol 2017 May;15(5):728-737.e3 [FREE Full text] [doi: 10.1016/j.cgh.2016.10.038] [Medline: 27856366]
31. Eng C, Jácome AA, Agarwal R, Hayat MH, Byndloss MX, Holowatyj AN, et al. A comprehensive framework for early-onset colorectal cancer research. Lancet Oncol 2022 Mar;23(3):e116-e128. [doi: 10.1016/S1470-2045(21)00588-X] [Medline: 35090673]
32. Bailey CE, Hu C, You YN, Bednarski BK, Rodriguez-Bigas MA, Skibber JM, et al. Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975-2010. JAMA Surg 2015 Jan;150(1):17-22 [FREE Full text] [doi: 10.1001/jamasurg.2014.1756] [Medline: 25372703]
33. Muller C, Ihionkhan E, Stoffel EM, Kupfer SS. Disparities in early-onset colorectal cancer. Cells 2021 Apr 26;10(5):1018 [FREE Full text] [doi: 10.3390/cells10051018] [Medline: 33925893]
34. Nguyen SP, Bent S, Chen Y, Terdiman JP. Gender as a risk factor for advanced neoplasia and colorectal cancer: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2009 Jun;7(6):676-681.e1. [doi: 10.1016/j.cgh.2009.01.008] [Medline: 19514116]
35. Newcomb PA, Zheng Y, Chia VM, Morimoto LM, Doria-Rose VP, Templeton A, et al. Estrogen plus progestin use, microsatellite instability, and the risk of colorectal cancer in women. Cancer Res 2007 Aug 01;67(15):7534-7539. [doi: 10.1158/0008-5472.CAN-06-4275] [Medline: 17671225]
36. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA 2002 Jul 17;288(3):321-333. [doi: 10.1001/jama.288.3.321] [Medline: 12117397]
37. Sung JJY, Lau JYW, Young GP, Sano Y, Chiu HM, Byeon JS, Asia Pacific Working Group on Colorectal Cancer. Asia Pacific consensus recommendations for colorectal cancer screening. Gut 2008 Aug;57(8):1166-1176 [FREE Full text] [doi: 10.1136/gut.2007.146316] [Medline: 18628378]
38. Abu Hassan MR, Ismail I, Mohd Suan MA, Ahmad F, Wan Khazim WK, Othman Z, et al. Incidence and mortality rates of colorectal cancer in Malaysia. Epidemiol Health 2016;38:e2016007 [FREE Full text] [doi: 10.4178/epih.e2016007] [Medline: 26971697]
39. Lee HP, Lee J, Shanmugaratnam K. Trends and ethnic variation in incidence and mortality from cancers of the colon and rectum in Singapore, 1968 to 1982. Ann Acad Med Singap 1987 Jul;16(3):397-401. [Medline: 3435003]
40. Wong MC, Ding H, Wang J, Chan PS, Huang J. Prevalence and risk factors of colorectal cancer in Asia. Intest Res 2019 Jul;17(3):317-329 [FREE Full text] [doi: 10.5217/ir.2019.00021] [Medline: 31085968]
41. Burkill S, Tan CS, Lau J, Koh GC, Tan KK. Colorectal cancer incidence in a multi-ethnic population: age-period-cohort analysis of trends in Singapore. Cancer Epidemiol 2021 Dec;75:102046. [doi: 10.1016/j.canep.2021.102046] [Medline: 34653790]
42. Muhammad Nawawi KN, Mokhtar NM, Wong Z, Mohd Azman ZA, Hsin Chew DC, Rehir R, et al. Incidence and clinicopathological features of colorectal cancer among multi-ethnic patients in Kuala Lumpur, Malaysia: a hospital-based retrospective analysis over two decades. PeerJ 2021;9:e12425 [FREE Full text] [doi: 10.7717/peerj.12425] [Medline: 34820182]
43. Zhang Y, Abdin E, Sambasivam R, Shafie S, Roystonn K, Vaingankar JA, et al. Changes in body mass index and its association with socio-demographic characteristics between 2010 and 2016 in Singapore. Front Public Health 2024;12:1374806 [FREE Full text] [doi: 10.3389/fpubh.2024.1374806] [Medline: 38601489]

44. Chan TK, Tan LWL, van Dam RM, Seow WJ. Cancer screening knowledge and behavior in a multi-ethnic Asian population: the Singapore Community Health Study. Front Oncol 2021;11:684917 [FREE Full text] [doi: 10.3389/fonc.2021.684917] [Medline: 34476210]
45. Key socio-economic outcomes across cohorts. Ministry of Finance Singapore. 2019 Oct. URL: https://www.mof.gov.sg/docs/default-source/default-document-library/news-and-publications/press-releases/mof-report.pdf?sfvrsn=5eadc128_2> [accessed 2024-01-02]
46. Jin Z, Dixon JG, Fiskum JM, Parekh HD, Sinicrope FA, Yothers G, et al. Clinicopathological and molecular characteristics of early-onset stage III Colon adenocarcinoma: an analysis of the ACCENT database. J Natl Cancer Inst 2021 Nov 29;113(12):1693-1704 [FREE Full text] [doi: 10.1093/jnci/djab123] [Medline: 34405233]
47. Fontana E, Meyers J, Sobrero A, Iveson T, Shields AF, Taieb J, et al. Early-onset colorectal adenocarcinoma in the IDEA Database: treatment adherence, toxicities, and outcomes with 3 and 6 months of adjuvant fluoropyrimidine and oxaliplatin. J Clin Oncol 2021 Dec 20;39(36):4009-4019 [FREE Full text] [doi: 10.1200/JCO.21.02008] [Medline: 34752136]
48. Wolf AMD, Fontham ETH, Church TR, Flowers CR, Guerra CE, LaMonte SJ, et al. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. CA Cancer J Clin 2018 Jul;68(4):250-281 [FREE Full text] [doi: 10.3322/caac.21457] [Medline: 29846947]
49. Ladabaum U, Mannalithara A, Meester RGS, Gupta S, Schoen RE. Cost-effectiveness and national effects of initiating colorectal cancer screening for average-risk persons at age 45 years instead of 50 years. Gastroenterology 2019 Jul;157(1):137-148 [FREE Full text] [doi: 10.1053/j.gastro.2019.03.023] [Medline: 30930021]
50. Wong RK, Wong ML, Chan YH, Feng Z, Wai CT, Yeoh KG. Gender differences in predictors of colorectal cancer screening uptake: a national cross sectional study based on the health belief model. BMC Public Health 2013 Jul 23;13:677 [FREE Full text] [doi: 10.1186/1471-2458-13-677] [Medline: 23879593]
51. National Population Health Survey 2023 Report. Ministry of Health Singapore. 2024 Aug. URL: https://www.moh.gov.sg/others/resources-and-statistics/nphs-2023> [accessed 2024-01-02]
52. Zaborowski AM, REACCT Collaborative. Colorectal cancer in the young: Research in Early Age Colorectal Cancer Trends (REACCT) Collaborative. Cancers (Basel) 2023 May 30;15(11):2979 [FREE Full text] [doi: 10.3390/cancers15112979] [Medline: 37296939]
53. Pearlman R, Frankel WL, Swanson B, Zhao W, Yilmaz A, Miller K, Ohio Colorectal Cancer Prevention Initiative Study Group. Prevalence and spectrum of germline cancer susceptibility gene mutations among patients with early-onset colorectal cancer. JAMA Oncol 2017 Apr 01;3(4):464-471 [FREE Full text] [doi: 10.1001/jamaoncol.2016.5194] [Medline: 27978560]
54. Mork ME, You YN, Ying J, Bannon SA, Lynch PM, Rodriguez-Bigas MA, et al. High prevalence of hereditary cancer syndromes in adolescents and young adults with colorectal cancer. J Clin Oncol 2015 Nov 01;33(31):3544-3549 [FREE Full text] [doi: 10.1200/JCO.2015.61.4503] [Medline: 26195711]
55. Goel A, Nagasaka T, Spiegel J, Meyer R, Lichliter WE, Boland CR. Low frequency of Lynch syndrome among young patients with non-familial colorectal cancer. Clin Gastroenterol Hepatol 2010 Nov;8(11):966-971 [FREE Full text] [doi: 10.1016/j.cgh.2010.06.030] [Medline: 20655395]

Abbreviations

APC: annual percentage change
ASR: age-specific incidence rate
ASIR: age-standardized incidence rate
CRC: colorectal cancer
EOCRC: early-onset colorectal cancer
HR: hazard ratio
ICD: International Classification of Diseases
IRR: incidence rate ratio
SCR: Singapore Cancer Registry
SingStat: Department of Statistics Singapore
TNM: tumor, node, and metastasis

Edited by A Mavragani; submitted 09.06.24; peer-reviewed by F Seow, L Zhu, A Schwartz; comments to author 11.09.24; revised version received 02.11.24; accepted 24.12.24; published 14.02.25.

Please cite as:
Chen HLR, Chong QD, Tay B, Zhou S, Wong EYT, Seow-En I, Tan KK, Wang Y, Seow A, Tan KWE, Tan BHI, Tan SH
Trends in Early-Onset Colorectal Cancer in Singapore: Epidemiological Study of a Multiethnic Population
JMIR Public Health Surveill 2025;11:e62835
URL: https://publichealth.jmir.org/2025/1/e62835
doi:10.2196/62835
PMID:

Original Paper

Building and Developing a Tool (PANDEM-2 Dashboard) to Strengthen Pandemic Management: Participatory Design Study

Carlos Tighe[1], MA; Lem Ngongalah[2], PhD; Alexis Sentís[3], MD, PhD; Francisco Orchard[4], BSc; Gheorghe-Aurel Pacurar[5], Google Agile Project Management Professional; Conor Hayes[6], PhD; Jessica S Hayes[7], PhD; Adrian Toader[8], BSc; Máire A Connolly[7], MB, MPH

[1]Insight, SFI Research Centre for Data Analytics, University of Galway, Galway, Ireland
[2]Trilateral Research Ltd, London, United Kingdom
[3]Epidemiology Department, Epiconcept, Paris, France
[4]IT Department, Epiconcept, Paris, France
[5]PMO, Modus Create, Cluj Napoca, Romania
[6]School of Computer Science, University of Galway, Galway, Ireland
[7]School of Health Sciences, University of Galway, Galway, Ireland
[8]Enterprise Engineering, Modus Create, Cluj Napoca, Romania

Corresponding Author:
Jessica S Hayes, PhD
School of Health Sciences, University of Galway
University of Galway, University Road, Galway, Ireland
Galway, H91 TK33
Ireland
Phone: 353 91524411
Email: jessica.hayes@universityofgalway.ie

Abstract

Background: The COVID-19 pandemic exposed challenges in pandemic management, particularly in real-time data sharing and effective decision-making. Data protection concerns and the lack of data interoperability and standardization hindered the collection, analysis, and interpretation of critical information. Effective data visualization and customization are essential to facilitate decision-making.

Objective: This study describes the development of the PANDEM-2 dashboard, a system providing a standardized and interactive platform for decision-making in pandemic management. It outlines the participatory approaches used to involve expert end users in its development and addresses key considerations of privacy, data protection, and ethical and social issues.

Methods: Development was informed by a review of 25 publicly available COVID-19 dashboards, leading to the creation of a visualization catalog. User requirements were gathered through workshops and consultations with 20 experts from various health care and public health professions in 13 European Union countries. These were further refined by mapping variables and indicators required to fulfill the identified needs. Through a participatory design process, end users interacted with a preprototype platform, explored potential interface designs, and provided feedback to refine the system's components. Potential privacy, data protection, and ethical and social risks associated with the technology, along with mitigation strategies, were identified through an iterative impact assessment.

Results: Key variables incorporated into the PANDEM-2 dashboard included case rates, number of deaths, mortality rates, hospital resources, hospital admissions, testing, contact tracing, and vaccination uptake. Cases, deaths, and vaccination uptake were prioritized as the most relevant and readily available variables. However, data gaps, particularly in contact tracing and mortality rates, highlighted the need for better data collection and reporting mechanisms. User feedback emphasized the importance of diverse data visualization formats combining different data types, as well as analyzing data across various time frames. Users also expressed interest in generating custom visualizations and reports, especially on the impact of government interventions. Participants noted challenges in data reporting, such as inconsistencies in reporting levels, time intervals, the need for standardization between member states, and General Data Protection Regulation concerns for data sharing. Identified risks included ethical concerns (accessibility, user autonomy, responsible use, transparency, and accountability), privacy and data protection (security and access controls and data reidentification), and social issues (unintentional bias, data quality and accuracy, dependency on technology, and collaborative development). Mitigation measures focused on designing user-friendly interfaces, implementing robust security protocols, and promoting cross-member state collaboration.

Conclusions: The PANDEM-2 dashboard provides an adaptable, user-friendly platform for pandemic preparedness and response. Our findings highlight the critical role of data interoperability, cross-border collaboration, and custom IT tools in strengthening future health crisis management. They also offer valuable insights into the challenges and opportunities in developing IT solutions to support pandemic preparedness.

(JMIR Public Health Surveill 2025;11:e52119) doi:10.2196/52119

KEYWORDS

pandemic preparedness and response; COVID-19; cross-border collaboration; surveillance; data collection; data standardization; data sharing; dashboard; IT system; IT tools

Introduction

Background

The COVID-19 pandemic posed considerable challenges to European countries with widespread variation in response capacity. The COVID-19 pandemic revealed shortcomings in coordination, information sharing, and response strategies among countries, leading to fragmented approaches and delayed actions [1-3]. The lack of standardized data collection and analysis methods hindered the ability to assess the situation comprehensively, to allow data sharing and strengthen or facilitate cross-border collaboration, and to make informed decisions [4,5]. In addition, limited resources and infrastructure strained health care systems, exacerbating the impact of the COVID-19 pandemic. Lessons learned from the COVID-19 experience emphasize the critical need for enhanced preparedness and response efforts [6].

In response to the above challenges, the European Union (EU) has mandated the development and implementation of robust measures for pandemic preparedness and response across member states [7]. This mandate highlights the significance of reinforcing coordination and information exchange, as well as strengthening early detection of threats, monitoring and forecasting of diseases and resources (eg, beds and vaccines), rapid risk assessments, and the capacity of evaluating medical countermeasures (eg, therapeutics) and nonpharmaceutical interventions (eg, social distancing measures such as school closures or lockdowns). The aim is that these activities will build capacity to enhance preparedness and response capabilities to future health threats, including pandemics. Recent studies have demonstrated the potential of digital health dashboards to transform public health policy by providing real-time data analytics for rapid decision-making [8].

The PANDEM-2 IT system provides an innovative solution to address these objectives by offering a standardized and interactive platform for pandemic preparedness training and response to support decision-making on pandemic management. It also offers adaptability for addressing other potential cross-border threats. The system, which is now developed to an advanced prototype level [9,10], enables pandemic managers and stakeholders to acquire the necessary skills, knowledge, and decision-making capabilities required to respond effectively to future pandemics. Furthermore, it is designed to serve as a prototype that can be further developed or adapted to cater to different training needs. The PANDEM-2 system was tested in a 2-day functional exercise to simulate a pandemic caused by a novel strain of avian influenza. The cross-border exercise was conducted between 2 public health emergency operating centers (National Institute of Public Health & Environment, Netherlands, and Robert Koch Institute, Germany) [11]. Several EU public health and first responder agencies, including the Directorate-General for Health and Food Safety's Early Warning Response System simulation platform, were involved in supporting roles. Qualitative and quantitative data were collected to evaluate the performance of the system and its potential as a cross-border preparedness and response tool for health emergencies. Results indicate that the PANDEM-2 IT system holds potential to become a robust training hub for pandemic management in the future, serving as a valuable resource for continuous learning and preparedness.

Objectives

This study aimed to describe the development and potential of the PANDEM-2 dashboard and to outline the methodologies used in obtaining valuable insights and feedback from end users to further refine its development. It also describes how privacy, data protection, and security risks were considered throughout the design of the system. The lessons learned from the development and implementation of the platform can pave the way for future research and development in pandemic management and response technologies and inform the development of similar IT solutions.

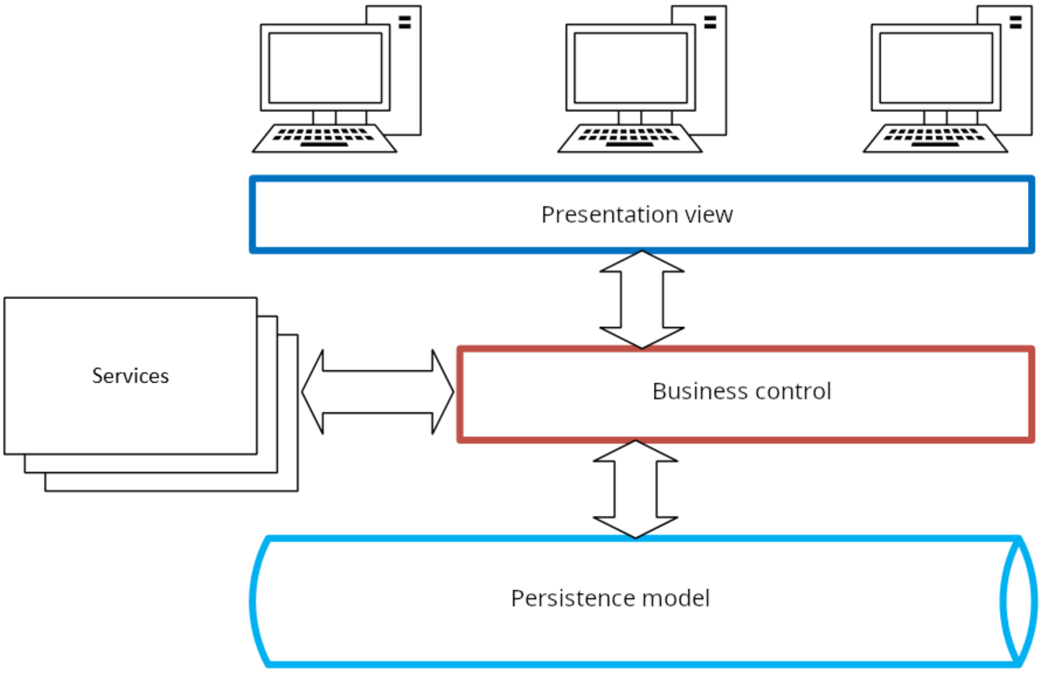
Methods

PANDEM-2 System Architecture

System Design and Components

The PANDEM-2 architecture (Figure 1) is an open-source framework comprising multiple modules and components that are accessed and integrated as web services. Its design and development were guided by the selection of widely supported libraries and adherence to industry-standard development practices and principles. The platform features a multilayered architecture that facilitates communication between its components. The persistence layer, built using MongoDB [12], ensures efficient data retrieval and storage. This layer is accessed through an exposed application programming interface (API) from the NodeJS business layer [13], which facilitates interaction between the MongoDB database and other modules or external systems. NodeJS is known for its scalability and efficiency and ensures the smooth functioning of the business layer [14]. The presentation layer is developed using the Angular framework [15], which manages the user interface, providing an intuitive and interactive experience for users. Angular was chosen for its robustness, scalability, and extensive library support, which enhances the platform's user-friendliness and maintainability. Angular is a well-maintained framework, with releases every 6 months. Its documentation is comprehensive, containing deprecation, compatibility, and breaking change policies alongside update path guides, ensuring clear upgrades and maintenance. Data visualization within the application is handled by the Highcharts JavaScript library [16]. The use of these stable and well-supported technologies underpins an architectural design that ensures the platform is scalable, adaptable to various presentation contexts, and easy to maintain.

Figure 1. Multilayered architecture of the PANDEM-2 dashboard illustrating the interaction between the different layers, including the persistence layer (MongoDB), business layer (NodeJS), and presentation layer (Angular).

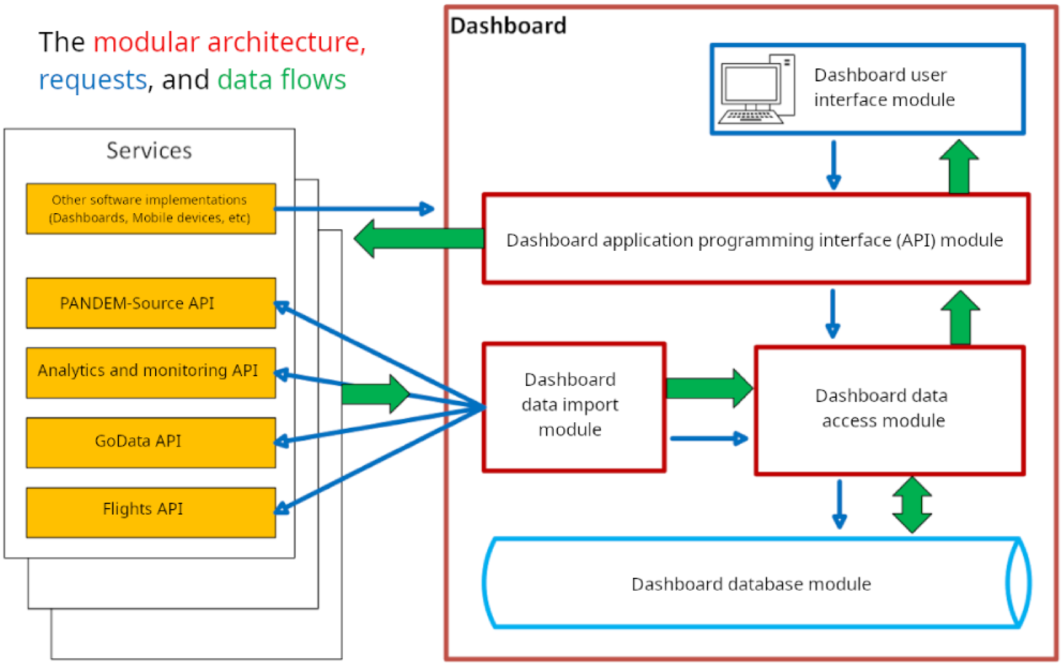


Data Collection and Integration

Data within PANDEM-2 are gathered from various sources via data collectors, which can either be human-operated or automated. The PANDEM-Source module [17], an extract, transform, and load (ETL) pipeline and database was specifically developed for PANDEM-2 to facilitate the continuous collection and integration of heterogeneous surveillance data into a coherent database. PANDEM-Source uses JSON (JavaScript Object Notation)-formatted source definition files to describe each source, including format, monitoring frequency, mapping to PANDEM-2 variables, and coding systems. It performs transcoding by downloading and applying mappings between coding systems, reporting any missing codes or integration errors to a data manager who can update mappings or monitor integration progress. A key challenge was standardizing geographic locations to the Nomenclature of Territorial Units for Statistics (NUTS), a hierarchical system developed by Eurostat that divides the territory of the EU into regions for statistical and reporting purposes [18]. For natural language processing used to detect locations in social media data, we used the GeoNames database, which contains local aliases for geographic locations [19]. To reduce methodological disparities between sources, we prioritized official European Centre for Disease Prevention and Control (ECDC) data where available and relied on existing data collectors like COVID-19 Data Hub [20] when ECDC data were not available. PANDEM-Source allows data managers to tag datasets, enabling the selection of relevant data combinations for export to the PANDEM-2 dashboard. This process, while facilitated by the tool, requires human oversight. To ensure the accuracy of the data acquisition process, manual validation was performed during both the data import and dashboard implementation stages. The declarative syntax used by the tool minimized the risk of errors in data processing. A more formal validation process is currently underway and described in greater detail in a related preprint [21].

The PANDEM-Source module functions as a service with an API providing relevant and preprocessed data to the persistence layer of the dashboard. Data are then processed, exposed through a specific API, and made accessible to the front-end of the dashboard, as well as other interested data publishers (Figure 2).

Figure 2. Dashboard modular architecture demonstrating the flow of data from external sources into the PANDEM-2 dashboard, showing how data is processed through various modules and application programming interfaces (APIs).



The business layer in the dashboard collects data from external modules, such as the World Health Organization GoData implementation [22] and also from the Flights module [23] and the Modelling and Analytics module developed internally as part of the project. It processes these data, performing validations, adaptations, and enrichment based on formulas, to prepare them for display in the presentation layer.

Security Aspects of the PANDEM-2 Dashboard

The PANDEM-2 dashboard infrastructure is secured using Amazon Web Services (AWS) Cloud Security [24]. The AWS environment is monitored using AWS CloudWatch and CloudTrail, ensuring General Data Protection Regulation (GDPR) compliance [25]. Several other AWS services were used to enhance privacy and control network access, including Amazon GuardDuty for continuous threat monitoring [26] and AWS Secret Manager and Key Management Services for data protection [27]. Data are encrypted in transit with Transport Layer Security across the AWS services, and secure sockets layer and transport layer security certificates, are managed with AWS Certificate Manager. A summary of the PANDEM-2 security aspects can be found in Multimedia Appendix 1. The project only hosted publicly available aggregated or simulated data, which addressed data protection concerns for the duration of the project.

Dashboard Design and Development

Exploratory Research and Design Rationale: Why a Dashboard Approach?

A dashboard is a visual display of key analytical indicators arranged on a single screen, enabling real-time monitoring and quick assessment [28]. The goal is to provide stakeholders with relevant analyses to anticipate issues and make decisions in domains with complex multivariate data. For example, a study by Franklin et al [29] modeled the real-time status of a hospital's emergency department as a dashboard to support rapid clinical decision-making and improve throughput. However, the growing use of dashboards has raised concerns about the cognitive effort required to extract actionable information, particularly for users with limited analytic skills [30]. Data storytelling has been proposed as a more intuitive way to communicate information that motivates decision-making [31]. Stakeholders, however, determined that maintaining and updating multiple specific decision narratives would be overly complex. Therefore, we chose a dashboard approach for its flexibility and adaptability across various public health scenarios, including surveillance and training. In addition, a key design requirement was the ability to customize the dashboard view and generate reports, which was seen as a more scalable and sustainable solution compared to data storytelling, given the diverse decision-making needs of different stakeholders. The primary users of the PANDEM-2 dashboard are public health officers and medics typically trained in data analysis.

Review of Existing Dashboards

Responses to the ongoing pandemic provided opportunities to review public health indicators, visualization approaches used, and the design decisions made by different organizations. To gain insights into these aspects, we conducted a review of publicly available dashboards in April and May 2021, operating under the assumption that these reflected the information needs of public health agencies and could be analyzed without an imposition on end users (Multimedia Appendix 2). This review included dashboards recommended by collaborating partners based on their initial user requirements, as well as publicly accessible COVID-19 projects under development (Multimedia Appendix 3). The review assessed the available dashboards, focusing on their aims, target users, and content. We also collated data sources for each dashboard, along with their breakdowns by population, time, location, granularity, and visualization approaches. This analysis provided insights into how organizations prioritized information and the data views expected by experts. It also highlighted different data interaction models and how data artifacts were aggregated, filtered, modified, and recorded. While the data sources identified during this review were considered for inclusion in the PANDEM-2 database, consortium members' own data eventually proved to be more suitable, especially as larger organizations and governments began openly sharing their information.

Data Visualization Catalog

Capturing and storing visualizations for viewing or referencing during the design process was challenging due to their volume and complexity. Initially, screenshots of each visualization, along with related data, were added to a document (Multimedia Appendix 4). However, as the document grew, this approach became impractical and difficult to reference. To overcome this challenge, a data visualization catalog was created. The catalog was specifically developed for annotating and uploading data visualizations during the design phase of the project. The visualizations became searchable using keywords related to graph types, data types, or indicator types. The catalog served as a directory of existing public health visualization approaches, which the project team could consult during the participatory design process. It is important to note that the Vizualisation Catalogue was used solely as a development tool during the design phase and is not part of the final application.

User Requirements and Participatory Design

Gathering of User Requirements

The design and development of the PANDEM-2 dashboard began with collecting user requirements from members of the project consortium, outlining the necessary functionalities and features to meet users' needs. Participants represented a wide range of domains, including universities, national public health institutes, the Italian Red Cross, the Austrian Red Cross, the National Institute of Medical Emergency (Portugal), a leading institute for trend and technological analysis, and MedTech health organizations. Each consortium partner appointed a contact person to gather internal requirements and communicate them to the project's technical team.

These designated representatives participated in initial user forums that were held during the project's inception, aimed at gathering user requirements from domain experts and capturing their needs and preferences for features in pandemic management software (Multimedia Appendix 5). The group was diverse in age, experience, and gender, with participants ranging from recent postdoctorates to professionals with >30 years of experience, and included software designers, data scientists, epidemiologists, microbiologists, physicians, first responders, and project managers. They also represented a wide geographic spread, including Romania, Italy, Portugal, France, Germany, the Netherlands, Belgium, Sweden, Finland, and Ireland. This diversity ensured a wide range of expertise and perspectives, thus reducing potential design biases.

The requirements gathering process involved 2 workshops held between February and April 2021. During these workshops, users described their ideal dashboard for pandemic management and identified essential visual features for the dashboard. These requirements were then compiled into a comprehensive list and organized into functional units. The participatory design surveys were conducted as part of this process, providing interactive design prototypes and collecting feedback from users on key functionalities (Multimedia Appendix 6).

Generation and Analysis of User Stories

User requirements were assigned to different groups within the technical team, who analyzed and converted them into user stories. A user story describes a feature from the end-user's perspective, detailing how it provides value [32] (eg, "As a policy advisor or health care professional, I want to see cases stratified by age, sex, time, and severity to monitor clusters and track the virus spread over time").

By exploring user requirements and generating user stories, the technical team identified common needs, consolidated similar requests, and clarified key functionalities for the PANDEM-2 system. Each user story generated focused on a primary indicator query and was linked to the initial requirements for traceability. It defined query resources for user interactions, enabling data filtering, aggregation, new data creation, and record management. In addition, each user story considered the time-related and geographic aspects of the indicator within the application. Multimedia Appendix 7 provides an example of a user story.

Prioritization and Refinement of Key Variables and Indicators

To refine and prioritize user requirements for the dashboard, a prioritization process was undertaken. This involved a data availability and prioritization survey, which was conducted with the consortium members using the initial requirements generated in the Gathering of User Requirements section. A matrix was created, mapping all necessary variables and indicators required to fulfill the identified needs. This was followed by an assessment of the priorities and availability of these variables in open and restricted data sources. Users were invited to provide insights into their institution's data availability, including details about the level of information, reporting periodicity, data format, and the geographic level at which the data were accessible (municipal, regional, or national). The questionnaire covered various aspects, such as level of priority, data availability within their institution, the level of detail of publicly available data, adherence to the European Surveillance System (TESSy) format, and the pathogens for which data were collected (Multimedia Appendix 8).

The list of requirements was further divided into 14 variable categories, and each item on the list was classified as either an observation, a characteristic, an indicator, a resource, a document, or a referential. These terminologies are defined in Textbox 1.

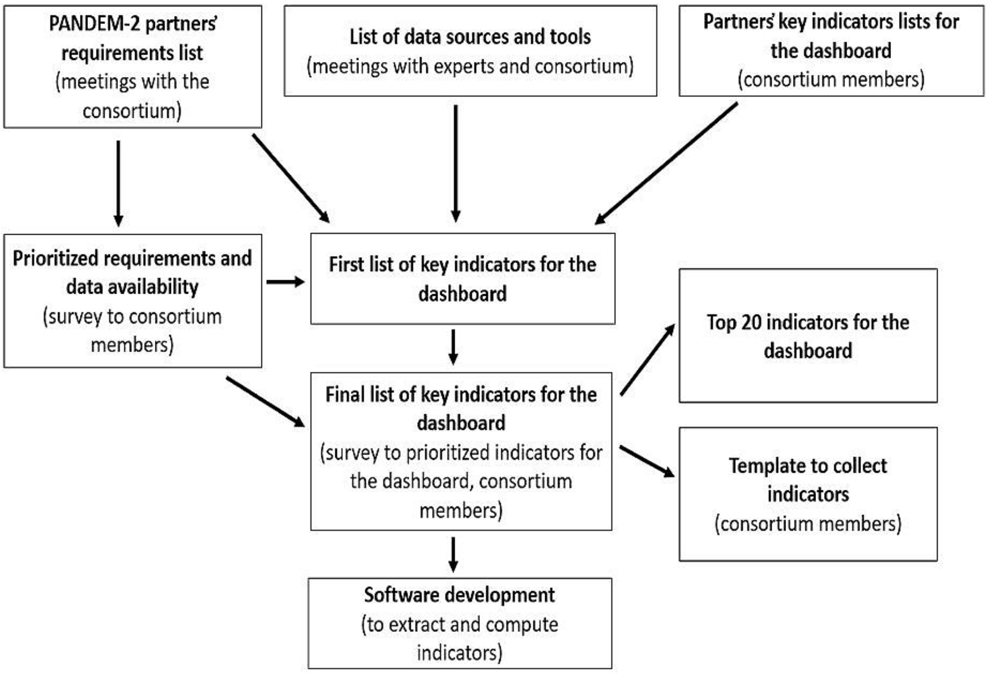
Textbox 1. Definitions of variable categories used in the PANDEM-2 dashboard.

- Observation: Any variable or item that reflects a measurement or observation reported by a data source. These variables are typically numeric and represent concepts like the number of people or resources, or could include nonnumeric items, such as statements (eg, tweets) published by individuals.
- Indicator: Specific observations used to monitor particular questions. They can be categorized as either calculated values with associated methodologies and formulas or directly reported by users.
- Characteristic: Any variable or item which describes or categorizes an observation. It is always linked to variables of the observation type, such as the country, date, or variant of the observed cases.
- Referential: A specific type of characteristic distinguished by having a unique code that can be mapped to other characteristics. For example, municipalities serve as a referential, each possessing a unique code and linking to a country.
- Resources: A specific type of referential that concerns resources needed to handle pandemics. They play an important role in PANDEM-2 for resource modeling features (eg, nurses per bed or material resources). They are associated with observations like the number of available resources.
- Document: Refers to any generic term for a record or written material.

The methods used for defining key indicators and data sources, as outlined earlier, culminated in a final variables list that encompasses both open data sources and data contributed by partners (Multimedia Appendix 9). Alongside this list, data formats and an ETL process were specified to facilitate seamless integration with PANDEM-Source. To support the ETL process, a collaborative effort led to the development of a Microsoft Excel template to accommodate specific data needs for most consortium members. Furthermore, relevant heterogeneous data sources for effective pandemic management, such as Next Generation Sequencing–Laboratory data, social media data, and participatory surveillance data, were identified and corresponding ETL processes were devised for inclusion in the PANDEM-2 database.

After receiving the initial list of requirements and the prioritization feedback, the technical team began further refinement of the requirements. The process was guided by the concept of a minimal viable product (MVP), which focuses on prioritizing essential features to maximize learning with minimal effort [33]. Although the PANDEM-2 system is a prototype, the MVP approach helped in determining where efforts should be concentrated and which requirements should be given priority during the implementation process. Each group in the technical team looked at the requirements list and the prioritization process and created an MVP list. Subsequently, with participation from all Public Health Agencies and first responder partners in the biweekly web-based workshops, the initial requirements, the prioritization list, and the MVPs were combined and deliberated upon. The process of gathering and refining user requirements is illustrated in Figure 3. This process underpinned the development of the PANDEM-Source software used to extract and compute indicators, which were visualized on the PANDEM-2 dashboard. At the end of this prioritization process, a list of key indicators was generated.

Figure 3. Process used to gather partners' requirements list and prioritize partners' key indicator lists for developing the dashboard’s core features.

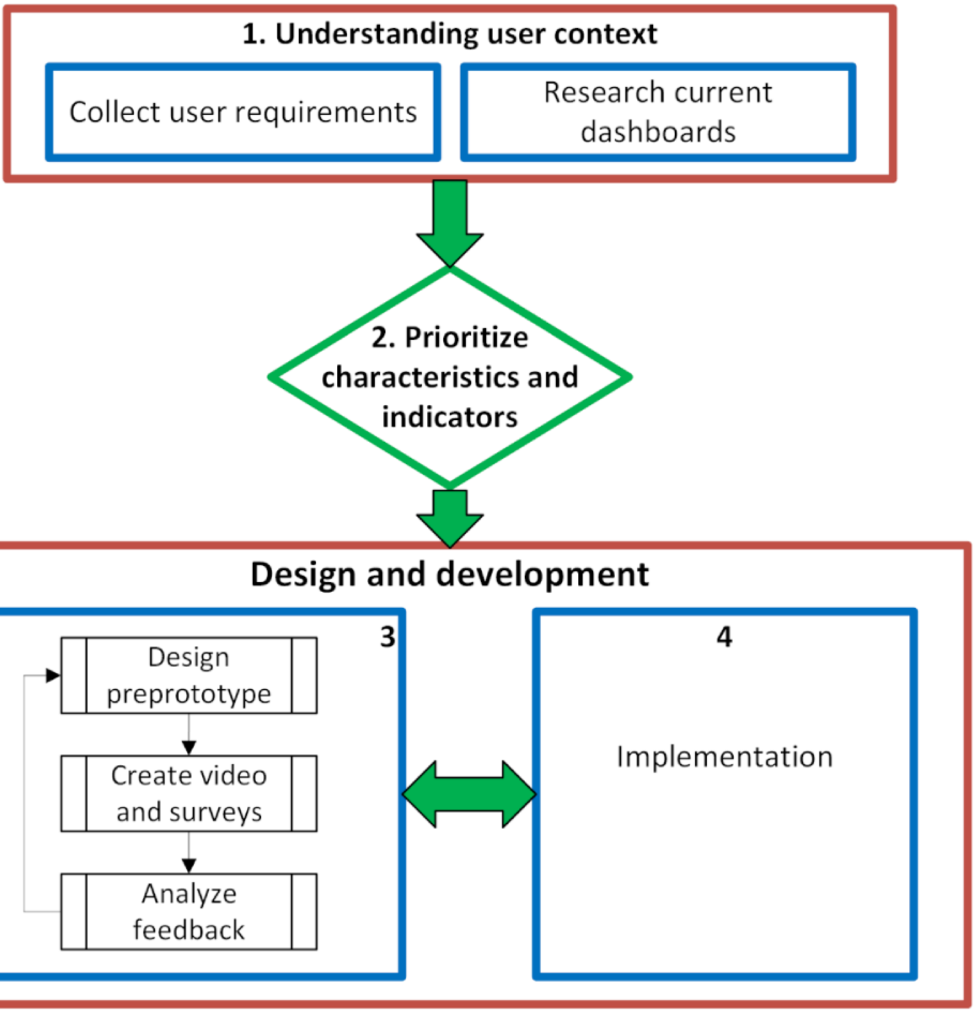


Prototype Design, Participatory Feedback, and Iterative Improvements

The design of the PANDEM-2 dashboard followed an iterative process involving a participatory design approach, with user feedback incorporated throughout the design stages. Participatory design represents a design approach in which the people destined to use the system being designed play a critical role in its design [34]. Generally, participatory design is conducted in face-to-face settings where users are engaged in recording requirements and teasing out functionalities that are necessary to implement them. However, due to the COVID-19 pandemic, these physical meetings were not possible to organize. The participatory design process was therefore adapted to the pandemic context and involved several asynchronous methods, as illustrated in Figure 4. The adapted process allowed relevant stakeholders to participate and enabled collaboration across multiple teams. The analytical components of the dashboard were then designed based on the outcomes of this collaboration and are presented as follows:

1. Understanding the context of use by gathering insights from various stakeholders to understand end-user requirements and identify COVID-19 public dashboards for analysis
2. Prioritization of dashboard characteristics and indicators
3. Asynchronous participatory design (APD) process that involved weekly asynchronous interactions with users. Interactive design prototypes along with video documentation and feedback forms were sent to consortium members, who explored the designs and provided feedback. The primary objective of the APD process was to enable users to interact with a preprototype platform, explore potential interface designs, and provide feedback. This process involved targeted questionnaires, which aimed to explore what graphical representations and interactions were beneficial to users for their respective roles. Participants were presented with a range of potential functionalities linked to priority indicators, which included visual analytics components and interaction models, such as the ability to filter and switch between different data views. Participants shared their feedback on the structure of the pages, the data presented, the data visualizations that were used on each page, and the interactions with those data visualizations. Through this process, users provided feedback on elements to be added, removed, or modified within different components in the system. This feedback was then integrated into subsequent designs.
4. In parallel to the above step, the design prototypes were sent to the software team, who carried out the implementation of designs and improvements toward the final prototype.

Figure 4. Development workflow, including asynchronous participatory design, detailing the iterative feedback process used to incorporate user input into the PANDEM-2 dashboard design. AVG: average.



Impact Assessment

The implementation of ethical, privacy, data protection, and social principles was an integral part of the PANDEM-2 project. The process was led by Trilateral Research, a research and innovation organization specializing in applied ethics, privacy, and data protection in technology development. The project team actively engaged with end users both within and outside the project consortium to identify and mitigate potential risks arising from the tool's development. Following relevant guidance, including the Information Commissioner's Office Privacy Impact Assessment Code of Practice [35] and International Organization for Standardization and International Electrotechnical Commission standard for privacy impact assessments [36], the process mapped information flows, assessed potential risks within the system, and provided recommendations to mitigate the risks. A diverse group of stakeholders were involved in the process, participating in 2 workshops conducted in November 2021 and May 2023. There were 20 participants in each of the workshops, representing a wide geographic spread, including Romania, Italy, Portugal, France, Germany, the Netherlands, Belgium, Sweden, Finland, and Ireland. Those that partook in the workshops came from universities (UCLouvain and University of Galway), national public health institutes (Institute of Public Health in Romania, Public Health Agency of Sweden, the National Institute for Public Health and the Environment in the Netherlands, the Robert Koch Institute, the Portuguese National Emergency Medicine Institute, and the Finnish Institute for Health and Welfare), as well as other organizations such as the Italian and Austrian Red Cross, the Fraunhofer Institute for Technological Trend Analysis, and software companies specializing in health, Epiconcept, and development, Clarisoft. Roles of those taking part included software design and development, data science, epidemiology, microbiology, infectious disease specialists, doctors, researchers, first responders, pandemic managers, and project managers. Notably, participants included a former adviser to the World Health Organization on health security and a former head of the Chemical, Biological, Radiological and Nuclear threat unit for Belgian defense. The initial workshop aimed to raise awareness of ethical, privacy, data protection, and social principles requiring consideration during the development and deployment of the PANDEM-2 system and tools. It also focused on mapping data flows within the PANDEM-2 system and identifying related risks or opportunities. In addition, the workshop explored privacy, social, and ethical issues and opportunities to be considered in different pandemic scenarios, both in the preparedness and response phases. Possible mitigations for the identified risks were also discussed. The second workshop involved an open and collaborative discussion, allowing stakeholders to contribute insights on mitigation measures, necessary actions, and final recommendations for both project stakeholders and end users.

Ethical Considerations

The PANDEM-2 project adhered to GDPR regulations and Horizon 2020 guidelines. Ethical approvals were obtained from the relevant institutional review boards of all consortium partners involved in data collection and processing. These approvals ensured that the project complied with the highest ethical standards, protecting participants' rights and ensuring the responsible handling of personal data. Informed consent was obtained from all participants before their involvement in the stakeholder engagement activities. The consent procedures ensured that participants clearly understood the nature of the study, the use of their data, and their rights, including the ability to withdraw at any time. Participants were also informed that their data would be used for research purposes only. All personal data were anonymized or pseudonymized to protect participants' identities. Data were securely stored on password-protected servers, accessible only to authorized personnel. Participant identifiers were removed or replaced with pseudonyms to ensure anonymity in the research outputs, preventing any risk of reidentification. No financial compensation was provided to participants. Participation was entirely voluntary, and participants had the right to withdraw from the study at any time without any negative consequences. No identifiable images of participants are included in the study materials or outputs.

Results

System Architecture and Data Integration

Implementation and Configuration

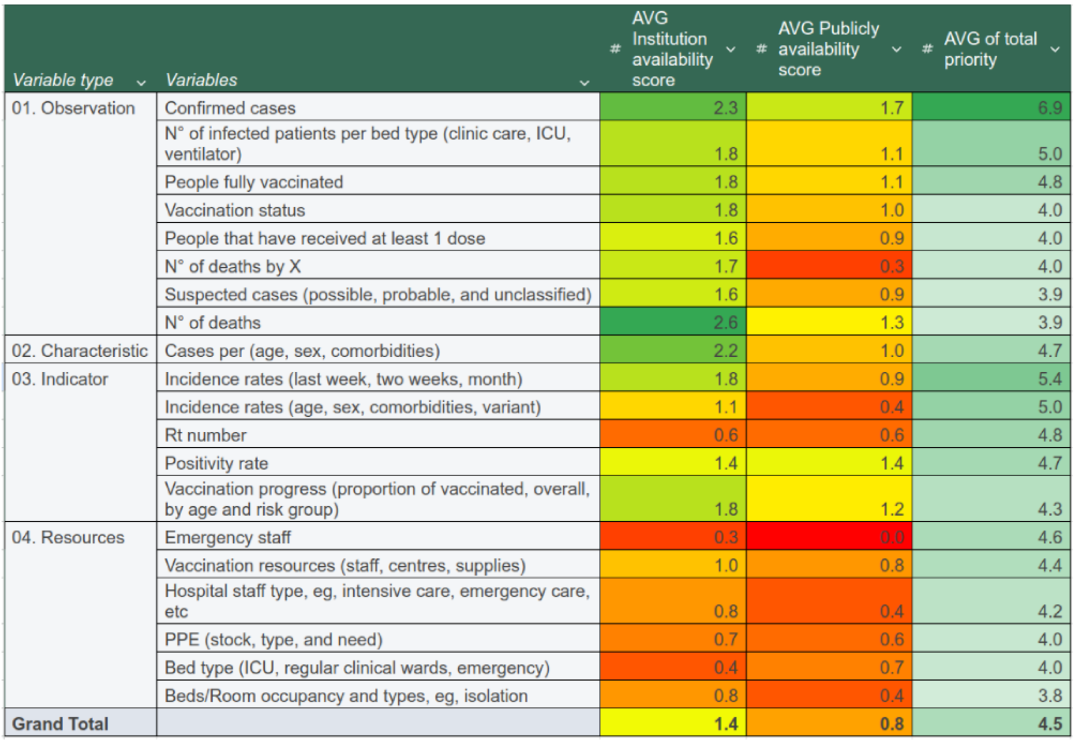
The dashboard was implemented and hosted on the Amazon Cloud as a central data management system at the EU level, accessible to all national health institutes across Europe. To address challenges with semitransparent data access, whereby some but not all data from different institutions can be shared with others, the architecture is not only open access but also adaptable. An adaptable architecture allows for both isolation of data between groups of countries and on-demand installation or deployment in a single health institute.

The dashboard offers flexibility in its configuration, allowing customization in 2 main areas. First, user experience and user-interface attributes, which encompass elements related to user experience and user interface. This involves tailoring aspects such as the language used for interface elements, the design of graphical features, and the overall visual style. Users can personalize the appearance of the dashboard to align with their preferences and needs by adjusting these attributes. Second, internal functional parameters that deal with the technical aspects of the dashboard's operation can be customized. This includes details on how different modules within the system access APIs and specific technical information relevant to the system's functioning. The administration module of PANDEM-2 serves as the control center for managing these parameters, granting users the ability to fine-tune the dashboard's functions and adapt it to their specific requirements.

Data Sources and Key Indicators

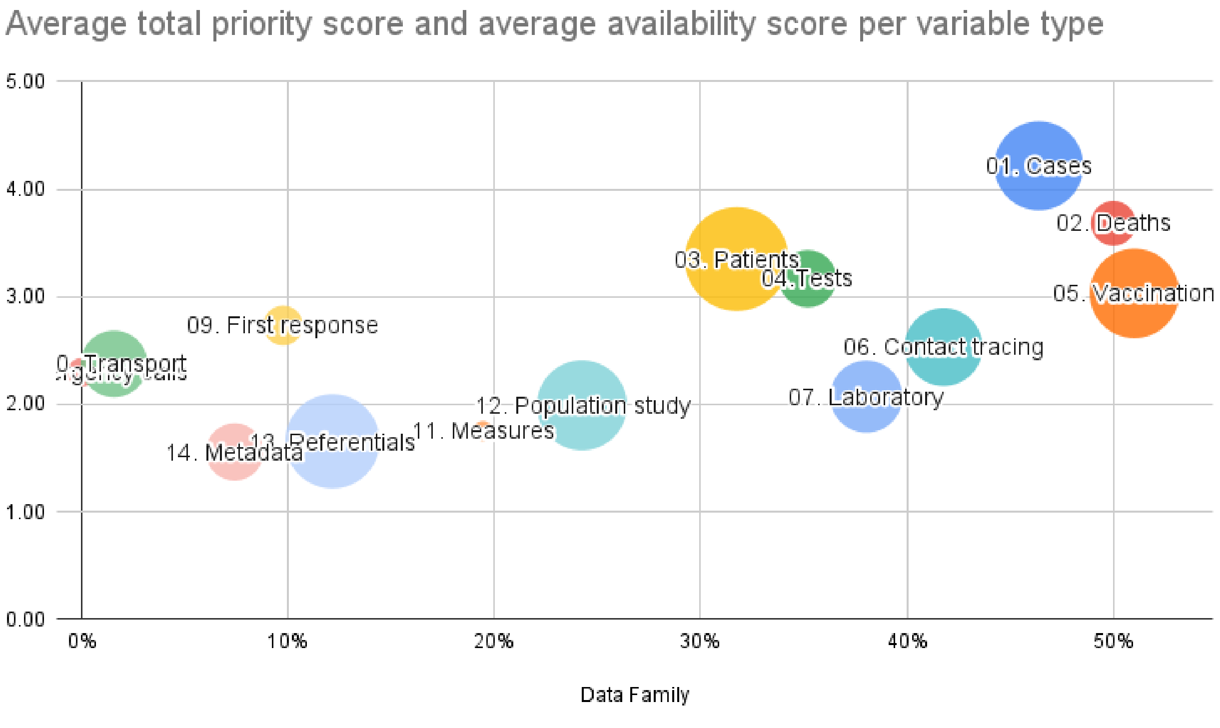
Responses to the data availability and priority survey are shown in Figures 5 and 6, and tables 1 and 2. The 9 participant institutions were questioned about 103 variables classified by groups. For each variable, the respondent answered a series of questions about availability, priority, and format of the associated information. On the basis of these answers, data availability scores were calculated and color-coded, ranging from low (red) to high (green; Figure 5). The variables identified as most relevant were cases, vaccination uptake, deaths, hospital resources, and hospital admissions. Figure 6 shows that the most important variable categories were also the categories with the most available data. Higher priority categories are toward the top of the graph, while high availability categories are toward the right. This observation highlights the importance of prioritizing and ensuring accessibility to critical data categories to support effective pandemic preparedness and response efforts.

Figure 5. Survey results highlighting the availability and prioritization of key data variables for the dashboard (AVG; X). AVG: average; ICU: intensive care unit; N°: number of; PPE: personal protective equipment; Rt:; X: specific pathogen of the pandemic.



Variable type	Variables	AVG Institution availability score	AVG Publicly availability score	AVG of total priority
01. Observation	Confirmed cases	2.3	1.7	6.9
	N° of infected patients per bed type (clinic care, ICU, ventilator)	1.8	1.1	5.0
	People fully vaccinated	1.8	1.1	4.8
	Vaccination status	1.8	1.0	4.0
	People that have received at least 1 dose	1.6	0.9	4.0
	N° of deaths by X	1.7	0.3	4.0
	Suspected cases (possible, probable, and unclassified)	1.6	0.9	3.9
	N° of deaths	2.6	1.3	3.9
02. Characteristic	Cases per (age, sex, comorbidities)	2.2	1.0	4.7
03. Indicator	Incidence rates (last week, two weeks, month)	1.8	0.9	5.4
	Incidence rates (age, sex, comorbidities, variant)	1.1	0.4	5.0
	Rt number	0.6	0.6	4.8
	Positivity rate	1.4	1.4	4.7
	Vaccination progress (proportion of vaccinated, overall, by age and risk group)	1.8	1.2	4.3
04. Resources	Emergency staff	0.3	0.0	4.6
	Vaccination resources (staff, centres, supplies)	1.0	0.8	4.4
	Hospital staff type, eg, intensive care, emergency care, etc	0.8	0.4	4.2
	PPE (stock, type, and need)	0.7	0.6	4.0
	Bed type (ICU, regular clinical wards, emergency)	0.4	0.7	4.0
	Beds/Room occupancy and types, eg, isolation	0.8	0.4	3.8
Grand Total		**1.4**	**0.8**	**4.5**

Figure 6. Comparison of priority versus availability of variable categories, showing the alignment between prioritized variables and their availability across institutions.



Most of the data were reported on a weekly or daily basis, indicating a commitment to timely and up-to-date information (Table 1). A significant proportion of the data were collected and recorded by various institutions and from data sources related to COVID-19. There was limited recording of pandemic-related data before the COVID-19 pandemic. There were also variations in data granularity, reporting frequency, geographic coverage, or other specific details related to the COVID-19 information being collected. On average, only 9.7% (10/103) variables were available in the TESSy [37] format and 26.2% (27/103) only in a different format, a format used by the TESSy (Multimedia Appendix 10).

Table 1. Data reporting frequency and types of pandemic-related data available, highlighting the frequency of reporting (daily, weekly, or monthly) and the prevalence of COVID-19 data from various sources (n=115).

Data category	Reporting variables or items, n (%)
Periodicity	
Weekly	17 (14.7)
Daily	14 (12.1)
Others[a] (specify in comments)	8 (6.9)
Monthly	2 (1.7)
Pathogens	
COVID-19	17 (14.7)
Others (specify in comments)	13 (11.3)
Influenza and COVID-19	8 (6.9)
Influenza	1 (0.8)

[a] The “others” option refers to any reporting frequencies that were neither daily, weekly, nor monthly, which participants were asked to provide qualitatively through comments.

The survey also identified several data gaps, as shown in Table 2. Among these, 2 areas of particular significance were contact tracing and deaths. It was evident that these areas lacked sufficient information despite their high priority in the survey. In addition, there was also a noticeable lack of information concerning resource availability, despite its recognized importance in pandemic management.

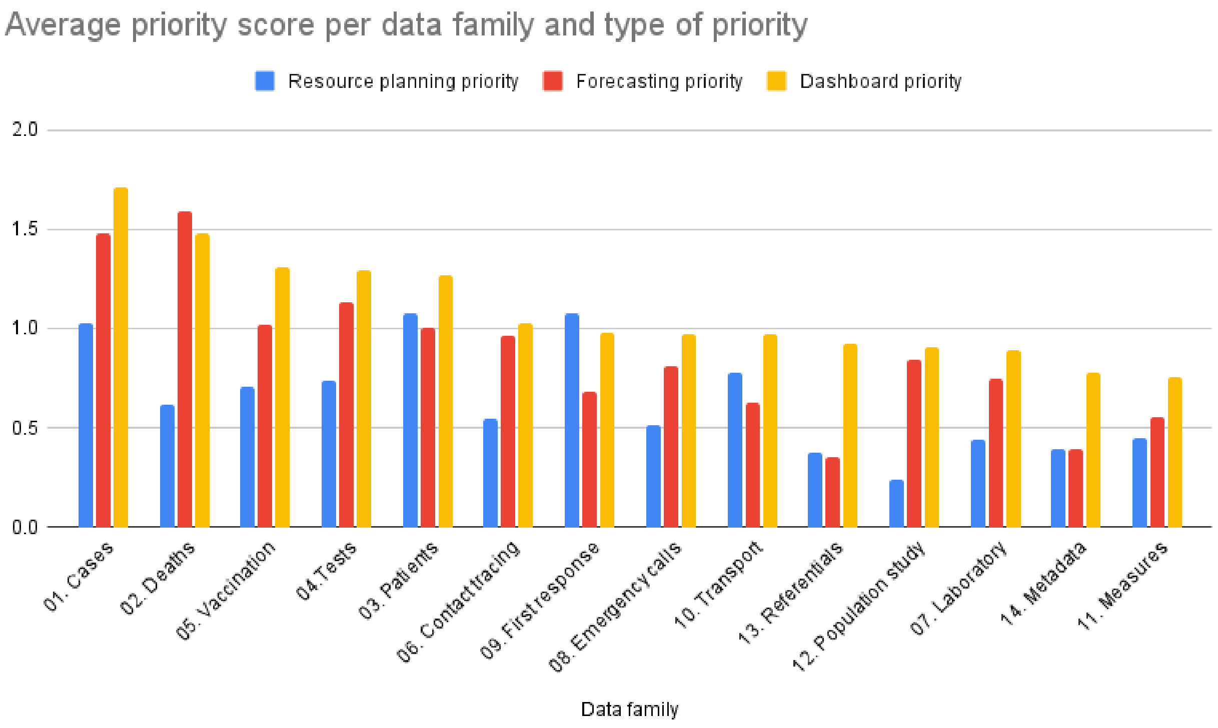
Table 2. Data gaps identified during data availability survey, showing critical pandemic-related data, such as contact tracing and deaths that were insufficiently available despite high priority.

Variable category	Average public availability gap
Deaths	1.1
Contact tracing	0.9
Cases	0.8
Laboratory	0.6
Patients	0.5
Vaccination	0.5
Tests	0.3
Referential	0.2
Population study	0.1
Metadata	0.1
First response	0.1
Transport	0.0
Emergency calls	0.0
Measures	0.0

The public availability gap is a metric quantifying the specific types of data accessible at a detailed level but not disclosed publicly. Each participant provided their estimation of variable availability within their institution and at the public level, to the best of their knowledge. The assigned scale attributes 4 as the highest value to individual level, 3 to municipality, 2 to regional, and 1 to country level. In instances where data were unavailable or unknown, a score of 0 was assigned. The data availability gap is determined by calculating the difference between the average availability score at the institution level and at the public level.

Figure 7 illustrates the findings regarding the variables that users prioritized for resource planning, forecasting, and the dashboard. To calculate the priority score for these purposes, participants were asked to rate the importance of each variable from 0 to 3 for the 3 purposes, with 0 being the lowest and 3 being the highest. The results were averaged by variable category. In most cases, dashboard variables held a higher level of importance compared to forecasting or resource planning variables. The spikes observed in the resource planning and forecasting areas highlight the data points of utmost priority for those domains, with particular significance placed on the ability to forecast deaths and to perform resource planning for patient care at the hospital and first response level.

Figure 7. Prioritization of variable categories for resource planning, forecasting, and the dashboard, showing importance ratings assigned to various data categories by participants.



At the end of the data and indicator prioritization process, a template to collect variables and indicators from nonopen sources (or partner-restricted data) was created.

Dashboard Design and User Engagement

Insights and Outcomes From Dashboard Review

The review of existing dashboards provided several insights that influenced the design of the PANDEM-2 dashboard. The data visualizations within the reviewed dashboards highlighted the prioritization of information by various organizations and identified the data views expected by experts. This review also revealed different data interaction models useful in public health domains and how data artifacts were aggregated, filtered, changed, and recorded. Insights gained from this review helped shape the decision to use consortium members' data as the primary data source, as these proved to be more comprehensive and reliable than many external sources. The review also informed the design features and visualization approaches adopted in the PANDEM-2 dashboard.

Prototype Design and Features

Multimedia Appendix 11 visually illustrates the design of a PANDEM-2 dashboard page. The key components on each page of the dashboard include the following:

1. Overview explanation component. This component includes a page title and provides an explanation of the indicator and the data presented on the page. It serves as a brief summary to provide context and understanding.
2. Indicator cards. These cards display high-priority data related to the indicator. They offer a quick snapshot of the current state of the pandemic, allowing users to grasp important information at a glance. The cards also feature an information button that provides additional details about the data being shown. The highlighted figures show a comparison between the most recent 7 days and the preceding 7 days. The highlighting provides an instant visual understanding of how important indicator variables such as admissions and intensive care unit admissions are changing over time.
3. Map. The map component offers a geographical perspective on the pandemic situation. Users can choose the specific information they wish to view on the map, such as cases, vaccination rates, or other relevant metrics. They can also segment the map by NUTS regions, providing a more localized view.
4. Graphs. The dashboard includes multiple graphs that present detailed information about the indicator. These graphs allow users to explore and filter the data according to their preferences. This feature enables a deeper analysis of the pandemic trends and patterns.
5. User-interface functionality. Users can access the PANDEM-2 system using the interface in the presentation layer. This interface offers multiple possibilities to display and filter information, build reports based on displayed graphical elements and comparisons between different data, and export these reports in many predefined formats, including images, tables, and web service-specific formats.

The final design framework of the PANDEM-2 dashboard was structured to provide an intuitive and interactive experience for users. Each indicator view was conceptualized as a dedicated page, comprising various relevant components. Abstract components were developed that could be adapted to each indicator page, ensuring a unified design and interaction experience across pages. This approach allowed users to intuitively understand the interactions and functionality across different pages once they had become familiar with the system on one page.

Together, these components enhance the accessibility and usability of the PANDEM-2 dashboard, empowering users to gain insights into the pandemic data through both high-level summaries and detailed visualizations. More illustrations of initial and final dashboard designs are shown in Multimedia Appendix 12.

Participatory Design Feedback

The participatory design process led to several modifications and improvements in the PANDEM-2 dashboard prototype. The number of responses to each of the APD feedback forms ranged from 4 to 9.

Feedback from the APD process highlighted the importance of data visualization in diverse formats, combining different data types, and analyzing data across various time frames. Users expressed their interest in generating custom reports and exploring data alongside government interventions (countermeasures):

> *The addition of countermeasures on the emotion/sentiment timelines is good. However, the countermeasures would be easier to see or understand, if they were all on the top part of the graph with the name of each intervention placed next to the line [on the left or right side of the graph]. Now it's difficult to check back to the figure legend below and tell one shade of grey from another. Now there's also a false impression that a countermeasure is somehow linked to the emotions on the same level of the graph* [for example, social distancing with anticipation and disgust]

Because of the feedback received, the final prototype incorporated various features that catered to the specific needs of the stakeholders, such as customizable dashboard views and report generation capabilities. The design also addressed the cognitive effort required to extract actionable information, ensuring that the dashboard remained user-friendly and accessible to all users, regardless of their analytical skills. This feedback proved invaluable for the technical team, offering insights into the preferred indicators and formats for their presentation. In addition, the feedback offered the technical team an understanding of how end users intended to use the data beyond the dashboard, including saving reports, downloading visualizations, and obtaining underlying data. When asked about saving outputs from the dashboard, participants expressed a desire for flexibility, with various preferences emerging. Users wished to save the current state of the page to share with others, download images or PDFs of the page, obtain the data within the page, or capture images of specific sections of the page.

Regarding epidemiological and resource modeling, participants expressed the desire to save the model used, the input parameters for scenario runs, visual outputs, data outputs in alternative formats, and high-level indicators.

Participants' feedback also shed light on data reporting challenges, particularly in terms of disparities in reporting levels and time intervals and the need for standardization across organizations. For instance, a participant highlighted how delays in testing workflows in Finland resulted in challenges when reporting cases and tests:

> *Due to unavoidable delays in testing workflow [sampling, sample shipping and processing, analysis, registration of results] it may not be realistic to report tests and cases 'this week.' In reality [in Finland], numbers of cases and tests from this week cannot be reported before Wednesday [COVID-19] or Friday [other viruses in infectious diseases register] next week. Sentinel surveillance may have even longer delays.*

The APD process further brought to light the complexity of data in public health situations, emphasizing the significance of considering interdependencies. It highlighted that the system developed serves as a decision support tool and that the role of public health officials remains crucial in making informed decisions based on the presented information. One participant's feedback from Germany highlighted a specific example of complexity in data representation related to counting beds versus assessing capacity:

> *In Germany we do not really count the beds but the capacity. This is because to be able to add a new patient you need a bed, a room, ventilation, staff etc, or do you mean capacities with the term "bed"?*

This inquiry highlights the importance of clarity and precision in defining data elements to ensure accurate interpretation and decision-making within the tool. Finally, through their engagement with an expanding collection of interactive designs, users were able to identify gaps in the available indicators and provide specific requests for additional information, such as excess mortality data. This iterative process was instrumental in ensuring continuous and active user involvement throughout the project, enabling their valuable contributions to the design and development of the PANDEM-2 system.

Impact Assessment Findings and Mitigation Strategies

The impact assessment process identified 10 potential risks, all of which were deemed to have a low to medium likelihood of occurring. In total, 4 (40%) of the risks were ethical, 2 (20%) regarded privacy and data protection, and 4 (40%) related to social factors. These are summarized in Textbox 2 and reported in more detail elsewhere [38].

Textbox 2. Summary of the identified risks categorized into ethical, privacy and data protection, and social factors.

Ethical risks

- Accessibility and inclusivity: risk of excluding individuals with certain needs.
- User autonomy: risk of limiting end-users' control over their use of the training platform.
- Responsible use: risk of incorrect or unethical use of the training platform.
- Transparency and accountability: risks associated with the ability to account for actions taken or decisions made during the development of the training platform.

Privacy and data protection risks

- Security and access controls: risks relating to the protection of digital information from unwanted actions or unauthorized users, based on commonly accepted security frameworks, such as the confidentiality, integrity, and availability model.
- Data reidentification: risk associated with the correlation or linking of anonymous datasets with other data sources—especially those that have been collected by the state

Social risks

- Unintentional bias: risks associated with biased algorithms, biased data sources, and bias in training content that may lead to systematic errors that create prejudiced or unintended outcomes.
- Data quality and accuracy: risks associated with the ability to make informed decisions based on accurate information.
- Dependency on technology: risk of heavy reliance on the IT platform for critical operations, processes, or decision-making.
- Collaborative development: risk of inadequate collaboration with end users or public health authorities in the development and postproject stages to understand training needs and enhance the platform's capabilities.

The risks were effectively mitigated through several measures throughout the project. Design considerations extended beyond functionality to inclusivity and accessibility. The user interface was designed to accommodate multilingual users, providing options in both English and Romanian with provisions for translation into additional languages. Accessibility for individuals who are visually impaired was ensured through the selection of a color-blind–friendly palette. To facilitate ease of use, user-friendly training materials and guides were developed, offering clear instructions for system navigation. Supplementary features such as tooltips and contextual help were integrated, further enhancing the user experience. In addition, transparency was prioritized through the provision of accessible data sources and modeling assumptions used in the development process. Security was a paramount concern, prompting the adoption of Amazon Cloud, a reputable cloud computing service known for its robust security measures. The commitment to open collaboration and adaptability was evident in the decision to offer all components of the dashboard as open-source software. This approach encourages wider use, fosters adaptability, and invites collaboration across diverse contexts. Furthermore, a comprehensive developer guide in English was provided, serving as a valuable resource for future developers interested in building upon the PANDEM-2 platform or creating similar tools. End-users' feedback was continuously obtained and integrated to ensure the platform's suitability to diverse user needs.

To enhance the accuracy and relevance of training content, public health authorities' expertise was used in content development, and active end users engaged in patient surveillance and treatment assessed the training materials. This collaborative approach guarantees that the training content accurately reflects current information, best practices, and public health guidelines. Training exercises were based on scripts generated by public health authorities, adding objectivity and minimizing subjective biases or arbitrary preferences. These actions collectively enhance the accuracy and scientific validity of the training content.

Discussion

Principal Findings

The PANDEM-2 project aimed to develop an IT prototype system that offers a standardized and interactive platform for pandemic preparedness training and response to support decision-making on pandemic management. This paper outlined the development and potential of the PANDEM-2 dashboard, including methodologies used to gather users' insights to enhance its design and steps taken to identify and address potential risks associated with the technology.

The COVID-19 pandemic brought about an unprecedented surge in pandemic-related data at national, EU, and global levels, presenting a valuable opportunity to identify crucial data needs for effective pandemic response. PANDEM-2 leveraged user surveys to identify essential data requirements, not only for the ongoing pandemic but also for potential future health threats. Despite this increase in data, significant challenges emerged due to data availability, diverse data sources, nonstandardized formats, varying granularity, and data sharing restrictions among EU member states. These obstacles hindered timely data exchange for a coordinated EU-wide response and hampered cross-border response efforts. Addressing these data challenges became a central concern throughout the PANDEM-2 project, and consequently it emerged that it is imperative to proactively anticipate and address data requirements for potential future pandemics.

The EU initiated measures to address data sharing gaps through Regulation 2022/2731 on serious cross-border threats to health [39]. The Commission's responsibility includes defining mechanisms for data exchange, adhering to personal data protection rules and information exchange security. In addition, Article 5 of Regulation 2022/2730 on serious cross-border threats to health specifically refers to the development of the Union health crisis and pandemic plan "to promote effective and coordinated response to cross-border health threats at Union level" [40]. Specifically, article 5, point 5, states, "In order to ensure the operation of the Union preparedness and response plan, the Commission shall conduct stress tests, exercises and in-action and after-action reviews with Member States and update the plan as necessary."

The development process of PANDEM-2 aligns with these measures, both in terms of data sharing and exercises. Lessons learned in the development of the platform offer insights into data analysis, sharing, and reporting methods, effectively supporting public health actions and control measures. The PANDEM-2 platform's design, built on an open architecture and open data sources, ensures security, stability, extensibility, and flexibility, allowing it to adapt to evolving data needs and data sharing practices.

The PANDEM-2 platform is also a training platform, designed to facilitate both national and cross-border simulation exercises. A simulation exercise is a form of training or evaluation of capabilities involving the simulation of an emergency [41]. Operationally, there is the potential for the PANDEM-2 dashboard to be deployed at regional, subnational, national, and EU levels to conduct simulation exercises aimed at stress-testing pandemic response capacity. A recent large-scale functional exercise involving 2 public health emergency operating centers within the PANDEM-2 consortium (the Robert Koch Institute in Germany and the National Institute for Public Health and the Environment in the Netherlands) supports this position. The aim of this exercise was to explore the coordinated response to a large-scale pandemic due to a novel strain of avian influenza in Europe with the objective of testing and evaluating the functionalities of the PANDEM-2 IT system and associated tools. During the 2-day exercise, participants received 44 injects to develop the scenario and provide specific tasks to complete within their country's context. A manuscript outlining the implementation of this exercise is currently under review [11]. Evaluation of the exercise comprised both qualitative (hot-wash group interviews and after-action interviews) and quantitative (questionnaires and interviews), and feedback indicated that the PANDEM-2 dashboard met or exceeded their technical requirements in terms of user experience. Moreover, participants agreed that the dashboard addresses a current gap in national and cross-border training capacity for pandemic preparedness and response.

Alongside addressing data requirements, effective methods for data presentation need to be considered. The PANDEM-2 platform's participatory design approach emphasizes the importance of involving end users in the design of a system. Through this interactive and iterative approach, users can define system requirements, share domain expertise, and identify high-priority indicators for pandemic management. This approach within the PANDEM-2 platform enabled the creation of a design framework that offers adaptability and scalability, ensuring consistent interactions even as indicators evolve.

The participatory design approach also supported the effective knowledge transfer among end users and the technical team. This was crucial to enable the design team to understand the requirements of diverse and specialized domains and to facilitate the design of novel technical solutions within those domains. Ongoing research endeavors should be directed toward enhancing user engagement with intricate software in specialized areas, such as epidemiological and resource models, balancing user-friendly interfaces with necessary complexity, and exploring efficient techniques for presenting voluminous datasets and findings.

Data storytelling has become one such technique for enhancing user engagement and is often seen as an intuitive way to drive decision-making and reduce the risk of information overload. The participatory design process revealed that managing multiple decision narratives would add unnecessary complexity. The dashboard approach chosen provides flexibility across public health scenarios like surveillance and training. The ability to customize views and generate reports makes it a more scalable solution, especially for our primary users, public health officers, and medics trained in data analysis, who require more in-depth exploration than storytelling offers. The dashboard includes filtering, tagging, and prioritization features to address information overload, but managing complex data while meeting the needs of expert users remains a challenge. The participatory design process helped refine key indicators to present concise, relevant information, but balancing exploration and clarity is an ongoing effort in systems like PANDEM-2.

The identification of high-priority indicators and the possibility of consolidating them into a singular variable that captures the status of a pandemic merits in-depth investigation. This strategy could streamline decision-making processes and yield immediate insights into the dynamic nature of the pandemic across its various stages.

Summary

During the project, challenges related to data standardization across consortium members became evident, particularly due to the lack of data in the TESSy format. In addition, many countries had regional divisions that differed from those used by the ECDC (NUTS). End-user experts also highlighted the critical role dashboards play in pandemic preparedness and response. The findings underscore the importance of making data for these indicators accessible for public health purposes.

The project's commitment to responsible innovation and user-centric principles was demonstrated through its robust impact assessment process, which prioritized key elements such as privacy-by-design and adherence to GDPR as well as regulations governing health technology development. This iterative process ensured that the platform remained relevant and user-friendly. Transparency stands as a core value of the PANDEM-2 project, demonstrating its commitment to openness and accountability. All project resources have been made publicly available for interested parties to access and review on the project website [42] and the European research platform Cordis [43], where stakeholders can find comprehensive information about data sources used, project procedures, and methods.

While the PANDEM-2 dashboard holds promise for effectively managing real-time data during future pandemics, successful integration with existing IT systems within public health agencies will be a necessary step. In the interim, the dashboard stands as an innovative training platform, providing member states the opportunity to assess their pandemic preparedness capacity through collaborative cross-border exercises developed by the PANDEM-2 project. Despite the limitations encountered during the project, our approach and findings provide valuable insights through open-access materials and tools to support public health agencies in EU member states to better prepare for upcoming pandemics.

To ensure comprehensive and transparent reporting of the development process, this study followed the Guidelines for Intervention Development checklist, documenting each stage of the dashboard's design and stakeholder engagement (Multimedia Appendix 13) [44].

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation program (grant 883285). The material presented and views expressed here are the responsibility of the authors only. The EU Commission takes no responsibility for any use made of the information set out. The authors would like to acknowledge the contribution of the PANDEM-2 public health agencies and first responder partners for their valuable input to the design and development process. This publication has emanated from research supported in part by a grant from Science Foundation Ireland (grant 12/RC/2289_P2; Insight SFI Centre for Data Analytics).

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Security and data protection aspects of the dashboard.
[DOCX File , 18 KB - publichealth_v11i1e52119_app1.docx]

Multimedia Appendix 2
Dashboards reviewed.
[XLSX File (Microsoft Excel File), 39 KB - publichealth_v11i1e52119_app2.xlsx]

Multimedia Appendix 3
List of COVID-19–related projects.
[XLSX File (Microsoft Excel File), 68 KB - publichealth_v11i1e52119_app3.xlsx]

Multimedia Appendix 4
Example of data visualization entered in document.
[PDF File (Adobe PDF File), 225 KB - publichealth_v11i1e52119_app4.pdf]

Multimedia Appendix 5
Checklist for reporting results of the initial user requirements survey.
[DOCX File , 16 KB - publichealth_v11i1e52119_app5.docx]

Multimedia Appendix 6
Checklist for reporting results of internet e-surveys: participatory design.
[DOCX File , 16 KB - publichealth_v11i1e52119_app6.docx]

Multimedia Appendix 7
Example user stories.
[DOCX File , 16 KB - publichealth_v11i1e52119_app7.docx]

Multimedia Appendix 8
Data availability and prioritization survey questions.
[DOCX File , 14 KB - publichealth_v11i1e52119_app8.docx]

Multimedia Appendix 9
Final list of variables by variable category and requirement type.
[DOCX File , 25 KB - publichealth_v11i1e52119_app9.docx]

Multimedia Appendix 10
Existing data in the European Surveillance System format.
[PNG File , 53 KB - publichealth_v11i1e52119_app10.png]

Multimedia Appendix 11
Example design of dashboard page illustrating key components: (1) explanation header; (2) indicator card; (3) map; and (4) graph. NUTS: Nomenclature of Territorial Units for statistics (a hierarchical system for dividing up the economic territory of the European Union). Implemented dashboard page in the PANDEM-2 prototype, reflecting design feedback with features like customizable reports and enhanced data visualization.
[PDF File (Adobe PDF File), 640 KB - publichealth_v11i1e52119_app11.pdf]

Multimedia Appendix 12
Illustrations of initial and final dashboard designs: progression of the map component.
[DOCX File , 541 KB - publichealth_v11i1e52119_app12.docx]

Multimedia Appendix 13
Guidelines for Intervention Development checklist.
[DOCX File , 30 KB - publichealth_v11i1e52119_app13.docx]

References

1. Jit M, Ananthakrishnan A, McKee M, Wouters OJ, Beutels P, Teerawattananon Y. Multi-country collaboration in responding to global infectious disease threats: lessons for Europe from the COVID-19 pandemic. Lancet Reg Health Eur 2021 Oct;9:100221 [FREE Full text] [doi: 10.1016/j.lanepe.2021.100221] [Medline: 34642675]
2. Haldane V, Jung AS, Neill R, Singh S, Wu S, Jamieson M, et al. From response to transformation: how countries can strengthen national pandemic preparedness and response systems. BMJ 2021 Nov 28;375:e067507 [FREE Full text] [doi: 10.1136/bmj-2021-067507] [Medline: 34840139]
3. Lencucha R, Bandara S. Trust, risk, and the challenge of information sharing during a health emergency. Global Health 2021 Feb 18;17(1):21 [FREE Full text] [doi: 10.1186/s12992-021-00673-9] [Medline: 33602281]
4. Badker R, Miller K, Pardee C, Oppenheim B, Stephenson N, Ash B, et al. Challenges in reported COVID-19 data: best practices and recommendations for future epidemics. BMJ Glob Health 2021 May;6(5):e005542 [FREE Full text] [doi: 10.1136/bmjgh-2021-005542] [Medline: 33958393]
5. Tacconelli E, Gorska A, Carrara E, Davis RJ, Bonten M, Friedrich AW, et al. Challenges of data sharing in European covid-19 projects: a learning opportunity for advancing pandemic preparedness and response. Lancet Reg Health Eur 2022 Oct;21:100467 [FREE Full text] [doi: 10.1016/j.lanepe.2022.100467] [Medline: 35942201]
6. Rosenthal A, Waitzberg R. The challenges brought by the COVID-19 pandemic to health systems exposed pre-existing gaps. Health Policy Open 2023 Dec;4:100088 [FREE Full text] [doi: 10.1016/j.hpopen.2022.100088] [Medline: 36536931]
7. Kokki M, Ammon A. Preparing Europe for future health threats and crises - key elements of the European Centre for Disease Prevention and Control's reinforced mandate. Euro Surveill 2023 Jan;28(3):2300033 [FREE Full text] [doi: 10.2807/1560-7917.ES.2023.28.3.2300033] [Medline: 36695487]
8. Katapally TR, Ibrahim ST. Digital health dashboards for decision-making to enable rapid responses during public health crises: replicable and scalable methodology. JMIR Res Protoc 2023 Jun 30;12:e46810 [FREE Full text] [doi: 10.2196/46810] [Medline: 37389905]
9. Tighe C. PANDEM-2 dashboard demo. Zenodo. 2023 Sep 20. URL: https://zenodo.org/records/8363810 [accessed 2024-09-14]
10. PANDEM-2 dashboard. PANDEM-2. URL: https://pandem-2.eu/pandem-2-dashboard/ [accessed 2024-09-14]
11. Houareau C, Spieker C, Grote U, Perseke K, An der Heiden M, Caglar R, et al. The PANDEM-2 Simulation Exercise: Training the Coordinated Response to a Large-Scale Pandemic in 2 European Public Health Emergency Operations Centers. Disaster Medicine and Public Health Preparedness 2025 Jan 14;19(14):e14 [FREE Full text] [doi: 10.1017/dmp.2024.298] [Medline: 39807047]

12. MongoDB homepage. MongoDB. URL: https://www.mongodb.com/ [accessed 2024-08-25]
13. NodeJS homepage. NodeJS. URL: https://nodejs.org/en [accessed 2024-08-28]
14. Chaniotis IK, Kyriakou KI, Tselikas ND. Is Node.js a viable option for building modern web applications? A performance evaluation study. Computing 2014 Mar 28;97:1023-1044. [doi: 10.1007/s00607-014-0394-9]
15. Angular homepage. Angular. URL: https://angular.io [accessed 2024-09-03]
16. Highcharts homepage. Highcharts. URL: https://www.highcharts.com/ [accessed 2024-09-10]
17. Pandem-source. GitHub. URL: https://github.com/Pandem2/pandem-source [accessed 2024-09-10]
18. Overview - Eurostat. Eurostat. URL: https://ec.europa.eu/eurostat/web/nuts/overview [accessed 2024-08-28]
19. GeoNames homepage. GeoNames. URL: https://www.geonames.org [accessed 2024-09-03]
20. COVID-19 data hub homepage. COVID-19 Data Hub. URL: https://covid19datahub.io [accessed 2024-09-10]
21. Chen H, Gala JL, Bonjean M, Ambroise J, Zayed O, Buitelaar P, et al. A COVID-19 European data set to support training in pandemic management. Zenodo. 2023. URL: https://zenodo.org/records/8339303 [accessed 2025-01-06]
22. Go.Data: managing complex data in outbreaks. World Health Organization. URL: https://www.who.int/godata [accessed 2024-09-10]
23. Flightradar24 homepage. Flightradar24. URL: https://www.flightradar24.com [accessed 2024-08-28]
24. Amazon web services cloud security homepage. Amazon Web Services. URL: https://aws.amazon.com/security [accessed 2024-09-03]
25. The AWS shared responsibility model and GDPR. Amazon Web Services. URL: https://aws.amazon.com/blogs/security/the-aws-shared-responsibility-model-and-gdpr/ [accessed 2024-09-10]
26. Amazon GuardDuty. Amazon Web Services. URL: https://aws.amazon.com/guardduty [accessed 2024-09-03]
27. AWS key management service. Amazon Web Services. URL: https://aws.amazon.com/kms/ [accessed 2024-08-28]
28. Few S. Information Dashboard Design: The Effective Visual Communication of Data. Sebastopol, CA: O'Reilly Media, Inc; 2006.
29. Franklin A, Gantela S, Shifarraw S, Johnson TR, Robinson DJ, King BR, et al. Dashboard visualizations: supporting real-time throughput decision-making. J Biomed Inform 2017 Jul;71:211-221 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.024] [Medline: 28579532]
30. Lennerholt C, Van Laere J, Söderström E. Implementation challenges of self-service business intelligence: a literature review. In: Proceedings of the 51st Hawaii International Conference on System Sciences. 2018 Presented at: HICSS 2018; January 3-6, 2018; Hilton Waikoloa Village, HI. [doi: 10.24251/hicss.2018.631]
31. Shao H, Martinez-Maldonado R, Echeverria V, Yan L, Gasevic D. Data storytelling in data visualisation: does it enhance the efficiency and effectiveness of information retrieval and insights comprehension? In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. 2024 Presented at: CHI '24; May 11-16, 2024; Honolulu, HI. [doi: 10.1145/3613904.3643022]
32. Cohn M. User Stories Applied: For Agile Software Development. Boston, MA: Addison-Wesley Publishing Company; 2004.
33. Lenarduzzi V, Taibi D. MVP explained: a systematic mapping study on the definitions of minimal viable product. In: Proceedings of the 42th Euromicro Conference on Software Engineering and Advanced Applications. 2016 Presented at: SEAA 2016; August 31-September 2, 2016; Limassol, Cyprus. [doi: 10.1109/seaa.2016.56]
34. Simonsen J, Robertson T. Routledge International Handbook of Participatory Design. Milton Park, UK: Routledge; 2013.
35. Conducting privacy impact assessments code of practice. Information Commissioner's Office. 2014. URL: https://ico.org.uk/media/about-the-ico/consultations/2052/draft-conducting-privacy-impact-assessments-code-of-practice.pdf [accessed 2024-09-03]
36. ISO/IEC 29134:2017(en) Information technology — security techniques — guidelines for privacy impact assessment. International Organization for Standardization. 2017. URL: https://www.iso.org/obp/ui/#iso:std:iso-iec:29134:ed-1:v1:en [accessed 2024-08-28]
37. The European Surveillance System (TESSy). European Centre for Disease Prevention and Control. 2017 May 17. URL: https://www.ecdc.europa.eu/en/publications-data/european-surveillance-system-tessy [accessed 2024-09-10]
38. Pandemic preparedness and response. European Commission. URL: https://cordis.europa.eu/project/id/883285/results [accessed 2024-09-10]
39. Regulation 2022/2371 - serious cross-border threats to health. EU Monitor. URL: https://www.eumonitor.eu/9353000/1/j4nvk6yhcbpeywk_j9vvik7m1c3gyxp/vlync5alazzz#:~:text=This%20Regulation%20should%20strengthen%20and,European%20Parliament%20and%20of%20the [accessed 2024-09-03]
40. Regulation (EU) 2022/2370 of the European parliament and of the council of 23 November 2022. Official Journal of the European Union. 2022 Dec 6. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R2370&from=EN [accessed 2024-08-28]
41. Simulation exercises. World Health Organization. URL: https://www.who.int/emergencies/operations/simulation-exercises [accessed 2024-09-03]
42. EU pandemic research. PANDEM-2. URL: https://pandem-2.eu/ [accessed 2024-09-10]

43. CORDIS - EU research results. European Commission. URL: https://cordis.europa.eu/ [accessed 2024-09-03]
44. Duncan E, O'Cathain A, Rousseau N, Croot L, Sworn K, Turner KM, et al. Guidance for reporting intervention development studies in health research (GUIDED): an evidence-based consensus study. BMJ Open 2020 Apr 08;10(4):e033516 [FREE Full text] [doi: 10.1136/bmjopen-2019-033516] [Medline: 32273313]

Abbreviations

APD: asynchronous participatory design
API: application programming interface
AWS: Amazon Web Services
ECDC: European Centre for Disease Prevention and Control
ETL: extract, transform, and load
EU: European Union
GDPR: General Data Protection Regulation
MVP: minimal viable product
NUTS: Nomenclature of Territorial Units for Statistics
TESSy: European Surveillance System
WHO: World Health Organization

Edited by A Mavragani; submitted 25.08.23; peer-reviewed by KS Sahu, O Leal Neto; comments to author 18.07.24; revised version received 16.09.24; accepted 23.09.24; published 05.03.25.

Please cite as:
Tighe C, Ngongalah L, Sentís A, Orchard F, Pacurar GA, Hayes C, Hayes JS, Toader A, Connolly MA
Building and Developing a Tool (PANDEM-2 Dashboard) to Strengthen Pandemic Management: Participatory Design Study
JMIR Public Health Surveill 2025;11:e52119
URL: https://publichealth.jmir.org/2025/1/e52119
doi:10.2196/52119
PMID:40053759

Suspected Mpox Symptoms and Testing in Men Who Have Sex With Men in the United States: Cross-Sectional Study

Kaitlyn Atkins[1,2], MPH, PhD; Thomas Carpino[2], MPH; Amrita Rao[2], ScM, PhD; Travis Sanchez[3], DVM, MPH; O Winslow Edwards[3], MPH; Marissa Hannah[3], MPH; Patrick S Sullivan[3], DVM, PhD; Yasmin P Ogale[4], MSPH, PhD; Winston E Abara[4], MD, PhD; Kevin P Delaney[4], MPH, PhD; Stefan D Baral[2], MPH, MD

[1]Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[2]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[3]Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA, United States

[4]Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:
Kaitlyn Atkins, MPH, PhD
Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

Abstract

Background: The 2022 mpox outbreak in the United States disproportionately affected gay, bisexual, and other men who have sex with men (GBMSM). Uptake of mpox testing may be related to symptomology, sociodemographic characteristics, and behavioral characteristics.

Objective: This study aimed to describe suspected mpox symptoms and testing uptake among a sample of GBMSM recruited via the internet in the United States in August 2022.

Methods: We conducted a rapid internet-based mpox survey from August 5 to 15, 2022, among cisgender men 15 years and older who had previously participated in the 2021 American Men's Internet Survey. We estimated the prevalence of suspected mpox symptoms (fever or rash or sores with unknown cause in the last 3 mo) and uptake of mpox testing. We calculated adjusted prevalence ratios (aPRs) and 95% CIs for associations between participant characteristics and suspected mpox symptoms and summarized characteristics of GBMSM reporting mpox testing. Among symptomatic GBMSM who did not receive mpox testing, we described testing self-efficacy, barriers, and facilitators.

Results: Of 824 GBMSM, 126 (15.3%) reported at least 1 mpox symptom in the last 3 months; 58/126 (46%) with rash or sores, 57 (45.2%) with fever, and 11 (8.7%) with both. Increased prevalence of suspected mpox symptoms was associated with condomless anal sex (CAS; aPR 1.53, 95% CI 1.06 - 2.20). Mpox testing was reported by 9/824 GBMSM (1%), including 5 with symptoms. Most GBMSM reporting mpox testing were non-Hispanic White men (7/9 vs 1 Black and 1 Hispanic or Latino man), and all 9 lived in urban areas. Most reported having an sexually transmitted infections test (8/9), 2 or more partners (8/9), CAS (7/9), and group sex (6/9) in the last 3 months. Of those tested, 3 reported living with HIV and all were on treatment, whereas the remaining 6 men without HIV reported current pre-exposure prophylaxis (PrEP) use. Of symptomatic GBMSM who did not report mpox testing, 47/105 (44.8%) had low mpox testing self-efficacy. Among those with low self-efficacy, the most common barriers to testing were not knowing where to get tested (40/47, 85.1%) and difficulty getting appointments (23/47, 48.9%). Among those with high testing self-efficacy (58/105, 55.2%), the most common facilitators to testing were knowing where to test (52/58, 89.7%), convenient site hours (40/58, 69%), and low-cost testing (38/58, 65.5%).

Conclusions: While all GBMSM who reported testing for mpox were linked to HIV treatment or PrEP, those with symptoms but no mpox testing reported fewer such links. This suggests targeted outreach is needed to reduce structural barriers to mpox services among GBMSM in rural areas, Black and Hispanic or Latino GBMSM, and GBMSM living with HIV. Sustaining and scaling community-tailored messaging to promote testing and vaccination represent critical interventions for mpox control among GBMSM in the United States.

(JMIR Public Health Surveill 2025;11:e57399) doi:10.2196/57399

KEYWORDS

mpox; emerging infectious diseases; men who have sex with men; sexual health; testing; monkeypox; epidemic; MSM; United States; rural; cross-sectional study; gay; bisexual; symptomology; sociodemographic; online sample; self-efficacy; rash; fever; HIV; HIV prevention; GBMSM; Black and Hispanic; LGBTQ; Latino; public health; surveillance; sores

Introduction

Human mpox is an orthopox virus first identified in the Democratic Republic of the Congo and endemic to regions of Central and Western Africa, which has caused localized outbreaks for decades [1]. mpox is classified into 2 clades; of the two, clade I mpox is associated with more severe symptoms and is highly lethal, with an estimated 10% mortality rate and several thousands of cases annually in Central Africa, including a 2023 - 2024 outbreak in the Democratic Republic of the Congo [2]. While clade I mpox has not been observed outside the African continent, clade II mpox has been occasionally reported internationally, often linked to travel [3]. Although clade II infections are associated with a much lower mortality rate, they have resulted in increasing global morbidity, including a global outbreak in 2022 [3]. From 2022 to 2024, there have been 95,000 confirmed cases of clade II mpox globally, including several deaths, predominantly affecting gay, bisexual, and other men who have sex with men (GBMSM) [1].

The first case in the 2022 mpox outbreak was identified in the United States on May 17, 2022, triggering a nationwide response to identify and monitor new cases and deliver the mpox vaccine [4]. As of March 5, 2024, approximately 32,000 mpox cases have been reported in the United States [1], and over 90% of cases in the United States have occurred among cisgender men, the majority of whom had recent sexual contact with a man [5].

As elsewhere, the 2022US mpox outbreak and its disproportionate impacts on GBMSM have been attributed to connected sexual networks. Moreover, the US mpox outbreak demonstrated disparities in infections associated with structural determinants of health such as racism: communities of color reported higher burdens of infections, but lower uptake of testing and vaccines [1,6]. In studies of testing for sexually transmitted infections (STIs) and COVID-19, lower socioeconomic status has been associated with reduced testing uptake [7,8]. However, few studies to date have examined how structural determinants may have shaped mpox testing in a large, national sample of MSM [9,10].

Beyond structural factors, engagement in mpox screening and testing services may be driven by individual factors, particularly risk perception and behavioral indication for mpox tests [11]. For example, GBMSM who report regular condom use have been found to have lower engagement in testing for HIV [12] and bacterial STIs [13]. Condom use and engagement in other sexual behaviors associated with mpox (such as group sex) [14,15] may therefore be associated with increased engagement in mpox testing.

New infections have slowed significantly since the peak of the outbreak [1,16], but equitable testing remains key to contract tracing, treatment, and linkage to vaccination programs. Understanding potential barriers and facilitators to mpox testing during the peak of the 2022 outbreak can inform strategies for future mpox or other STI outbreaks. The current work aims to fill knowledge gaps around these barriers and facilitators by describing suspected mpox symptoms and testing from an mpox study among a sample of prior participants in the American Men's Internet Survey (AMIS).

Methods

Recruitment

The AMIS mpox study was an internet-based survey conducted from August 5 to 15, 2022 in GBMSM aged 15 years and older who had participated in AMIS 2021. Methods for the annual AMIS have been described elsewhere [17]. Annual AMIS participants were recruited via social media and deemed eligible if they identified as cisgender male, aged ≥15 years, had ever had sex with a man, and lived in the United States. For the mpox study, the sampling frame included GBMSM who had completed the AMIS 2021 survey cycle, had sex with a man in the past year, and had consented to be recontacted for future research (N=2999).

Ethical Considerations

This study received ethical approval from the Emory University Institutional Review Board (IRB00047676). All activities were conducted consistent with US Centers for Disease Control and Prevention policies and regulations. All participants provided informed consent to participate and were not compensated. All study data were deidentified.

Procedures

Potential participants (ie, all AMIS 2021 participants who had consented to be recontacted for future research, N=2999) were contacted via email with an invitation to participate and a secure survey link. Potential participants were contacted up to twice and provided electronic informed consent after eligibility screening. The internet-based survey included questions in English about demographics, sexual behavior, substance use, HIV and STI testing and diagnosis, HIV pre-exposure prophylaxis (PrEP) use, and mpox knowledge, symptoms, testing, diagnosis, vaccination, stigma, and behavior change.

Measures

We measured participant age, race or ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic or Latino, other), US census region (Northeast, Midwest, South, or West), urbanicity based on 2013 National Center for Health Statistics classification [18], health insurance (none, private, public, other), self-reported HIV status, current antiretroviral therapy use (among people living with HIV, [19]) and current PrEP use (among those without HIV).

Participants reported, in the past 3 months, their number of sexual partners (1 or ≥2), participation in group sex, bathhouses or sex clubs, or sex parties, and any illicit or injection drug use. They also reported condomless anal sex (CAS) in the past 12 months, CAS with partners of a different HIV status in the past 3 months, HIV testing in the past 12 months, and STI testing in the past 3 months.

Outcomes

We assessed symptoms suggesting mpox by asking: "In the past three months, have you had a (1: fever or 2: new rash/sores on the skin) and didn't know the cause?" For GBMSM reporting rashes, we asked their location and pain severity (rated 0 - 10), and whether, after GBMSM first noticed the rash or sores, they changed in appearance, number, or location.

We measured mpox testing by self-report. For GBMSM reporting no mpox testing, we assessed agreement (4-point Likert scale) with the statement: “If I thought I had mpox, I could get a test if I wanted it.” GBMSM who strongly or somewhat agreed were considered to have high mpox testing self-efficacy, and those who strongly or somewhat disagreed as having low self-efficacy. We also asked participants what made it easy (for those with high self-efficacy) or difficult (for those with low self-efficacy) to get a test; parallel response options included location, eligibility, appointments, using one’s own provider, site times, privacy, and cost.

Statistical Analysis

We calculated frequencies for participant characteristics, symptoms suggesting mpox, and mpox testing. We calculated unadjusted prevalence ratios (PRs) and 95% CIs for associations between participant characteristics and symptoms suggesting mpox. Factors found to have a significant association (α<.05) with symptoms were included in a multivariable model to calculate adjusted prevalence ratios (aPRs). The final model included age, race or ethnicity, PrEP use, STI testing, and CAS. We assessed model fit using Hosmer-Lemeshow test, which has been recommended for use with binomially distributed data [20]. We used frequencies to summarize the characteristics of GBMSM reporting mpox testing. Among symptomatic GBMSM who did not report mpox testing, we summarized testing self-efficacy, barriers, and facilitators. All analyses were performed in Stata 15.1 (StataCorp) [21].

Results

Sample Characteristics

Of 2999 GBMSM in the sampling frame, 824 completed the mpox study survey for a response rate of 27.5%. The majority of mpox study participants were aged ≥40 years (478/824, 58%), non-Hispanic White (581/824, 70.9%), residents of urban areas (770/824, 93.4%), and had private health insurance (564/824, 68.6%, Table 1). In the last 3 months, a minority reported CAS with a partner of a different HIV status (220/824, 26.7%), group sex (238/824, 29.1%), or attending a sex club (147/824, 19.7%), or sex party (111/824, 14.9%). Just over two-thirds (575/824, 69.8%) reported past-year HIV testing, and around half (408/824, 49.6%) reported STI testing in the last 3 months. Most (720/824, 87.4%) were not people living with HIV and 44.7% of these (316/720, 44.7%) were currently using PrEP; of 104 people living with HIV in the sample, 99 (96.1%) were currently on HIV treatment.

Table . Characteristics of participants in the American Men’s Internet Survey mpox substudy, United States, August 2022 (n=824).

Characteristics		Participants (n=824), n (%)
Age (years)		
	15 - 24	46 (5.6)
	25 - 29	86 (10.4)
	30 - 39	214 (26)
	40 or older	478 (58)
Race or ethnicity		
	Non-Hispanic White	581 (70.9)
	Non-Hispanic Black	87 (10.6)
	Hispanic or Latino	88 (10.7)
	Other	64 (7.8)
US census region		
	Northeast	174 (21.1)
	Midwest	141 (17.1)
	South	305 (37)
	West	204 (24.8)
Urbanicity		
	Rural	54 (6.6)
	Urban	770 (93.4)
Health insurance		
	None	24 (2.9)
	Private	564 (68.6)
	Public	164 (20)
	Other or multiple	70 (8.5)
HIV test, last 12 months		
	No	249 (30.2)
	Yes	575 (69.8)
Bacterial STI[a] test, last 3 months		
	No	410 (49.8)
	Yes	409 (49.6)
HIV status		
	Not living with HIV	720 (87.4)
	Living with HIV	104 (12.6)
Current ART[b] use (n=103)[c]		
	No	4 (3.9)
	Yes	99 (96.1)
Current PrEP[d] use (n=707)		
	No	391 (55.3)
	Yes	316 (44.7)
Group sex, last 3 months		
	No	580 (70.9)
	Yes	238 (29.1)

Characteristics		Participants (n=824), n (%)
Sex club, last 3 months		
	No	599 (80.3)
	Yes	147 (19.7)
Sex party, last 3 months		
	No	634 (85.1)
	Yes	111 (14.9)
CAS[e] with serodifferent partner, last 3 months		
	No	604 (73.3)
	Yes	220 (26.7)

[a] STI: sexually transmitted infection.

[b] ART: antiretroviral therapy.

[c] One participant with HIV did not provide information on ART use.

[d] PrEP: HIV pre-exposure prophylaxis.

[e] CAS: condomless anal sex.

Prevalence of Symptoms Suggesting mpox and Associated Factors

In total, 126/824 (15.3%) participants reported at least 1 mpox symptom in the last 3 months (Table 2): 58/126 (46%) with rash or sores, 57 (45%) with fever, and 11 (9%) with both. In unadjusted analyses, there was decreased prevalence of symptoms suggesting mpox among participants aged ≥40 years versus 15 - 24 years (PR 0.48, 95% CI 0.28 - 0.82) and among Hispanic or Latino versus non-Hispanic White participants (PR 0.44, 95% CI 0.20 - 0.97), and increased prevalence among participants reporting current PrEP use (PR 1.48, 95% CI 1.05 - 2.09), recent STI tests (PR 1.40, 95% CI 1.01 - 1.95), and CAS in the last 12 months (PR 1.64, 95% CI 1.14 - 2.35). In our adjusted model, we failed to reject the Hosmer-Lemeshow null hypothesis (Hosmer-Lemeshow $\chi_8^2$5.01; *P*=.76), indicating adequate model fit [20]. In adjusted analyses, prevalence of symptoms suggesting mpox remained lower among GBMSM aged ≥40 years versus 15 - 24 years (aPR 0.49, 95% CI 0.28 - 0.83), and higher among those reporting CAS (aPR 1.53, 95% CI 1.06 - 2.20).

Table . Symptoms suggesting mpox and testing among participants in the American Men’s Internet Survey mpox substudy, United States, August 2022 (n=824). Bold text indicates measures of association found to be statistically significant at α<.05; that is, confidence intervals that did not cross 1.00.

	At least 1 symptom suggesting mpox[a], last 3 months			Symptoms, no mpox testing	Ever tested for mpox
	Participants (n=126), n (%)	PR[b] (95% CI)	aPR[c] (95% CI)	Participants (n=114), n (%)	Participants (n=9), n (%)
Symptoms suggesting mpox					
No fever or rash	—[d]	—	—	—	4 (44.4)
Fever only	57 (45.2)	—	—	53 (46.5)	0 (0)
Rash or sores only	58 (46)	—	—	52 (45.6)	4 (44.4)
Fever and rash or sores	11 (8.7)	—	—	9 (7.9)	1 (11.1)
Age (years)					
15 - 24	12 (9.5)	REF[e]	REF	10 (8.8)	1 (11.1)
25 - 29	14 (11.1)	0.61 (0.31 - 1.21)	0.64 (0.33 - 1.26)	12 (10.5)	2 (22.2)
30 - 39	39 (31)	0.68 (0.39 - 1.20)	0.64 (0.37 - 1.12)	36 (31.6)	3 (33.3)
40 or older	61 (48.4)	0.48 (0.28 - 0.82)	0.49 (0.28 - 0.83)	56 (49.1)	3 (33.3)
Race or ethnicity					
Non-Hispanic White	90 (71.4)	REF	REF	81 (71.1)	7 (77.8)
Non-Hispanic Black	14 (11.1)	1.03 (0.62 - 1.73)	1.01 (0.60 - 1.71)	12 (10.5)	1 (11.1)
Hispanic or Latino	6 (4.8)	0.44 (0.20 - 0.97)	0.41 (0.19 - 0.91)	5 (4.4)	1 (11.1)
Other	16 (12.7)	1.63 (1.02 - 2.59)	1.44 (0.90 - 2.30)	16 (14)	0 (0)
Urbanicity					
Nonurban	11 (8.7)	REF	—	10 (8.8)	0 (0)
Urban	115 (91.3)	0.74 (0.42 - 1.28)	—	104 (91.2)	9 (100)
HIV status					
Not living with HIV	111 (88.1)	REF	—	102 (89.5)	6 (66.7)
Living with HIV	15 (11.9)	0.93 (0.56 - 1.53)	—	12 (10.5)	3 (33.3)
Current ART[f] use (n=15)					
No	0 (0)	—	—	0 (0)	0 (0)
Yes	15 (100)	—	—	12 (100)	3 (100)
Current PrEP[g] use (n=707)[h], [i]					
No	50 (45.5)	REF	REF	47 (46.5)	0 (0)
Yes	60 (54.5)	1.48 (1.05 - 2.09)	1.15 (0.75 - 1.75)	54 (53.5)	6 (100)
STI[j] test, last 3 months					
No	52 (41.3)	REF	REF	45 (39.5)	1 (11.1)

		At least 1 symptom suggesting mpox[a], last 3 months			Symptoms, no mpox testing	Ever tested for mpox
		Participants (n=126), n (%)	PR[b] (95% CI)	aPR[c] (95% CI)	Participants (n=114), n (%)	Participants (n=9), n (%)
	Yes	74 (58.7)	1.40 (1.01 - 1.95)	1.33 (0.96 - 1.85)	69 (60.5)	8 (88.9)
Number of partners, last 3 months						
	One	25 (21)	REF	—	23 (21.3)	0 (0)
	Two or more	94 (79%)	1.44 (0.95 - 2.17)	—	85 (78.7)	8 (100)
Group sex, last 3 months						
	No	81 (64.3%)	REF	—	76 (66.7)	3 (33.3)
	Yes	45 (35.7)	1.35 (0.97 - 1.88)	—	38 (33.3)	6 (66.7)
Sex club, last 3 months						
	No	88 (74)	REF	—	81 (75)	5 (62.5)
	Yes	31 (26)	1.44 (1.00 - 2.08)	—	27 (25)	3 (37.5)
Sex party, last 3 months						
	No	96 (80.7)	REF	—	89 (82.4)	5 (62.5)
	Yes	23 (19.3)	1.36 (0.91 - 2.05)	—	19 (17.6)	3 (37.5)
CAS[k], last 12 months						
	No	36 (28.6)	REF	REF	33 (28.9)	2 (22.2)
	Yes	90 (71.4)	1.64 (1.14 - 2.35)	1.53 (1.06 - 2.20)	81 (71.1)	7 (77.8)

[a] mpox: mpox.

[b] PR: prevalence ratio.

[c] aPR: adjusted prevalence ratio.

[d] Not applicable.

[e] REF: reference.

[f] ART: antiretroviral therapy.

[g] PrEP: HIV pre-exposure prophylaxis.

[h] The model with PrEP use was restricted to individuals without HIV, and included all listed covariates except HIV status.

[i] No PR reported given 100% of symptomatic people living with HIV reported current antiretroviral therapy use.

[j] STI: sexually transmitted infection.

[k] CAS: condomless anal sex.

Uptake of mpox Testing

Among 126 symptomatic GBMSM, 114 (90.5%) did not report mpox testing; their characteristics mirrored those of the overall sample (Table 2). mpox testing was reported by 9 participants (1%, Table 2), including 5 with reported symptoms. Most GBMSM reporting mpox testing were non-Hispanic White (7/9 vs 1 Black and 1 Hispanic or Latino man) and all 9 lived in urban areas. Most reported STI testing (8/9), 2 or more partners (8/9), CAS (7/9), and group sex (6/9) in the last 3 months. Three were people living with HIV, all on treatment; the remaining 6 without HIV reported current PrEP use.

Among 69 GBMSM experiencing recent rashes or sores with unknown cause, most located them on hands or arms, feet or legs, and genitals or pelvic area (Table 3). Over half (31/50, 62%) rated their pain as ≤2 on a 10-point scale. Most (38/69, 55%) reported no changes over time; the most common reported change was worsening appearance of rash or sores (22/69, 32%).

Table . Rash or sore location, severity, and changes among men having sex with men in the United States reporting new rashes or sores in the last 3 months (n=69).

		Experienced new rash or sores with unknown cause, last 3 months (n=69), n (%)
Location of rash		
	Hands or arms	26 (37.7)
	Feet or legs	24 (34.8)
	Genitals or pelvic area	19 (27.5)
	Chest	14 (20.3)
	Face	13 (18.8)
	Anus	10 (14.5)
	Mouth	7 (10.1)
Severity of rash or sores		
	0: no pain at all	10 (20)
	1 - 2	21 (42)
	3 - 4	11 (22)
	5 - 6	4 (8)
	7 - 8	4 (8)
	9 - 10: the worst pain possible	0 (0)
Changes in rash or sores		
	No changes	38 (55.1)
	Worsening appearance	22 (31.9)
	Increased number of rashes or sores in the same location	11 (15.9)
	Rashes or sores in new locations	7 (10.1)

Of the 126 symptomatic GBMSM, 105 (83%) did not report mpox testing and were offered questions about testing self-efficacy and barriers and facilitators to testing (Table 4). Among GBMSM with high testing self-efficacy (58/105, 55%), common facilitators to testing were knowing where to test (90%), convenient site hours (69%), and low-cost testing (38/58, 66%). Among those with low testing self-efficacy (47/105, 45%), the most common barriers to testing were not knowing where to get tested (40/47, 85%) and difficulty getting appointments (23/47, 49%).

Table . Characteristics of gay, bisexual, and other men who have sex with men with symptoms suggesting mpox who did not take up mpox testing (n=105).

Barriers or facilitators to testing	Low testing self-efficacy[a] (n=47), n (%)	High testing self-efficacy (n=58), n (%)
(Not) knowing where to get tested	40 (85.1)	52 (89.7)
Getting an appointment	23 (48.9)	29 (50)
Getting tested at own doctor	17 (36.2)	24 (41.4)
Convenience or testing site hours	11 (23.4)	40 (69)
(Not) knowing who is eligible	10 (21.3)	29 (50)
Cost considerations	7 (14.9)	38 (65.5)
Privacy considerations	6 (12.8)	19 (32.8)

[a] Assessed using the statement: “If I thought I had mpox, I could get a test if I wanted it.” Participants who strongly or somewhat agreed were considered to have high mpox testing self-efficacy, and those who strongly or somewhat disagreed as having low self-efficacy.

Discussion

Principal Findings

Our findings show that a large proportion of GBMSM with symptoms suggesting mpox did not access mpox testing and mpox testing was low or zero among GBMSM in rural areas and GBMSM who were younger, Black, Hispanic or Latino, or people living with HIV.

Compared with non-Hispanic White participants, we observed low levels of mpox testing among Black and Hispanic or Latino GBMSM in this study. While this likely reflects our overall sample, which was primarily non-Hispanic White, it is worth noting that this finding aligns with broader disparities not only in mpox cases and vaccine delivery, but also with the majority of US vaccine doses administered to date to non-Hispanic White individuals [1]. Our findings may reflect structural challenges with communication and knowledge of access to and locations of testing, laboratory capacity, and stigma [22]. As others have noted, additional data are needed to ensure equitable access to and communication on mpox testing, including routine collection of demographic data through testing programs [23,24].

There was no reported testing among GBMSM in rural or nonurban areas in our study. While few participants in our study were from rural areas, this parallels findings from an internet-based study also conducted in August 2022, which found that GBMSM in rural areas had lower perceived susceptibility to and severity of mpox, compared with urban GBMSM [25]. The authors of the earlier manuscript found further disparities in rural GBMSM's perceptions of mpox vaccine benefits and in their intention to be vaccinated, amidst potentially inadequate vaccine supply. Our findings likely reflect broader challenges with public health infrastructure in rural areas [26] and increased availability of HIV and STI services in urban areas, which may have facilitated linkages to mpox testing. While there are opportunities to leverage these services for mpox testing, given only half of potentially symptomatic GBMSM were engaged with PrEP and STI services, additional outreach may be needed.

Testing uptake was also low among GBMSM living with HIV. While only about 1 in 5 participants in our study were people living with HIV, we agree with others that tailored interventions can better serve this community with services, including mpox testing, postexposure prophylaxis for mpox exposure, and vaccination [27]. Current US surveillance data suggest that around 40% of mpox diagnoses were among people living with HIV [1]. Given recent evidence that suggests severe mpox in the context of advanced immunosuppression may be an AIDS-defining condition [28], structural interventions to improve access to mpox testing among people living with HIV, such as integrating mpox testing into routine care for people living with HIV, are critical to ensure early identification and treatment of new mpox cases.

We found that the most symptoms suggesting mpox in our sample were rashes or sores, most commonly in the arms, legs, and genitals. Rashes in the arms and legs were reported at a similar rate in our study as has been shown in broader US surveillance data; however, we saw less frequent rashes in the genitals and face [5]. In our study, rashes and sores were reported to be generally mild and commonly did not change after they were first observed. This may explain the relatively low uptake of mpox testing among potentially symptomatic GBMSM, because GBMSM with mild symptoms suggesting mpox may not have felt they needed a test or been eligible for testing at that time [5]. During clinical encounters with GBMSM (eg, STI testing), providers should continue to counsel GBMSM to seek mpox testing and vaccine if they develop unexplained rash or other symptoms consistent with mpox.

Limitations

A key limitation of these analyses is the small sample of GBMSM in our study who reported mpox testing (n=9), precluding advanced statistical analyses comparing the characteristics of GBMSM who tested for mpox with those who did not. Furthermore, we are unable to draw generalizable conclusions based on the characteristics of these 9 individuals. We recruited a convenience sample that is subject to selection bias; our respondents were predominantly non-Hispanic White, elderly, and privately insured and may have increased access to mpox testing services—the sample is thus unlikely to represent all GBMSM in the United States. Our use of self-reported data and recall period of 3 months for most outcomes may have introduced recall bias; furthermore, we are unable to confirm whether those with symptoms suggesting mpox indeed had acquired mpox. Given symptoms were reported in the last 3 months, in some cases, participant responses may have reflected symptoms before the first US mpox case on May 17, 2022. We were also not able to place individuals' locations at the time of symptom onset or testing, which may have affected testing uptake given the nonuniform roll-out of testing nationwide. Regarding symptoms, we only included items assessing the presence of fever and rash or sores; we did not inquire about other symptoms commonly reported among mpox cases such as malaise, headache, nausea or vomiting, pain, itching, or bleeding. Low mpox testing prevalence in this sample may be attributed to perceptions that symptoms, such as fever or rash, were unrelated to mpox and therefore did not require testing. Our findings also reflect the timing of data collection (August 2022); access to mpox testing and other services likely changed as the US outbreak response evolved. For example, testing uptake could have been limited due to sustained increases of telehealth services after the COVID-19 pandemic, particularly among wealthier, insured patients such as the GBMSM who comprised the majority of our sample [29]. Furthermore, access to testing over the course of the outbreak likely varied by geographic location, something we were unable to fully assess in this analysis.

Conclusions

Similar structural barriers that have long been known to increase the risk of other STIs were rapidly replicated during the mpox outbreak in 2022. Sustained disparities fostered by racism, stigma, and limited access to health care in rural settings contribute to ongoing morbidity and mortality for GBMSM and predispose many communities to unexpected new outbreaks and epidemics such as mpox. While the epidemiology of mpox continues to evolve, transmission risk remains present and characterizing optimal public health interventions for mpox is critical to both prevent future outbreaks and more rapidly respond if they occur.

Acknowledgments

This work was supported by the National Institute of Allergy and Infectious Disease (P30AI050409), National Institute of Mental Health (R01MH132150; R01MH110358; R25MH083620; P30MH136919), and National Institute of Nursing Research (R01NR020437).

Disclaimer

The results presented in this manuscript do not necessarily reflect the views of the National Institutes of Health or the US government.

Data Availability

The datasets analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

TS, KPD, and SDB developed the research questions and designed the study. KA conceived and conducted data analyses for this manuscript, and wrote the manuscript. All authors reviewed the manuscript and provided critical feedback.

Conflicts of Interest

None declared.

References

1. Mpox. Centers for Disease Control and Prevention (CDC). 2023. URL: https://www.cdc.gov/poxvirus/monkeypox/index.html [accessed 2024-12-12]
2. Kibungu EM, Vakaniaki EH, Kinganda-Lusamaki E, et al. Clade I-associated mpox cases associated with sexual contact, the Democratic Republic of the Congo. Emerg Infect Dis 2024 Jan;30(1):172-176. [doi: 10.3201/eid3001.231164] [Medline: 38019211]
3. Van Dijck C, Hoff NA, Mbala-Kingebeni P, et al. Emergence of mpox in the post-smallpox era-a narrative review on mpox epidemiology. Clin Microbiol Infect 2023 Dec;29(12):1487-1492. [doi: 10.1016/j.cmi.2023.08.008] [Medline: 37574113]
4. Minhaj FS, Ogale YP, Whitehill F, et al. Monkeypox outbreak - nine states, May 2022. MMWR 2022 Jun 10;71(23):764-769. [doi: 10.15585/mmwr.mm7123e1] [Medline: 35679181]
5. Kava CM, Rohraff DM, Wallace B, et al. Epidemiologic features of the monkeypox outbreak and the public health response - United States, May 17-October 6, 2022. MMWR Morb Mortal Wkly Rep 2022 Nov 11;71(45):1449-1456. [doi: 10.15585/mmwr.mm7145a4] [Medline: 36355615]
6. Alavian N, Mourad A, Woodhouse EW, et al. Disparities in mpox vaccination among priority populations during the 2022 outbreak. Open Forum Infect Dis 2023 Sep;10(9):fad434. [doi: 10.1093/ofid/ofad434] [Medline: 37662451]
7. Bien-Gund CH, Shaw PA, Agnew-Brune C, et al. HIV self-testing and risk behaviors among men who have sex with men in 23 US cities, 2017. JAMA Netw Open 2022 Dec 1;5(12):e2247540. [doi: 10.1001/jamanetworkopen.2022.47540] [Medline: 36534398]
8. Zimba R, Romo ML, Kulkarni SG, et al. Patterns of SARS-CoV-2 testing preferences in a national cohort in the United States: latent class analysis of a discrete choice experiment. JMIR Public Health Surveill 2021 Dec 30;7(12):e32846. [doi: 10.2196/32846] [Medline: 34793320]
9. Le Forestier JM, Page-Gould E, Chasteen A. Identity concealment may discourage health-seeking behaviors: evidence from sexual-minority men during the 2022 global mpox outbreak. Psychol Sci 2024 Feb;35(2):126-136. [doi: 10.1177/09567976231217416] [Medline: 38215021]
10. Hong C. Mpox on Reddit: a thematic analysis of online posts on mpox on a social media platform among key populations. J Urban Health 2023 Dec;100(6):1264-1273. [doi: 10.1007/s11524-023-00773-4] [Medline: 37580545]
11. Walsh-Buhi ML, Houghton RF, Valdez D, Walsh-Buhi ER. A theory-based assessment of mpox: findings from a nationally representative survey of U.S. adults. PLoS One 2024;19(3):e0299599. [doi: 10.1371/journal.pone.0299599] [Medline: 38489274]
12. Nelson KM, Pantalone DW, Gamarel KE, Carey MP, Simoni JM. Correlates of never testing for HIV among sexually active internet-recruited gay, bisexual, and other men who have sex with men in the United States. AIDS Patient Care STDS 2018 Jan;32(1):9-15. [doi: 10.1089/apc.2017.0244] [Medline: 29232170]

13. de Voux A, Bernstein KT, Kirkcaldy RD, Zlotorzynska M, Sanchez T. Self-reported extragenital chlamydia and gonorrhea testing in the past 12 months among men who have sex with men in the United States-American Men's Internet Survey, 2017. Sex Transm Dis 2019 Sep;46(9):563-570. [doi: 10.1097/OLQ.0000000000001032] [Medline: 31415039]
14. Copen CE, Delaney KP, Agnew-Brune C, et al. Modifications to sexual behaviors associated with mpox (monkeypox) virus transmission among persons presenting for mpox vaccination, Washington, DC, August–October, 2022. Sex Transm Dis 2024 Jan 1;51(1):54-60. [doi: 10.1097/OLQ.0000000000001889] [Medline: 37889944]
15. Sönmez İ, Martínez Riveros H, Folch C, et al. Egocentric sexual network analysis among gay and bisexual men who have sex with men with and without mpox infection. Sex Transm Infect 2023 Dec;99(8):541-547. [doi: 10.1136/sextrans-2023-055901] [Medline: 37704363]
16. Adams C, Kirby AE, Bias M, et al. Detecting mpox cases through wastewater surveillance — United States, August 2022–May 2023. MMWR Morb Mortal Wkly Rep 2024 Jan 18;73(2):37-43. [doi: 10.15585/mmwr.mm7302a3] [Medline: 38236784]
17. Wiatrek S, Zlotorzynska M, Rai R, Sullivan P, Sanchez T. The annual American Men's Internet Survey of behaviors of men who have sex with men in the United States: Key Indicators Report 2018. JMIR Public Health Surveill 2021 Mar 4;7(3):e21812. [doi: 10.2196/21812] [Medline: 33496669]
18. NCHS Urban-Rural Classification Scheme for counties. Centers for Disease Control and Prevention (CDC). 2014. URL: https://www.cdc.gov/nchs/data-analysis-tools/urban-rural.html#:~:text=A%20key%20feature%20of%20the,%E2%80%9D%20metro%20(akin%20to%20suburbs) [accessed 2025-01-06]
19. Carrico AW, Hunt PW, Neilands TB, et al. Stimulant use and viral suppression in the era of universal antiretroviral therapy. J Acquir Immune Defic Syndr 2019 Jan 1;80(1):89-93. [doi: 10.1097/QAI.0000000000001867] [Medline: 30272634]
20. Blizzard L, Hosmer DW. Parameter estimation and goodness-of-fit in log binomial regression. Biom J 2006 Feb;48(1):5-22. [doi: 10.1002/bimj.200410165] [Medline: 16544809]
21. New in Stata 15. StataCorp. 2017. URL: https://www.stata.com/stata15/ [accessed 2025-01-06]
22. Torres TS, Silva MST, Coutinho C, et al. Evaluation of mpox knowledge, stigma, and willingness to vaccinate for mpox: cross-sectional web-based survey among sexual and gender minorities. JMIR Public Health Surveill 2023 Jul 17;9:e46489. [doi: 10.2196/46489] [Medline: 37459174]
23. HHS monkeypox letter. Infectious Diseases Society of America. 2022 Jul 22. URL: https://www.idsociety.org/globalassets/bb-complex-pages/hhs-monkeypox-letter-july-2022-final.pdf [accessed 2025-01-06]
24. Nsoesie EO, Vu C. Monkeypox: where are the racial disparities? Lancet Reg Health Am 2023 Jan;17:100372. [doi: 10.1016/j.lana.2022.100372] [Medline: 36249074]
25. Owens C, Hubach RD. Rural-urban differences in monkeypox behaviors and attitudes among men who have sex with men in the United States. J Rural Health 2023 Mar;39(2):508-515. [doi: 10.1111/jrh.12726] [Medline: 36394371]
26. Lakhani HV, Pillai SS, Zehra M, Sharma I, Sodhi K. Systematic review of clinical insights into novel coronavirus (CoVID-19) pandemic: persisting challenges in U.S. rural population. Int J Environ Res Public Health 2020 Jun 15;17(12):4279. [doi: 10.3390/ijerph17124279] [Medline: 32549334]
27. Maan I, Kohli M, Gilson R. Mpox in people living with HIV. Curr Opin Infect Dis 2024 Feb 1;37(1):1-7. [doi: 10.1097/QCO.0000000000000994] [Medline: 38112084]
28. Mitjà O, Alemany A, Marks M, et al. Mpox in people with advanced HIV infection: a global case series. Lancet 2023 Mar 18;401(10380):939-949. [doi: 10.1016/S0140-6736(23)00273-8] [Medline: 36828001]
29. Wiefels MD, Gmunder KN, Ruiz JW. Patient income level and health insurance correlate with differences in health care utilization during the COVID-19 pandemic. J Public Health Res 2023 Jan;12(1):22799036231160624. [doi: 10.1177/22799036231160624] [Medline: 36911537]

Abbreviations

AMIS: American Men's Internet Survey
aPR: adjusted prevalence ratio
CAS: condomless anal sex
GBMSM: gay, bisexual, and other men who have sex with men
PR: prevalence ratio
PrEP: pre-exposure prophylaxis
STI: sexually transmitted infection

Edited by A Mavragani; submitted 15.02.24; peer-reviewed by M Xiridou, S Chen; revised version received 16.07.24; accepted 06.08.24; published 16.01.25.

Please cite as:
Atkins K, Carpino T, Rao A, Sanchez T, Edwards OW, Hannah M, Sullivan PS, Ogale YP, Abara WE, Delaney KP, Baral SD
Suspected Mpox Symptoms and Testing in Men Who Have Sex With Men in the United States: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e57399
URL: https://publichealth.jmir.org/2025/1/e57399
doi:10.2196/57399

Effect of a Short, Animated Storytelling Video on Transphobia Among US Parents: Randomized Controlled Trial

Doron Amsalem[1], MD; Merlin Greuel[2], MD; Shuyan Liu[3,4], PhD; Andrés Martin[5,6], MD, PhD; Maya Adam[2,7,8], MD, PhD

[1]Department of Psychiatry, Columbia University, New York, NY, United States
[2]University Hospital, Heidelberg Institute of Global Health, Heidelberg University, Heidelberg, Germany
[3]Department of Psychiatry and Psychotherapy, Charité—Universitätsmedizin, Campus Charité Mitte, Berlin, Germany
[4]German Center for Mental Health, Deutsche Zentrum für Psychische Gesundheit, Berlin, Germany
[5]Child Study Center, Yale School of Medicine, Yale University, New Haven, CT, United States
[6]Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
[7]Stanford Department of Pediatrics, Stanford University School of Medicine, Stanford University, 453 Quarry Rd, Palo Alto, CA, United States
[8]Stanford Center for Digital Health, Stanford University School of Medicine, Stanford University, Stanford, CA, United States

Corresponding Author:
Maya Adam, MD, PhD
University Hospital, Heidelberg Institute of Global Health, Heidelberg University, Heidelberg, Germany

Abstract

Background: Parents play a pivotal role in supporting transgender and gender diverse (TGD) youth. Yet only 35% of TGD youth describe their home as a gender-affirming place. Lack of parental support contributes to recent findings that TGD youth are approximately three times more likely to attempt suicide than their cisgender peers. In contrast, parents' affirmation of their children's gender identity significantly improves their mental health outcomes, by reducing anxiety, depression, and suicidality.

Objective: Addressing the urgent need for effective, scalable interventions, this study evaluates a novel digital approach: short, animated storytelling videos. We hypothesized that our 2.5-minute video intervention would reduce antitransgender stigma, or transphobia, and improve attitudes toward gender diverse children among US parents.

Methods: We recruited 1267 US parents, through the Prolific Academic (Prolific) online research platform, and randomized them into video intervention or control groups. We measured transphobia using the Transgender Stigma Scale, and attitudes toward transgender children using the gender thermometer, before and after watching the video. We compared outcomes between the two groups using 2 × 3 ANOVA. Both groups were invited to return 30 days later for follow-up assessment, before being offered posttrial access to the intervention video, which portrayed an authentic conversation between a mother and her transgender child.

Results: Single exposure to a short, animated story video significantly reduced transphobia and improved attitudes toward transgender children among US parents, immediately post intervention. We observed a significant group-by-time interaction in mean Transgender Stigma Scale scores ($F_{2,1}=3.7$, $P=.02$) and significant between-group changes when comparing the video and control groups from baseline to post intervention ($F_1=27.4$, $P<.001$). Effect sizes (Cohen *d*) indicated small to moderate immediate changes in response to the 2.5-minute video, though the effect was no longer observed at the 30-day follow-up. Gender thermometer scores revealed significant immediate improvements in the attitudes of participants in the video intervention arm, and this improvement was sustained at the 30-day time point.

Conclusions: Short, animated storytelling is a novel digital approach with the potential to boost support and affirmation of transgender children, by offering authentic insights into the lived experiences of TGD youth. Repeated exposures to such interventions may be necessary to sustain improvements over time. Future studies could test a series of short, animated storytelling videos featuring the lived experiences of several TGD youth. Evaluating the effect of such a series could contribute to the fields of digital health communication and transgender health. Digital approaches, such as short, animated storytelling videos, that support empathy and acceptance of TGD youth could foster a more inclusive society in which every child can thrive.

Trial Registration: AsPredicted.org 159248; https://aspredicted.org/ptmd-3kfs.pdf

(JMIR Public Health Surveill 2025;11:e66496) doi:10.2196/66496

KEYWORDS

public health communication; vulnerable population; stigma reduction; stigma; transphobia; transgender; gender diverse; LGBTQ; parent; mental health; mental illness; transgender children; children; youth; adolescent; storytelling; animation

Introduction

Optimal health outcomes for transgender and gender diverse (TGD) youth depend on the support and affirmation of their parents, as well as other adults with whom they interact [1-5]. Yet, even in their own homes, only 35% of TGD youth feel they are in a gender-affirming place [6]. Lack of parental support contributes to recent findings that TGD youth are 2.7 - 3.5 times more likely to attempt suicide than their cisgender peers [7,8]. Simple, gender-affirming acts, such as consistently respecting a child's pronouns in the home, can dramatically reduce the percent of TGD youth who attempt suicide from 21% to 12% as in a 2023 US national survey [6]. However, reducing stigma toward TGD youth, among adults, remains challenging [9,10]. Stigma acts as a broader social mechanism that contributes to transphobia (prejudice against individuals who defy traditional gender norms) [9,11]. Recent findings indicate that, even within their own homes, only 35% of TGD youth report feeling that they are in a gender-affirming place, with parental transphobia being a key contributing factor [6]. Lack of adult support and affirmation remains one of the leading causes of poor mental health outcomes in TGD youth [2-5,12,13].

This public health problem is aggravated by a growing climate of general societal intolerance in the United States. A record number of anti-LGBTQ (lesbian, gay, bisexual, transgender, queer) bills (more than 400) were introduced across US state legislatures in 2023 [14]. These environmental factors contribute to an alarmingly high rate of suicidality among TGD youth: the 2023 US National Survey on the Mental Health of LGBTQ+ Young People reported that 48% of transgender girls, 56% of transgender boys, and 48% of nonbinary youth had seriously considered attempting suicide in the past year [6]. Overall, TGD youth are 3.5 times more likely to attempt suicide than their cisgender peers [8]. Simple gender-affirming acts from adults, such as consistently respecting a child's pronouns, have been associated with a dramatic reduction in the percent of TGD youth who attempt suicide from 21% to 12% [6]. Even before the COVID-19 pandemic, a scoping review published in the *Journal of Adolescent Health* identified an urgent need for innovative interventions to reduce transphobia toward TGD youth [15]. Pandemic lockdowns, through a loss of community and school support, further threatened the mental health of TGD youth, who found themselves isolated and unsupported [16-18].

Research on transphobia reduction in schools highlights the need for interventions to reduce stigma among all parents, not just those of TGD youth [19]. Not only does affirmation from the general parent community positively impact TGD children directly, it also makes it easier for the parents of TGD children to adopt gender-affirming attitudes toward their children, when these prevail in the broader parent community [20,21]. Importantly, a parent's journey toward acceptance and support of their TGD child begins with them as a member of the general parent community [22]. Only once they have truly accepted the TGD identity of their child will they identify as "a parent of a TGD child." So, efforts to promote acceptance and reduce transphobia in the general parent population supports TGD youth in two ways: (1) a more accepting general parent population will model acceptance for the parents of TGD youth, and (2) a more accepting general parent population results in gender-affirming interactions with TGD youth in school or social settings, thereby creating a more accepting environment in which all children can thrive [19-22].

A significant body of research suggests that social contact-based interventions may be the most effective way to reduce stigma [23-26]. Grounded in intergroup contact theory, this research suggests that exposure to the stories and experiences of transgender individuals can effectively reduce transphobia toward them [27-31]. A known stumbling block of this approach is the selection bias that limits exposure to the stories of TGD youth, since prejudiced individuals generally avoid intergroup contact [30]. Disseminating the stories of TGD youth to the general public, on the platforms where they readily consume information, still presents a major challenge.

In the postpandemic era, social media has become a primary source of health information for the public [32]. A 2021 scoping review in the *Lancet Digital Health* described social media as a "crucial communication tool" for public health [33]. Social media has proven effective for reaching broader and more varied audiences, including hard-to-reach populations [32]. As such, social media presents a powerful, emerging pathway for scaling the stories of TGD youth to the public.

Short, animated storytelling videos, disseminated via social media, emerged as a powerful digital approach to engaging diverse audiences around the world, during the COVID-19 pandemic [34]. Harnessing the appeal of human stories, enhanced with compelling soundtracks, this novel digital approach is easily shareable and thus readily scalable via social media [34,35].

This study investigated the effect of a short, animated storytelling video on transphobia in a sample of US parents. Our hypothesis was that exposure to the authentic narratives of TGD youth, animated to protect their identities, could effectively reduce transphobia and improve attitudes toward gender diverse children, thereby boosting affirmation and support for TGD youth, among parents in the general public.

Methods

Participants and Recruitment Procedure

From January to February 2024, we enrolled parents from the general US population using Prolific Academic, a crowdsourcing platform frequently used to recruit participants in online psychology and medical research [36]. Prolific supports high-quality data collection by ensuring consistency in demographic responses over time, screening for bots, blocking participants who hide their location, and assigning anonymous unique participant IDs [37,38]. For this trial, we recruited

English-speaking parents, aged 18 - 50 years, with at least one child, who were residing in the United States. Participants were recruited through the Prolific Academic platform. Before the initiation of this study, all participants reviewed a detailed information sheet, outlining the rationale for this study as well as all study procedures, and potential risks and benefits of taking part in this study. Those who agreed to participate were directed to complete this study's procedures on Qualtrics, a secure online data-collection platform.

Ethical Considerations

This study, which involved human participants, was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and was approved by the Stanford University internal review board (IRB-#72761) on January 21, 2024. We followed the CONSORT (Consolidated Standards of Reporting Trials) guidelines [39], and this study was registered (#159248) with AsPredicted, a clinical trial registry created in 2015 by the Wharton School at the University of Pennsylvania [40].

Informed consent was obtained from all participants prior to enrollment in this study. Participants were provided with an information sheet that contained a detailed description of this study's purpose, procedures, potential risks, and benefits. The information sheet was reviewed and approved by the Stanford IRB. Participants were informed that participation was voluntary and that they could withdraw at any time without penalty. The consent process was facilitated by Prolific Academic, the academic recruitment platform through which participants were recruited. Participants were paid US $2.40 for participation in this short study, at 2 time points and in accordance with the payment recommendations of Prolific Academic. Participants were assured that their responses would remain confidential and that any identifying data would not be shared with the research team by Prolific. The informed consent process was reviewed and approved by the Stanford IRB (#72761) ensuring compliance with ethical research standards.

Randomization and Study Design

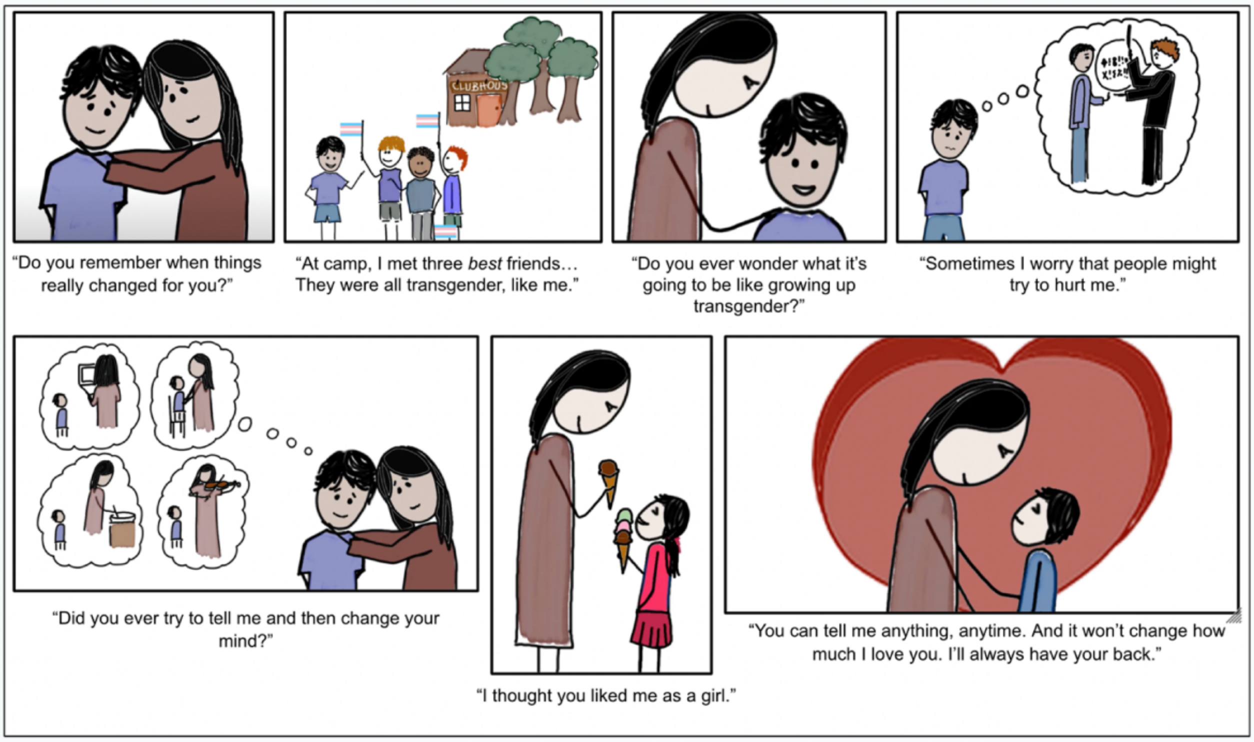
After consenting to take part in this study, participants were randomly assigned to one of two groups: short, animated storytelling video intervention group or a control group that received a simple fact sheet about TGD youth. We used a baseline survey to collect demographic information and we administered validated questionnaires to assess transphobia and attitudes toward TGD youth. We conducted the first round of surveys immediately post exposure to the short, animated storytelling intervention. Follow-up assessments, using the same surveys, were conducted 30 days later. To ensure accuracy and validity of the results, we included attention checks (eg, "For this question, please select the answer 'Neither agree nor disagree.'") and excluded participants who failed these. We also used a timer to confirm that participants in the control group read the TGD fact sheet and that those in the intervention group viewed the short, animated storytelling video before proceeding to the next screen. Each participant received US $2.40 for completing this study.

Intervention

Our novel, digital intervention used a short, animated storytelling video to capture an authentic prerecorded conversation between a mother and her transgender child. We chose this video because it follows the principles of social contact-based interventions, which foster identification and emotional engagement. A significant body of research suggests that these principles have proven effective in reducing stigma and transphobia by exposing viewers to the real-life stories and experiences of marginalized individuals [23-30,41]. In the conversation featured in our intervention, a mother asks her son to recall experiences related to his gender identity, including his early hesitations to confide in her that he was transgender. The child then asks his mother about her worries, and she shares these candidly. The identities of the speakers are concealed at all times, and we received written permission to use the audio recording from the speakers. The scenes and characters used to represent the speakers were drawn by hand on a Wacom tablet, by our coauthor (MA), and edited by the Educational Technology Group at Stanford Medicine. Figure 1 shows selected scenes and quotations from the intervention video.

The short, animated video tested in this trial can be previewed on YouTube [42].

Figure 1. Selected scenes and quotations from the short, animated storytelling intervention video shown to reduce transphobia in US parents in this online, randomized controlled trial.



Instruments

The primary outcome was transphobia toward TGD youth, which was measured using the total score of the Transgender Stigma Scale (TSS) developed by Madera et al [43]. We adjusted the wording of the scale items to assess attitudes toward transgender children (eg, "Boys who act like girls should be ashamed of themselves," and "Children should play with toys appropriate to their own sex"). The TSS is scored along a 5-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (5). Higher scores indicate greater transphobia. The TSS is highly reliable and has a Cronbach α of 0.89 [43].

As secondary outcome, we measured attitudes toward TGN children, using the "gender thermometer," a tool we previously developed to assess attitudes regarding sexual orientation and gender diversity [27]. The thermometer provides the following prompt: "Using a scale from zero to 100, please tell us about your personal feelings toward transgender boys or girls. As you do this task, think of an imaginary thermometer. The warmer or more favorable you feel toward the group, the higher the number you should give it. The colder or less favorable you feel, the lower the number. If you feel neither warm nor cold toward the group, rate it 50." We asked respondents about their attitudes toward (1) boys or girls and (2) transgender boys or girls. Higher ratings indicate more positive attitudes toward the group and lower ratings indicate less positive attitudes. Researchers in the fields of political science and psychology have frequently used feeling thermometers to assess attitudes toward diverse groups, including sexual minorities [44].

Analysis

The main outcome measures were changes in total TSS scores and gender thermometer scores, from baseline to post intervention and from baseline to 30-day follow-up. We calculated our sample size based on our previous studies [27,45-48]. We used Pearson chi-square and one-way ANOVA tests to compare sociodemographic characteristics between the short, animated storytelling video intervention and control groups. We used repeated-measures ANOVA tests to compare the mean transphobia score between the groups at 3 time points. Next, we used a one-way ANOVA test to compare changes from baseline scores, at post intervention, and at 30-day follow-up. We used independent *t* tests (2-tailed) to compare gender differences at baseline. Data were analyzed using SPSS (version 29.0; IBM Corp) [49].

Results

Sample Characteristics

We recruited 1267 US parents, through the Prolific Academic online research platform. In total, 1177 parents completed the preintervention assessment after we excluded 40 (3%) who failed attention tests and 50 (4%) who left this study. Of these, 1159 (98%) completed the postintervention assessment and 976 (83%) completed the 30-day follow-up assessment. Figure 2 shows the design and flow of this study.

Sociodemographic characteristics and completion rates did not differ significantly between groups (Table 1), nor did baseline characteristics differ between follow-up completers and noncompleters. Mean participant age was 39 (SD 6.9; range 20 - 50) years. More than half of participants self-identified as female (n=653, 56%), 510 (43%) male, and 14 (1%) transgender or nonbinary. Overall, 110 (9%) participants self-identified as Hispanic, 192 (16%) non-Hispanic Black, 789 (67%) non-Hispanic White, 57 (5%) non-Hispanic Asian, 8 (1%) non-Hispanic Native American, and 18 (2%) other (Table 1).

Figure 2. Study design and flow for our online, randomized controlled trial testing the effect of a short, animated storytelling video intervention on transphobia in US adults.

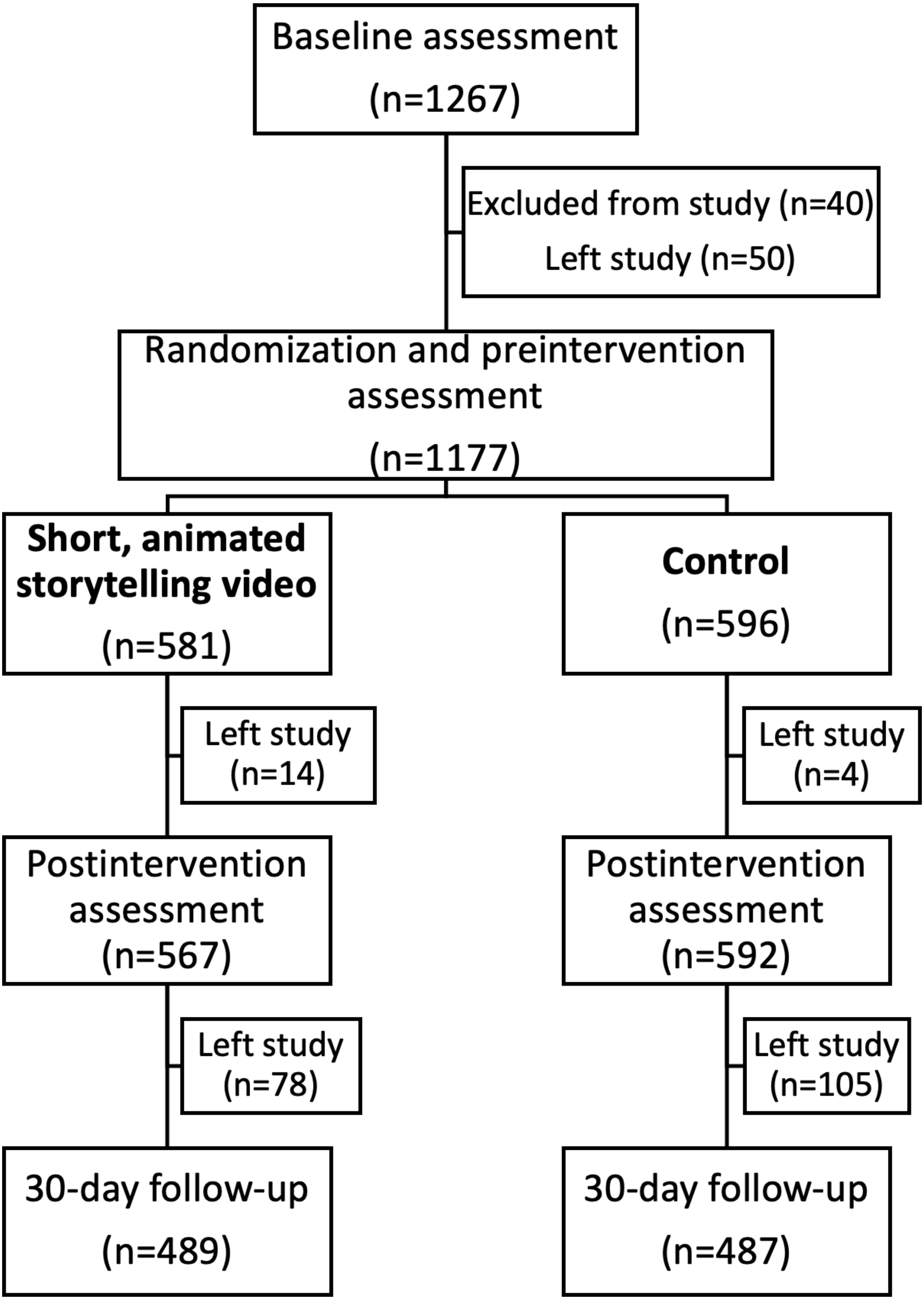


Table . Demographic characteristics of study participants.

Items		Brief animation (n=581)	Narrative (n=596)	Total (N=1177)	Statistic	
					Pearson χ^2 (*df*)	*P* value
Age (years), mean (SD)		39.1 (7.1)	39 (6.8)	39 (6.9)	0.18 (*1175*)[a]	.86
Gender, n (%)					2.62(*2*)	.27
	Female	336 (58)	317 (53)	653 (56)		
	Male	238 (41)	272 (46)	510 (43)		
	Transgender or nonbinary	7 (1)	7 (1)	14 (1)		
	Prefer not to answer	0 (0)	0 (0)	0 (0)		
Race and ethnicity, n (%)					9.35 (*6*)	.16
	Hispanic	62 (11)	48 (8)	110 (9)		
	Non-Hispanic Asian	32 (6)	25 (4)	57 (5)		
	Non-Hispanic Black	97 (17)	95 (16)	192 (16)		
	Non-Hispanic Native American	3 (1)	5 (1)	8 (1)		
	Non-Hispanic White	378 (65)	411 (69)	789 (66)		
	Non-Hispanic other[b]	6 (1)	12 (2)	18 (2)		
	Prefer not to answer	3 (1)	0 (0)	3 (0)		

[a] Independent *t* tests.

[b] Non-Hispanic other: multiracial (n=15), Native Hawaiian (n=1), Jewish (n=1), and South Asian (n=1).

Intervention Effects

We observed significant differences in outcomes between this study's groups. Figure 3 presents mean TSS scores of this study arms at 3 time points, showing that the control arm changed minimally across time points, in contrast to the short, animated storytelling video intervention arm, which showed a strongly significant decrease in transphobia immediately post intervention. This effect was no longer evident at 30 days post intervention. A 2 × 3 group-by-time ANOVA showed a significant group-by-time interaction in mean TSS scores ($F_{2,1}$=3.7, *P*=.02). One-way ANOVAs revealed significant between-group changes between the short, animated storytelling video and control groups from baseline to post intervention (F_1=27.4, *P*<.001), but not from baseline to the 30-day follow-up.

Figure 4 presents mean gender thermometer scores, by study arm, over time. Figure 4A shows changing attitudes toward transgender boys or girls. Attitudes of participants in the short, animated storytelling video intervention arm improved significantly over time and this improvement was sustained at the 30-day time point. In contrast, we observed a slight improvement in the control arm, which was not significant, and showed a subsequent rebound at the 30-day follow-up.

We did not find a 2 × 3 group-by-time interaction. However, independent *t* tests showed a significant difference between the changes from baseline to post intervention between short, animated storytelling video intervention and control groups (change of 2.2, SD 9.7 vs 0.8, SD 6.5; *t*=3, t_{df}=1151 *P*=.003). We did not observe an effect of the intervention on attitudes toward cisgender boys or girls (Figure 4B).

Figure 3. Mean scores on the TSS between short, animated storytelling video intervention (n=588) and control (n=598) groups at 3 time points in our online RCT measuring the effect of a short, animated storytelling intervention on transphobia in US adults. Follow-up surveys completed after 30 days; higher scores indicate greater transphobia; TSS (range 10 - 50). Repeated measure ANOVA (*P*=.02). RCT: randomized controlled trial; TSS: Transgender Stigma Scale.

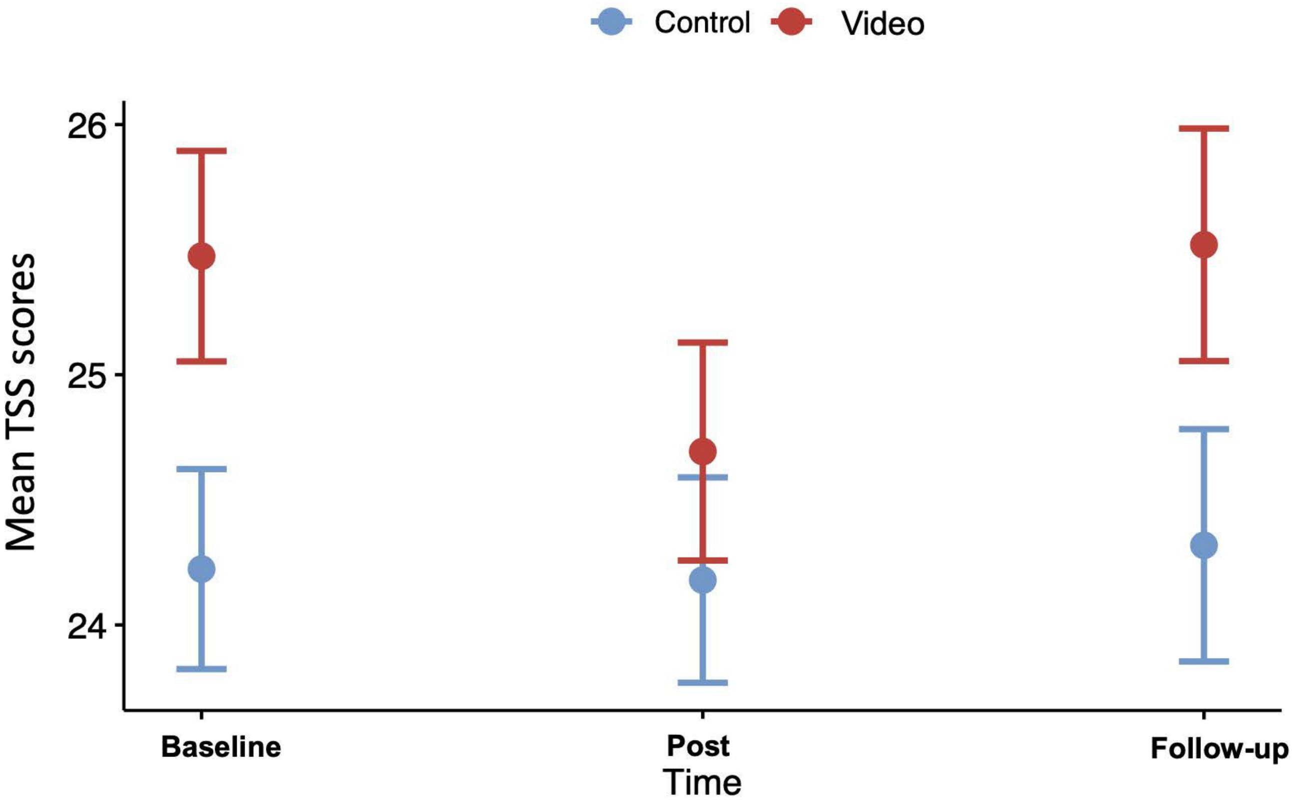
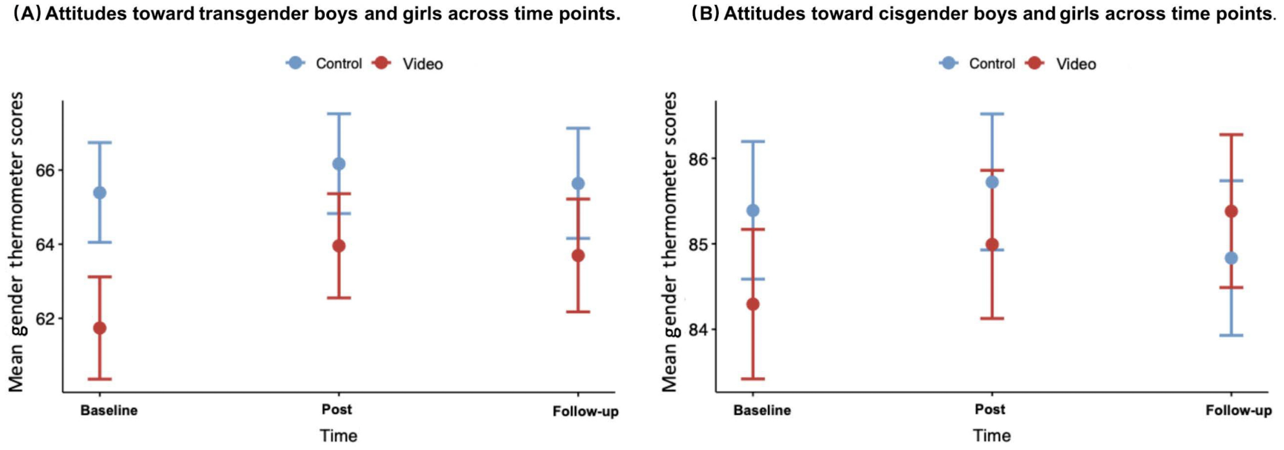


Figure 4. Gender thermometer scores toward (A) transgender children and (B) cisgender children, at 3 time points in our online RCT measuring the effect of a short, animated storytelling intervention on transphobia in US adults. Follow-up surveys completed after 30 days; higher scores indicate more favorable attitudes; gender thermometer (range 0 - 100). RCT: randomized controlled trial.

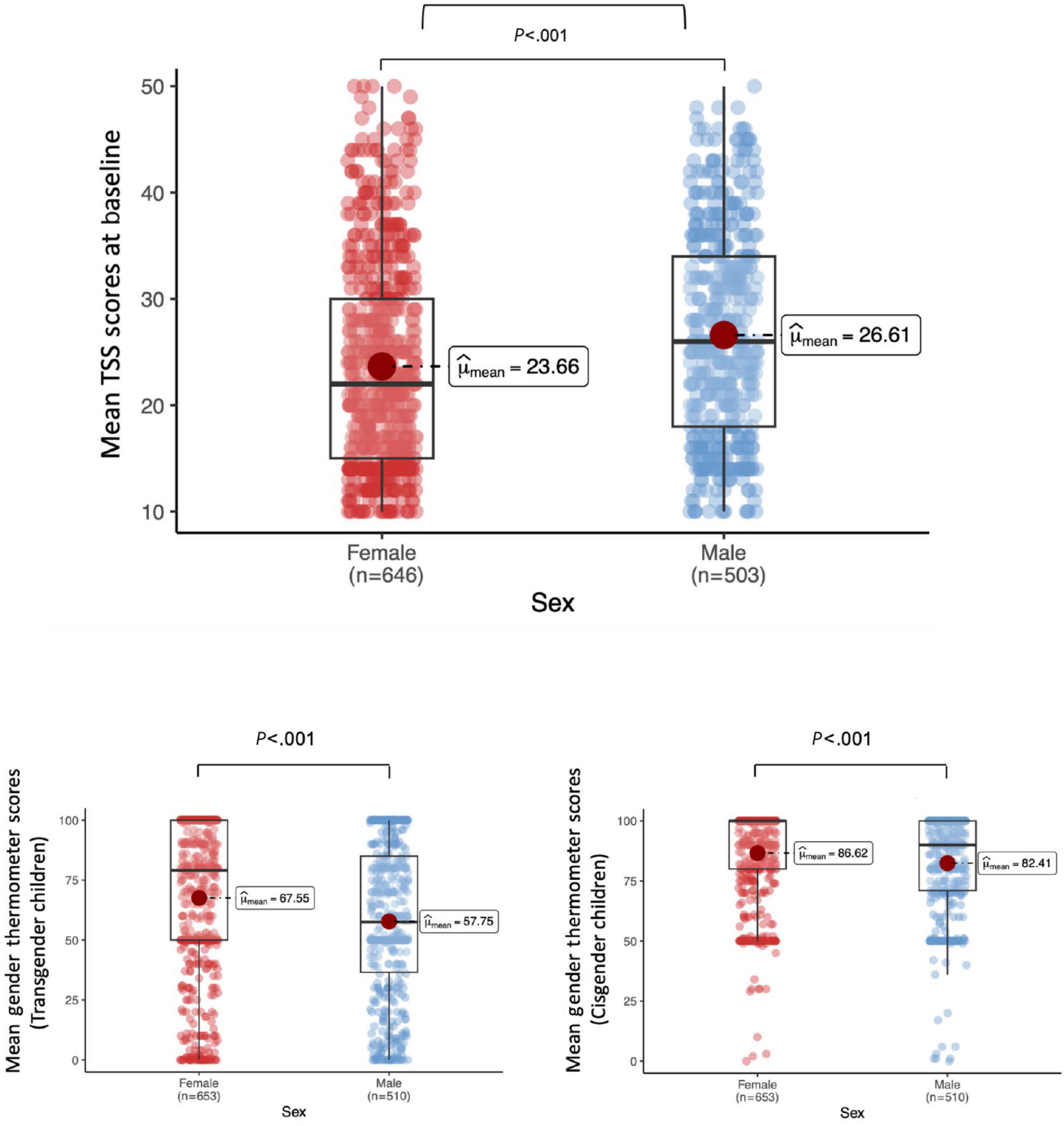

(A) Attitudes toward transgender boys and girls across time points.

(B) Attitudes toward cisgender boys and girls across time points.

Participant Gender Differences

Figure 5 presents baseline differences between men and women on the TSS (26.6, SD 9.7 vs 23.7, SD 9.8; independent *t* tests*:* *t*=5.1, *t*df=1161, *P*<.001) and on the gender thermometer both for transgender boys or girls (57.7, SD 32 vs 67.5, SD 33.1, *t*=5.1, *t*df=1161, *P*<.001) and cisgender boys or girls (82.4.7, SD 21 vs 86.6, SD 19.7, *t*=3.5, *t*df=1161 *P*<.001). The observed baseline differences between men and women remained constant across time points in both study groups.

Figure 5. Gender differences (TSS and gender thermometer) at baseline in our online RCT measuring the effect of a short, animated storytelling intervention on transphobia in US adults. Higher scores indicate greater transphobia; TSS (range 10 - 50). Higher scores indicate more positive attitudes; GT (range 0 - 100). GT: gender thermometer; RCT: randomized controlled trial; TSS: Transgender Stigma Scale.



Discussion

In this study, we conducted a large-scale, randomized, controlled experiment to document the potential of a novel, digital approach—short, animated storytelling—to reduce transphobia and improve attitudes toward TGD youth among US parents. Within our study population, we observed significantly reduced transphobia and significantly improved attitudes toward TGD youth, immediately after viewing the 2.5-minute short, animated storytelling video intervention. The effect of this short video was no longer evident one month after the intervention, suggesting that repeated, short exposures to similar short, animated storytelling video content may be necessary to sustain transphobia reductions and improvements in attitudes over time. One potential approach is to develop and disseminate a series of short, animated storytelling videos featuring the authentic lived experiences of several TGD young people. This scalable, digital approach could be implemented through a smartphone-based ecological momentary intervention. Evaluating the effect of such a series could be a meaningful future research direction in the fields of both digital health communication and transgender health.

During the COVID-19 pandemic, our coinvestigator (MA) developed the short, animated storytelling approach, a novel digital modality for scaling health information worldwide. Early short, animated storytelling COVID prevention videos reached more than 15 million people via social media within 4 months [35]. High voluntary public engagement with short, animated storytelling content was also documented in randomized, controlled trials [50,51]. As short, animated storytelling videos have proven effective for engaging diverse adult audiences, they present a potentially promising new method for engaging the public in the stories of TGD youth. The findings of this study support the potential for using short, animated storytelling to reduce transphobia and elicit empathy. Animation also protects the identities of the TGD storytellers—an important consideration, since many of these young people still live in “stealth” [52].

Intergroup contact theory, and recent research grounded on this theory, support the stigma-reducing potential of social contact with TGD individuals [27,28,30,31,53,54]. To date, interventions designed to reduce transphobia have been difficult to scale broadly to the public because (1) they involve face-to-face or online synchronous interactions, (2) they involve a significant time commitment and some form of active “opting in,” thereby self-selecting for people who already have low stigma, or (3) they lack the production value needed to make them highly engaging—one of the keys to scalability [55]. The short, animated storytelling video tested in this study packages a theory-driven approach to stigma reduction (the sharing of human stories) in a readily scalable form of health message (short, animated videos that can be shared via social media) to overcome a critical health challenge (lack of support and affirmation for TGD youth). Anecdotally, and supporting the potential for easy scalability, one participant in our study sent a message to the research team:

> *As the father of a child with 2 trans friends, this video is amazing and I showed it to my daughter, who is 8 and she completely loved it. She asked if we could send it to her friends’ parents, who called us and thanked us. So I wanted to let you know this was helpful and well thought. I can’t thank you enough!*

In this study, we also observed significant differences in baseline transphobia between men and women. Aligned with prior research, men in our study scored significantly higher on transphobia than did women, and these differences persisted across time points [11,56]. Other researchers have proposed that threatening gender norms can trigger a perceived loss of social status among men more than women [57]. Men may also be more likely to feel that transgender individuals pose a threat to their own gender identity and masculinity [58]. Future interventions could include story-based content that is tailored toward men. Featuring the stories of fathers, who are supporting of and affirming to their transgender children, may be particularly important. We also note that, to date, much of the research on family support of TGD youth focuses on parents, but there is an emerging body of research on the attitudes of other family members (eg, grandparents and siblings) toward transgender youth. If parents are not supportive, these other family members may be able to play a vital role in protecting transgender youth and the approach described here could also work with a wider range of target audiences.

A limitation of this study is the fact that our sample was recruited via an online research platform, suggesting that the sample may not be entirely representative of the general population in the United States. However, considering the intended, digital dissemination route for our intervention (social media), we feel this population is similar enough to our target population to allow us to gain some meaningful insights and draw relevant conclusions.

Our findings highlight the important role of authentic storytelling, packaged in scalable, accessible digital formats, for reducing transphobia toward TGN youth. These results also add to the growing field of research on short, animated storytelling, an innovative digital approach to rapidly scaling health information worldwide. Finally, these findings underscore the potential for light-touch, scalable video content to improve attitudes toward TGN youth within our society. Simply put, adults can be critical determinants of better health outcomes for TGD youth, simply by affirming their right to live and express themselves. Using authentic storytelling to engage and elicit empathy, and social media to broadly reach adults across demographics, we have the potential to catalyze a shift toward affirmation and support within society—a shift that could help children of all gender identities to survive and thrive.

Acknowledgments

This study was funded by the Department of Psychiatry, Columbia University Vagelos College of Physicians & Surgeons and the Stanford Center for Digital Health. The funders played no role in study design, data collection, analysis and interpretation of data, or the writing of this paper. We wish to thank the parent and child, who courageously shared their story in the intervention video, with the hope of contributing to a more accepting environment for gender diverse children.

Data Availability

The datasets collected and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

DA and MA led the design, implementation, data collection, and writing of this paper. MG contributed to the design, interpretation of results, and writing of this paper. DA and SL led the analysis, interpretation, and presentation of the results. AM contributed to the design, interpretation, presentation of results, and editing of this paper.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File, 798 KB - publichealth_v11i1e66496_app1.pdf]

References

1. Ghassabian A, Suleri A, Blok E, Franch B, Hillegers MHJ, White T. Adolescent gender diversity: sociodemographic correlates and mental health outcomes in the general population. J Child Psychol Psychiatry 2022 Nov;63(11):1415-1422. [doi: 10.1111/jcpp.13588] [Medline: 35147218]
2. Malpas J, Pellicane MJ, Glaeser E. Family-based interventions with transgender and gender expansive youth: systematic review and best practice recommendations. Transgend Health 2022 Feb;7(1):7-29. [doi: 10.1089/trgh.2020.0165] [Medline: 36644030]
3. Olezeski CL, Kamody RC. Parents matter: considering evidence-based relationship variables within families when working with gender expansive youth. Pract Innov (Wash D C) 2020;5(3):218-229. [doi: 10.1037/pri0000131]
4. Westwater JJ, Riley EA, Peterson GM. What about the family in youth gender diversity? A literature review. Int J Transgend 2019;20(4):351-370. [doi: 10.1080/15532739.2019.1652130] [Medline: 32999622]
5. Kidd KM, Didden E, Harman H, et al. Parents of gender diverse youth: support sought, received, and still needed. J Adolesc Health 2024 Jun;74(6):1078-1087. [doi: 10.1016/j.jadohealth.2023.08.001] [Medline: 37715763]
6. 2023 U.S. National Survey on the Mental Health of LGBTQ+ Young People. The Trevor Project. 2023. URL: https://www.thetrevorproject.org/survey-2023/ [accessed 2024-12-18]
7. Jackman KB, Caceres BA, Kreuze EJ, Bockting WO. Suicidality among gender minority youth: analysis of 2017 youth risk behavior survey data. Arch Suicide Res 2021;25(2):208-223. [doi: 10.1080/13811118.2019.1678539] [Medline: 31642392]
8. Toomey RB, Syvertsen AK, Shramko M. Transgender adolescent suicide behavior. Pediatrics 2018 Oct;142(4):e20174218. [doi: 10.1542/peds.2017-4218] [Medline: 30206149]
9. Hatzenbuehler ML, Pachankis JE. Stigma and minority stress as social determinants of health among lesbian, gay, bisexual, and transgender youth: research evidence and clinical implications. Pediatr Clin North Am 2016 Dec;63(6):985-997. [doi: 10.1016/j.pcl.2016.07.003] [Medline: 27865340]
10. Wittlin NM, Kuper LE, Olson KR. Mental health of transgender and gender diverse youth. Annu Rev Clin Psychol 2023 May 9;19:207-232. [doi: 10.1146/annurev-clinpsy-072220-020326] [Medline: 36608332]
11. Nagoshi JL, Adams KA, Terrell HK, Hill ED, Brzuzy S, Nagoshi CT. Gender differences in correlates of homophobia and transphobia. Sex Roles 2008 Oct;59(7-8):521-531. [doi: 10.1007/s11199-008-9458-7]
12. Rivas-Koehl M, Valido A, Espelage DL, et al. Understanding protective factors for suicidality and depression among U.S. sexual and gender minority adolescents: implications for school psychologists. Sch Psych Rev 2022 May 4;51(3):290-303. [doi: 10.1080/2372966X.2021.1881411]
13. Katz-Wise SL, Sarda V, Line EC, et al. Longitudinal family functioning and mental health in transgender and nonbinary youth and their families. J Child Fam Stud 2024 Apr;33(4):1321-1335. [doi: 10.1007/s10826-023-02738-2] [Medline: 39220671]
14. Choi A. Record number of anti-LGBTQ bills were introduced in 2023. CNN Online.: CNN; 2024. URL: https://edition.cnn.com/politics/anti-lgbtq-plus-state-bill-rights-dg/index.html [accessed 2025-01-14]
15. Parker CM, Hirsch JS, Philbin MM, Parker RG. The urgent need for research and interventions to address family-based stigma and discrimination against lesbian, gay, bisexual, transgender, and queer youth. J Adolesc Health 2018 Oct;63(4):383-393. [doi: 10.1016/j.jadohealth.2018.05.018] [Medline: 30146436]
16. Hawke LD, Hayes E, Darnay K, Henderson J. Mental health among transgender and gender diverse youth: an exploration of effects during the COVID-19 pandemic. Psychol Sex Orientat Gend Divers 2021;8(2):180-187. [doi: 10.1037/sgd0000467]
17. Jones BA, Bowe M, McNamara N, Guerin E, Carter T. Exploring the mental health experiences of young trans and gender diverse people during the COVID-19 pandemic. Int J Transgend Health 2023;24(3):292-304. [doi: 10.1080/26895269.2021.1890301] [Medline: 37519915]
18. Kidd JD, Jackman KB, Barucco R, et al. Understanding the impact of the COVID-19 pandemic on the mental health of transgender and gender nonbinary individuals engaged in a longitudinal cohort study. J Homosex 2021 Mar 21;68(4):592-611. [doi: 10.1080/00918369.2020.1868185] [Medline: 33502286]
19. Bartholomaeus C, Riggs DW. Whole-of-school approaches to supporting transgender students, staff, and parents. Int J Transgenderism 2017;18:361-366. [doi: 10.1080/15532739.2017.1355648]
20. Mangin MM. Transgender Students in Elementary School: Creating an Affirming and Inclusive School Culture: Harvard Education Press Cambridge; 2020.

21. Mariotto M, Albanesi C, Lorusso MM. Encountering transgender and gender-expansive children in school: exploring parents' and teachers' practices through the lens of cisgenderism and adultism. Child Youth Serv Rev 2024 Sep;164:107819. [doi: 10.1016/j.childyouth.2024.107819]
22. Sansfaçon AP, Kirichenko V, Holmes C, et al. Parents' journeys to acceptance and support of gender-diverse and trans children and youth. J Fam Issues 2020 Aug;41(8):1214-1236. [doi: 10.1177/0192513X19888779]
23. Jankowski SE, Yanos P, Dixon LB, Amsalem D. Reducing public stigma towards psychosis: a conceptual framework for understanding the effects of social contact based brief video interventions. Schizophr Bull 2023 Jan 3;49(1):99-107. [doi: 10.1093/schbul/sbac143] [Medline: 36190348]
24. Rani N, Samuel AA. Reducing transphobia: comparing the efficacy of direct and indirect contact. Industrial and Commercial Training 2019 Oct 7;51(7/8):445-460. [doi: 10.1108/ICT-12-2018-0102]
25. Reinke RR, Corrigan PW, Leonhard C, Lundin RK, Kubiak MA. Examining two aspects of contact on the stigma of mental illness. J Soc Clin Psychol 2004 Jun;23(3):377-389. [doi: 10.1521/jscp.23.3.377.35457]
26. Thornicroft G, Mehta N, Clement S, et al. Evidence for effective interventions to reduce mental-health-related stigma and discrimination. The Lancet 2016 Mar;387(10023):1123-1132. [doi: 10.1016/S0140-6736(15)00298-6]
27. Amsalem D, Halloran J, Penque B, Celentano J, Martin A. Effect of a brief social contact video on transphobia and depression-related stigma among adolescents: a randomized clinical trial. JAMA Netw Open 2022 Feb 1;5(2):e220376. [doi: 10.1001/jamanetworkopen.2022.0376] [Medline: 35212749]
28. Boccanfuso E, White FA, Maunder RD. Reducing transgender stigma via an e-contact intervention. Sex Roles 2021 Mar;84(5-6):326-336. [doi: 10.1007/s11199-020-01171-9]
29. Hughto JMW, Reisner SL, Pachankis JE. Transgender stigma and health: a critical review of stigma determinants, mechanisms, and interventions. Soc Sci Med 2015 Dec;147:222-231. [doi: 10.1016/j.socscimed.2015.11.010] [Medline: 26599625]
30. Pettigrew TF. Intergroup contact theory. Annu Rev Psychol 1998;49:65-85. [doi: 10.1146/annurev.psych.49.1.65] [Medline: 15012467]
31. Walch SE, Ngamake ST, Francisco J, Stitt RL, Shingler KA. The attitudes toward transgendered individuals scale: psychometric properties. Arch Sex Behav 2012 Oct;41(5):1283-1291. [doi: 10.1007/s10508-012-9995-6] [Medline: 22810996]
32. Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res 2021 May 12;23(5):e17917. [doi: 10.2196/17917] [Medline: 33978589]
33. Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. Lancet Digit Health 2021 Mar;3(3):e175-e194. [doi: 10.1016/S2589-7500(20)30315-0] [Medline: 33518503]
34. Adam M, Bärnighausen T, McMahon SA. Design for extreme scalability: a wordless, globally scalable COVID-19 prevention animation for rapid public health communication. J Glob Health 2020 Jun;10(1):010343. [doi: 10.7189/jogh.10.010343] [Medline: 32411360]
35. Vandormael A, Adam M, Greuel M, et al. The effect of a wordless, animated, social media video intervention on COVID-19 prevention: online randomized controlled trial. JMIR Public Health Surveill 2021 Jul 27;7(7):e29060. [doi: 10.2196/29060] [Medline: 34174778]
36. Peer E, Brandimarte L, Samat S, Acquisti A. Beyond the Turk: alternative platforms for crowdsourcing behavioral research. J Exp Soc Psychol 2017 May;70:153-163. [doi: 10.1016/j.jesp.2017.01.006]
37. Douglas BD, Ewell PJ, Brauer M. Data quality in online human-subjects research: comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. PLoS One 2023;18(3):e0279720. [doi: 10.1371/journal.pone.0279720] [Medline: 36917576]
38. Palan S, Schitter C. Prolific.ac—a subject pool for online experiments. J Behav Exp Finance 2018 Mar;17:22-27. [doi: 10.1016/j.jbef.2017.12.004]
39. Moher D. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials (Chinese version). J Chin Integr Med 2010 Aug 15;340:701-741. [doi: 10.3736/jcim20100801]
40. Olevska A, Bert B, Ebrahimi L, Schonfelder G, Heinl C. Ensuring reproducible research requires a support infrastructure: the value of public registries to publishers. Sci Editor 2021 Apr 23;44:4-7. [doi: 10.36591/SE-D-4401-4]
41. Walch SE, Sinkkanen KA, Swain EM, Francisco J, Breaux CA, Sjoberg MD. Using intergroup contact theory to reduce stigma against transgender individuals: impact of a transgender speaker panel presentation. J Appl Soc Psychol 2012 Oct;42(10):2583-2605 [FREE Full text] [doi: 10.1111/j.1559-1816.2012.00955.x]
42. You amaze me every day. Maya Adam YouTube Page. 2017 Mar 26. URL: https://www.youtube.com/watch?v=aDihlx776vM [accessed 2024-12-18]
43. Madera SL, Díaz NV, Padilla M, et al. "Just like any other patient": transgender stigma among physicians in Puerto Rico. J Health Care Poor Underserved 2019;30(4):1518-1542. [doi: 10.1353/hpu.2019.0089] [Medline: 31680112]
44. Norton AT, Herek GM. Heterosexuals' attitudes toward transgender people: findings from a national probability sample of U.S. adults. Sex Roles 2013 Jun;68(11-12):738-753. [doi: 10.1007/s11199-011-0110-6]

45. Amsalem D, Jankowski SE, Pagdon S, et al. Selfie videos to reduce stigma and increase treatment seeking among youths: two noninferiority randomized controlled trials. Psychiatr Serv 2023 Mar 1;74(3):229-236. [doi: 10.1176/appi.ps.20220168] [Medline: 36254455]
46. Amsalem D, Markowitz JC, Jankowski SE, et al. Sustained effect of a brief video in reducing public stigma toward individuals with psychosis: a randomized controlled trial of young adults. Am J Psychiatry 2021 Jul;178(7):635-642. [doi: 10.1176/appi.ajp.2020.20091293] [Medline: 33900809]
47. Amsalem D, Martin A. Reducing depression-related stigma and increasing treatment seeking among adolescents: randomized controlled trial of a brief video intervention. J Child Psychol Psychiatry 2022 Feb;63(2):210-217. [doi: 10.1111/jcpp.13427] [Medline: 33821507]
48. Amsalem D, Yang LH, Jankowski S, Lieff SA, Markowitz JC, Dixon LB. Reducing stigma toward individuals with schizophrenia using a brief video: a randomized controlled trial of young adults. Schizophr Bull 2021 Jan 23;47(1):7-14. [doi: 10.1093/schbul/sbaa114] [Medline: 33484269]
49. Pallant J. SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS: Routledge; 2020. [doi: 10.4324/9781003117445]
50. Favaretti C, Adam M, Greuel M, et al. Participant engagement with a short, wordless, animated video on COVID-19 prevention: a multi-site randomized trial. Health Promot Int 2023 Aug 1;38(4):daab179. [doi: 10.1093/heapro/daab179] [Medline: 35137068]
51. Favaretti C, Vandormael A, Hachaturyan V, et al. Participant engagement and reactance to a short, animated video about added sugars: web-based randomized controlled trial. JMIR Public Health Surveill 2022 Jan 24;8(1):e29669. [doi: 10.2196/29669] [Medline: 35072639]
52. Olson KR, Durwood L, DeMeules M, McLaughlin KA. Mental health of transgender children who are supported in their identities. Pediatrics 2016 Mar;137(3):e20153223. [doi: 10.1542/peds.2015-3223] [Medline: 26921285]
53. Cipolletta S, Votadoro R, Faccio E. Online support for transgender people: an analysis of forums and social networks. Health Soc Care Community 2017 Sep;25(5):1542-1551. [doi: 10.1111/hsc.12448] [Medline: 28329916]
54. Massey ZB, Wong NCH, Barbati JL. Meeting the (trans)parent: test of parasocial contact with transgender characters on reducing stigma toward transgender people. Commun Stud 2021 Mar 4;72(2):232-250. [doi: 10.1080/10510974.2021.1876125]
55. Guadagno RE, Rempala DM, Murphy S, Okdie BM. What makes a video go viral? An analysis of emotional contagion and internet memes. Comput Hum Behav 2013 Nov;29(6):2312-2319. [doi: 10.1016/j.chb.2013.04.016]
56. Hatch HA, Warner RH, Broussard KA, Harton HC. Predictors of transgender prejudice: a meta-analysis. Sex Roles 2022 Dec;87(11-12):583-602. [doi: 10.1007/s11199-022-01338-6]
57. Warriner K, Nagoshi CT, Nagoshi JL. Correlates of homophobia, transphobia, and internalized homophobia in gay or lesbian and heterosexual samples. J Homosex 2013;60(9):1297-1314. [doi: 10.1080/00918369.2013.806177] [Medline: 23952924]
58. Harrison BF, Michelson MR. Gender, masculinity threat, and support for transgender rights: an experimental study. Sex Roles 2019 Jan;80(1-2):63-75. [doi: 10.1007/s11199-018-0916-6]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
IRB: internal review board
LGBTQ: lesbian, gay, bisexual, transgender, queer
TGD: transgender and gender diverse
TSS: Transgender Stigma Scale

Edited by A Mavragani; submitted 14.09.24; peer-reviewed by JD Bres, S Haim-Nachum; revised version received 21.11.24; accepted 24.11.24; published 20.01.25.

Please cite as:
Amsalem D, Greuel M, Liu S, Martin A, Adam M
Effect of a Short, Animated Storytelling Video on Transphobia Among US Parents: Randomized Controlled Trial
JMIR Public Health Surveill 2025;11:e66496
URL: https://publichealth.jmir.org/2025/1/e66496
doi:10.2196/66496

Exploring Social-Ecological Pathways From Sexual Identity to Sleep Among Chinese Women: Structural Equation Modeling Analysis

Chanchan Wu, BSc, MSc, PhD; Pui Hing Chau, BSocSci, PhD; Edmond Pui Hang Choi, BNurs, MPhil, PhD

School of Nursing, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 5/F, Academic Building, 3 Sassoon Road, Pok Fu Lam, Hong Kong, China (Hong Kong)

Corresponding Author:
Edmond Pui Hang Choi, BNurs, MPhil, PhD
School of Nursing, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 5/F, Academic Building, 3 Sassoon Road, Pok Fu Lam, Hong Kong, China (Hong Kong)

Abstract

Background: Women and sexual minority individuals have been found to be at higher risk for experiencing poor sleep health compared to their counterparts. However, research on the sleep health of sexual minority women (SMW) is lacking in China.

Objective: This study aimed to examine sleep quality and social support for Chinese women with varied sexual identities, and then investigate the in-depth relationships between sexual identity and sleep.

Methods: This was a cross-sectional web-based survey. All participants completed a structured questionnaire containing a set of sociodemographic items referring to the social-ecological model of sleep health, the Pittsburgh Sleep Quality Index, the Social Support Rating Scale, and social relationships and environment domains of the World Health Organization Quality of Life-abbreviated short version. Pearson correlation coefficients were used to examine the relationship between sleep quality and social support as well as the two domains of quality of life. Structural equation modeling analysis was used to explore the social-ecological relationships.

Results: A total of 250 cisgender heterosexual women (CHW) and 259 SMW were recruited from July to September 2021. A total of 241 (47.3%) women experienced poor sleep quality and the rate was significantly higher in SMW than in CHW (55.2% vs 39.2%, $P<.001$). Around one-fifth of SMW reported low levels of social support, which was significantly higher than that of CHW (21.6% vs 5.6%, $P<.001$). Pearson correlations showed that overall sleep quality was significantly negatively associated with social support with weak correlations ($r=-0.26$, $P<.001$). The final structural equation modeling analysis with satisfactory fit indices identified 6 social-ecological pathways, showing that alcohol use, objective support, utilization of support, and perceived social relationship and environment quality of life played important roles in the sleep quality of individuals from their sexual identity.

Conclusions: SMW experienced poorer sleep quality compared to CHW. Further research is recommended to address the modifiable factors affecting sleep and then implement tailored sleep improvement programs.

(JMIR Public Health Surveill 2025;11:e53549) doi:10.2196/53549

KEYWORDS

sleep; social support; sexual minority women; social-ecological model; quality of life; structural equation model; Chinese women; China; women; structural equation modeling analysis; sleep quality; sexual identity; survey; heterosexual; cisgender

Introduction

Sleep is a crucial element that significantly impacts health across all populations. As a modifiable behavior, sleep is strongly linked to health and well-being [1], and abnormal sleep has been proven to be significantly associated with an extensive range of adverse health-related outcomes [2,3]. Consequently, sleep health is increasingly recognized as a public health issue that concerns everyone [4,5]. Sleep quality is widely used to evaluate sleep health and includes both objective and subjective aspects of sleep [6]. Many studies have shown that women generally experience poorer sleep quality than age-matched men, possibly due to the influence of sex steroids over sleep [7,8]. Although research on sleep health among vulnerable populations has gradually increased, most studies have been limited to racial or ethnic minorities [7,9-12], and knowledge of sleep health among sexual minority populations is relatively scarce [13].

To date, only 3 reviews have been published on the sleep health of sexual minorities, with all being narrative reviews that contain a very limited number of studies [14-16], reflecting the fact that relevant research is still in its infancy. Overall, these reviews suggest that sleep health constitutes an unmet health need for sexual minorities, and sleep health disparities related to sexual identity have been widely documented. Consistently, existing studies have recognized that sexual minority individuals experience significantly worse sleep quality than their cisgender and heterosexual peers. Furthermore, women and sexual minority individuals reported more sleep difficulties than did men or heterosexual participants [17], and sexual minority women (SMW) were more likely to report a higher prevalence of poor sleep quality than heterosexual women regardless of race/ethnicity [18-20], especially those in less supportive environments [21]. Additionally, SMW were more vulnerable to experiencing sleep disturbances than both sexual minority and heterosexual men [22], indicating that the sleep health of SMW deserves more attention.

The current evidence on sleep health among sexual minority Chinese is very limited; it remains an understudied area of research. So far, there are only 3 relevant studies, 2 of which are based on the analysis of data from nationwide school-based surveys of adolescents [23,24] and college students [25], indicating that sexual minority status was significantly associated with poor sleep quality. Notably, there is only 1 study that specifically investigated the sleep and discrimination experienced by lesbian, gay, and bisexual individuals in Hong Kong [26], and the results showed that discrimination experienced was associated with greater sleep disturbance, which in turn led to poorer physical and mental health conditions.

Health-related research on sexual minority Chinese also exhibits significant gender disparities, manifested by a substantially higher focus on the male sexual minority population than on female groups [27]. A recently published scoping review that mapped all the scientific literature and gray reports on the health needs of women with same-sex attraction in mainland China determined that Chinese sexual minority women have multiple unmet health needs [28], such as substance abuse, concerning mental health, and sexual and reproductive health. However, little is known about their sleep health, which is a crucial health issue that deserves more attention.

Research on factors affecting sleep has received more attention in recent years [29,30]. Grandner et al [31,32] initially proposed the social-ecological model of sleep health that integrated possible determinants of sleep and sleep-induced health outcomes from a global perspective. This model has been applied to studies in different populations [9,33,34] but has not yet been applied to sexual minority populations. Regarding the determinants of sleep, this model considers from a socio-ecological perspective that sleep may be determined by multiple causes at 3 levels, including the individual level, social level, and societal level.

Individual-level factors contain all aspects directly related to individuals’ sleep such as genetics and sleep-related behaviors. A growing body of literature suggests that alcohol use has detrimental effects on sleep [35] by reducing sleep quality [36]. Meanwhile, there is also comprehensive evidence that the pooled association between alcohol use and sleep disorders was not significant [37], suggesting that the relationship between alcohol use and sleep needs further investigation. Similarly, smoking has been identified as another sleep determinant [38], with smokers being more vulnerable to poorer sleep quality [39]. In China, a review targeting sexual minority women concluded that the prevalence of smoking and alcohol use among minority women was much higher than among general women [28]. Therefore, alcohol use and smoking status need to be considered along with sexual identity when exploring individual-level factors that affect sleep quality.

Social-level factors influencing sleep could be socioeconomic status, ethnicity, and social support. Ethnic minorities and individuals from disadvantaged economic backgrounds generally report lower sleep quality [12,13]. In addition, a review of sleep research concluded that financial hardship is associated with poor sleep health in the general population [16]. Notably, many studies have confirmed that social support is a key factor affecting sleep [40,41], with greater social support associated with better sleep outcomes [42-45], while having strained relationships was linked to more troubled sleep [46]. Compared with heterosexual populations, social support has unique functions in sexual minority individuals, but they reported receiving less support [47]. Therefore, when exploring social-level factors affecting sleep quality, conditions such as ethnicity, education, employment, and economic status need to be considered in addition to social support.

Regarding societal-level factors affecting sleep, geography, physical environment, and other aspects of the environment could be grouped into this category [32], and environmental factors may positively or negatively impact sleep [48]. A cohort study also found a link between declining social relationship quality and poor sleep quality [49]. The societal-level factors are closely linked to everyone’s livelihood, and therefore it is worth exploring whether individuals are local residents or new migrants to their current residence and whether they are cohabitating with others. The abovementioned factors involved in the social-ecological model might have direct or indirect impacts on sleep, but the associations among these factors have not been fully studied.

Despite the evident vulnerability of women and sexual minorities to poor sleep, there has been little research on their sleep health. Meanwhile, in China, research on the sleep health of sexual minorities, although gradually receiving attention in recent years, is still very limited, and there is currently no research specifically on the sleep of sexual minority women. To help eliminate inequalities in sleep health, a more diverse and inclusive sample, including women of different sexual identities, is greatly needed. Therefore, this study aimed to examine the sleep quality and social support in Chinese adult women, compare them between SMW and cisgender heterosexual women (CHW), then investigate the in-depth relationships between sexual identity and sleep using structural equation modeling (SEM) in the Chinese context.

Methods

Participants and Recruitment

This was a web-based cross-sectional study. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist [50] was used for reporting the study findings. Eligible participants were Chinese women who: (1) were at least 18 years old, (2) self-identified as female or were assigned female gender at birth, and (3) were able to read and understand Mandarin Chinese.

Generally, the minimum sample size for SEM is 200 [51,52], and the sample size needs to be more than 25 times the number of parameters as a rule of thumb [53]. In this study, based on the hypothesized model, it is evident that 10 variables are expected to be entered into the SEM model validation, thus the minimum sample size is 250.

Convenience sampling and respondent-driven sampling methods were used to recruit this relatively hidden population [54,55], by releasing the study poster via 4 popular nongovernmental organizations and encouraging respondents to help recruit potential peers through their network of connections. Details regarding the recruitment procedure have been previously reported [56].

In this study, 524 questionnaires were collected between July and September 2021, of which 15 were excluded for being cisgender men, adolescents, or invalid data. Finally, a total of 509 (97.1%) adult women were enrolled in the study with no missing data.

Ethical Considerations

This study was approved by the Human Research Ethics Committee of the University of Hong Kong (reference number EA210325) on July 8, 2021. Electronic written consent was obtained for each study participant. Data were collected using an online survey platform (Wenjuanxing), and no compensation was provided. All information collected during the study was kept anonymous and strictly confidential.

Measures

Pittsburgh Sleep Quality Index

The Pittsburgh Sleep Quality Index (PSQI) is a commonly used measure of individuals' sleep quality [14,57]. It includes 19 self-rated items, and they generate 7 component scores with an average weight of 0 - 3 points. The Chinese version of the PSQI has been previously validated [58,59] and was used in this study. Specifically, the seven components are: (1) subjective sleep quality, (2) sleep latency, (3) sleep duration, (4) habitual sleep efficiency, (5) sleep disturbances, (6) sleeping medication use, and (7) daytime dysfunction. The sum of the 7 component scores yields 1 global score, ranging from 0 to 21 points, and higher scores represent poorer sleep quality. A global PSQI score >5 indicates poor sleep quality [6,57].

Social Support Rating Scale

The Social Support Rating Scale (SSRS) is a 10-item scale specially designed for the Chinese population with sound reliability and validity [60]. It includes three dimensions of social support: (1) subjective support, (2) utilization of support, and (3) objective support. The full-scale score and dimension scores are the sum of the scores of each item, with higher scores indicating higher levels of social support [61,62]. In addition, a total score of 22 or below could be classified as low level, 23 - 44 as medium level, and 45 - 66 as a high level of social support [63].

World Health Organization Quality of Life – Abbreviated Short Version

The World Health Organization Quality of Life – Abbreviated Short Version (WHOQOL-BREF) is a generic quality of life (Qol) measure comprising 4 domains, and the psychometric properties of its Chinese version have been confirmed [64-66]. According to the social-ecological model of sleep [32], social relationships and environment could be considered as the societal-level factors affecting sleep, but there is little evidence of associations between sleep and these factors. Thus, this study used the social relationship domain and environment domain of the WHOQOL-BREF as societal-level factors of sleep for investigation. Each domain score can be transformed into a score ranging from 4 to 20, with higher scores representing better Qol [64,67].

Sociodemographic Information

All participants completed a set of sociodemographic items that were widely reported in previous studies referring to the social-ecological model of sleep health [31,32], including individual-level factors (sexual identity, age, smoking status, alcohol use status, drug use status), social-level factors (ethnicity, education, employment, income, number of friends, social support, relationship status, bed-sharing situation), and societal-level factors (local resident or not, duration in current residence, cohabitation situation, social relationship and environment Qol).

Statistical Analysis

Descriptive statistics were reported on participants' demographic characteristics, sleep quality, and social support. Independent *t* tests and χ^2 tests were conducted for comparisons between SMW and CHW. Cohen *d* effect size was calculated [68].

Social-Ecological Factors Identification

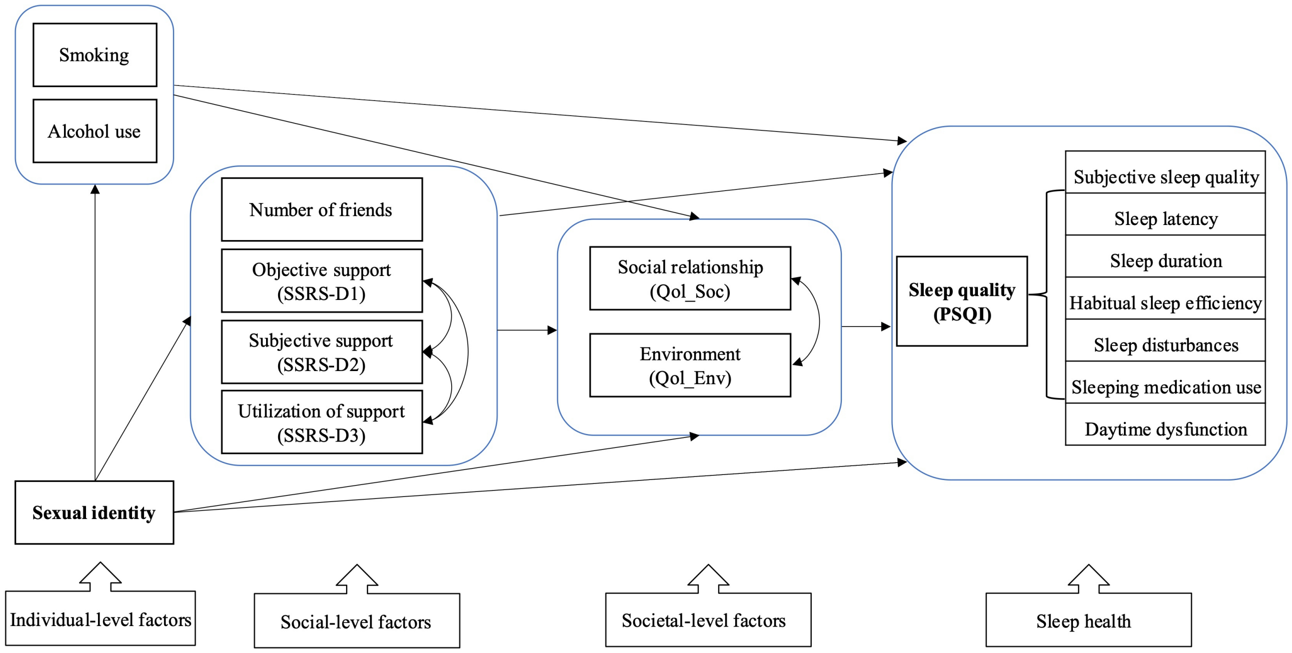
Pearson correlation coefficients were used to examine the relationship between sleep quality and social support as well as the 2 domains of quality of life, and the correlations were defined as strong (≥0.5), moderate (≥0.3 and <0.5), or weak (<0.3) [68]. Then, independent *t* tests, 1-way ANOVA, and multiple linear regressions were performed to identify factors significantly associated with sleep quality.

Structural Equation Modeling Analysis

Based on the factors identified previously (which will be further detailed in the Results section), with reference to the socio-ecological model of sleep, the following factors with significant coefficients were simultaneously entered into the SEM analysis, including sexual identity (binary), alcohol use (binary), and smoking (binary) (identified as individual-level factors); the number of friends (categorical) and social support (continuous) (identified as social-level factors); and social relationship Qol (continuous) and environment Qol (continuous) (identified as societal-level factors). Sleep quality was treated as the latent construct using the 7 domains of the PSQI measure.

The hypothesized model for socioecological influences on sleep based on the social-ecological model of sleep is shown in Figure 1.

Figure 1. Hypothesized model for socioecological influences on sleep (based on the social-ecological model of sleep). PSQI: Pittsburgh Sleep Quality Index; Qol: Quality of life; Qol_Env: Environment domain of Qol; Qol_Soc: Social relationship of Qol; SSRS: Social Support Rating Scale; SSRS-D1: objective support; SSRS-D2: subjective support; SSRS-D3: utilization of support.



All identified variables and the 7 components of sleep quality were included in the SEM model to test the hypothesized relationships among the previously listed variables. Standardized coefficients (β), standard errors, and associated *P* values are reported for all paths of the final model. Model goodness of fit was evaluated using chi-square (χ^2), root mean square error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis index (TLI), and standardized root mean square residual (SRMR). The model fit was considered adequate [69,70] when RMSEA (90% CI) ≤0.08, CFI ≥0.90, TLI ≥0.90, and SRMR ≤0.05.

Data were analyzed using SPSS version 28.0 (IBM Corp), and the SEM analysis was performed using Mplus 8.6 software [71]. All significance tests were 2-tailed, and findings with *P* values <.05 were considered statistically significant.

Results

Participant Characteristics

Table 1 presents the sociodemographic characteristics of the 509 women. Approximately half of the participants (259/509, 50.9%) were SMW, and 250 (49.1%) were CHW. The mean age of the overall sample was 25.57 years (SD 5.77), ranging from 18 to 56 years old. Details regarding the sexual orientation and gender identity of participants have been previously reported [56]. Notably, SMW had higher rates of substance use than CHW, and the rates of smoking and alcohol use showed significant differences (27.4% vs 4%, *P*<.001; 73.7% vs 56%, *P*<.001; respectively). Nine participants in total (1.8%) had a history of recreational drug use (4 used cannabis and 5 used methamphetamine), 7 of whom were SMW.

Table . Sociodemographic characteristics of the study sample with comparisons (N=509).

Characteristics	Overall (N=509)	SMW[a] (n=259)	CHW[b] (n=250)
Smoking, n (%); *P* <.001 [c]			
Never smoked or have quit smoking	428 (84.1)	188 (72.6)	240 (96)
Current smoker	81 (15.9)	71 (27.4)	10 (4)
Seldom (<1 time/week)[d]	48 (59.3)	41 (57.7)	7 (70)
Usually (1-7 times/week)[d]	9 (11.1)	9 (12.7)	0
Almost smoke every day[d]	24 (29.6)	21 (29.6)	3 (30)
Alcohol use, n (%); *P* <.001			
Never drank or have quit	178 (35)	68 (26.3)	110 (44)
Current alcohol user	331 (65)	191 (73.7)	140 (56)
Seldom (<1 time/week)[e]	220 (43.2)	111 (42.9)	109 (43.6)
Occasionally (2 - 4 times/month)[e]	88 (17.3)	64 (24.7)	24 (9.6)
Usually (2 - 4 times/week)[e]	17 (3.3)	13 (5)	4 (1.6)
Almost use alcohol every day[e]	6 (1.2)	3 (1.2)	3 (1.2)
Drug use, n (%); *P* =.10			
Never used before	500 (98.2)	252 (97.3)	248 (99.2)
Have used drugs before	9 (1.8)	7 (2.7)	2 (0.8)
Ethnicity, n (%); *P* =.12			
Han people	466 (91.6)	242 (93.4)	224 (89.6)
Others (Muslim, etc)	43 (8.4)	17 (6.6)	26 (10.4)
Education, n (%); *P* <.001			
High school and below	33 (6.5)	21 (8.1)	12 (4.8)
College/bachelor	284 (55.8)	167 (64.5)	117 (46.8)
Graduate degree and above	192 (37.7)	71 (27.4)	121 (48.4)
Monthly income (Chinese yuan)[f], n (%); *P* <.001			
≤1000	163 (32)	88 (54)	75 (46)
1001 - 3000	79 (15.5)	51 (64.6)	28 (35.4)
3001 - 5000	40 (7.9)	23 (57.5)	17 (42.5)
5001 - 7000	33 (6.5)	17 (51.5)	16 (48.5)
7001 - 9000	45 (8.8)	24 (53.3)	21 (46.7)
9001 - 11,000	38 (7.5)	16 (42.1)	22 (57.9)
More than 11,000	111 (21.8)	40 (36)	71 (64)
Relationship, n (%); *P* <.001			
Have a steady partner	251 (49.3)	115 (44.4)	136 (54.4)
Have no steady partner(s)	258 (50.7)	144 (55.6)	114 (45.6)
Bed sharing status, n (%); *P*= .12			
Separate bed in separate room	293 (57.6)	160 (61.8)	133 (53.2)
Separate bed in shared room	76 (14.9)	37 (14.3)	39 (15.6)
Sharing same bed with partner	140 (27.5)	62 (23.9)	78 (31.2)
Local resident or not, n (%); *P*= .52			

Characteristics	Overall (N=509)	SMW[a] (n=259)	CHW[b] (n=250)
Local resident	181 (35.6)	96 (37.1)	85 (34)
Nonlocal resident (migrant)	328 (64.4)	163 (62.9)	165 (66)
Duration in current residence, n (%); *P*= .59			
<3 months	50 (9.8)	27 (10.4)	23 (9.2)
3 - 6 months	17 (3.3)	11 (4.2)	6 (2.4)
7 - 12 months	33 (6.5)	15 (5.8)	18 (7.2)
>1 year	409 (80.4)	206 (79.5)	203 (81.2)
Cohabitation status, n (%); *P* <.001			
Live alone	114 (22.4)	60 (23.2)	54 (21.6)
Live with same-sex partner	41 (8.1)	34 (13.1)	7 (2.8)
Live with opposite-sex partner	75 (14.7)	18 (7)	57 (22.8)
Live with friends	91 (17.9)	42 (16.2)	49 (19.6)
Live with family	162 (31.8)	93 (35.9)	69 (27.6)
Other	26 (5.1)	12 (4.6)	14 (5.6)
Quality of life, mean (SD); *P* <.001			
Social relationship domain	13.62 (3.07)	13.02 (3.35)	14.23 (2.61)
Environment domain	13.74 (2.68)	13.32 (2.90)	14.18 (2.35)

[a] SMW: sexual minority women. The percentages for this column were all calculated with 259 SMW as the denominator.

[b] CHW: cisgender heterosexual women. The percentages for this column were all calculated with 250 CHW as the denominator.

[c] All *P* values were reported by conducting comparisons performing χ^2 tests or independent *t* tests.

[d] The denominator for this row is the number of current smokers (n=81).

[e] The denominator for this row is the number of current alcohol users (n=331).

[f] 1 CNY=US $0.136612.

Sleep Quality and Social Support

A total of 241 (47.3%) women experienced poor sleep quality and the rate was significantly higher in SMW than in CHW (55.2% vs 39.2%, *P*<.001). Similarly, independent *t* tests showed that SMW experienced significantly worse overall sleep quality than CHW (*P*<.001). There were statistically significant differences in subjective sleep quality, sleep latency, habitual sleep efficiency, sleep disturbances, and daytime dysfunction between CHW and SMW (Table 2). In addition, 81.5% (415/509) of all participants reported medium levels of social support. Around one-fifth of SMW reported low levels of social support, which was significantly higher than that of CHW (21.6% vs 5.6%, *P*<.001). Table 2 shows that SMW reported significantly lower support in all aspects than CHW (*P*<.01).

Table . Sleep quality and social support of the study sample with comparisons (N=509).

	Overall (N=509)	SMW[a] (n=259)	CHW[b] (n=250)	*P* value	Chi-square (*df*) or Cohen *d*
Sleep quality (PSQI [c]), n (%)					
Good (total score ≤5)	268 (52.7)	116 (44.8)	152 (60.8)	<.001	13.08 (1)
Poor (total score >5)	241 (47.3)	143 (55.2)	98 (39.2)		
Sleep quality (PSQI), mean (SD)					
Subjective sleep quality	1.13 (0.74)	1.22 (0.74)	1.04 (0.72)	.005	0.25
Sleep latency	1.24 (0.98)	1.33 (1.01)	1.14 (0.93)	.02	0.20
Sleep duration	0.42 (0.71)	0.46 (0.75)	0.37 (0.66)	.13	0.14
Habitual sleep efficiency	0.38 (0.78)	0.45 (0.86)	0.30 (0.67)	.03	0.19
Sleep disturbances	1.06 (0.53)	1.12 (0.53)	1.00 (0.53)	.009	0.23
Sleeping medication use	0.17 (0.59)	0.21 (0.66)	0.12 (0.50)	.06	0.17
Daytime dysfunction	1.53 (0.98)	1.68 (0.97)	1.37 (0.96)	<.001	0.32
Total score	5.92 (3.29)	6.48 (3.32)	5.33 (3.16)	<.001	0.36
Social support (SSRS [d]), n (%)					
Low level (total score ≤22)	70 (13.8)	56 (21.6)	14 (5.6)	<.001	36.42 (2)
Medium level (total score 23-44)	415 (81.5)	199 (76.8)	216 (86.4)		
High level (total score >45)	24 (4.7)	4 (1.6)	20 (8)		
Social support (SSRS), mean (SD)					
Objective support	7.55 (2.34)	7.23 (2.35)	7.89 (2.28)	.001	−0.28
Subjective support	16.17 (5.59)	14.54 (5.09)	17.85 (5.60)	<.001	−0.62
Utilization of support	7.49 (1.97)	7.04 (1.88)	7.96 (1.95)	<.001	−0.48
Total score	31.21 (7.96)	28.81 (7.41)	33.70 (7.75)	<.001	−0.65

[a] SMW: sexual minority women.

[b] CHW: cisgender heterosexual women.

[c] PSQI: Pittsburgh Sleep Quality Index.

[d] SSRS: Social Support Rating Scale.

Associations Between Social-Ecological Factors and Sleep

The correlations among sleep quality (PSQI total score and seven component scores) and social support (SSRS total score and three domain scores), social relationship Qol, and environment Qol are reported in Table 3. The total sleep quality was significantly negatively associated with social support and all domains with weak correlations (r ranged from −0.15 to −0.26, $P<.001$), and it was significantly negatively linked with social relationship Qol ($r=-0.38$, $P<.001$) and environment Qol ($r=-0.39$, $P<.001$) with moderate correlations.

Table . Correlations between sleep quality, social support, and quality of life.

Sleep quality (PSQI[a])	Social support (SSRS[b])				Quality of life	
	Objective	Subjective	Utilization	Sum score	Social	Environment
Subjective sleep quality						
r	−0.07	−0.17	−0.13	−0.17	−0.23	−0.26
P value	.14	<.001	.005	<.001	<.001	<.001
Sleep latency						
r	−0.13	−0.12	−0.16	−0.16	−0.21	−0.19
P value	.004	.005	<.001	<.001	<.001	<.001
Sleep duration						
r	−0.12	−0.15	−0.13	−0.17	−0.26	−0.28
P value	.007	<.001	.003	<.001	<.001	<.001
Habitual sleep efficiency						
r	−0.12	−0.15	−0.18	−0.18	−0.25	−0.26
P value	.007	<.001	<.001	<.001	<.001	<.001
Sleep disturbances						
r	−0.09	−0.12	−0.13	−0.14	−0.26	−0.29
P value	.04	.009	.003	.001	<.001	<.001
Sleeping medication use						
r	−0.12	−0.12	−0.09	−0.14	−0.10	−0.15
P value	.007	.007	.04	.001	.02	<.001
Daytime dysfunction						
r	−0.02	−0.18	−0.08	−0.15	−0.32	−0.25
P value	.66	<.001	.09	<.001	<.001	<.001
Total score						
r	−0.15	−0.24	−0.21	−0.26	−0.38	−0.39
P value	<.001	<.001	<.001	<.001	<.001	<.001

[a] PSQI: Pittsburgh Sleep Quality Index.

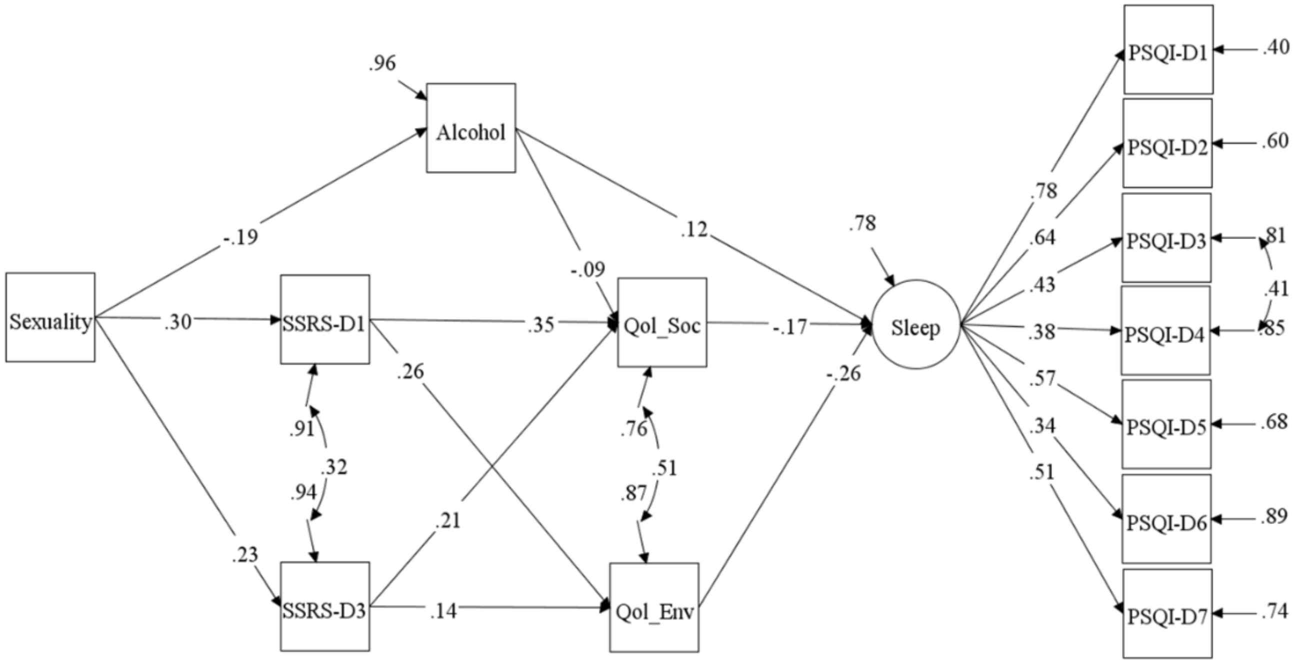
[b] SSRS: Social Support Rating Scale.

Comparisons of overall sleep quality between different socioecological groups are shown in Table S1 in Multimedia Appendix 1, with significant differences in sleep quality between people with different sexual identities, smoking status, alcohol use status, number of friends, and perceived social support (all $P<.01$). Table S2 in Multimedia Appendix 1 presents the multiple linear regression model of all the significant variables on sleep quality, and SEM was then performed among all these significant variables.

Model Testing

The full initial model including smoking and number of friends failed because these 2 variables were neither directly nor indirectly correlated to sleep in the model and were therefore not included in subsequent model analyses. Contrary to the hypothesized model, none of the aspects of social support were found to have a direct relationship with sleep quality. Furthermore, the model involving the subjective support domain of social support (SSRS-D2) showed no direct or indirect association of this domain with sleep and it was therefore also removed. The final model was then proposed and illustrated in Figure 2, exhibiting robust fit indices (χ^2_{51}=118.80; RMSEA=0.051, 90% CI 0.039-0.063; CFI=0.948, TLI=0.921, SRMR=0.040).

Figure 2. Final structural equation model of this study. Correlations and paths with nonsignificant *P* values are not depicted. PSQI: Pittsburgh Sleep Quality Index; PSQI-D1: subjective sleep quality; PSQI-D2: sleep latency; PSQI-D3: sleep duration; PSQI-D4: habitual sleep efficiency; PSQI-D5: sleep disturbance; PSQI-D6: sleeping medication use; PSQI-D7: daytime dysfunction; Qol: quality of life; Qol_Soc: social relationship of Qol; Qol_Env: environment domain of Qol; SSRS: Social Support Rating Scale; SSRS-D1: objective support; SSRS-D3: utilization of support.



All model standardized coefficients (β), their standard errors, and associated *P* values are reported in Table 4. There were 6 significant paths from sexual identity (sexuality) to sleep quality in the final SEM model. Specifically, alcohol use (β=0.12, *P*=.01), social relationship Qol (β=−0.17, *P*=.004), and environment Qol (β=−0.26, *P*<.001) were direct predictors of sleep quality. Although social support did not directly influence sleep, it may indirectly influence sleep by affecting the perceived social relationship Qol and environment Qol (β range 0.14 - 0.35, *P*<.01).

Table . Standardized path coefficients and standard errors for all pathways in the final model.

Pathways	β[a]	SE	*P* value
Sexuality → Sleep quality	−0.07	0.05	.17
Sexuality → Alcohol use	−0.19	0.04	<.001
Sexuality → SSRS-D1[b]	0.30	0.04	<.001
Sexuality → SSRS-D3[c]	0.23	0.04	<.001
Sexuality → Qol_Soc[d]	0.03	0.04	.51
Sexuality → Qol_Env[e]	0.05	0.04	.23
Alcohol use → Sleep quality	0.12	0.05	.01
Alcohol use → Qol_Soc	−0.09	0.04	.02
Alcohol use → Qol_Env	0.00	0.04	.98
SSRS-D1 → Sleep quality	−0.03	0.05	.65
SSRS-D1 → Qol_Soc	0.35	0.04	<.001
SSRS-D1 → Qol_Env	0.26	0.05	<.001
SSRS-D3 → Sleep quality	−0.06	0.05	.25
SSRS-D3 → Qol_Soc	0.21	0.04	<.001
SSRS-D3 → Qol_Env	0.14	0.05	.002
Qol_Soc → Sleep quality	−0.17	0.06	.004
Qol_Env → Sleep quality	−0.26	0.06	<.001

[a] β: standardized coefficient.

[b] SSRS-D1: objective support domain of social support.

[c] SSRS-D3: utilization of support domain of social support.

[d] Qol_Soc: social relationship domain of quality of life.

[e] Qol_Env: environment domain of quality of life.

Discussion

Principal Findings

This study was the first to explore the social-ecological pathways from sexual identity to sleep quality. Few studies worldwide have examined the sleep health of SMW, and this investigation was the first to examine the sleep health of Chinese women with diverse sexual identities. This study showed SMW reported significantly worse overall sleep quality relative to CHW, which is consistent with existing reviews [14-16] and findings among Chinese youth [24,25]. Compared with the pooled PSQI mean score of the general Chinese population from a recent meta-analysis [72], our study populations reported higher total scores on the same standardized PSQI scale (5.92 vs 4.32), implying that the female population in this study experienced noticeably poorer sleep quality. Furthermore, compared with the percentage of Korean sexual minority adults (lesbian, gay, and bisexual) reporting poor sleep quality (33.8%) [73], the rate of SMW reporting poor sleep quality in our study (55.2%) was higher, indicating that sleep in SMW requires more attention and improvement.

This study also found that the proportions of current smokers and alcohol users were significantly higher among SMW than those among CHW, which is consistent with findings in China and overseas [28,74]. A scoping review summarized that some Chinese SMW regard smoking and drinking as a means of interacting with other SMW in specific social gatherings. It is also reported that smoking and drinking were the main strategies they used as a sexual minority to cope with marital stress and feeling isolated [28]. These also echo our findings on the mental health of this sample population, where we have found that SMW indeed significantly experienced more psychological symptoms [56]. A national survey conducted in the United States revealed higher rates of substance use disorders among sexual minority adults compared to heterosexual adults, with SMW showing the highest rates [75], suggesting that substance use among women, especially SMW, needs to be given attention as a public health issue.

Alcohol is one of the most commonly used psychoactive substances in different social communities around the world. A growing body of literature has confirmed that alcohol use has deleterious effects on sleep by reducing sleep duration and sleep quality [35,36], which has also been documented in sexual minority men [76] and is consistent with this study. However, a meta-analysis of cohort studies indicates that the relationship between alcohol consumption and sleep disorders is conditional, with pooled analyses finding that general drinking and the incidence of sleep disorder were significantly correlated, while heavy drinking was not [37]. Another review concluded that alcohol use has a bidirectional relationship with sleep continuity disturbance [77]. For instance, insomnia symptoms were associated with subsequent heavy drinking; conversely, heavy drinking was associated with subsequent insomnia symptoms [78]. Nevertheless, both comparative analyses and the final SEM in this study revealed that drinking alcohol was significantly associated with poor sleep quality compared to not drinking alcohol, suggesting that both alcohol use and sleep in women deserve more attention and tailored intervention. Considering the limitations of the cross-sectional design used in this study, which could not provide evidence of causal relationships between drinking and sleep quality, more longitudinal studies with further subdivision of alcohol use are warranted in the future.

The adverse effects of smoking on sleep are well documented, as evidenced by the fact that smokers are more likely to experience sleeping difficulties, longer sleep latency, and poorer sleep quality [38,39,79], and similar findings were observed in the all-female population included in this study. The possible mechanism regarding the effects of smoking on sleep may be that smoking triggers depression or sleep-related breathing problems, which can lead to poor sleep quality [79]. However, a meta-analysis that included only cohort studies showed that, despite the significant correlation between smoking and the prevalence of insomnia, the difference between smokers and nonsmokers was very small (odds ratio 1.07) [80]. This may partly explain why the smoking variable that differed in between-group comparisons in our study was no longer associated with sleep quality in our final SEM. Future studies should record and subdivide smoking status and use prospective study designs to further explore the relationship between smoking and sleep across sexual identities.

The correlation between social support and sleep has been extensively studied in different populations [42,43,45] but its exploration in sexual minorities remains scarce. Therefore, the strength of this study is to fill this gap by examining the association between social support and sleep quality among Chinese women with diverse sexual identities. Consistent with previous research [49], overall social support was significantly and negatively associated with each component of sleep quality in the current study sample. Moreover, sexual minority individuals generally experience a higher level of poor social relationships, especially in the relatively traditional Chinese culture, where such groups are often invisible and therefore perceive more discrimination and stress [27,28,56]. However, findings from a study of 3 generations showed that feeling stigma was statistically significant in predicting a sleep disorder diagnosis among sexual minority individuals [81].

A mediation analysis of Chinese college students nationwide found that the relationship between sexual orientation and sleep quality was independently and in series mediated by interpersonal relationships and depressive symptoms, and this effect was more robust in men than women [25]. These interpersonal difficulties were related to their perceived stress associated with sleep difficulties. Our findings align with previous studies that found that sexual minority status is significantly associated with poor sleep quality [23], and social support has both direct and indirect correlations with sleep quality [42]. The relationships manifested in this study were that sexual identity was significantly associated with inadequate levels of social support, which in turn was significantly associated with poor levels of quality of life in terms of social relationships and environment, all of which were associated with poor sleep quality. Therefore, the lack of social support for sexual minority women is a significant public health issue, as is their poor sleep health, which deserves more attention.

Our correlation findings suggest that there may be more complex pathways from sexual identity to sleep among Chinese women with different sexual identities. Therefore, based on the holistic perspective proposed by the social-ecological model of sleep [32], this study explored and demonstrated the socioecological pathways from sexual identity to sleep quality for the first time using the multivariate SEM analysis method, providing a comprehensive evidence reference for promoting sleep health. The results of the final fitted model with satisfactory fit indices showed that starting from sexual identity, the following factors played important roles in women’s sleep quality: alcohol use; objective social support and utilization of support; and social relationship and environment Qol. Our findings demonstrated the important roles of socioecological factors on sleep quality among women with different sexual identities. Therefore, future research should consider not only modifiable individual-level factors but also the potential influence of social- and societal-level factors when developing and implementing sleep promotion interventions. In particular, future interventions targeting sleep quality in sexual minority women may benefit from incorporating strategies that improve their social support, alcohol use, and the quality of their living environment.

This study has several limitations. First, due to the cross-sectional design, it was not possible to determine the explicit causal relationships among variables, so further research using longitudinal designs is needed. Second, despite the diversity of sexual identities, our sample was an all-female population (self-identified as female or assigned female at birth), so the results might not be generalizable to the broader Chinese sexual minority population. There has also been research indicating that bisexual people [82], those with less education [83], and those who are economically disadvantaged [16] may be more likely to suffer from sleep disorders than other groups. Hence, future studies with more diverse and representative samples including individuals from varied socioeconomic backgrounds are warranted. Third, the social-ecological model of sleep used in this study is only the upstream end of sleep research, and the downstream effects, namely the effects of sleep on specific health outcomes, have not been expanded upon in this study. Therefore, further research is recommended to comprehensively explore sleep and holistic health across diverse populations.

Conclusion

Overall, this study contributes to the existing literature by including SMW and expanding knowledge on sleep health in this population. By exploring socioecological pathways from sexual identity to sleep quality, this study provides a comprehensive understanding of factors that affect sleep health at different levels. These contextual and modifiable factors, based on the socioecological model of sleep, are layered and may interact with one another, providing guidance for future interventions. In addition, the pathways from sexual identity to sleep identified in this study are complex and multidisciplinary; thus, all stakeholders and sleep professionals across disciplines are encouraged to collaborate and contribute to sleep improvement programs tailored for sexual minority individuals.

Acknowledgments

We are grateful to all the participants for sharing their experiences with us. We also thank TrueSelf Shenzhen, Period Pride China, Wuhan LGBT Center, and r&B bisexual community, as well as all those who shared this survey with their friends, for their contribution to the questionnaire distribution.

Data Availability

The datasets generated during and/or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Table S1: Sleep quality of the study sample with comparisons; Table S2: Multiple linear regressions of sleep quality.

[DOCX File, 78 KB - publichealth_v11i1e53549_app1.docx]

References

1. Czeisler CA. Duration, timing and quality of sleep are each vital for health, performance and safety. Sleep Health 2015 Mar;1(1):5-8. [doi: 10.1016/j.sleh.2014.12.008] [Medline: 29073414]
2. Gao C, Guo J, Gong TT, et al. Sleep duration/quality with health outcomes: an umbrella review of meta-analyses of prospective studies. Front Med (Lausanne) 2021;8:813943. [doi: 10.3389/fmed.2021.813943] [Medline: 35127769]
3. Sivakumaran K, Ritonja JA, Palmer N, et al. Effect of sleep disturbance on biomarkers related to the development of adverse health outcomes: a systematic review of the human literature. J Sleep Res 2023 Jun;32(3):e13775. [doi: 10.1111/jsr.13775] [Medline: 36330773]
4. Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014 Jan 1;37(1):9-17. [doi: 10.5665/sleep.3298] [Medline: 24470692]
5. Weaver MD, Barger LK. Sleep health as an issue of public safety. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:489-499.
6. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
7. van de Langenberg SCN, Kocevska D, Luik AI. The multidimensionality of sleep in population-based samples: a narrative review. J Sleep Res 2022 Aug;31(4):e13608. [doi: 10.1111/jsr.13608] [Medline: 35429087]
8. Mong JA, Cusmano DM. Sex differences in sleep: impact of biological sex and sex steroids. Philos Trans R Soc Lond B Biol Sci 2016 Feb 19;371(1688):20150110. [doi: 10.1098/rstb.2015.0110] [Medline: 26833831]
9. Grandner MA, Williams NJ, Knutson KL, Roberts D, Jean-Louis G. Sleep disparity, race/ethnicity, and socioeconomic position. Sleep Med 2016 Feb;18:7-18. [doi: 10.1016/j.sleep.2015.01.020] [Medline: 26431755]
10. Patel NP, Grandner MA, Xie D, Branas CC, Gooneratne N. “Sleep disparity” in the population: poor sleep quality is strongly associated with poverty and ethnicity. BMC Public Health 2010 Aug 11;10:475. [doi: 10.1186/1471-2458-10-475] [Medline: 20701789]
11. Baldwin CM, Ervin AM, Mays MZ, et al. Sleep disturbances, quality of life, and ethnicity: the Sleep Heart Health Study. J Clin Sleep Med 2010 Apr 15;6(2):176-183. [Medline: 20411696]
12. Blanc J, Nunes J, Williams N, Robbins R, Seixas AA, Jean-Louis G. Sleep health equity. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:473-480. [doi: 10.1016/B978-0-12-815373-4.00035-6]
13. Williams N, Jean-Louis G, Blanc J, Wallace DM. Race, socioeconomic position and sleep. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:57-76. [doi: 10.1016/B978-0-12-815373-4.00006-X]
14. Butler ES, McGlinchey E, Juster RP. Sexual and gender minority sleep: a narrative review and suggestions for future research. J Sleep Res 2020 Feb;29(1):e12928. [doi: 10.1111/jsr.12928] [Medline: 31626363]
15. Dolsen EA, Byers AL, Flentje A, et al. Sleep disturbance and suicide risk among sexual and gender minority people. Neurobiol Stress 2022 Nov;21:100488. [doi: 10.1016/j.ynstr.2022.100488] [Medline: 36164391]
16. Patterson CJ, Potter EC. Sexual orientation and sleep difficulties: a review of research. Sleep Health 2019 Jun;5(3):227-235. [doi: 10.1016/j.sleh.2019.02.004] [Medline: 30975620]

17. Patterson CJ, Potter EC. Sleep difficulties of sexual minority adults: evidence from the 2018 National Health Interview Study. Psychol Sex Orientat Gend Divers 2022;9(2):244-253. [doi: 10.1037/sgd0000473]
18. Caceres BA, Hickey KT, Heitkemper EM, Hughes TL. An intersectional approach to examine sleep duration in sexual minority adults in the United States: findings from the Behavioral Risk Factor Surveillance System. Sleep Health 2019 Dec;5(6):621-629. [doi: 10.1016/j.sleh.2019.06.006] [Medline: 31377249]
19. Trinh MH, Agénor M, Austin SB, Jackson CL. Health and healthcare disparities among U.S. women and men at the intersection of sexual orientation and race/ethnicity: a nationally representative cross-sectional study. BMC Public Health 2017 Dec 19;17(1):964. [doi: 10.1186/s12889-017-4937-9] [Medline: 29258470]
20. Galinsky AM, Ward BW, Joestl SS, Dahlhamer JM. Sleep duration, sleep quality, and sexual orientation: findings from the 2013-2015 National Health Interview Survey. Sleep Health 2018 Feb;4(1):56-62. [doi: 10.1016/j.sleh.2017.10.004] [Medline: 29332681]
21. Martin-Storey A, Prickett KC, Crosnoe R. Disparities in sleep duration and restedness among same- and different-sex couples: findings from the American Time Use Survey. Sleep 2018 Aug 1;41(8):zsy090. [doi: 10.1093/sleep/zsy090] [Medline: 29726972]
22. Fricke J, Sironi M. Dimensions of sexual orientation and sleep disturbance among young adults. Prev Med Rep 2017 Dec;8:18-24. [doi: 10.1016/j.pmedr.2017.07.008] [Medline: 28831369]
23. Huang Y, Li P, Lai Z, et al. Role of sleep quality in mediating the relationship between sexual minority status and suicidal behavior among Chinese adolescents. Psychol Res Behav Manag 2018;11:607-615. [doi: 10.2147/PRBM.S186586] [Medline: 30573996]
24. Li P, Huang Y, Guo L, et al. Is sexual minority status associated with poor sleep quality among adolescents? Analysis of a national cross-sectional survey in Chinese adolescents. BMJ Open 2017 Dec 26;7(12):e017067. [doi: 10.1136/bmjopen-2017-017067] [Medline: 29282258]
25. Wu R, Wang W, Li W, et al. Sexual orientation and sleep problem among Chinese college students: mediating roles of interpersonal problems and depressive symptoms. J Affect Disord 2021 Dec 1;295:569-577. [doi: 10.1016/j.jad.2021.08.075] [Medline: 34509072]
26. Chan KKS, Fung WTW. Differential impact of experienced and anticipated discrimination on sleep and health among sexual minorities. Arch Sex Behav 2021 Oct;50(7):3053-3063. [doi: 10.1007/s10508-021-01981-2] [Medline: 34617190]
27. Wu C, Choi EPH, Chau PH. The holistic health status of Chinese homosexual and bisexual adults: a scoping review. Front Public Health 2021;9:710575. [doi: 10.3389/fpubh.2021.710575] [Medline: 34504828]
28. Wei TH, Jiang Y. Mapping the scientific literature on health needs of women with same-sex attraction in Mainland China: a scoping review. LGBT Health 2023 May;10(4):263-277. [doi: 10.1089/lgbt.2022.0119] [Medline: 36595507]
29. Billings ME, Hale L, Johnson DA. Physical and social environment relationship with sleep health and disorders. Chest 2020 May;157(5):1304-1312. [doi: 10.1016/j.chest.2019.12.002] [Medline: 31870910]
30. Etindele Sosso FA, Holmes SD, Weinstein AA. Influence of socioeconomic status on objective sleep measurement: a systematic review and meta-analysis of actigraphy studies. Sleep Health 2021 Aug;7(4):417-428. [doi: 10.1016/j.sleh.2021.05.005] [Medline: 34266774]
31. Grandner MA, Hale L, Moore M, Patel NP. Mortality associated with short sleep duration: the evidence, the possible mechanisms, and the future. Sleep Med Rev 2010 Jun;14(3):191-203. [doi: 10.1016/j.smrv.2009.07.006] [Medline: 19932976]
32. Grandner MA. Social-ecological model of sleep health. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:45-53. [doi: 10.1016/B978-0-12-815373-4.00005-8]
33. Fischer AR, Green SRM, Gunn HE. Social-ecological considerations for the sleep health of rural mothers. J Behav Med 2021 Aug;44(4):507-518. [doi: 10.1007/s10865-020-00189-4] [Medline: 33083923]
34. Li S. Nonrestorative sleep: its evaluation, associated factors and impact. Thesis.: The University of Hong Kong; 2021 URL: https://hub.hku.hk/handle/10722/306966 [accessed 2025-01-13]
35. He S, Taylor BV, Thakur NP, Chakravorty S. Sleep and alcohol use. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:269-281. [doi: 10.1016/B978-0-12-815373-4.00021-6]
36. Ebrahim IO, Shapiro CM, Williams AJ, Fenwick PB. Alcohol and sleep I: effects on normal sleep. Alcohol Clin Exp Res 2013 Apr;37(4):539-549. [doi: 10.1111/acer.12006] [Medline: 23347102]
37. Hu N, Ma Y, He J, Zhu L, Cao S. Alcohol consumption and incidence of sleep disorder: a systematic review and meta-analysis of cohort studies. Drug Alcohol Depend 2020 Dec 1;217:108259. [doi: 10.1016/j.drugalcdep.2020.108259] [Medline: 32927195]
38. Patterson F, Grandner MA, Lozano A, Satti A, Ma G. Transitioning from adequate to inadequate sleep duration associated with higher smoking rate and greater nicotine dependence in a population sample. Addict Behav 2018 Feb;77:47-50. [doi: 10.1016/j.addbeh.2017.09.011] [Medline: 28950118]
39. Patterson F, Ashare R. Improved sleep as an adjunctive treatment for smoking cessation. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:283-301. [doi: 10.1016/B978-0-12-815373-4.00022-8]
40. Gordon AM, Carrillo B, Barnes CM. Sleep and social relationships in healthy populations: a systematic review. Sleep Med Rev 2021 Jun;57:101428. [doi: 10.1016/j.smrv.2021.101428] [Medline: 33596514]

41. Nomura K, Yamaoka K, Nakao M, Yano E. Social determinants of self-reported sleep problems in South Korea and Taiwan. J Psychosom Res 2010 Nov;69(5):435-440. [doi: 10.1016/j.jpsychores.2010.04.014] [Medline: 20955862]
42. Kent de Grey RG, Uchino BN, Trettevik R, Cronan S, Hogan JN. Social support and sleep: a meta-analysis. Health Psychol 2018 Aug;37(8):787-798. [doi: 10.1037/hea0000628] [Medline: 29809022]
43. Krause N, Rainville G. Exploring the relationship between social support and sleep. Health Educ Behav 2020 Feb;47(1):153-161. [doi: 10.1177/1090198119871331] [Medline: 31452389]
44. Jin Y, Ding Z, Fei Y, et al. Social relationships play a role in sleep status in Chinese undergraduate students. Psychiatry Res 2014 Dec 15;220(1-2):631-638. [doi: 10.1016/j.psychres.2014.08.029] [Medline: 25200188]
45. Mesas AE, Peppard PE, Hale L, Friedman EM, Nieto FJ, Hagen EW. Individuals' perceptions of social support from family and friends are associated with lower risk of sleep complaints and short sleep duration. Sleep Health 2020 Feb;6(1):110-116. [doi: 10.1016/j.sleh.2019.08.013] [Medline: 31740375]
46. Ailshire JA, Burgard SA. Family relationships and troubled sleep among U.S. adults: examining the influences of contact frequency and relationship quality. J Health Soc Behav 2012;53(2):248-262. [doi: 10.1177/0022146512446642] [Medline: 22653715]
47. Frost DM, Meyer IH, Schwartz S. Social support networks among diverse sexual minority populations. Am J Orthopsychiatry 2016 Jan;86(1):91-102. [doi: 10.1037/ort0000117] [Medline: 26752447]
48. Jackson CL, Gaston SA. The impact of environmental exposures on sleep. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:85-103. [doi: 10.1016/B978-0-12-815373-4.00008-3]
49. Stafford M, Bendayan R, Tymoszuk U, Kuh D. Social support from the closest person and sleep quality in later life: evidence from a British birth cohort study. J Psychosom Res 2017 Jul;98:1-9. [doi: 10.1016/j.jpsychores.2017.04.014] [Medline: 28554363]
50. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
51. Crowley SL, Fan X. Structural equation modeling: basic concepts and applications in personality assessment research. J Pers Assess 1997 Jun;68(3):508-531. [doi: 10.1207/s15327752jpa6803_4] [Medline: 9170298]
52. Jackson DL. Revisiting sample size and number of parameter estimates: some support for the N:q hypothesis. Struct Equ Modeling 2003 Jan;10(1):128-141. [doi: 10.1207/S15328007SEM1001_6]
53. Violato C, Hecker KG. How to use structural equation modeling in medical education research: a brief guide. Teach Learn Med 2007;19(4):362-371. [doi: 10.1080/10401330701542685] [Medline: 17935466]
54. Crawford FW, Wu J, Heimer R. Hidden population size estimation from respondent-driven sampling: a network approach. J Am Stat Assoc 2018;113(522):755-766. [doi: 10.1080/01621459.2017.1285775] [Medline: 30828120]
55. Hequembourg AL, Panagakis C. Maximizing respondent-driven sampling field procedures in the recruitment of sexual minorities for health research. SAGE Open Med 2019;7:2050312119829983. [doi: 10.1177/2050312119829983] [Medline: 30800299]
56. Wu C, Chau PH, Choi EPH. Quality of life and mental health of Chinese sexual and gender minority women and cisgender heterosexual women: cross-sectional survey and mediation analysis. JMIR Public Health Surveill 2023 Feb 22;9:e42203. [doi: 10.2196/42203] [Medline: 36811941]
57. Choi EPH, Wan EYF, Kwok JYY, Chin WY, Lam CLK. The mediating role of sleep quality in the association between nocturia and health-related quality of life. Health Qual Life Outcomes 2019 Dec 11;17(1):181. [doi: 10.1186/s12955-019-1251-5] [Medline: 31829192]
58. Tsai PS, Wang SY, Wang MY, et al. Psychometric evaluation of the Chinese version of the Pittsburgh Sleep Quality Index (CPSQI) in primary insomnia and control subjects. Qual Life Res 2005 Oct;14(8):1943-1952. [doi: 10.1007/s11136-005-4346-x] [Medline: 16155782]
59. Liu X, Tang M, Hu L, et al. Reliability and validity of the Pittsburgh sleep quality index. Chin J Neuro Psychiatry 1996;29(2):103-107.
60. Xiao S. The theoretical basis and application of social support rating scale (SSRS). J Clin Psychol Med (Chin) 1994;4(2):98-100.
61. Zhang JY, Shu T, Xiang M, Feng ZC. Learning burnout: evaluating the role of social support in medical students. Front Psychol 2021;12:625506. [doi: 10.3389/fpsyg.2021.625506] [Medline: 33692725]
62. Hou T, Zhang T, Cai W, et al. Social support and mental health among health care workers during coronavirus disease 2019 outbreak: a moderated mediation model. PLoS ONE 2020;15(5):e0233831. [doi: 10.1371/journal.pone.0233831] [Medline: 32470007]
63. Sun J, Sun R, Jiang Y, et al. The relationship between psychological health and social support: evidence from physicians in China. PLoS One 2020;15(1):e0228152. [doi: 10.1371/journal.pone.0228152] [Medline: 31995601]
64. Xia P, Li N, Hau KT, Liu C, Lu Y. Quality of life of Chinese urban community residents: a psychometric study of the mainland Chinese version of the WHOQOL-BREF. BMC Med Res Methodol 2012 Mar 27;12:37. [doi: 10.1186/1471-2288-12-37] [Medline: 22452994]

65. Choi EPH, Wong JYH, Fong DYT. Mental health and health-related quality of life of Chinese college students who were the victims of dating violence. Qual Life Res 2017 Apr;26(4):945-957. [doi: 10.1007/s11136-016-1413-4] [Medline: 27660071]
66. Bonomi AE, Patrick DL, Bushnell DM, Martin M. Validation of the United States' version of the World Health Organization Quality of Life (WHOQOL) instrument. J Clin Epidemiol 2000 Jan;53(1):1-12. [doi: 10.1016/s0895-4356(99)00123-7] [Medline: 10693897]
67. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
68. Cohen J. Statistical Power Analysis for the Behavioral Sciences: Routledge; 2013.
69. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
70. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods 1996;1(2):130-149. [doi: 10.1037//1082-989X.1.2.130]
71. Muthén B, Muthén L. Mplus 86 software. URL: https://www.statmodel.com/index.shtml [accessed 2025-01-13]
72. Chen P, Lam MI, Si TL, et al. The prevalence of poor sleep quality in the general population in China: a meta-analysis of epidemiological studies. Eur Arch Psychiatry Clin Neurosci 2024 Oct;274(7):1-14. [doi: 10.1007/s00406-024-01764-5] [Medline: 38429554]
73. Choo S, Kim R, Lee H, Yi H, Kim R, Kim SS. Association between discrimination and poor sleep health outcomes among 2192 South Korean gay, lesbian, and bisexual adults: a nationwide cross-sectional survey. Sleep Health 2022 Dec;8(6):587-592. [doi: 10.1016/j.sleh.2022.09.006] [Medline: 36220734]
74. Johns MM, Pingel ES, Youatt EJ, Soler JH, McClelland SI, Bauermeister JA. LGBT community, social network characteristics, and smoking behaviors in young sexual minority women. Am J Community Psychol 2013 Sep;52(1-2):141-154. [doi: 10.1007/s10464-013-9584-4] [Medline: 23783884]
75. Kahle EM, Veliz P, McCabe SE, Boyd CJ. Functional and structural social support, substance use and sexual orientation from a nationally representative sample of US adults. Addiction 2020 Mar;115(3):546-558. [doi: 10.1111/add.14819] [Medline: 31599027]
76. Duncan DT, Goedel WC, Mayer KH, et al. Poor sleep health and its association with mental health, substance use, and condomless anal intercourse among gay, bisexual, and other men who have sex with men. Sleep Health 2016 Dec;2(4):316-321. [doi: 10.1016/j.sleh.2016.07.003] [Medline: 29073390]
77. He S, Hasler BP, Chakravorty S. Alcohol and sleep-related problems. Curr Opin Psychol 2019 Dec;30:117-122. [doi: 10.1016/j.copsyc.2019.03.007] [Medline: 31128400]
78. Haario P, Rahkonen O, Laaksonen M, Lahelma E, Lallukka T. Bidirectional associations between insomnia symptoms and unhealthy behaviours. J Sleep Res 2013 Feb;22(1):89-95. [doi: 10.1111/j.1365-2869.2012.01043.x] [Medline: 22978579]
79. Amiri S, Behnezhad S. Smoking and risk of sleep-related issues: a systematic review and meta-analysis of prospective studies. Can J Public Health 2020 Oct;111(5):775-786. [doi: 10.17269/s41997-020-00308-3] [Medline: 32185746]
80. Hu N, Wang C, Liao Y, Dai Q, Cao S. Smoking and incidence of insomnia: a systematic review and meta-analysis of cohort studies. Public Health (Fairfax) 2021 Sep;198:324-331. [doi: 10.1016/j.puhe.2021.07.012] [Medline: 34507139]
81. Segovia M, Sparks PJ. Sexual minorities and sleep health: how does discrimination and stigma influence a sleep disorder diagnosis? Sleep Health 2024 Feb;10(1):41-47. [doi: 10.1016/j.sleh.2023.11.003] [Medline: 38151378]
82. Caceres BA, Hickey KT. Examining sleep duration and sleep health among sexual minority and heterosexual adults: findings from NHANES (2005-2014). Behav Sleep Med 2020;18(3):345-357. [doi: 10.1080/15402002.2019.1591410] [Medline: 30916580]
83. Chu Y, Aune D, Yu C, et al. Temporal trends in sleep pattern among Chinese adults between 2010 and 2018: findings from five consecutive nationally representative surveys. Public Health (Fairfax) 2023 Dec;225:360-368. [doi: 10.1016/j.puhe.2023.10.004] [Medline: 37981452]

Abbreviations

CFI: comparative fit index
CHW: cisgender heterosexual women
PSQI: Pittsburgh Sleep Quality Index
Qol: quality of life
Qol_Env: environment domain of quality of life
Qol_Soc: social relationship domain of quality of life
RMSEA: root mean square error of approximation
SEM: structural equation modeling
SMW: sexual minority women
SRMR: standardized root mean square residual

SSRS: Social Support Rating Scale
TLI: Tucker-Lewis index
WHOQOL-BREF: World Health Organization Quality of Life–abbreviated short version

Edited by A Mavragani; submitted 10.10.23; peer-reviewed by A AL-Asadi, Q Chen; revised version received 22.10.24; accepted 29.10.24; published 21.01.25.

Please cite as:
Wu C, Chau PH, Choi EPH
Exploring Social-Ecological Pathways From Sexual Identity to Sleep Among Chinese Women: Structural Equation Modeling Analysis
JMIR Public Health Surveill 2025;11:e53549
URL: https://publichealth.jmir.org/2025/1/e53549
doi:10.2196/53549

Strategies to Increase Response Rate and Reduce Nonresponse Bias in Population Health Research: Analysis of a Series of Randomized Controlled Experiments during a Large COVID-19 Study

Christina J Atchison[1], PhD; Nicholas Gilby[2], BA; Galini Pantelidou[2], MSc; Sam Clemens[2], BA; Kevin Pickering[2], MSc; Marc Chadeau-Hyam[1], PhD; Deborah Ashby[1], PhD; Wendy S Barclay[3], PhD; Graham S Cooke[3], PhD; Ara Darzi[4], MD; Steven Riley[1], PhD; Christl A Donnelly[1,5], ScD; Helen Ward[1], PhD; Paul Elliott[1], PhD

[1]School of Public Health, Imperial College London, London, United Kingdom
[2]Ipsos, London, United Kingdom
[3]Department of Infectious Disease, Imperial College London, Norfolk Place, London, United Kingdom
[4]Institute of Global Health Innovation, Imperial College London, South Kensington Campus, London, United Kingdom
[5]Department of Statistics, University of Oxford, Oxford, United Kingdom

Corresponding Author:
Paul Elliott, PhD
School of Public Health, Imperial College London, London, United Kingdom

Abstract

Background: High response rates are needed in population-based studies, as nonresponse reduces effective sample size and bias affects accuracy and decreases the generalizability of the study findings.

Objective: We tested different strategies to improve response rate and reduce nonresponse bias in a national population–based COVID-19 surveillance program in England, United Kingdom.

Methods: Over 19 rounds, a random sample of individuals aged 5 years and older from the general population in England were invited by mail to complete a web-based questionnaire and return a swab for SARS-CoV-2 testing. We carried out several nested randomized controlled experiments to measure the impact on response rates of different interventions, including (1) variations in invitation and reminder letters and SMS text messages and (2) the offer of a conditional monetary incentive to return a swab, reporting absolute changes in response and relative response rate (95% CIs).

Results: Monetary incentives increased the response rate (completed swabs returned as a proportion of the number of individuals invited) across all age groups, sex at birth, and area deprivation with the biggest increase among the lowest responders, namely teenagers and young adults and those living in more deprived areas. With no monetary incentive, the response rate was 3.4% in participants aged 18 - 22 years, increasing to 8.1% with a £10 (US $12.5) incentive, 11.9% with £20 (US $25.0), and 18.2% with £30 (US $37.5) (relative response rate 2.4 [95% CI 2.0-2.9], 3.5 [95% CI 3.0-4.2], and 5.4 [95% CI 4.4-6.7], respectively). Nonmonetary strategies had a modest, if any, impact on response rate. The largest effect was observed for sending an additional swab reminder (SMS text message or email). For example, those receiving an additional SMS text message were more likely to return a completed swab compared to those receiving the standard email-SMS approach, 73.3% versus 70.2%: percentage difference 3.1% (95% CI 2.2%-4.0%).

Conclusions: Conditional monetary incentives improved response rates to a web-based survey, which required the return of a swab test, particularly for younger age groups. Used in a selective way, incentives may be an effective strategy for improving sample response and representativeness in population-based studies.

(JMIR Public Health Surveill 2025;11:e60022) doi:10.2196/60022

KEYWORDS

study recruitment; response rate; population-based research; COVID-19; SARS-CoV-2; web-based questionnaires

Introduction

In population-based studies, a high response rate from a representative sample may reduce nonparticipation bias, increase the generalizability, and improve the accuracy of study estimates [1]. However, achieving this goal is challenging, both due to the difficulty in contacting and then engaging eligible participants [2]. For example, UK Biobank, a population-based cohort study with stored biological samples from half a million participants aged 40 - 69 years in the United Kingdom, achieved an overall response rate of 5.5% [3], which was lower in men, younger people, and those living in more deprived areas [4]. The impact of nonresponse and nonrepresentativeness on the generalizability of disease prevalence and incidence rates in the UK Biobank has been widely debated [5,6].

It is important to address low or falling response rates to reduce the likelihood of systematic biases that may affect study estimates [7]. While weighting is commonly applied to correct for differential participation, it may fail to correct bias if the responders in a particular subgroup of the population are not representative of that subgroup as a whole. Furthermore, weighting to correct for observed biases worsens precision (reducing the effective sample size) [8].

Systematic reviews that have evaluated interventions to increase response rates in surveys have concluded that monetary incentives are more effective than nonmonetary incentives [9-11], although findings were inconsistent concerning web-based surveys in educational research [12]. Some studies have found incentives can increase response among under-represented sociodemographic groups, such as those with low incomes, those with low education, single parents, and minority ethnic groups, potentially reducing nonresponse bias [13], while others show mixed results [9].

Other strategies that have been shown to improve response rates in surveys have included the use of SMS text message reminders to enhance the contact method of letters and emails [14,15], using alternative motivational statements in invitation letters [16], and changing the font color of text [17]. In a United Kingdom–based study investigating the effects of augmenting the contact strategy of letters and emails with SMS text messages for a web questionnaire, the findings indicated that SMS text messages did not help to significantly increase response rates overall, although some subgroups benefited from them, such as younger panel members and those with an irregular response pattern [15].

The Real-time Assessment of Community Transmission-1 (REACT-1) study was one of the largest population surveillance studies in the world. Across 19 rounds between May 1, 2020 and March 31, 2022, it provided timely prevalence estimates of SARS-CoV-2, the virus that causes COVID-19, from random cross-sectional samples of the population in England [18,19].

Response rate varied between 11.7% and 30.5% and, like in many population surveys, varied across demographic groups [19]. For financial reasons, we could not issue more than 845,000 invitation letters by mail, so we could only achieve the minimum desired sample size adopted from round 12 (May 20 to June 7, 2021) of 100,000 by improving response [18]. The observed nonresponse biases meant REACT-1 was under-representing groups with lower vaccination rates and where COVID-19 prevalence was highest; thus, we were likely underestimating the true population prevalence despite our attempts to correct for such biases by use of weighting on known demographic variables [20,21]. Here we present results of experiments nested within the REACT-1 study to test the effectiveness of different strategies to increase response rates and participation of groups with a lower propensity to take part.

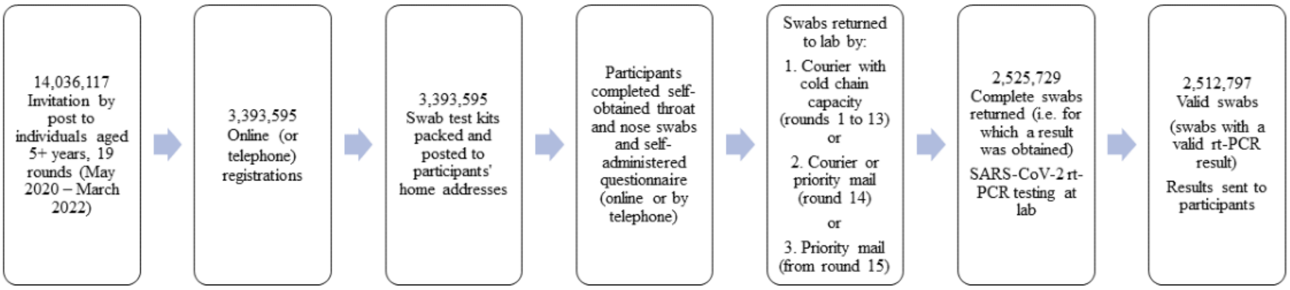
Methods

The REACT-1 Study

Methods for the study, including sample size calculations, are described in detail elsewhere [18,19]. In summary, at approximately monthly intervals, between 395,020 and 841,227 people were sent personalized invitations by mail to take part. For children (5-17 years old), the invitation was sent to or via the parents or guardians. Individuals aged 5 years and older were randomly sampled from the National Health Service (NHS) list of patients in England (with near-universal population coverage) across all 316 Lower Tier Local Authorities [18,19]. This list includes the name, address, date of birth, and sex of everyone registered with a general practitioner in England. Invitees who registered (most digitally, some by telephone) for the study received a kit by mail with instructions on how to take a throat and nose swab and send it for SARS-CoV-2 testing using reverse transcriptase polymerase chain reaction (rt-PCR). Swabs were transferred to laboratories for processing, initially being picked up by courier with cold chain capacity (rounds 1-13 and part of 14) or sent by priority mail (part of round 14 and subsequent rounds). Participants were also asked to complete a self-administered web-based or telephone questionnaire [18,19].

Over the 19 study rounds, we sent out 14,036,117 invitations, 3,393,595 registrations were made, and 2,525,729 completed swabs were returned (ie, for which a laboratory result was obtained) (Figure 1). Of these swabs, 2,512,797 (99.5% of completed swabs returned) were considered valid for analysis in REACT-1 (swabs with a valid rt-PCR result) [19]. A swab with a valid rt-PCR result was a swab for which a “cycle threshold” (Ct) value could be obtained. Therefore, not all swabs tested by the lab were considered valid. Overall, 12,932 (0.5% of completed swabs returned) were considered invalid and rejected. Reasons included inadequate sample volume, contamination during sample collection, inappropriate sample storage, or inappropriate sample transportation. All analyses in this paper are based on completed swabs returned (ie, for which a laboratory result was obtained, n=2,525,729), thus including swabs for which a Ct value could not be obtained but excluding swabs returned unused.

Figure 1. REACT-1 study process over 19 rounds of data collection: England, May 1, 2020 to March 31, 2022. Overall, across 19 rounds, we report the number of invitations sent, the number of participants registered, the number of swab test kits sent out, the number of completed swabs returned (ie, for which a laboratory result was obtained) and the number of valid swabs (swabs with a valid rt-PCR result). REACT-1: REal-time Assessment of Community Transmission-1; rt-PCR: reverse transcriptase Polymerase Chain Reaction; SARS-CoV-2: Severe Acute Respiratory Syndrome CoronaVirus 2.



All the experiments carried out to improve response rate were randomized trials, enabling an unbiased assessment of the impact of the changed survey procedure compared to a control group. Due to funding constraints, initial experiments focused on approaches which would not materially affect the survey budget, before turning to an experiment with monetary incentives.

Swab Reminder and Tailored Letter or SMS Experiments

In each round of REACT-1, those registering for a swab test were, where necessary, sent at least one reminder to complete the swab test and return it, to maximize the number of swabs returned. In round 3 (July 24 to August 11, 2020) we conducted an experiment to establish the optimal use of email and SMS text message swab return reminders, with participants randomly allocated to the experimental conditions (Table 1).

Table . Round 3 swab reminder experimental conditions, England, July 24 to August 11, 2020.

Condition	Reminder on day 4 after swab test kit received	Reminder on day 6 after swab test kit received	Reminder on day 8 after swab test kit received	Sample, n
Control group	Email (SMS if no email address)	SMS	None	11,194
Experimental group A	SMS	Email (SMS if no email address)	None	11,154
Experimental group B	Email (SMS if no email address)	SMS	Email (SMS if no email address)	96,337
Experimental group C	SMS	Email (SMS if no email address)	SMS	96,305

The tailored letters or SMS experiments are summarized in Table 2. Further details are available in Multimedia Appendix 1. The experiments tested whether it was possible to increase participation by different types of conditions (Textbox 1).

Table . Rounds 9, 11, and 12 registration invitation letter experimental conditions and rounds 10 and 12 SMS registration reminder experimental conditions, England, February 4 to June 7, 2021.

Age and letter or SMS type	Additional content for experiment (actual additional content used in bold text)	Sample, n
Round 9[a] (≥70 years)		
Standard invitation letter Adult	None	37,037
Experiment invitation letter A	"**It is still important to take part in this study if you have received a vaccination from COVID-19 or expect to be vaccinated in the near future**. Your participation will help DHSC assess the impact of the vaccines on COVID-19 infection rates." As well as a new sub-heading "**COVID-19 Testing Study: Take part to help measure COVID-19 infection rates among those aged 70 and over**."	37,037
Experiment invitation letter B	As per Experiment Letter A with additional line "**Older people are a vulnerable group, so we need your help to monitor prevalence**. It is still important to take part in this study if you have received a vaccination from COVID-19 or expect to be vaccinated in the near future."	37,036
Round 9 (5 - 12 years)		
Standard invitation letter Child (addressed to parent)	None	24,009
Experiment invitation letter C	**"We need to know how many children and young people have COVID-19, and how easily the new variant spreads amongst them."**	24,009
Experiment invitation letter D	As per Experiment Letter C with new sub-heading **"COVID-19 Testing Study: Take part to help measure how easily COVID-19 spreads among children and young people."**	24,008
Round 9 (all)		
Standard registration reminder letter	Blue text used	306,012
Experiment registration reminder letter E	Red text used	305,041
Round 11[b] (≥18 years)		
Standard invitation reminder letter	None	178,828
Experiment invitation reminder letter A	New content asking participants to take a test to help prevent the spread of COVID-19 and explaining that taking part would help the Government work out the best way to manage the pandemic. Also mentioned testing for new variants, that the study compared people who had been vaccinated with those who had not, and that taking part would help inform the vaccine strategy and help to avoid lockdowns.	178,809
Round 12[c] (all)		
Standard invitation final reminder letter	Double-sided	169,845
Shorter invitation final reminder letter	Single-sided	342,191
Round 10[d] (all)		
Standard SMS first reminder	Unchanged **"The study is closing soon, please register by 18 March if you want to take part."**	50,000

Age and letter or SMS type	Additional content for experiment (actual additional content used in bold text)	Sample, n
Experiment first SMS reminder	New SMS content “**Taking part will help inform decisions about the best time to lift restrictions**.”	430,283
Round 11 (all)		
Standard SMS second reminder	Unchanged “**The study is closing soon, please register by 3pm on 22 April if you want to take part**.”	127,028
Experiment second SMS reminder	New SMS content “**Taking part will help monitor infection rates and new variants of the virus**.”	127,028
Round 12 (all)		
Standard SMS first reminder	Unchanged “**Taking part will help inform decisions about the best time to lift restrictions**.”	321,042
Experiment first SMS reminder	New SMS Content **“Taking part will help monitor infection rates and new variants of the virus.”**	155,683
Standard SMS second reminder	Unchanged **“Taking part will help monitor infection rates and new variants of the virus.”**	272,836
Experiment second SMS reminder	New SMS Content **“Last chance to help monitor variants in your area.”**	136,026

[a] Round 9 (Feb 4-23, 2021)

[b] Round 11 (Apr 15 to May 3, 2021)

[c] Round 12 (May 20 to Jun 7, 2021)

[d] Round 10 (Mar 11-30, 2021)

Textbox 1. Types of conditions tested in the tailored letters or SMS experiments.

- Using additional content in the invitation letter, tailored for the oldest and youngest age groups.
- Using color, additional content, and varying the length of the reminder letter.
- Using additional content in the SMS reminder.

Incentives Experiment

In round 15 (October 19 to Nov 5, 2021), conditional incentives (£10 [US $12.5], £20 [US $25.0], or £30 [US $37.5] gift vouchers for returning a completed swab test) were tested in a randomized controlled trial for all age groups except 5- to 12-year-olds.

The process for obtaining consent in REACT-1 for children was undertaken differently based on participant age at the time of the invitation [18]. For 5- to 12-year-olds, the parent or guardian was contacted via letter and asked to consent on behalf of the child. Therefore, we did not include the 5- to 12-year-olds in the trial, as the sampled child would not be making the decision to take the swab test, and their parent would be incentivized, raising ethical and reputational concerns. For 13- to 17-year-olds, the parent or guardian received a letter addressed to them, asking them to pass on an enclosed invitation letter addressed to their sampled child if they agreed for their child to take part in the study. As such, children aged 13-17 years were able to decide whether to consent to the study and take the swab test. In addition, those aged 13 to 15 years were asked at registration to confirm the name of the parent or guardian who had given them permission to take part. This was not required for those aged 16-17 years, as in UK health research, the Health Research Authority states that young people over 16 are presumed capable of giving consent on their own behalf [22].

Participants were randomly allocated to experimental and control groups: (1) £10 (US $12.5) conditional incentive (n=10,900), (2) £20 (US $25.0) conditional incentive (n=10,900), (3) £30 (US $37.5) conditional incentive (only for 18- to 32-year-olds) (n=1750), and (4) control group (n=23,500). Further details of the sample size calculations are available in Multimedia Appendix 1.

The £30 (US $37.5) incentive was limited to the 18- to 32-year-olds because the response rate in REACT-1 was lowest among this age group. Also, there is evidence that incentives can be more effective among younger age groups [23]. Those in this age group were of particular interest as they were less likely to be vaccinated, had more social contact (and therefore were more likely to be at risk of infection), and had been particularly impacted by the pandemic (in terms of well-being, education, and employment) [21,24]. It was decided to test offering a larger (£30 [US $37.5]) incentive to this age group (and not the other age groups) to overcome their higher reluctance to take part and to better represent this group in the achieved sample. Based on the same rationale, we oversampled younger age groups to maximize the statistical power we had to detect an increase in the response rate due to the use of incentives among these groups.

The primary outcome was overall swab response rate, ie, the number of completed swabs returned (referred to as swabs returned forthwith) as a proportion of the number of invitations sent. For those invited, we knew age, sex at birth, and score from an area-level index of multiple deprivation, the Index of Multiple Deprivation (IMD) 2019 [25]. Participants were classified by quintiles of the deprivation score based on their residential postcode.

We used COVID-19 vaccination status (the proportion who had received at least one vaccine dose) as a proxy for attitudes to health behaviors and health care access, hypothesizing that REACT-1 responders would be more likely to be vaccinated than those who did not, indicating a responder bias. Thus, the difference in vaccination status at registration between the experimental and control groups was used as a crude indicator of how incentives might improve response rates in individuals less likely to participate in research, beyond sociodemographic characteristics. We also compared the COVID-19 vaccination status of those who returned a swab with the achieved population vaccination rate for that age group as a whole. To obtain information on dates of received COVID-19 vaccine doses, participant study data were linked to their NHS records from NHS Digital (now NHS England) on COVID-19 vaccination events [26] using their unique NHS number and other personal identifiers. This was only possible for study participants who had consented to data linkage. The source of vaccination data for the population vaccination rates was the NHS National Immunization Management System [27].

Ethical Considerations

The study was ethically approved by the South Central-Berkshire B Research Ethics Committee (IRAS ID: 283787). Participants provided informed consent when they registered for the study, and all data were handled securely in accordance with a detailed privacy notice. Collected data were deidentified; the data used in this study were anonymous and did not contain any personally identifiable information. Participants had the ability to opt out anytime during the research period. The study did not provide any specific compensation other than the monetary gift vouchers for returning a completed swab test as set out in the study's incentives experiment described above.

Statistical Analysis

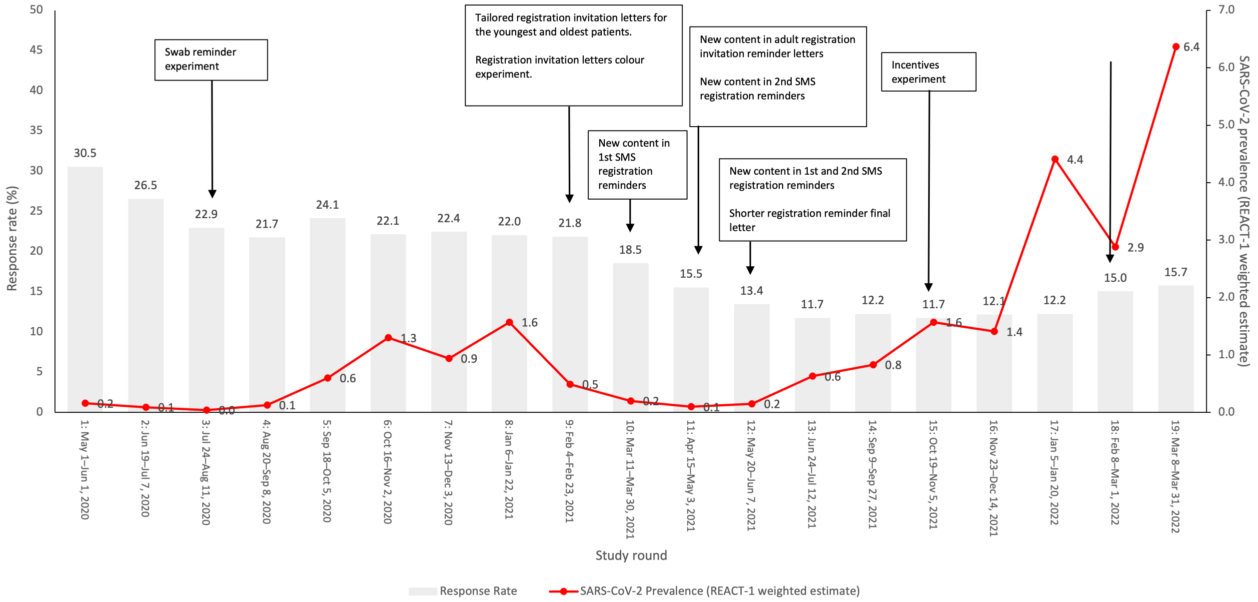
Data analysis was conducted using IBM SPSS Statistics (version 28). As the incentives experiment was skewed toward younger age groups, swab response rates for sex at birth and area deprivation (IMD) were calculated with age-standardized weighting using 2021-based population estimates for England [28]. The percentage point difference (95% CI) and independent 2-tailed *t*-tests were used to show the absolute difference in swab response rates between the experimental and control groups and were also used to show the absolute difference in vaccination rates at registration between the experimental and control groups. Using multivariable logistic regression, we tested the impact of each of the incentive conditions on swab response rate by age, sex at birth, and area deprivation (relative response rate [RRR] with 95% CI). The reference group was the no-incentive condition—eg, the response rates for females in the £10 (US $12.5), £20 (US $25.0), and £30 (US $37.5) incentives groups were compared to females in the no-incentive group (£0 [US $0.0]). We tested interaction terms for age, sex at birth, and area deprivation by incentive (incentive*age, incentive*sex at birth, and incentive*IMD), which can be interpreted as testing whether the estimated effects of incentives on swab response rates differ by each of these 3 covariates.

Results

Overview

Overall, 24.2% (3,393,595/14,036,117) of invitees registered for the study, and 74% (2,512,797/3,393,595) of those registered returned valid swabs, giving an overall response rate for the REACT-1 study (number of valid swabs/number of invitations) of 17.9% (2,512,797/14,036,117) [19]. Whilst the rate at which registered participants returned valid swabs remained relatively stable across rounds (range 67.2%-78.9%), response rates varied more widely, ranging from 11.7% in rounds 13 (98,233/841,227) (June 24 to July 12, 2021) and 15 (100,112/859,184) (October 19 to November 5, 2021) to 30.5% in round 1 (120,620/395,020) (May 1 to June 1, 2020, during the first lockdown in England) (Figure 2). The following groups were relatively underrepresented: younger people, men, ethnic minorities, and those living in the most deprived areas (comparing achieved sample profiles with population profiles) (Table S1 in Multimedia Appendix 2).

Figure 2. REACT-1 study timeline over 19 rounds of data collection showing response rates, SARS-CoV-2 prevalence (weighted), and timing of experiments to improve response. England, May 1, 2020 to March 31, 2022. REACT-1 Response Rate: number of valid swabs returned/number of invitations. We report weighted SARS-CoV-2 swab-positivity prevalence for individuals aged 5 years and older from all rounds of the REACT-1 study. REACT-1: REal-time Assessment of Community Transmission-1; SARS-CoV-2: Severe Acute Respiratory Syndrome CoronaVirus 2.



Swab Reminder and Tailored Letter or SMS Experiments

Table S2 in Multimedia Appendix 2 summarizes the results of the swab reminder and tailored letter or SMS experiments. Sending an additional reminder (email or SMS) to those who registered resulted in a small increase in response rate: those receiving a third swab reminder (experimental groups B and C) were more likely to return a completed swab compared to those receiving the standard Email-SMS approach (group B vs control: 73% vs 70.2%, percentage difference 2.8% [95% CI 1.9%-3.7%]; group C vs control 73.3% vs 70.2%, percentage difference 3.1% [95% CI 2.2%-4%]).

In round 9 (February 4-23, 2021), both experimental invitation letters A and B had a small but positive impact on response rate in participants aged ≥70 years of 0.9% (95% CI 0.2%-1.5%) and 1.2% (95% CI 0.6%-1.8%) percentage difference, respectively, compared to the standard invitation letter. For participants aged 5 - 12 years, experiment letter C generated a slightly higher response rate compared to the standard letter (16.6% vs 15.9%; percentage difference 0.7% (95% CI 0.1%-1.4%)). In round 11 (April 15 to May 3, 2021) and round 12 (May 20 to June 7, 2021), the experimental invitation reminder letters had a small positive impact on response rate compared to the standard letters: round 11 (new content), 5.6% vs 5.4%, percentage difference 0.2% (95% CI 0%-0.3%); round 12 (shorter), 2.3% vs 1.6%, percentage difference 0.8% (95% CI 0.7%-0.8%). We saw no effect on response rate for any of the other nonmonetary strategies (Table S2 in Multimedia Appendix 2).

Incentives Experiment

The conditional monetary incentives increased the response rate across all age groups but were particularly effective among the lowest responding groups, those aged 13 - 17 years and 18 - 22 years (Figure 3 and Tables S3 and S4 in Multimedia Appendix 2). Table 3 shows the RRR for each incentive level by age, sex at birth, and area deprivation. The higher the monetary value of the incentive, the higher the response rate. For example, in participants aged 18 - 22 years, the response rate in the control group was 3.4% (95% CI 2.9%-3.8%), increasing to 8.1% (95% CI 7.0%-9.2%), 11.9% (95% CI 10.6%-13.2%), and 18.2% (95% CI 15.4%-21.1%) with £10 (US $12.5), £20 (US $25.0), and £30 (US $37.5) incentives, respectively. The largest relative increase was with the £30 (US $37.5) incentive in 18 - to 22-year-olds (RRR 5.4 [95% CI 4.4-6.7]) (Table 3). All incentive conditions led to a greater increase in response rate in younger age groups. The £20 (US $25.0) incentive led to a greater increase in the more deprived areas, RRR 2.7 (95% CI 2.2-3.3) for the most deprived quintile and RRR 1.8 (95% CI 1.6-2.1) for the least deprived.

Figure 3. Swab response rates and 95% CIs for the intervention and control groups in the incentives experiment in round 15, England, October 19 to November 5, 2021. Note: Participants randomly allocated to experimental and control groups. (1) £10 (US $12.5) conditional incentive (n=10,900), (2) £20 (US $25.0) conditional incentive (n=10,900), (3) £30 (US $37.5) conditional incentive (only for 18- to 32-year-olds) (n=1750), and (4) control group (n=23,500).

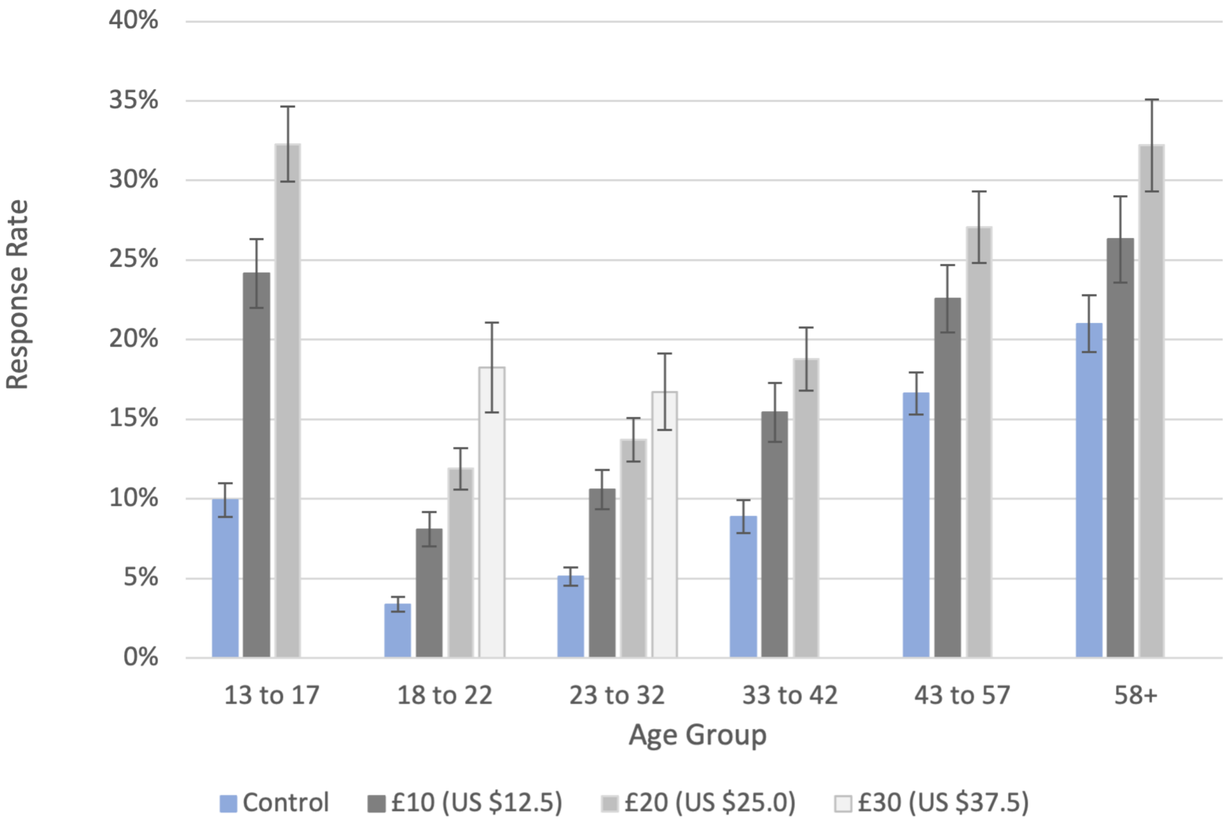


Table . Variation in relative response rates (RRR) and 95% CI for the interventions compared to the control group by age, sex at birth, and area deprivation (IMD[a]) in the incentives experiment in round 15, England, October 19 to November 5, 2021.

	£10 (US $12.5) RRR (95% CI)	£20 (US $25.0) RRR (95% CI)	£30 (US $37.5) RRR (95% CI)
Age[b] (years)			
13-17	2.4 (2.1-2.8)	3.3 (2.6-3.7)	N/A
18-22	2.4 (2.0-2.9)	3.5 (3.0-4.2)	5.4 (4.4-6.7)
23-32	2.1 (1.7-2.4)	2.7 (2.3-3.1)	3.3 (2.7-3.9)
33-42	1.7 (1.5-2.0)	2.1 (1.8-2.5)	—[c]
43-57	1.4 (1.2-1.5)	1.6 (1.5-1.8)	—
58+	1.3 (1.1-1.4)	1.5 (1.4-1.7)	—
P value for interaction between incentive and age	<.001	<.001	<.001
Sex at birth[bd]			
Male	1.4 (1.3-1.6)	1.8 (1.7-2.1)	3.7 (3.0-4.7)
Female	1.5 (1.4-1.6)	1.8 (1.6-2.0)	3.7 (3.1-4.4)
P value for interaction between incentive and sex at birth	.37	.96	.68
IMD[bd]			
1—most deprived	1.8 (1.5-2.3)	2.7 (2.2-3.3)	4.8 (3.3-7.0)
2	1.6 (1.3-1.9)	1.9 (1.6-2.3)	4.0 (2.9-5.5)
3	1.4 (1.2-1.6)	1.5 (1.3-1.8)	3.4 (2.5-4.8)
4	1.3 (1.1-1.5)	1.7 (1.5-1.9)	3.2 (2.4-4.3)
5—least deprived	1.5 (1.3-1.7)	1.8 (1.6-2.1)	3.5 (2.6-4.8)
P value for interaction between incentive and IMD	.38	.01	0.72

[a] IMD: Index of Multiple Deprivation

[b] Reference group, ie, the reference group for each row is the no incentive condition. For example, the RRR for female £10 (US $12.5), female £20 (US $25.0), and female £30 (US $37.5) is versus female £0 (US $0.0). *P* value for main effect of incentive on response rate for all row comparisons <.001.

[c] Not applicable

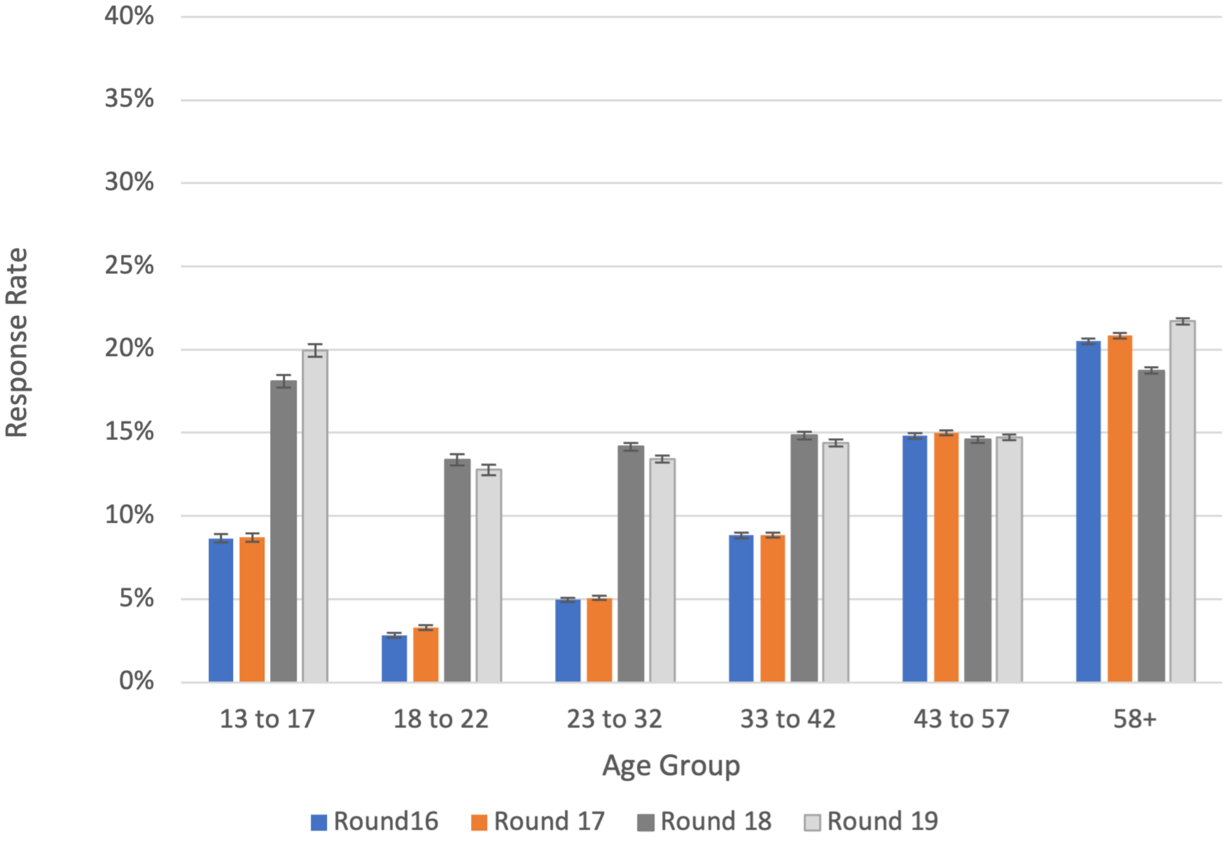
[d] Age-standardized weighting applied to calculate swab response rate with the control group totals used as the sample profiles.

Following the results of the selective use of incentives in round 15, they were introduced in rounds 18 (Feb 8-Mar 1, 2022) and 19 (Mar 8-Mar 31, 2022). For returning their completed test, those aged 13-17 and 35-44 years were offered a gift voucher worth £10 (US $12.5), while those aged 18-34 years were offered a voucher worth £20 (US $25.0). In these final 2 rounds, this had the effect of increasing the swab response rate in these groups and was associated with less variation in response rate by age (Figure 4), suggesting that the selective use of incentives reduced participation bias by age.

Table S5 in Multimedia Appendix 2 shows the effective sample sizes and sample efficiency for each round of REACT-1. The effective sample size measures the size of a (unweighted) simple random sample that would achieve the same precision (standard error) as the design used. The efficiency of a sample is given by the ratio of the effective sample size to the actual sample size. Rounds 18 and 19, where selective use of incentives was used, saw the fourth and second highest (respectively) effective sample sizes of any REACT-1 round, and the highest sample efficiency for any REACT-1 round.

Overall, vaccination rates were higher in REACT-1 participants than in the general population (Tables S6 and S7 in Multimedia Appendix 2). For example, by October 24, 2021, just over 3 quarters of 18- to 22-year-olds had received at least one vaccine dose nationally (Table S6 in Multimedia Appendix 2) [29], lower than the 84.0% (95% CI 78.1%-88.6%) in the round 15 (October 19-November 5, 2021) control group for that age (Table S7 in Multimedia Appendix 2). With the incentives that proportion declined to 82.1% (95% CI 76.0%-86.8%), 73.9% (95% CI 68.6%- 78.7%), and 75.9% (95% CI 67.9%-82.5%) for £10 (US $12.5), £20 (US $25.0), and £30 (US $37.5), respectively, suggesting that the selective use of incentives reduced participation bias in relation to vaccination status as a proxy for health behaviors.

Figure 4. Swab response rates and 95% CIs for round 16 (November 23 to December 14, 2021), round 17 (January 5-20, 2022), and rounds 18 (February 8 to March 1, 2022) and 19 (March 8-31, 2022) in which incentives were used selectively, England. Note: Incentive amounts used in rounds 18 and 19: £10 (US $12.5) for 13- to 17-year-olds and 35- to 44-year-olds, £20 (US $25.0) for 18- to 34-year-olds, and no incentives for other age groups.



Discussion

Principal Findings

In this large population-based study of the prevalence of SARS-CoV-2 infection in England, we tested several measures to increase response rates and reduce nonresponder bias. We found that changes to the wording of letters, timing, and numbers of reminders made only limited differences to response rates, with a maximum increase in response rate of 3.1 percentage points for additional swab reminders sent to people who had already registered for the study. Sending an additional reminder, regardless of its form (SMS or email), increased response. This is consistent with other studies in the literature [30]. These reminder strategies may have helped slow, but did not halt, the decline in response rates over time observed during REACT-1. Nonetheless, these findings informed the swab reminder strategy and invitation letter wording in later rounds. In contrast, the offer of a financial incentive conditional on the return of a completed swab made a more substantial difference of up to 22.3 percentage points and was particularly effective in those with a lower propensity to respond: younger age groups and those living in more deprived areas. Similarly, incentives increased the return of completed swabs by unvaccinated individuals so that COVID-19 vaccination rates were more in keeping with those in the general population at the time. Thus, the selective use of incentives may reduce nonresponder bias in relation to factors of interest in population health research beyond sociodemographic characteristics.

The selective use of incentives was subsequently adopted from round 18, making the achieved sample more representative by age, with a reduction in age-based variation in response rates. Previous research suggests that ethnic minorities [31], individuals living in more deprived areas [32], those in urban areas [33], and the youngest and oldest age groups [34,35] are the least likely to respond in general population surveys. Using incentives selectively allowed us, at modest cost, to increase recruitment among such groups and hence increase the effective sample size; thus, in round 18, the effective sample size was over 10,000 greater compared to round 17, even though we received circa 7000 fewer swabs. We were able to reduce the number of invitations sent out while achieving a similar number of completed swabs returned as in earlier rounds when response rates were higher.

Using incentives selectively has been tried in UK social surveys previously and is common practice in the United States, where studies show they are cost-effective, improve response, and reduce bias [11,13,23,36]. From an ethical perspective, in the selective use of (versus universal) incentives, it was important to consider not only issues of equity but also cost and the public interest in continuing to obtain high-quality data, covering all sectors of society, to monitor the spread of a serious disease. This needs to be balanced against the possible disappointment of some participants who learn others are being offered a (larger) incentive. These considerations might apply to many population-based surveys. We accept that the argument for using incentives selectively may have been more persuasive in the context of REACT-1, a study to measure the spread of SARS-CoV-2 during the pandemic, the policy responses to which had far-reaching consequences for the way of life of every person in England.

Both unconditional and conditional financial incentives have been shown to significantly increase response rates to both postal and web-based surveys [37,38]. Although unconditional incentives appear to have the largest effect, the conditional approach is more cost-effective [37,38]. Unconditional incentives have been used in social surveys in the United Kingdom, and in experiments in how to increase response rates [39,40]. Unconditional incentives were not an option for REACT-1 due to the constraints of the survey budget.

Limitations

In terms of limitations, it was not possible to ascertain the extent to which noncontact (ie, the intended recipient did not receive the invitation letter) accounted for nonresponse. Such situational factors, for example, not informing their general practitioner of a change in address or having moved with no forwarding address (shown to be greater for young adults and lower socioeconomic groups) [32] will not be affected by the experimental conditions; therefore, our estimates of effect are likely conservative, as invitations sent out do not necessarily mean that invitations were received. In addition, the unique circumstances of carrying out such assessments of response rates during a global pandemic may not “read across” to other less pressing issues.

Conclusions

We achieved small improvements in response rates by varying the number, order, and content of invitations and reminders but much larger effects were seen through the use of monetary incentives. Lessons learnt from the REACT-1 study may help inform the design and implementation of future population-based surveys where the intent is to obtain as representative a sample as possible and to reduce nonresponse bias at reasonable cost. The results suggest selectively using incentives with younger and more deprived individuals may be justifiable to achieve these ends.

Acknowledgments

This work was supported by the Department of Health and Social Care in England. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of this article. Infrastructure support for this research was provided by the National Institute for Health (NIHR) Imperial Biomedical Research Centre (BRC). PE is director of the Medical Research Council (MRC) Centre for Environment and Health (MR/L01341X/1 and MR/S019669/1). PE acknowledges support from Health Data Research UK, the NIHR Imperial Biomedical Research Centre; NIHR Health Protection Research Units in Chemical and Radiation Threats and Hazards, and Environmental Exposures and Health; the British Heart Foundation Centre for Research Excellence at Imperial College London (RE/18/4/34215); and the UK Dementia Research Institute at Imperial College London (MC_PC_17114). HW acknowledges support from an NIHR Senior Investigator Award, the Wellcome Trust (205456/Z/16/Z), the NIHR Imperial Biomedical Research Centre, the NIHR School of Public Health Research, and the NIHR Applied Research Collaboration North West London. CAD and SR acknowledge support from MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1), jointly funded by the UK MRC and the UK Foreign, Commonwealth & Development Office (FCDO), under the MRC/FCDO Concordat agreement. CAD acknowledges support from the NIHR Health Protection Research Unit (HPRU) in Emerging and Zoonotic Infections (NIHR200907), and the NIHR funded Vaccine Efficacy Evaluation for Priority Emerging Diseases (PR-OD-1017-20007). SR acknowledges support from the NIHR HPRU in Respiratory Infections, Wellcome Trust (200861/Z/16/Z, 200187/Z/15/Z), and Centres for Disease Control and Prevention (US, U01CK0005-01-02). MCH acknowledges support from Cancer Research UK, Population Research Committee Project grant “Mechanomics” (grant 22184). MCH acknowledges support from the H2020-EXPANSE (Horizon 2020 grant 874627) and H2020-LongITools (Horizon 2020 grant 874739).We thank key collaborators on this work: K Beaver, S Finlay, J Kennedy, D Peskett, G Welch, and K Ward at Ipsos. We also thank the REACT Public Advisory Panel who provided input which changed the wording of the invitation letters and reminders and provided advice on the design and equity concerns in relation to the incentives experiment.

Data Availability

The datasets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

NG, GP, SC, SR, HW, and PE designed the study. NG conducted the analyses. KP provided statistical support. CJA, NG, HW, and PE designed the analytical strategy and helped to interpret the findings. CJA, NG, and HW conducted the literature review. CJA, NG, HW and PE drafted the manuscript. HW, GSC, CJA, AD, CAD, MCH, SR, DA, WSB, and PE provided study oversight. AD and PE obtained funding. All authors have reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary methods.
[DOCX File, 19 KB - publichealth_v11i1e60022_app1.docx]

Multimedia Appendix 2
Supplementary tables.
[DOCX File, 55 KB - publichealth_v11i1e60022_app2.docx]

References

1. Groves RM, Peytcheva E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. Public Opin Q 2008;72(2):167-189. [doi: 10.1093/poq/nfn011]
2. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
3. Allen N, Sudlow C, Downey P, et al. UK Biobank: current status and what it means for epidemiology. Health Policy Technol 2012 Sep;1(3):123-126. [doi: 10.1016/j.hlpt.2012.07.003]
4. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Allen NE. OP41: the representativeness of the UK biobank cohort on a range of sociodemographic, physical, lifestyle and health-related characteristics. J Epidemiol Community Health 2016 Sep;70(Suppl 1):A26. [doi: 10.1136/jech-2016-208064.41]
5. Swanson JM. The UK Biobank and selection bias.. Lancet 2012 Jul;380(9837):110. [doi: 10.1016/S0140-6736(12)61179-9]
6. Keyes KM, Westreich D. UK Biobank, big data, and the consequences of non-representativeness. Lancet 2019 Mar 30;393(10178):1297. [doi: 10.1016/S0140-6736(18)33067-8] [Medline: 30938313]
7. Andrade C. Sample size and its importance in research. Indian J Psychol Med 2020;42(1):102-103. [doi: 10.4103/IJPSYM.IJPSYM_504_19] [Medline: 31997873]
8. Gelman A, Carlin J. Poststratification and weighting adjustments. In: Groves R, Dillman D, Eltinge J, Little R, editors. Survey Nonresponse: Wiley; 2000:289-302.
9. Pforr K, Blohm M, Blom AG, et al. Are Incentive effects on response rates and nonresponse bias in large-scale, face-to-face surveys generalizable to Germany? Evidence from ten experiments. Public Opin Q 2015;79(3):740-768. [doi: 10.1093/poq/nfv014]
10. Lange C, Finger JD, Allen J, et al. Implementation of the European health interview survey (EHIS) into the German health update (GEDA). Arch Public Health 2017;75:40. [doi: 10.1186/s13690-017-0208-6] [Medline: 28936356]
11. Singer E, Hoewyk JV, Gebler N, Raghunathan TE, Mcgonagle KA. The effect of incentives on response rates in interviewer-mediated surveys. J Off Stat 1999;15(2):217-230.
12. Wu MJ, Zhao K, Fils-Aime F. Response rates of online surveys in published research: a meta-analysis. Comput Hum Behav Rep 2022 Aug;7:100206. [doi: 10.1016/j.chbr.2022.100206]
13. Nicolaas G CE, Davies B. The use of incentives to recruit and retain hard-to-get populations in longitudinal studies. NatCen Soc Res 2019 May 31.
14. Mavletova A, Couper MP. Mobile web survey design: scrolling versus paging, SMS versus e-mail invitations. J Surv Stat Methodol 2014 Dec;2(4):498-518. [doi: 10.1093/jssam/smu015]
15. Cabrera-Álvarez P, Lynn P. Text messages to facilitate the transition to web-first sequential mixed-mode designs in longitudinal surveys. J Surv Stat Methodol 2024 Jun 1;12(3):651-673. [doi: 10.1093/jssam/smae003]
16. Lynn P. Targeted appeals for participation in letters to panel survey members. PUBOPQ 2016;80(3):771-782. [doi: 10.1093/poq/nfw024]
17. Taylor KSM, Counsell CE, Harris CE, Gordon JC, Fonseca SCC, Lee AJ. In a randomized study of envelope and ink color, colored ink was found to increase the response rate to a postal questionnaire. J Clin Epidemiol 2006 Dec;59(12):1326-1330. [doi: 10.1016/j.jclinepi.2006.03.011] [Medline: 17098576]
18. Riley S, Atchison C, Ashby D, et al. REal-time Assessment of Community Transmission (REACT) of SARS-CoV-2 virus: study protocol. Wellcome Open Res 2020;5(200):200. [doi: 10.12688/wellcomeopenres.16228.2] [Medline: 33997297]
19. Elliott P, Whitaker M, Tang D, et al. Design and implementation of a National SARS-CoV-2 Monitoring Program in England: REACT-1 study. Am J Public Health 2023 May;113(5):545-554. [doi: 10.2105/AJPH.2023.307230] [Medline: 36893367]
20. Ward H, Whitaker M, Flower B, et al. Population antibody responses following COVID-19 vaccination in 212,102 individuals. Nat Commun 2022;13(1):907. [doi: 10.1038/s41467-022-28527-x]
21. Eales O, Wang H, Haw D, et al. Trends in SARS-CoV-2 infection prevalence during England's roadmap out of lockdown, January to July 2021. PLoS Comput Biol 2022 Nov;18(11):e1010724. [doi: 10.1371/journal.pcbi.1010724] [Medline: 36417468]
22. Adolescent consent. NHS Health Research Authority. 2022. URL: https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/adolescent-consent/

#:~:text=Young%20people%20over%20the%20age%20of%2016%20are,consent%20from%20their%20parents%20provided%20they%20are%20competent [accessed 2024-12-31]

23. Wong E. Incentives in Longitudinal Studies, CLS Working Paper 2020/1: UCL Centre for Longitudinal Studies; 2020.
24. Ward H, Atchison C, Whitaker M, et al. SARS-CoV-2 antibody prevalence in England following the first peak of the pandemic. Nat Commun 2021 Feb 10;12(1):905. [doi: 10.1038/s41467-021-21237-w] [Medline: 33568663]
25. Comparisons of all-cause mortality between european countries and regions: january to june 2020. Office for National Statistics. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/comparisonsofallcausemortalitybetweeneuropeancountriesandregions/januarytojune2020 [accessed 2024-12-31]
26. COVID-19 vaccine data. NHS Digital. 2022. URL: https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/data-provision-notices-dpns/covid-19-vaccine-data [accessed 2022-12-31]
27. Tessier E, Edelstein M, Tsang C, et al. Monitoring the COVID-19 immunisation programme through a national immunisation management system - England’s experience. Int J Med Inform 2023 Feb;170:104974. [doi: 10.1016/j.ijmedinf.2022.104974] [Medline: 36577202]
28. Population and household estimates, england and wales: census 2021. Office for National Statistics. 2022. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationandhouseholdestimatesenglandandwales/census2021 [accessed 2024-12-31]
29. COVID-19 vaccinations archive. 2021. URL: https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-vaccinations/covid-19-vaccinations-archive [accessed 2024-12-31]
30. Sammut R, Griscti O, Norman IJ. Strategies to improve response rates to web surveys: a literature review. Int J Nurs Stud 2021 Nov;123:104058. [doi: 10.1016/j.ijnurstu.2021.104058] [Medline: 34454334]
31. D. Royal K. Surveying Black or African American populations: challenges and solutions in medicine and education. Educ Med J 2019 Jun;11(2):59-61. [doi: 10.21315/eimj2019.11.2.8]
32. Goodman A, Gatward R. Who are we missing? Area deprivation and survey participation. Eur J Epidemiol 2008;23(6):379-387. [doi: 10.1007/s10654-008-9248-0] [Medline: 18409005]
33. Inpatient experience survey 2018: technical report. Scottish Government. 2018 Aug 28. URL: https://www.gov.scot/publications/inpatient-experience-survey-2018-technical-report/ [accessed 2024-12-31]
34. Harrison S, Alderdice F, Henderson J, Redshaw M, Quigley MA. Trends in response rates and respondent characteristics in five National Maternity Surveys in England during 1995-2018. Arch Public Health 2020;78(46):46. [doi: 10.1186/s13690-020-00427-w] [Medline: 32509303]
35. Klein DJ, Elliott MN, Haviland AM, et al. Understanding nonresponse to the 2007 Medicare CAHPS survey. Gerontologist 2011 Dec;51(6):843-855. [doi: 10.1093/geront/gnr046] [Medline: 21700769]
36. Yu S, Alper HE, Nguyen AM, et al. The effectiveness of a monetary incentive offer on survey response rates and response completeness in a longitudinal study. BMC Med Res Methodol 2017 Apr 26;17(1):77. [doi: 10.1186/s12874-017-0353-1] [Medline: 28446131]
37. Young JM, O’Halloran A, McAulay C, et al. Unconditional and conditional incentives differentially improved general practitioners’ participation in an online survey: randomized controlled trial. J Clin Epidemiol 2015 Jun;68(6):693-697. [doi: 10.1016/j.jclinepi.2014.09.013] [Medline: 25450450]
38. Abdelazeem B, Hamdallah A, Rizk MA, et al. Does usage of monetary incentive impact the involvement in surveys? A systematic review and meta-analysis of 46 randomized controlled trials. PLoS One 2023;18(1):e0279128. [doi: 10.1371/journal.pone.0279128] [Medline: 36649255]
39. Office for National Statistics. Surveys and provision of monetary voucher for participation. 2018. URL: https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/surveysandprovisionofmonetaryvoucherforparticipation [accessed 2024-12-31]
40. Office for National Statistics. Crime survey for england and wales: technical note (Y/E march 2022). 2023. URL: https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/crimeandjustice/methodologies/crimeandjusticemethodology/csewtechnicalnoteyemarch2022.pdf [accessed 2024-12-31]

Abbreviations

Ct: cycle threshold
IMD: Index of Multiple Deprivation
NHS: National Health Service
REACT-1: Real-time Assessment of Community Transmission-1
RRR: relative response rate
rt-PCR: reverse transcriptase polymerase chain reaction

Edited by A Mavragani; submitted 29.04.24; peer-reviewed by D Falk, P Lynn; revised version received 12.11.24; accepted 13.11.24; published 09.01.25.

Please cite as:
Atchison CJ, Gilby N, Pantelidou G, Clemens S, Pickering K, Chadeau-Hyam M, Ashby D, Barclay WS, Cooke GS, Darzi A, Riley S, Donnelly CA, Ward H, Elliott P
Strategies to Increase Response Rate and Reduce Nonresponse Bias in Population Health Research: Analysis of a Series of Randomized Controlled Experiments during a Large COVID-19 Study
JMIR Public Health Surveill 2025;11:e60022
URL: https://publichealth.jmir.org/2025/1/e60022
doi:10.2196/60022

Contribution of Travelers to Plasmodium Vivax Malaria in South West Delhi, India: Cross-Sectional Survey

Deepali Savargaonkar[1], MBBS, MPH; Bina Srivastava[1], MSc; Chander Prakash Yadav[2], PhD; Mrigendra Pal Singh[3], PhD; Anup Anvikar[1,4], MBBS, MD; Amit Sharma[5], PhD; Himmat Singh[1,4*], PhD; Abhinav Sinha[1,4*], MBBS, MRES, MSc, MD, PhD

[1]ICMR-National Institute of Malaria Research, Sector 8, Dwarka, New Delhi, India
[2]ICMR-National Institute of Cancer Prevention and Research, Noida, India
[3]ICMR-National Institute for Research in Tribal Health, Jabalpur, India
[4]Academy of Scientific and Innovative Research, Ghaziabad, India
[5]International Centre for Genetic Engineering and Biotechnology, New Delhi, India
[*]these authors contributed equally

Corresponding Author:
Abhinav Sinha, MBBS, MRES, MSc, MD, PhD
ICMR-National Institute of Malaria Research, Sector 8, Dwarka, New Delhi, India

Abstract

Background: India is committed to malaria elimination by the year 2030. According to the classification of malaria endemicity, the National Capital Territory of Delhi falls under category 1, with an annual parasite incidence of <1, and was targeted for elimination by 2022. Among others, population movement across states is one of the key challenges for malaria control, as it can result in imported malaria, thus introducing local transmission in an area nearing elimination.

Objective: This descriptive study attempts to assess the contribution of such imported *Plasmodium vivax* cases to the malaria burden in South West Delhi (SWD).

Methods: A cross-sectional study was carried out at the fever clinic of the Indian Council of Medical Research-National Institute of Malaria Research in SWD from January 2017 to December 2019. Demographic and travel history data were recorded for all *P vivax* confirmed malaria cases diagnosed at the fever clinic. Vector and fever surveys along with reactive case detection were conducted in SWD and Bulandshahr district of Uttar Pradesh, 1 of the 6 geographical sources for a high number of imported malaria cases.

Results: A total of 355 *P vivax* malaria cases were reported during the study period. The proportion of imported cases was 63% (n=222). Of these, 96% (n=213) of cases were from Uttar Pradesh. The distribution of malaria cases revealed that imported cases were significantly associated with travel during the transmission season compared with that in the nontransmission season. Entomological and fever surveys and reactive case detection carried out in areas visited by imported *P vivax* malaria cases showed the presence of adults and larvae of *Anopheles* species and *P vivax* parasitemia.

Conclusions: Population movement is a key challenge for malaria elimination. Although additional *P vivax* infections and vector mosquitoes were detected at places visited by the imported malaria cases, the inability to detect the parasite in mosquitoes and the possibility of relapses associated with *P vivax* limit the significance of malaria associated with the travel. However, there remains a need to address migration malaria to prevent the introduction and re-establishment of malaria in areas with very low or 0 indigenous cases.

(JMIR Public Health Surveill 2025;11:e50058) doi:10.2196/50058

KEYWORDS

malaria; Plasmodium vivax; imported malaria; population movement; transmission; elimination; India

Introduction

Malaria is a parasitic disease transmitted by the bites of infected female *Anopheles* mosquitoes. The estimated number of malaria cases worldwide in 2021 was 247 million spread over 84 countries with 619,000 deaths. Although the World Health Organization (WHO) South East Asia region comprised ≈2% of the estimated global burden, India contributed to ≈79% of these cases, with a preponderance of *Plasmodium vivax* over *Plasmodium falciparum* [1].

Migration, both international and within-country, is a recognized social health determinant of multiple diseases across the globe, and malaria is no exception. Various factors put the migrating population at risk of contracting malaria, and these include their socioeconomic, living, working, and transit conditions [2]. The risk of malaria is also to the host communities that provide shelter to the migrants, particularly when the migration is along an epidemiological gradient from a high-burden to a low-burden or nonendemic area, putting malaria elimination efforts at risk [3-5]. Several pieces of evidence of this have been documented in the context of international migration [6-8] and within-country migration [9,10], including that in India [11].

India is committed to malaria elimination by 2030 and has formulated the National Framework for Malaria Elimination that classifies Indian states and union territories into 4 categories from 0 to 3, with category 3 being the highest-burden areas with an annual parasite incidence (API) of ≥1 per thousand persons at risk [12]. To achieve the elimination goal in the desired time frame, special focus needs to be given to the identified challenges by the National Center for Vector Borne Diseases Control (NCVBDC). Population size and migration are recognized as important challenges for malaria elimination, apart from asymptomatic parasite reservoirs, low-density infections, and parasite- and vector-resistance [13-15]. The movement of populations across and within Indian states is one of the key challenges in malaria control [12], and particularly, the migration of workers in large numbers from rural areas to cities has been reported in India [16]. Similarly, malaria cases among the mobile population contribute to a high percentage of total malaria cases in many countries [17] and have been a recognized challenge for malaria elimination [18]. Hence, it is important to carry out the screening and treatment for malaria in mobile populations for control and elimination of malaria in endemic areas, especially National Framework for Malaria Elimination category 1 areas (with API<1), and for prevention of the re-establishment of local transmission of malaria [12] in areas that have eliminated malaria (category 0 with 0 indigenous cases).

Despite being the capital of India that attracted >100,000 migrants each year since 2012 [19], no study on malaria in the migrant population in Delhi has yet been reported. Delhi falls under category 1 (API<1 per 1000 persons at risk), and its neighboring state, Uttar Pradesh (UP), is in category 2 with an API of less than 1 but with some districts having an API of 1 or more. With such a magnitude of migration, there remains a sustained threat of the introduction of *Plasmodium* infection by infected migrant populations from high-burden areas to the areas in categories 0 and 1 [16].

As Delhi and its neighboring states have *P vivax* as the predominant *Plasmodium* species causing malaria, this study was therefore carried out to assess the contribution of imported *P vivax* cases to the *P vivax* malaria burden in South West Delhi (SWD) by tracking the travel history of infected patients diagnosed at the fever clinic of Indian Council of Medical Research-National Institute of Malaria Research (ICMR-NIMR). Additionally, the study also aimed to identify mosquito breeding habitats and the presence and types of mosquito species in areas where these imported malaria cases resided/visited, and to detect additional *Plasmodium* infections/malaria cases through reactive case detection (RACD) and fever surveys in selected areas visited by these patients. These entomological and fever surveys and RACD were done to further identify any ongoing transmission in the areas visited by patients with imported malaria to support the hypothesis that these are indeed the cases likely to import malaria from areas with ongoing transmission to SWD.

Methods

Ethical Considerations

The study was approved by the Institutional Ethics Committee of ICMR-NIMR (ECR/NIMR/EC/2015/507 and ECR/NIMR/EC/2019/175). Informed consent was obtained from all human participants who were involved in the study. The participant identifier data were anonymized. No compensation was provided to the participants.

Study Sites, Samples, and Definitions

The study was initiated at the fever clinic at ICMR-NIMR, SWD, and later expanded to include the prominent catchment areas of the fever clinic of SWD and 6 villages of 1 selected district (Bulandshahr) of UP. Incoming febrile patients were screened for *Plasmodium* infection at the fever clinic of ICMR-NIMR from January 2017 to December 2019. Basic demographic data including age, gender, history of fever, and travel details (if any) were recorded using a paper-based structured questionnaire. The parasitological diagnosis at the clinic was performed by microscopy. Thick and thin blood smears were prepared, stained using the Jaswant Singh–Bhattacharji stain [20] and examined under 100× magnification. Microscopy was performed independently by 2 trained microscopists. In case of a discrepancy, a third trained microscopist examined the smears, and consensus observation by 2 trained microscopists was considered final. Those diagnosed with malaria were treated as per the national drug policy [21].

History of travel during the preceding 30 days of fever, including the places visited and duration of stay, was verbally elicited through a calendar-based recall method among all patients with malaria. Obtained travel history was used to classify the patients into 3 nonoverlapping categories: no travel history (patients neither traveled in or out of their residence in SWD), returning travelers (patients who were residing in SWD and traveled outside Delhi but came back), and incoming travelers (patients who were not residing in SWD but transiently traveled to SWD). Returning travelers were further classified based on the duration of stay outside Delhi into those returning to SWD within 7 days and those returning between 7 and 30 days. Similarly, incoming travelers were also classified into those who came to SWD within and beyond 30 days of fever.

Since the incubation period for malaria is 7 to 30 days [22-25], returning travelers with malaria who returned to SWD between 7 and 30 days of fever onset and incoming travelers who entered SWD within 30 days of fever onset were classified as imported malaria cases (acquired infection outside Delhi) for this study. The rest of the patients with malaria were considered to have indigenous infections.

Imported malaria cases were further investigated for the exact village and district of travel based on their recorded travel history. Thus, to further investigate whether the malaria cases were imported or indigenous, malaria vector (anopheline) surveys were carried out both in SWD and in the villages the imported malaria cases traveled to. These surveys were carried out by field workers adequately trained in entomology in catchment areas (Raj Nagar and Bagdola; every month from September 2018 to December 2019) of the fever clinic of ICMR-NIMR predominantly reporting malaria and also in villages (once in October 2019) of Bulandshahr district of UP state. Fever surveys and RACD were also carried out in these villages of Bulandshahr, UP. Bulandshahr district was preferred out of the 6 districts that showed significant sources of imported malaria in UP based on the burden of imported cases, logistic convenience, and operational feasibility. All 6 villages of Bulandhahr districts that had epidemiologically relevant travel connections with the imported malaria cases were surveyed.

The vector survey included the collection of mosquitoes (adults and larvae) from the houses of reported cases and their surrounding houses, species identification, and enumeration of mosquito breeding habitats. Resting adult mosquito collection was conducted in households of 6 villages of Bulandshahr district during early morning (6 AM to 8 AM) using hand aspirators. The larval collection was also done in each village from all water-bearing sites, that is, ponds, ditches, large cement tanks, drains, and seepages, in peridomestic and domestic areas of each village. In the Raj Nagar catchment locality of SWD, the houses were searched from 7 AM to 9 AM, and larval collection was also conducted simultaneously if searched houses and containers were found positive for larval presence from domestic and peridomestic water bodies and other sites including overhead tanks, large open water bodies, tires, coolers, bird pots, flowerpots, iron containers, and solid wastes in urban catchment areas of the clinic and from domestic and peridomestic containers in the houses of reported cases. The collected larvae were reared in an insectary separately up to their emergence to identify the mosquito species. Identification of species was done following the standard taxonomic key as described by Christophers [26]. Adult mosquito collection was also done using the hand catch method [27], and the collected *Anopheles* mosquitoes were screened for the presence of malaria parasites (*P falciparum*, *P vivax*, *Plasmodium malariae*, and *Plasmodium ovale*), through polymerase chain reaction (PCR) [28], in pools of mosquitoes made village-wise and species-wise. To estimate the critical density of malaria vectors, per man-hour density (PMHD) was calculated as the number of anophelines collected per hour by an insect collector using the formula:

$$\text{PMHD}=\text{No. of mosquitoes collected Time spent (in hours) x no. of insect collectors}$$

Fever surveys and RACD were carried out in October 2019 in 6 villages of the Bulandshahr (UP) with support from the local health personnel (Accredited Social Health Activists and Health Inspectors). Fever camps were organized at a central location in each village, and the local health personnel informed the villagers about the camp and motivated them to visit. Incoming febrile cases were screened for malaria by using a rapid diagnostic test (SD Bio Line Malaria Ag P.f / P.v, Standard Diagnostics, Inc, Republic of Korea) as per the manufacturer's instructions. All febrile cases were treated symptomatically, and patients with malaria were treated as per the national drug policy. RACD was done as described by the WHO [29], and blood smears were prepared from apparently healthy individuals in and around the household of the index cases. The smears were examined for the presence of malaria parasites at ICMR-NIMR, Delhi (as described previously), and the results were communicated to the concerned health personnel for further management.

Data Entry and Statistical Analysis

All the collected data were entered in a Microsoft Excel 2016 spreadsheet and presented as proportions (percentages), medians, and ranges, where appropriate. The strength of association was estimated using a chi-square test, and a *P* value of less than .05 was considered statistically significant.

Results

Overview

A total of 14,748 fever cases were screened for malaria by microscopy from January 2017 to December 2019. The 3-year period prevalence of malaria was 2.4% (364/14,748). Out of these 364 cases, 355 (97.5%) were *P vivax* mono-infections, 8 *P falciparum* mono-infections (2.1%), and 1 mixed infection of *P vivax* and *P falciparum*. There was male predominance among patients with fever (59%) as well as patients with *P vivax* malaria (71%). More than half of the patients with *P vivax* malaria were in the age group of 15 - 29 years (183/355, 52%) with a median age of 22 years. The parasite burden ranged from 63 to 206,187 parasites per microliter of blood.

Imported Malaria Burden

Out of the 250 *P vivax* cases with a travel history (250/355, 70%), 186 (74%) cases were returning travelers and the remaining 64 (26%) cases were incoming travelers (Figure 1). However, relevant travel history to be able to label them as imported cases was available from 63% (222/355) of patients. Out of these imported *P vivax* cases, 173 (78%) cases were returning and 49 (22%) cases were incoming travelers. Ninety-five percent (212/222) of the imported cases were from UP and 142 of them (142/212, 67%) had traveled to 1 of the 6 districts of UP viz Bareilly, Badaun, Aligarh, Hathras, Bulandshahr, and Mainpuri. The remaining 10 imported cases had traveled to Uttarakhand (3; Dehradun and Hardwar), Rajasthan (3; Bikaner, Bundi, and Sawai Madhopur), Madhya Pradesh (2; Gwalior), Haryana (1; Gurugram), and Punjab (1; Sri Muktsar Sahib), as shown in Figure 2. Out of 9 patients with *P falciparum* malaria (including 1 mixed infection), 8 had a travel history.

Figure 1. Travel history among 355 patients with *Plasmodium vivax* malaria. Based on the epidemiologically relevant travel history, cases associated with travel were categorized into "imported" (222; shown in red font) and indigenous (133; gold colored font) cases. The geographical distribution (states) of imported cases is also mentioned.

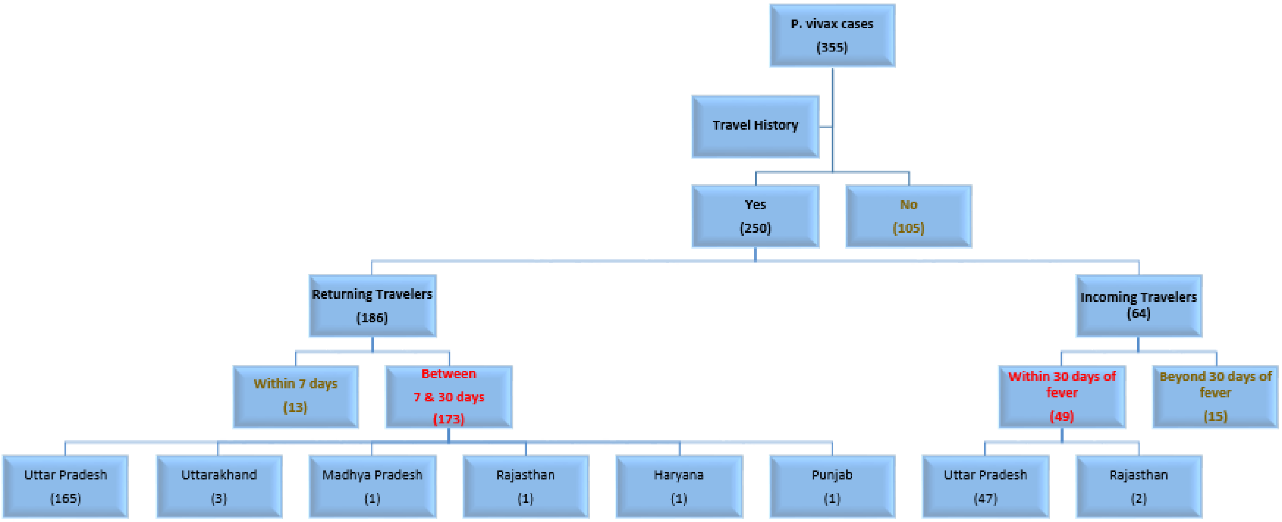
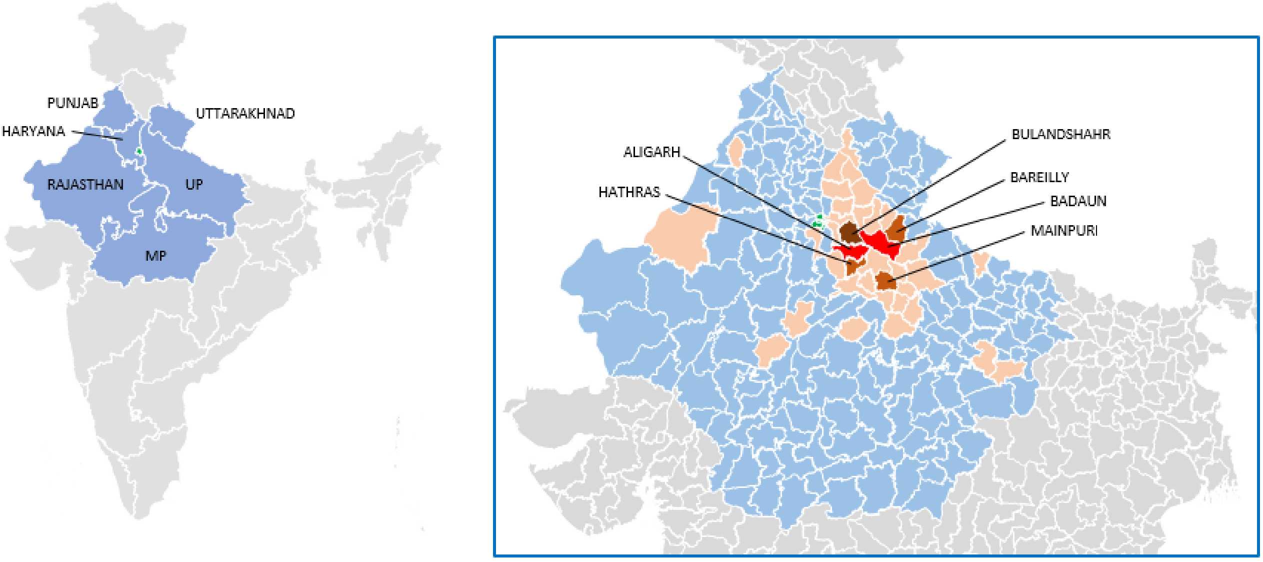


Figure 2. Areas traveled to by *Plasmodium vivax* malaria imported cases. The areas (states and union territories of India) are shown in blue (Punjab, Chandigarh, Haryana, Uttarakhand, UP, MP, and Rajasthan) whereas Delhi (study site) is shown in green. The zoomed-in image of the map in the inset shows further administrative breakdown of these 7 states and union territories (as districts) in blue. The districts within these 7 states and union territories, which are associated with the travel history of imported cases, are colored based on the number of imported cases contributed by each district: light orange (1 - 5 cases); dark orange (5 - 15 cases); darker orange (15 - 25 cases), and red (>25 cases). It is evident that UP has 3 dark orange districts: Hathras (10 cases), Mainpuri (12 cases), and Bareilly (15 cases); 1 darker orange district: Bulandshahr (25 cases); and 2 red districts: Aligarh (37 cases) and Badaun (47 cases). MP: Madhya Pradesh; UP: Uttar Pradesh.

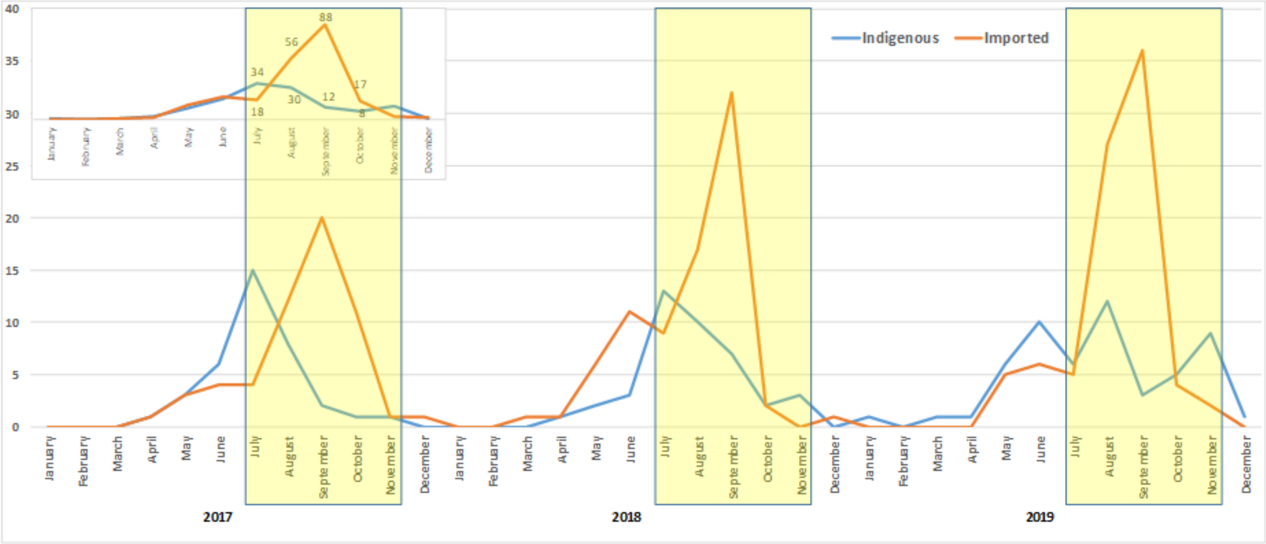


The reasons for travel in *P vivax* malaria cases included visiting their native residence in various states (mainly for returning travelers); education or employment (for incoming travelers); and visiting relatives, family, and friends during festivals (for returning and incoming travelers) as many Indian and regional festivals temporally coincide with the malaria transmission season. The minimum length of stay outside Delhi among travelers was 1 day while the maximum stay was of 111 days. Travel history was reported by the family members of the patients with malaria as well. There were 24 families with at least 2 members probably acquiring malaria after traveling.

A majority of the *P vivax* cases (279/355, 79%) were detected during the transmission season, that is, July to November. The nontransmission season (December to June) contributed to the remaining 21% (76/355) of cases.

During the transmission season (July to November; 2017 - 2019), the proportion of imported *P vivax* malaria cases diagnosed at the fever clinic was higher (65%) than that in the nontransmission season (53%; December to June), as shown in Figure 3, and the difference was statistically significant (χ^2_1=4.04; *P*=.031) at a 95% confidence level.

Figure 3. *Plasmodium vivax* malaria cases reported in Delhi (2017 - 2019). The figure shows the number of *P vivax* cases, indigenous (blue) and imported (imported), as reported each month and each year during the study period (January 2017 to December 2019) at the Indian Council of Medical Research-National Institute of Malaria Research fever clinic. The cumulative month-wise data from 2017 to 2019 are shown in the inset. The yellow semitransparent rectangles show the transmission season.



Entomological Surveys

Adult mosquito collections from the draining and catchment areas (Bagdola and Raj Nagar areas) of the fever clinic in SWD showed 4 mosquito species, with a low prevalence of *Anopheles* mosquitoes during the survey (September 2018 to December 2019). Out of 573 adult mosquitoes collected from 3395 houses, only 7 (1%) were *Anopheles stephensi* (Table 1), and all of them were found during the malaria transmission season (September). No other species of *Anopheles* were found. The majority of mosquitoes (443/573, 78%) were *Aedes aegypti* in the households.

Table . Month-wise adult mosquito collection in Bagdola and Raj Nagar, South West Delhi, between September 2018 and December 2019. Transmission season is highlighted in gold. The proportion of mosquito species identified out of the total is shown as a percentage (in parentheses).

		Houses checked, n	Mosquito species identified				
			Aedes aegypti, n (%)	*Aedes albopictus*, n (%)	*Anopheles stephensi*, n (%)	*Culex quinquefasciatus*, n (%)	Total, n
Month and year							
	September 2018[a]	316	28 (37)	0 (0)	1 (1)	46 (62)	75
	October 2018[a]	97	14 (70)	0 (0)	0 (0)	6 (30)	20
	November 2018[a]	402	63 (73)	0 (0)	0 (0)	23 (27)	86
	December 2018	220	16 (84)	1 (5)	0 (0)	2 (11)	19
	January 2019	112	1 (100)	0 (0)	0 (0)	0 (0)	1
	February 2019	70	0 (0)	0 (0)	0 (0)	0 (0)	0
	March 2019	179	2 (40)	0 (0)	0 (0)	3 (60)	5
	April 2019	60	13 (100)	0 (0)	0 (0)	0 (0)	13
	May 2019	344	10 (63)	0 (0)	0 (0)	6 (37)	16
	June 2019	367	89 (87)	0 (0)	0 (0)	13 (13)	102
	July 2019[a]	352	53 (75)	0 (0)	0 (0)	18 (25)	71
	August 2019[a]	270	61 (91)	0 (0)	6 (9)	0 (0)	67
	September 2019[a]	301	44 (90)	5 (10)	0 (0)	0 (0)	49
	October 2019[a]	211	40 (100)	0 (0)	0 (0)	0 (0)	40
	November 2019[a]	33	9 (100)	0 (0)	0 (0)	0 (0)	9
	December 2019	61	0 (0)	0 (0)	0 (0)	0 (0)	0
Total		3395	443 (78)	6 (1)	7 (1)	117 (20)	573

[a] Transmission season.

On the other hand, a total of 9 species of mosquitoes (5 anopheline, 2 *Aedes*, 1 *Culex*, and 1 *Armigeres* species) were collected from 6 villages in the Bulandshahr district (Table 2) of UP. Out of the 416 adult mosquitoes collected, 126 (30%) were *Anopheles* with *Anopheles subpictus* (21%) and *Anopheles culicifacies* (5%) dominating the anopheline burden. *Culex quinquefasciatus* was the major mosquito species found in rural Bulandshahr (235/416, 57%). Among the anophelines, *A subpictus* (PMHD 16.61) was the most abundant species followed by the main rural vector *A culicifacies* (PMHD 4.25) and *A stephensi* (PMHD 1.35).

Table . Adult mosquito collection and species identification in the surveyed 6 villages of Bulandshahr district of Uttar Pradesh during October 2019. The proportion of different mosquito species identified out of the total is shown as a percentage (in parentheses).

		Mosquito species identified									
		Aedes aegypti, n (%)	*Aedes albopictus*, n (%)	*Anopheles stephensi*, n (%)	*Anopheles annularis*, n (%)	*Anopheles culicifacies*, n (%)	*Anopheles nigerrimus*, n (%)	*Anopheles subpictus*, n (%)	*Culex quinquefasciatus*, n (%)	*Armigeris subalbatus*, n (%)	Total, n
Villages											
	Adauli	1 (0.7)	0 (0)	0 (0)	1 (0.7)	9 (7)	4 (3)	30 (22)	86 (63)	6 (4)	137
	Lakhaoti	0 (0)	1 (2)	0 (0)	4 (6)	5 (8)	0 (0)	12 (18)	26 (40)	18 (27)	66
	Shikarpur (Kot Kalan 1)	4 (6)	2 (3)	5 (8)	0 (0)	4 (6)	0 (0)	6 (9)	45 (68)	0 (0)	66
	Kutubpur	0 (0)	0 (0)	0 (0)	1 (4)	0 (0)	0 (0)	5 (19)	20 (77)	0 (0)	26
	Mustafabad Daduwa	0 (0)	0 (0)	2 (5)	0 (0)	2 (5)	1 (2)	0 (0)	18 (41)	21 (47)	44
	Dinoul	0 (0)	0 (0)	0 (0)	0 (0)	2 (3)	0 (0)	33 (43)	40 (52)	2 (3)	77
Total		5 (1)	3 (0.7)	7 (2)	6 (1)	22 (5)	5 (1)	86 (21)	235 (57)	47 (11)	416

A total of 24 collected anopheline mosquito pools (22 mosquitoes in 6 pools of *A culicifacies*, 7 mosquitoes in 2 pools of *A stephensi*, 3 mosquitoes in 3 pools of *Anopheles annularis*, 86 mosquitoes in 11 pools of *A subpictus*, and 5 mosquitoes in 2 pools of *Anopheles nigerrimus*) were tested by PCR for the presence of malaria parasites; however, none of the pools was found positive for the presence of malaria parasites.

The vector survey to identify mosquito breeding habitats in Raj Nagar and Bagdola catchment localities revealed that out of 14,333 containers (including large containers, cemented tanks, and underground tanks) checked, *Anopheles* breeding was found only in 8 containers that included coolers, overhead tanks, cement tanks, and iron containers. There were no large water bodies in the surrounding area of the survey.

A mosquito breeding habitat survey in the 6 villages of Bulandshahr found that out of the 203 water-holding containers and water bodies, 51 (25%) had *Anopheles* breeding. Major breeding sites included drains (4/7, 57%), canals (1/2, 50%), ponds (3/6, 50%), and pits (9/20, 45%). Other sites where breeding was found included domestic and peridomestic water bodies (26/123, 21%), cemented ground tanks (cattle tanks; 5/26, 19%), and rice fields (3/19, 16%). *A culicifacies* were found mostly in canal banks and village ponds whereas *C quinquefasciatus* in sewages.

Fever Surveys

Camp-based fever surveys in the 6 villages identified 86 persons with fever, with only 1 person testing positive for *P vivax* by the rapid diagnostic tests. RACD from 22 asymptomatic persons around the households of the index cases revealed 5 additional cases (5/22, 23%) of *P vivax* by microscopy.

Discussion

Principal Findings

Between January 2017 and December 2019, 355 monoinfected *P vivax* cases were reported, and out of them, 63% (n=223) could be categorized as possible imported malaria cases based on relevant travel history, thus forming a major burden of reported malaria cases in SWD. The study also detected 5 additional *P vivax* cases through RACD done in villages visited by the imported cases and identified malaria vectors of anopheline species and their breeding habitats in such areas.

The distribution of malaria cases reported in the fever clinic at ICMR-NIMR revealed that the malaria cases were more likely to be imported than indigenous and occur in transmission season. The period July to November is considered to be the malaria transmission season in Delhi, while December to June is considered a nontransmission season [30].

Although 67% of the *P vivax* cases were imported, being associated with relevant travel history, the remaining 37% of indigenous cases could be associated with possible local transmission of *P vivax* in SWD, as suggested by the presence of anophelines in Delhi (this study) and the reported presence of malaria vectors in Delhi [31,32]. *P vivax* malaria cases during the nontransmission season or in nontravelers might also be recurrences or relapses due to the activation of hypnozoites from the liver [30].

Recent travel within the country is associated with malaria in various studies [22,23]. This study showed that the proportion of males was more than females among imported as well as indigenous malaria cases and a similar trend was seen in patients with fever as well. In similar studies, men traveling away from home in the last 30 days were reported to be strongly associated with malaria in Ethiopia [24,25].

Although Delhi shares its borders with the state of UP in the east and the state of Haryana in the remaining directions, we observed that ≈96% of the imported cases were from UP. Data highlights of the census of India in 2001 and 2011 show that Delhi receives a higher number of migrants (≈50% of the total in-migrants) from UP versus that from Haryana (≈10% of in-migrants) [33]. With >20-fold higher malaria burden in UP (than in Haryana), the findings of >95% of cases being imported from UP are explainable [12]. Reasons for migration to Delhi are cited to be due to employment, business, education, marriage, etc [33,34]. The reasons for travel reported during this study were festivals, farming, and visits to relatives. Those visiting friends and relatives in malaria-endemic areas have been reported to be at high risk of contracting malaria [35,36].

Many districts in UP contributed to the imported *P vivax* cases in SWD (Figure 2); however, 6 UP districts contributed 10 or more cases: Hathras (10 cases), Mainpuri (12 cases), Bareilly (15 cases), Bulandshahr (25 cases), Aligarh (37 cases), and Badaun (47 cases). Further investigations (vector and fever surveys) were carried out in 6 villages of Bulandshahr district only due to reasons explained earlier. Bulandshahr district of UP, located southeast of Delhi, is situated between the Ganga and Jamuna rivers, which are the major rivers in North India. The soil is very fertile where mainly sugarcane, and wheat are grown. Irrigation facilities are also well-developed and this area is canal-irrigated as well [37] which makes the area highly mosquitogenic.

During vector surveillance in 6 villages of Bulandshahr, 25% of the water bodies were positive for anopheline larval breeding, and 5 species of adult *Anopheles* mosquito were found during adult mosquito collections. Unlike Bulandshahr, where almost every village had ponds, canals, and ample water in surrounding areas providing sufficient opportunities for the breeding of anophelines, Delhi is highly urbanized and lands are not available for ponds and crop fields. In comparison to Bulandshahr, the catchment areas of the fever clinic (Raj Nagar and Bagdola localities) of SWD had a very low prevalence of *Anopheles*. Only 1 species, that is, *A stephensi* was present in these localities in contrast to Bulandshahr where 5 species of *Anopheles* were collected out of which 2 were major malaria vectors, that is, *A stephensi* and *A culicifacies*. Larval surveys suggested that urban and rural areas have different breeding habitats. In villages, natural water bodies like ponds, canals, pits, and crop fields were more prominent and harbored more breeding than the peridomestic and domesticated containers in contrast to the urban areas where natural breeding sites are limited and were confined to peridomestic and domestic containers only. Mosquito species like *A stephensi* and *A aegypti* are adapted to breed in such urban areas whereas *A culicifacies* mostly breed in outdoor natural water habitats like canal banks, village ponds, etc and *C quinquefasciatus* is found in sewage water.

Among malaria vectors, *A culicifacies* was found to be the dominating mosquito species along with an efficient malaria-transmitting vector, *A stephensi*. However, the month of the survey (October) had low vector density, which may be due to the low ambient temperature (20 - 25 °C) during the survey period. Further, the mosquitoes that were collected from the villages of Bulandshahr district did not show parasite positivity by PCR. This may be due to multiple factors, including the very short period of vector survey (30 d), a limited number of vectors collected toward the end of transmission (October), and the difference in time of mosquito collection and case reporting in the clinic, as the vector survey was carried out as a response for tracking ongoing transmission in areas previously visited by imported malaria cases.

The 6 districts of UP that contributed most to imported cases in SWD had API (2018) of 0.06 (Bulandshahr and Mainpuri), 0.1 (Hathras and Aligarh), 5.5 (Badaun), and 7.3 (Bareilly), whereas the API of Delhi during this period was 0.02 [38]. A survey was therefore, carried out in Bulandshahr wherein 1 out of 86 febrile cases (fever survey) and 5 out of 22 afebrile persons (RACD) were identified with *P vivax* infections which signifies that further studies are needed to assess the extent of asymptomatic *Plasmodium* infection and its role in transmission in such areas.

The prevalence of malaria was found lower in the camp-based fever surveys compared with the prevalence reported from the fever clinic in SWD. This may be because the camp-based fever surveys were carried out during October, which marks the end of the transmission season and therefore may have had a lower number of cases. Further, the catchment area of the camp included a village whereas the fever clinic at SWD has a much larger and densely populated catchment area.

The regions nearing malaria elimination tend to have a heterogeneous endemicity, with foci of high burden and areas with no endogenous malaria transmission. For eliminable diseases such as malaria, within-country migration is a recognized but understudied challenge in such geographically heterogeneous transmission to sustain zero-burden and prevent reintroduction and re-establishment of transmission [23,39]. Such regions often lack a robust surveillance system to deal with imported cases besides treating them, and there appears to be a lack of documented cross-reporting and targeted intervention in the foci where the infections probably originated.

This study is therefore important, as it attempted to comprehensively investigate imported malaria cases, and may be adopted and locally adapted as an implementation model in similar areas with no or few locally acquired malaria cases.

Limitations

There is an obvious limitation of this study that a limited geographical area for fever and vector surveillance was selected. Nevertheless, the study shows the presence of malaria transmission in areas where patients with malaria reporting to the fever clinic had traveled. The study was also limited by the possibility of recall bias of study participants correctly recalling the exact dates of travel for both imported and indigenous patients with malaria. The investigators, however, tried to extract the near-exact dates by relating travel to the locally relevant cultural events, festivals, and other contextual events. Misclassification bias (incorrect classification of imported malaria) could have stemmed from the possibilities of recurrences and relapses of *P vivax* infections acquired before the study period. The study did not use available molecular methods to differentiate recurrent versus new infection and therefore could not account for *P vivax* relapses. However, the possibility that only up to 40% of *P vivax* infections in the study area (that too in the nontransmission season) could be due to possible relapses [30], nonavailability of molecular methods to confidently differentiate recurring infections from new infections, and random possibilities of recurrence in both the imported and indigenous patients with malaria may have compensated for this limitation.

Last, only 2 methods for mosquito collection were used. The hand catch method using an aspirator was the only method adopted for the estimation of vector density. For larval collection, dips were taken from water bodies for assessment of breeding. No other method was adopted for mosquito collection, and this might have underestimated the frequency of vectors and their possible infection with *Plasmodium*, because the PCR results did not show any vector positivity.

Conclusions

A significant burden (63%) of *P vivax* malaria reported in SWD was found to be imported from UP. Malaria transmission possibilities (multiple breeding sites suggesting stable breeding ground of anophelines) were higher in Bulandshahr than in SWD. Indigenous cases in SWD are also a concern, as adult vectors were also found in the area. Despite the detection of additional *P vivax* cases following RACD in Bulandshahr and vector-breeding sites being identified, the conclusion that the imported cases really acquired infections from the surveyed areas in Bulandshahr is limited by the correct recall of travel or fever dates and the possibility of relapses due to *P vivax*.

Way Forward

The study reiterates that population movement is a key challenge for malaria elimination, particularly in areas with very low or 0 indigenous malaria cases, and investigations of the potential role of travelers in introducing malaria and its further spread are definitely needed. Since the epidemiology of migration malaria is contextual, appropriate tailor-made measures are needed, both at the sites where imported cases are detected and in areas where these infections might have been acquired. In addition, effective “information, education, and communication” activities to educate travelers regarding the potential risks of travel-associated malaria and its prevention should be undertaken.

Acknowledgments

The authors acknowledge the support provided by the state programme officer of Uttar Pradesh and district malaria officer of Bulandshahr. The authors are thankful to the technical staff of fever clinic and vector biology division of ICMR-NIMR for their field support. The study did not receive any specific funding from a dedicated grant and was funded by Indian Council of Medical Research-National Institute of Malaria Research (ICMR-NIMR), New Delhi, out of its intramural funds.

Disclaimer

The views expressed in the submitted article are our own and not an official position of the institution or funder. No generative artificial intelligence was used in any portion of the manuscript writing.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

DS, HS, and A Sinha conceived the idea, carried out clinical work and fieldwork, and wrote the first draft. BS carried out laboratory work. CPY and MPS carried out analysis. AA and A Sinha analyzed and interpreted the data and reviewed the manuscript. A Sharma reviewed the manuscript. HS and A Sinha contributed equally as co-corresponding authors.

Conflicts of Interest

None declared.

References

1. World malaria report 2022. : World Health Organization; 2022 URL: https://www.who.int/publications/i/item/9789240064898 [accessed 2023-06-01]
2. Chapter 2 – migration and migrants: a global overview. World Migration Report 2018. : International Organization for Migration (IOM); 2017 URL: https://onlinelibrary.wiley.com/doi/10.1002/wom3.2 [accessed 2024-04-12] [doi: 10.1002/wom3.2]
3. Lu G, Zhang D, Chen J, et al. Predicting the risk of malaria re-introduction in countries certified malaria-free: a systematic review. Malar J 2023 Jun 6;22(1):175. [doi: 10.1186/s12936-023-04604-4] [Medline: 37280626]
4. Requena-Méndez A, Cattaneo P, Bogale RT, et al. Malaria parasite prevalence among migrants: a systematic review and meta-analysis. Clin Microbiol Infect 2023 Dec;29(12):1528-1537. [doi: 10.1016/j.cmi.2023.09.010] [Medline: 37739263]
5. Cohen JM, Smith DL, Cotter C, et al. Malaria resurgence: a systematic review and assessment of its causes. Malar J 2012 Apr 24;11:122. [doi: 10.1186/1475-2875-11-122] [Medline: 22531245]

6. Pousibet-Puerto J, Lozano-Serrano AB, Soriano-Pérez MJ, et al. Migration-associated malaria from Africa in southern Spain. Parasit Vectors 2021 May 7;14(1):240. [doi: 10.1186/s13071-021-04727-0] [Medline: 33962647]
7. Rodrigues PT, Valdivia HO, de Oliveira TC, et al. Human migration and the spread of malaria parasites to the new world. Sci Rep 2018 Jan 31;8(1):1993. [doi: 10.1038/s41598-018-19554-0] [Medline: 29386521]
8. Wångdahl A, Wyss K, Saduddin D, et al. Severity of Plasmodium falciparum and non-falciparum malaria in travelers and migrants: a nationwide observational study over 2 decades in Sweden. J Infect Dis 2019 Sep 13;220(8):1335-1345. [doi: 10.1093/infdis/jiz292] [Medline: 31175365]
9. Sorichetta A, Bird TJ, Ruktanonchai NW, et al. Mapping internal connectivity through human migration in malaria endemic countries. Sci Data 2016 Aug 16;3:160066. [doi: 10.1038/sdata.2016.66] [Medline: 27529469]
10. Prothero RM. Migration and malaria risk. Health Risk Soc 2001 Mar 1;3(1):19-38. [doi: 10.1080/13698570123413]
11. Tyagi BK. Trans-border migration and malaria in desert populations. In: Desert Malaria: Springer; 2023. [doi: 10.1007/978-981-19-7693-3_21]
12. Country Office for India, World Health Organization. WHO Country Office for India. 2016. URL: https://iris.who.int/handle/10665/246096 [accessed 2023-05-31]
13. Sharma S, Verma R, Yadav B, Kumar A, Rahi M, Sharma A. What India can learn from globally successful malaria elimination programmes. BMJ Glob Health 2022 Jun;7(6):e008431. [doi: 10.1136/bmjgh-2022-008431] [Medline: 35760440]
14. Rahi M, Sharma A. Malaria control initiatives that have the potential to be gamechangers in India's quest for malaria elimination. Lancet Reg Health Southeast Asia 2022 Jul;2:100009. [doi: 10.1016/j.lansea.2022.04.005] [Medline: 37383297]
15. Khan N, Awasthi G, Das A. How can the complex epidemiology of malaria in India impact its elimination? Trends Parasitol 2023 Jun;39(6):432-444. [doi: 10.1016/j.pt.2023.03.006] [Medline: 37031071]
16. Kumar A. Some considerable issues concerning malaria elimination in India. J Vector Borne Dis 2019;56(1):25-31. [doi: 10.4103/0972-9062.257770] [Medline: 31070162]
17. Malaria in migrants and mobile populations. World Health Organization. URL: https://www.who.int/malaria/areas/high_risk_groups/migrants_mobile_populations/en [accessed 2023-05-31]
18. Karim MA, Kabir MM, Siddiqui MA, Laskar MSI, Saha A, Naher S. Epidemiology of imported malaria in Netrokona district of Bangladesh 2013-2018: analysis of surveillance data. Malar Res Treat 2019;2019:6780258. [doi: 10.1155/2019/6780258] [Medline: 31312425]
19. Migrated population in delhi from 2003 to 2022. Statista. URL: https://www.statista.com/statistics/1400863/india-migration-in-delhi/ [accessed 2023-05-31]
20. Singh J, Bhattacharji LM. Rapid staining of malarial parasites by a water soluble stain. Ind Med Gaz 1944 Mar;79(3):102-104. [Medline: 29012018]
21. Anvikar AR, Arora U, Sonal GS, et al. Antimalarial drug policy in India: past, present & future. Indian J Med Res 2014 Feb;139(2):205-215. [Medline: 24718394]
22. Poumerol G, Wilder-Smith A, editors. Vaccines for routine and selective use: rabies. In: International Travel and Health: World Health Organization; 2012. URL: https://www.who.int/publications/i/item/9789241580472 [accessed 2023-05-31]
23. Lynch CA, Bruce J, Bhasin A, Roper C, Cox J, Abeku TA. Association between recent internal travel and malaria in Ugandan highland and highland fringe areas. Trop Med Int Health 2015 Jun;20(6):773-780. [doi: 10.1111/tmi.12480] [Medline: 25689689]
24. Alemu K, Worku A, Berhane Y, Kumie A. Men traveling away from home are more likely to bring malaria into high altitude villages, northwest Ethiopia. PLoS One 2014;9(4):e95341. [doi: 10.1371/journal.pone.0095341] [Medline: 24748159]
25. Yukich JO, Taylor C, Eisele TP, et al. Travel history and malaria infection risk in a low-transmission setting in Ethiopia: a case control study. Malar J 2013 Jan 24;12:33. [doi: 10.1186/1475-2875-12-33] [Medline: 23347703]
26. Christophers SR. The Fauna of British India, Including Ceylon and Burma. Diptera, Vol. Iv, Family Culicidae, tribe Anophelini: Taylor & Francis; 1933.
27. Manual on Practical Entomology in Malaria / Prepared by the WHO Division of Malaria and Other Parasitic Diseases: World Health Organization Division of Malaria and Other Parasitic Diseases; 1995. URL: https://iris.who.int/handle/10665/42481 [accessed 2024-12-04]
28. Tassanakajon A, Boonsaeng V, Wilairat P, Panyim S. Polymerase chain reaction detection of Plasmodium falciparum in mosquitoes. Trans R Soc Trop Med Hyg 1993;87(3):273-275. [doi: 10.1016/0035-9203(93)90124-9] [Medline: 8236389]
29. Malaria surveillance, monitoring & evaluation: a reference manual. : World Health Organization; 2018 URL: https://www.who.int/publications/i/item/9789241565578 [accessed 2024-12-04]
30. Adak T, Sharma VP, Orlov VS. Studies on the Plasmodium vivax relapse pattern in Delhi, India. Am J Trop Med Hyg 1998 Jul;59(1):175-179. [doi: 10.4269/ajtmh.1998.59.175] [Medline: 9684649]
31. Savargaonkar D, Nagpal BN, Srivastava B, Anvikar AR, Valecha N. The footprints of relapsing malaria in southwest Delhi, India. J Vector Borne Dis 2015 Dec;52(4):287-292. [Medline: 26714507]
32. Singh RK, Mittal PK, Kumar G, Dhiman RC. Insecticide susceptibility status of Aedes aegypti and Anopheles stephensi larvae against temephos in Delhi, India. Int J Mosq Res 2014;1(3):69-73.

33. Census of India 2001. Data highlights: migration tables (D1, D1 (Appendix), D2 and D3 tables). The Hindu Centre For Politics and Public Policy. 2001. URL: https://www.thehinducentre.com/publications/policy-watch/article33219541.ece/binary/data_highlights_D1D2D3.pdf [accessed 2023-05-31]
34. Sarkar P. An overview of out-migration from Uttar Pradesh using census 2011. J Migr Aff 2020;2(2):58. [doi: 10.36931/jma.2020.2.2.58-66]
35. Casuccio A, D'Angelo C, Casuccio N, Di Carlo P, Immordino P. Visiting friends and relatives (VFRs) travelers and imported malaria in the Palermo district (Sicily). Ann Ist Super Sanita 2014;50(4):369-374. [doi: 10.4415/ANN_14_04_13] [Medline: 25522079]
36. Heywood AE, Zwar N, Forssman BL, et al. The contribution of travellers visiting friends and relatives to notified infectious diseases in Australia: state-based enhanced surveillance. Epidemiol Infect 2016 Dec;144(16):3554-3563. [doi: 10.1017/S0950268816001734] [Medline: 27574034]
37. Preliminary draft report. District survey report for (planning & execution of) minor mineral excavation. : Directorate of Geology & Mining; 2016 URL: https://cdn.s3waas.gov.in/s3091d584fced301b442654dd8c23b3fc9/uploads/2019/02/2019021130.pdf [accessed 2021-01-11]
38. Directorate of National Vector Borne Diseases Control Programme, 22 Sham Nath Marg, DELHI-110054. Annual report 2018. National Center for Vector Borne Diseases Control (NCVBDC). 2018. URL: https://nvbdcp.gov.in/Doc/Annual-Report-2018.pdf [accessed 2024-12-04]
39. World malaria report 2023. : World Health Organization; 2023 URL: https://www.who.int/publications/i/item/9789240086173 [accessed 2024-12-04]

Abbreviations

API: annual parasite incidence
ICMR-NIMR: Indian Council of Medical Research-National Institute of Malaria Research
NCVBDC: National Center for Vector Borne Diseases Control
PCR: polymerase chain reaction
PMHD: per man-hour density
RACD: reactive case detection
SWD: South West Delhi
SWD: South West Delhi
UP: Uttar Pradesh
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 18.06.23; peer-reviewed by J Campbell, P Sahu; revised version received 16.05.24; accepted 23.05.24; published 08.01.25.

Please cite as:
Savargaonkar D, Srivastava B, Yadav CP, Singh MP, Anvikar A, Sharma A, Singh H, Sinha A
Contribution of Travelers to Plasmodium Vivax Malaria in South West Delhi, India: Cross-Sectional Survey
JMIR Public Health Surveill 2025;11:e50058
URL: https://publichealth.jmir.org/2025/1/e50058
doi:10.2196/50058

Consistency of Daily Number of Reported COVID-19 Cases in 191 Countries From 2020 to 2022: Comparative Analysis of 2 Major Data Sources

Han Liu[1], MPH; Huiying Zong[2], MPH; Yang Yang[3], PhD; David C Schwebel[4], PhD; Bin Xie[1], BM; Peishan Ning[1], PhD; Zhenzhen Rao[1], PhD; Li Li[1], PhD; Guoqing Hu[1], PhD

[1]Department of Epidemiology and Health Statistics, Hunan Provincial Key Laboratory of Clinical Epidemiology, Xiangya School of Public Health, Central South University, 172 Tongzipo Road, Yuelu District, Changsha, Hunan, China

[2]Department of Medical Records Management, Xiangya Hospital, Changsha, China

[3]Department of Statistics, Franklin College of Arts and Sciences, University of Georgia, Athens, GA, United States

[4]Department of Psychology, University of Alabama at Birmingham, Birmingham, AL, United States

Corresponding Author:
Guoqing Hu, PhD
Department of Epidemiology and Health Statistics, Hunan Provincial Key Laboratory of Clinical Epidemiology, Xiangya School of Public Health, Central South University, 172 Tongzipo Road, Yuelu District, Changsha, Hunan, China

Abstract

Background: The COVID-19 pandemic represents one of the most challenging public health emergencies in recent world history, causing about 7.07 million deaths globally by September 24, 2024. Accurate, timely, and consistent data are critical for early response to situations like the COVID-19 pandemic.

Objective: This study aimed to evaluate consistency of daily reported COVID-19 cases in 191 countries from the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) and the World Health Organization (WHO) dashboards during 2020 - 2022.

Methods: We retrieved data concerning new daily COVID-19 cases in 191 countries covered by both data sources from January 22, 2020, to December 31, 2022. The ratios of numbers of daily reported cases from the 2 sources were calculated to measure data consistency. We performed simple linear regression to examine significant changes in the ratio of numbers of daily reported cases during the study period.

Results: Of 191 WHO member countries, only 60 displayed excellent data consistency in the number of daily reported COVID-19 cases between the WHO and JHU CSSE dashboards (mean ratio 0.9-1.1). Data consistency changed greatly across the 191 countries from 2020 to 2022 and differed across 4 types of countries, categorized by income. Data inconsistency between the 2 data sources generally decreased slightly over time, both for the 191 countries combined and within the 4 types of income-defined countries. The absolute relative difference between the 2 data sources increased in 84 countries, particularly for Malta (R^2=0.25), Montenegro (R^2=0.30), and the United States (R^2=0.29), but it decreased significantly in 40 countries.

Conclusions: The inconsistency between the 2 data sources warrants further research. Construction of public health surveillance and data collection systems for public health emergencies like the COVID-19 pandemic should be strengthened in the future.

(JMIR Public Health Surveill 2025;11:e65439) doi:10.2196/65439

KEYWORDS

COVID-19; pandemic; data consistency; World Health Organization; data quality

Introduction

The COVID-19 pandemic represents one of the greatest global public health challenges of this century. According to World Health Organization (WHO) statistics, by September 24, 2024, COVID-19 led to over 7.04 million deaths globally [1]. Critical flaws in timely data collection and reporting have hindered and misled the prevention and control of the pandemic over the past 3 years [2]. For example, during the early phase of the pandemic, the cumulative number of COVID-19 cases in the United States released by the Johns Hopkins University dashboard quickly exceeded 1 million individuals by the end of April 2020, substantially higher than governmental statistics [3]. Such discrepancies created considerable chaos for US governmental responses to the COVID-19 pandemic [4]. To better respond to future pandemics and other global health crises, it is helpful to examine and summarize important lessons we learned.

One key aspect of public health crisis and pandemic preparedness is the need for high-quality data to guide decision-making, practice, and research about prevention and control. Furthermore, 2 open-access data sources were widely used in research and decision-making about the COVID-19 pandemic [5], the COVID-19 dashboards released by the WHO and the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE). The 2 data sources collected data in different ways and have been reported as having somewhat inconsistent data in several countries [6]. However, a comprehensive evaluation of the consistency between the 2 data sources has not been conducted across multiple countries and over the key time period of the pandemic, from 2020 to 2022.

Therefore, this study systematically examined the consistency of the number for daily reported COVID-19 cases in 191 countries from 2 data sources, WHO and JHU CSSE, from 2020 to 2022.

Methods

Database

Given their wide use and numerous citations, we selected the publicly accessible WHO COVID-19 dashboard and the JHU CSSE COVID-19 dashboard as the major sources to investigate data patterns concerning COVID-19 pandemic infection rates [6]. The WHO dashboard was created by the WHO. In January 2020, the WHO developed the case definition of COVID-19 infection and requested all member countries to report the number of daily reported cases through the International Health Regulations national focal points. It also collected COVID-19 daily cases and death counts from internet-based public sources (public dashboards and social media) as a supplement [7]. The JHU CSSE dashboard was created by Johns Hopkins University in early January 2020. It gathered data from publicly accessible sources referenced in a version-controlled README file on the repository [5].

For this study, we retrieved publicly accessible numbers of daily reported COVID-19 cases in the 191 United Nations member countries from both data sources between January 22, 2020 and December 31, 2022. In total, there were 6 days with missing values in the 2 data sources for all 191 countries and all study days together, comprising 0.00003% (6/205,325) of all study days. Considering the small number of missing values and the fact that they were unlikely to create significant impact on our results, we excluded those days from our analysis. Data before January 22, 2020, were also not included, as the outbreak had just initiated and the JHU CSSE COVID-19 dashboard was not yet established.

Data Analysis

In total, 2 measures were used to assess consistency between the 2 data sources. One measure was the ratio of the daily number of reported cases from the JHU CSSE dashboard to that from the WHO dashboard, aligned by calendar time. As an example, in Afghanistan on March 31, 2022, the JHU CSSE dashboard reported 264 newly reported cases and the WHO dashboard reported 220. The ratio of JHU CSSE over WHO dashboards was calculated as "264/220=1.2" for that day in Afghanistan. To facilitate our description, we refer to the ratio of daily numbers of reported cases between the 2 data sources as "daily ratio" or, simply, "ratio."

Based on the mean ratio of the 191 countries during the full study time period, we classified the daily ratios for each country into 5 categories of consistency: excellent (0.9≤mean ratio≤1.1), high (0.8≤mean ratio<0.9 or 1.1<mean ratio≤1.2), moderate (0.7≤mean ratio<0.8 or 1.2<mean ratio≤1.3), low (0.4≤mean ratio<0.7 or 1.3<mean ratio≤1.6), and poor (mean ratio<0.4 or 1.6<mean ratio). Countries where the 2 dashboards had higher consistency yielded ratios closer to 1.0, of course, and countries where there were greater discrepancies between the 2 dashboards had ratios either far below or far above 1.0.

To address certain unusual cases, we empirically built the following rules to calculate the daily ratio in certain unusual cases: (1) when the numerator and denominator were both 0, we defined the ratio as 1, as this represented perfect agreement between the dashboards; (2) when only the numerator was 0 and the denominator was 10 or greater, we defined the ratio as 0.5, and when the denominator was less than 10, we used the fourth criterion listed below to define the value of the ratio; (3) when the denominator was 0 and numerator was 10 or greater, we defined the ratio as 2.0, and when the numerator was less than 10, we used the fourth criterion below to define the value of the ratio; and (4) when the numerator and denominator were both less than 10 and their differences were less than 3 (explaining 0.1% of all data), between 4 and 6 (explaining 0.3%), or ≥7 (explaining 0.5%), we defined the ratio as 0.9 or 1.1, 0.7 or 1.3, 0.5 or 1.5, respectively, thus avoiding extremely large or small ratios due to small numerators or denominators. We mapped the mean and coefficient of variation (CV) of the daily ratios between the 2 data sources for each of the 191 countries.

We graphed box plots and performed the Kruskal-Wallis rank sum test to display and compare differences in the mean and CV of the daily ratios across groups of countries defined by income levels. The 191 countries were classified into 4 income-based categories according to the World Bank Analytical Classifications of 2023 [8]: low-income countries, lower middle-income countries, upper middle-income countries, and higher-income countries. In addition, simple linear regression, with the absolute relative difference in daily number used as the dependent variable and time (date) as the independent variable, was performed to examine trends in the daily differences during the study period. A significant and positive regression coefficient indicates worsening data consistency over time, while a significant and negative regression coefficient indicates improved data consistency.

SPSS statistics 26 (IBM Corp), R version 4.3.0 (R Core Team), and Microsoft Excel spreadsheet 2021 were used for data analysis. $P<.05$ was regarded as statistically significant.

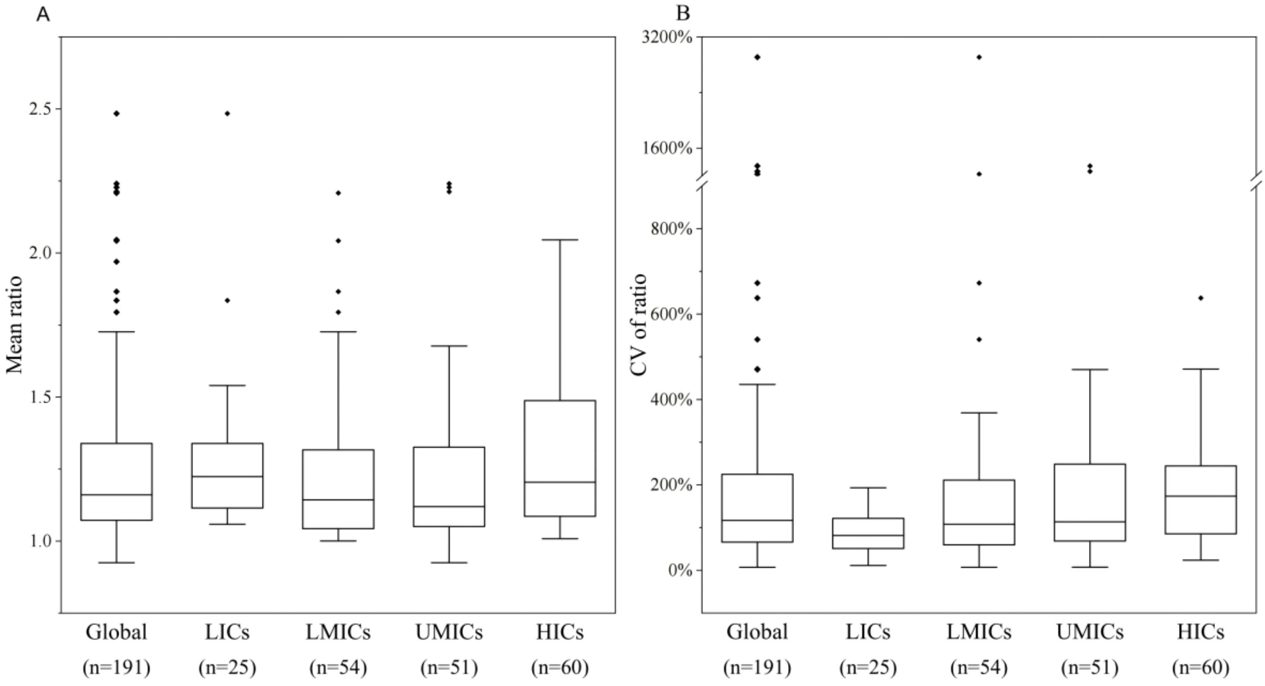
Ethical Considerations

This study does not cause harm to humans and does not involve individual information or commercial interest information. The data used in this study cannot be used to identify individual information, either directly or indirectly. Based on the given declarations, this study complies with the fundamental principles expressed in the Declaration of Helsinki and meets the ethical exemption requirements of the Ethical Review Measures for Life Sciences and Medical Research Involving Humans promulgated by China. Therefore, this research was exempted from ethical review.

Results

For all observed dates combined, the median ratio between the daily number of reported COVID-19 cases in the 2 data sources for the 191 countries was 1.14 (IQR 1.07-1.34), with the lowest mean ratio in Turkey (0.92) and the highest mean ratio in Sudan (2.48) (Figure 1A). The ratios did not differ significantly across the 4 categories of countries organized by income (*P*=.12). The median CV of ratio between the daily number of reported COVID-19 cases from the 2 data sources was 1.16 (IQR 0.66-2.25) for the 191 countries. The 4 types of countries had significantly different CVs of ratio (*H*=10.08, *P*=.02) (Figure 1B).

Figure 1. Ratio of number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard by country income category, from January 22, 2020, to December 31, 2022. CV: coefficient of variation; HIC: higher-income country; LIC: low-income country; LMIC: lower middle-income country; UMIC: upper middle-income country.



In Figure 1, the ratio was calculated as the number of daily new cases from the JHU CSSE COVID-19 dashboard divided by the number from the WHO COVID-19 dashboard, for each country. Figure 2 and Figure 3 display the variations in mean and CV of ratio between the daily number of reported COVID-19 cases from the 2 data sources across the countries. Of the 191 countries, 60 had excellent consistency (mean ratio: 0.9 - 1.1), while 37 had low consistency (mean ratio: 0.4 - 0.7 or 1.3 - 1.6) and 23 had poor consistency (mean ratio:<0.4 or >1.6) (Figure 2). Most African and European countries had low or poor data consistency and most Asian countries had high or excellent data consistency. Strikingly, the CV of ratio between the number of daily reported COVID-19 cases from the 2 data sources during the study time period was higher than 100% in 110 countries (110/191, 57%), and 20% or less in only 8 countries (8/191, 4%) (Figure 3).

Figure 2. Mean ratio of number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard across 191 countries from January 22, 2020, to December 31, 2022. The ratio was calculated as the number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard divided by the number from the World Health Organization COVID-19 dashboard, for each country. NA: not available.

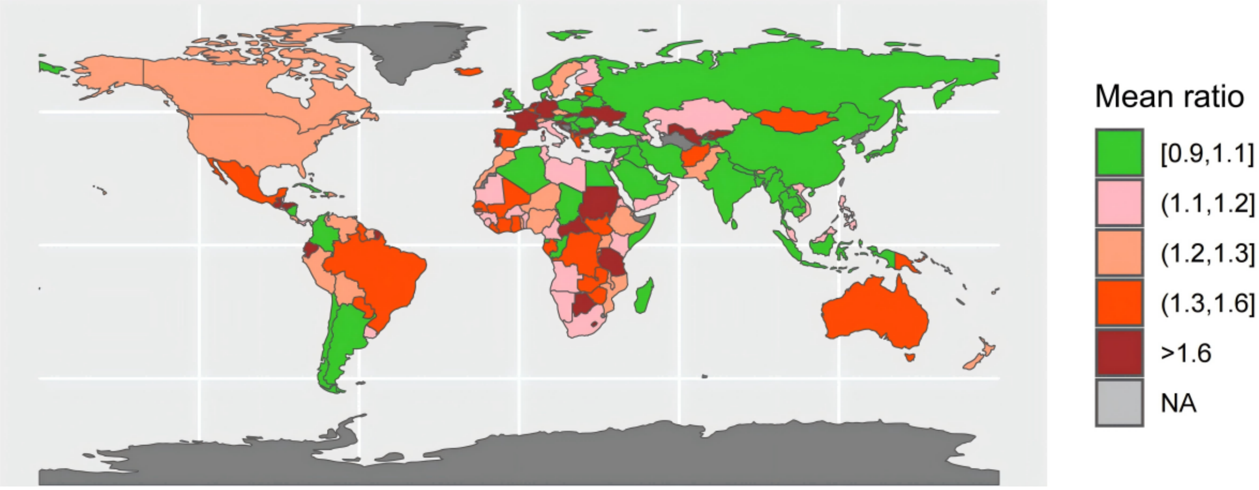
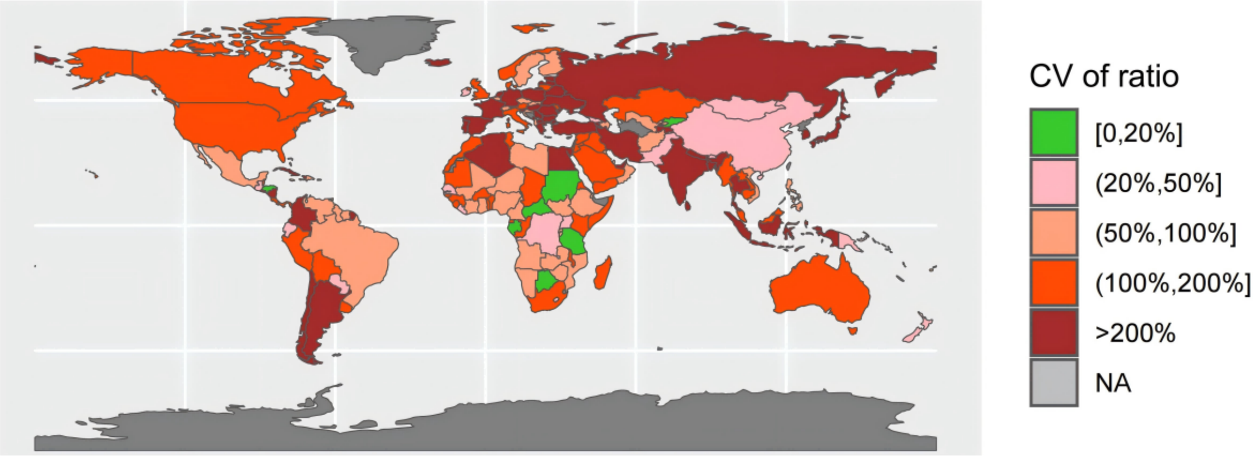


Figure 3. CV of ratio of number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard across 191 countries from January 22, 2020, to December 31, 2022. The ratio was calculated as the number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard divided by the number from the World Health Organization COVID-19 dashboard, for each country. CV: coefficient of variation; NA: not available.



For the 191 countries combined and within all 4 income-defined types of countries, the absolute relative difference between daily number of reported COVID-19 cases from the 2 data sources all increased slightly over the study time period, with R^2 values of 0.16, 0.08, 0.06, 0.14, and 0.14, respectively (Table 1). Country-specific analyses showed that the absolute relative difference increased in 84 countries but decreased in 40 countries significantly (Table 2). Notably, data inconsistency between the 2 data sources increased significantly over time in 3 countries (Malta: R^2=0.25; Montenegro: R^2=0.30; United States: R^2=0.29).

Table . Linear trends in ratio of number of daily reported cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard from January 22, 2020, to December 31, 2022, by country income category (dependent and independent variables of linear regression model were “the absolute value of (ratio-1)” and “time (date),” respectively).

Country income	Linear regression model	Coefficient of determination (R^2)	*P* value
Global	$\hat{y}=-713.80+0.00029x$	0.16	<.001
LICs[a]	$\hat{y}=-1110.18+0.00045x$	0.08	<.001
LMICs[b]	$\hat{y}=-1025.15+0.00041x$	0.06	<.001
UMICs[c]	$\hat{y}=-781.22+0.00031x$	0.14	<.001
HICs[d]	$\hat{y}=-1453.78+0.00059x$	0.14	<.001

[a] LIC: low-income country.

[b] LMIC: lower middle-income country.

[c] UMIC: upper middle-income country.

[d] HIC: higher-income country.

Table . Linear trends in ratio of country-specific number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard from January 22, 2020, to December 31, 2022. Based on coefficient of determination (R^2) of linear regression models with statistical significance, we classified linear trends of ratio into 3 grades: substantially ($R^2 \geq 0.25$), moderately ($0.10 < R^2 < 0.25$), and slightly ($R^2 \leq 0.10$). In addition, we did not detect significant changes in data consistency for 67 countries).

Linear trend		Country
Increased (n=84)		
	Substantially (n=3)	Malta (R^2=0.25), Montenegro (R^2=0.30), United States (R^2=0.29)
	Moderately (n=4)	Australia (R^2=0.18), Bulgaria (R^2=0.11), Iran (R^2=0.13), Slovenia (R^2=0.21)
	Slightly (n=77)	Albania, Algeria, Antigua and Barbuda, Azerbaijan, Bahamas, Bangladesh, Barbados, Belgium, Bhutan, Bolivia, Bosnia and Herzegovina, Brazil, Brunei, Burundi, Cambodia, Canada, Chile, Comoros, Croatia, Cyprus, Czechia, Democratic Republic of Congo, Denmark, Dominica, Dominican Republic, Ecuador, Ethiopia, Fiji, Finland, Georgia, Grenada, Haiti, India, Iraq, Ireland, Italy, Kazakhstan, Kenya, Kiribati, Kuwait, Laos, Latvia, Liechtenstein, Lithuania, Maldives, Marshall Islands, Mauritius, Mexico, Micronesia, Monaco, Mongolia, Morocco, Namibia, Nauru, Norway, Palau, Panama, Papua New Guinea, Peru, Romania, Saint Kitts and Nevis, Samoa, San Marino, Saudi Arabia, Seychelles, Slovakia, Solomon Islands, South Africa, Sweden, Switzerland, Thailand, Timor, Tonga, Ukraine, United Kingdom, Vanuatu, Venezuela
Decreased (n=40)		
	Moderately (n=2)	Poland (R^2=0.17), Spain (R^2=0.11)
	Slightly (n=38)	Andorra, Argentina, Bahrain, Belarus, Burkina Faso, Cameroon, Chad, Congo, Djibouti, Egypt, El Salvador, Estonia, Germany, Greece, Guinea, Hungary, Japan, Jordan, Madagascar, Malaysia, Mali, Mauritania, Moldova, Myanmar, Netherlands, North Macedonia, Portugal, Russia, Sierra Leone, Somalia, Togo, Tunisia, Turkey, Tuvalu, United Arab Emirates, Uruguay, Yemen, Zambia
No significant change (n=67)		Afghanistan, Angola, Armenia, Austria, Belize, Benin, Botswana, Cape Verde, Central African Republic, China, Colombia, Costa Rica, Cote d’Ivoire, Cuba, Equatorial Guinea, Eritrea, Eswatini, France, Gabon, Gambia, Ghana, Guatemala, Guinea-Bissau, Guyana, Honduras, Iceland, Indonesia, Israel, Jamaica, Kyrgyzstan, Lebanon, Lesotho, Liberia, Libya, Luxembourg, Malawi, Mozambique, Nepal, New Zealand, Nicaragua, Niger, Nigeria, Oman, Pakistan, Paraguay, Philippines, Qatar, Republic of Korea, Rwanda, Saint Lucia, Saint Vincent and the Grenadines, Sao Tome and Principe, Senegal, Serbia, Singapore, South Sudan, Sudan, Suriname, Sri Lanka, Syria, Tajikistan, Tanzania, Trinidad and Tobago, Uganda, Uzbekistan, Vietnam, Zimbabwe

Discussion

Main Findings

This study systematically examined the consistency of number of daily reported COVID-19 cases and number-based epidemic cycles in 191 countries from the 2 most cited data sources between 2020 and 2022. We generated 2 key findings. First, there was a difference in global number of daily reported COVID-19 cases between the WHO and JHU CSSE data, and data consistency differed between the 4 categories of income-defined countries and across the 191 countries. Second, data inconsistency between the 2 data sources generally decreased slightly over time, both for the 191 countries combined and within the 4 types of income-defined countries, but it changed differently over time across the 191 individual countries.

Interpretation of Findings

Inconsistencies in the number of daily reported COVID-19 cases between the WHO and JHU CSSE data likely reflect differences in data collection and data release strategies across the 2 data sources. The WHO data came primarily from official daily counts reported by WHO member states, territories, and areas [7], while the JHU CSSE data were derived from a combination of more than 400 sources for over 3500-point locations [5]. Compared with the WHO data, the JHU CSSE data were collected from a higher number of sources, potentially explaining why the number of daily reported COVID-19 cases from the JHU CSSE data was greater than that from the WHO data for the whole world, the 4 types of income-defined countries, and most individual countries during most dates [9]. The 2 data sources also adopted different data release strategies. The JHU CSSE COVID-19 dashboard used automated methods to extract data from each source every half hour and therefore updated data hourly [5]. In contrast, the WHO COVID-19 dashboard relied on official reports from all WHO member countries, territories, and areas, which were more prone to be delayed and less frequently updated than the JHU dashboard data [10].

Variations in data inconsistency between the WHO and JHU CSSE data might also be due to diverse data publishing policies and the evolution of data-releasing agencies (eg, delayed reporting, data entry errors, and data collection strategies), particularly at the local level across the 191 countries. As the pandemic eased, some countries like Sweden reduced the frequency of data release and even stopped publishing epidemic data, particularly at the local level [11]. These practices likely affected the original data collected by the WHO and the JHU CSSE COVID-19 dashboards in different ways. In particular, sources indexed by the JHU CSSE dashboard were more extensive and were more likely to change over time than those covered by the WHO dashboard [12].

Implications and Limitations

Our findings have 2 important implications. First, we do not have clear empirical basis to explain the drivers behind inconsistencies in the number of COVID-19 cases between the WHO and JHU CSSE dashboards or the changes in data inconsistency over time. A variety of internal and external factors likely could have contributed to the magnitude and pattern of the inconsistency, including but not limited to changes in government intervention policies or testing capacities.

Policy makers, researchers, and public health practitioners should recognize that the 2 data sources generate somewhat different and even occasionally conflicting results when they use the historical COVID-19 data collected by the WHO and JHU CSSE dashboards to summarize pandemic control experiences and lessons, conduct historical trend analyses, and evaluate intervention strategies [13-15]. Because strong evidence and detailed guidance are lacking concerning which of the 2 data sources is more valid for which specific time periods and countries, we recommend users consider both, recognizing each may have flaws. It also would be valuable to conduct further research assessing the quality of the data sources. For example, comparisons of the data sources with known cycles of the COVID-19 epidemic may yield additional information about the validity of each data source [16,17]. Another useful approach to bridge the gaps in high-level data from mainstream sources is to cross-validate them with data provided by local health authorities whenever and wherever available. This may be particularly useful for data on hospitalization and on fatalities, which are less sensitive to testing policies, health care–seeking behavior, and reporting bias [18].

Second, our results highlight the importance of gathering consistent and accurate data to fight against future pandemics or other public health emergencies [19]. Inconsistent data can lead to undesired consequences in decision-making, research, and prevention efforts [20,21]. Development of standardized data collection and release protocols, along with infrastructure- and competency-building to help all countries gather accurate health data, will prepare us not just for current health situations but also for future pandemics and broad public health emergencies [22].

Our research is primarily limited by the unavailability of a gold standard criterion. We assessed the consistency of 2 data sources and cannot comprehensively assess the reliability or validity of either, or identify which might be more accurate. No alternative or recognized valid measure of COVID-19 infections or fatalities exists worldwide. In addition, due to the unavailability of detailed information about COVID-19 data collection methods or data release policies across the 191 countries, we cannot reasonably interpret exactly why data consistency variations emerged across the 191 countries or why there were significant increases or decreases in data consistency over time for some countries. Future research may consider innovative strategies to identify key factors contributing to the observed data inconsistencies and significant changes in data consistency. Hospitalization data might be considered as a surrogate gold standard, for example, to approximate accuracy of other data sources [23], perhaps providing valuable clues. Advanced data analysis strategies, such as data harmonization or deep machine learning, might also help uncover reasons behind the results we report and be useful to address data accuracy in future public health emergencies. Finally, evidence-based data collection and release guidelines should be established by the WHO or other governing bodies, perhaps supplemented with financial and technical support to those countries with limited resources to follow international guidelines. With strong international collaboration, we should be able to improve transparency, consistency and overall quality of data emerging from public health crises in the future.

Conclusions

We report a moderate difference in global average daily number of reported COVID-19 cases between 2020 and 2022 in the WHO and JHU CSSE datasets. The COVID-19 data inconsistency changed significantly across the countries studied and across time, both globally and in 124 individual countries. Data users should be cautious in interpreting the data, given inconsistency across the 2 most commonly cited COVID-19 data sources. Further research is recommended to develop guidelines to support valid interpretation of existing COVID-19 data, explore possible reasons for discrepancies, and understand factors such as governmental policies and data collection systems that may lead to data inconsistency. With those efforts, society can be best prepared to conduct valid health-related data collection in response to future pandemics and public health emergencies.

Acknowledgments

The study was supported by Philosophy and Social Science Foundation of China (number 20&ZD120).

Data Availability

Both databases are freely accessible for the public. Throughout the data utilization process, we strictly adhered to the data usage requirements of the source databases, ensuring no modifications were made to the metadata. The attribution of the data sources has been explicitly stated in the manuscript.

Conflicts of Interest

None declared.

References

1. WHO COVID-19 dashboard. World Health Organization. URL: https://data.who.int/dashboards/covid19/data?n=c [accessed 2023-12-31]
2. Nageshwaran G, Harris RC, Guerche-Seblain CE. Review of the role of big data and digital technologies in controlling COVID-19 in Asia: public health interest vs. privacy. Dig Health 2021;7:20552076211002953. [doi: 10.1177/20552076211002953] [Medline: 33815815]
3. Pollard CA, Morran MP, Nestor-Kalinoski AL. The COVID-19 pandemic: a global health crisis. Physiol Genomics 2020 Nov 1;52(11):549-557. [doi: 10.1152/physiolgenomics.00089.2020] [Medline: 32991251]
4. Lessons learned from COVID-19 pandemic: virtual regional consultation with informal expert group, 19 October 2021. World Health Organization. URL: https://www.who.int/publications/i/item/sea-whe-6 [accessed 2025-01-02]
5. Dong E, Ratcliff J, Goyea TD, et al. The Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard: data collection process, challenges faced, and lessons learned. Lancet Infect Dis 2022 Dec;22(12):e370-e376. [doi: 10.1016/S1473-3099(22)00434-0] [Medline: 36057267]
6. Alamo T, Reina D, Mammarella M, Abella A. Covid-19: Open-data resources for monitoring, modeling, and forecasting the epidemic. Electronics (Basel) 2020;9(5):827. [doi: 10.3390/electronics9050827]
7. Allan M, Lièvre M, Laurenson-Schafer H, et al. The World Health Organization COVID-19 surveillance database. Int J Equity Health 2022 Nov 23;21(Suppl 3):167. [doi: 10.1186/s12939-022-01767-5] [Medline: 36419127]
8. World Bank country and lending groups. The World Bank. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519 [accessed 2025-01-28]
9. Wu Z, Huang NE, Long SR, Peng CK. On the trend, detrending, and variability of nonlinear and nonstationary time series. Proc Natl Acad Sci U S A 2007 Sep 18;104(38):14889-14894. [doi: 10.1073/pnas.0701020104] [Medline: 17846430]
10. Zuo X, Chen Y, Ohno-Machado L, Xu H. How do we share data in COVID-19 research? A systematic review of COVID-19 datasets in PubMed Central articles. Brief Bioinform 2021 Mar 22;22(2):800-811. [doi: 10.1093/bib/bbaa331] [Medline: 33757278]
11. Hu T, Guan WW, Zhu X, et al. Building an open resources repository for COVID-19 research. Data Inf Manag 2020 Sep 1;4(3):130-147. [doi: 10.2478/dim-2020-0012] [Medline: 35382104]
12. Lindström M. The COVID-19 pandemic and the timing of government response: a comparison of four Nordic countries March–June 2020. Scand J Public Health 2023 Jul;51(5):754-758. [doi: 10.1177/14034948231171201] [Medline: 37118947]
13. Badker R, Miller K, Pardee C, et al. Challenges in reported COVID-19 data: best practices and recommendations for future epidemics. BMJ Glob Health 2021 May;6(5):e005542. [doi: 10.1136/bmjgh-2021-005542] [Medline: 33958393]
14. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. Lancet Infect Dis 2020 Nov;20(11):1247-1254. [doi: 10.1016/S1473-3099(20)30553-3]

15. Guerrier C, McDonnell C, Magoc T, Fishe JN, Harle CA. Understanding health care administrators' data and information needs for decision making during the COVID-19 pandemic: a qualitative study at an academic health system. MDM Policy Pract 2022;7(1):23814683221089844. [doi: 10.1177/23814683221089844] [Medline: 35368410]
16. Dowell SF, Ho MS. Seasonality of infectious diseases and severe acute respiratory syndrome–what we don't know can hurt us. Lancet Infect Dis 2004 Nov;4(11):704-708. [doi: 10.1016/S1473-3099(04)01177-6]
17. Chemaitelly H, Tang P, Hasan MR, et al. Waning of BNT162b2 vaccine protection against SARS-CoV-2 infection in Qatar. N Engl J Med 2021 Dec 9;385(24):e83. [doi: 10.1056/NEJMoa2114114] [Medline: 34614327]
18. Huang AA, Huang SY. Hospitalized COVID-19 patients with diabetes have an increased risk for pneumonia, intensive care unit requirement, intubation, and death: a cross-sectional cohort study in Mexico in 2020. Health Sci Rep 2023 Apr;6(4):e1222. [doi: 10.1002/hsr2.1222] [Medline: 37081996]
19. Johnson B. Five discoveries about COVID-19 made since the public health emergency ended. Nat Med 2023 Dec;29(12):2974-2976. [doi: 10.1038/d41591-023-00084-w]
20. Kung S, Doppen M, Black M, et al. Underestimation of COVID-19 mortality during the pandemic. ERJ Open Res 2021 Jan;7(1):00766-02020. [doi: 10.1183/23120541.00766-2020]
21. TC J, Barbieri M, et al. Differences in COVID-19 mortality: Implications of imperfect and diverse data collection systems. Popul Health Metr 2021;76:35-72. [doi: 10.3917/popu.2101.0035]
22. Suthar AB, Schubert S, Garon J, Couture A, Brown AM, Charania S. COVID-19 case definitions, diagnostic testing criteria, and surveillance across the pandemic's 25 highest burden countries. medRxiv. Preprint posted online on May 14, 2021. [doi: 10.1101/2021.05.11.21257047]
23. Huang AA, Huang SY. Diabetes is associated with increased risk of death in COVID-19 hospitalizations in Mexico 2020: a retrospective cohort study. Health Sci Rep 2023 Jul;6(7):e1416. [doi: 10.1002/hsr2.1416] [Medline: 37415678]

Abbreviations

CV: coefficient of variation
JHU CSSE: Johns Hopkins University Center for Systems Science and Engineering
WHO: World Health Organization

Edited by A Mavragani; submitted 15.08.24; peer-reviewed by R Ibrahim, SY Huang; revised version received 13.11.24; accepted 24.11.24; published 06.02.25.

Please cite as:
Liu H, Zong H, Yang Y, Schwebel DC, Xie B, Ning P, Rao Z, Li L, Hu G
Consistency of Daily Number of Reported COVID-19 Cases in 191 Countries From 2020 to 2022: Comparative Analysis of 2 Major Data Sources
JMIR Public Health Surveill 2025;11:e65439
URL: https://publichealth.jmir.org/2025/1/e65439
doi:10.2196/65439

Meeting Global Health Needs via Infectious Disease Forecasting: Development of a Reliable Data-Driven Framework

Ravikiran Keshavamurthy[1,2], BVSc, MVSc, PhD; Karl T Pazdernik[1,3], PhD; Colby Ham[1], BS; Samuel Dixon[1], MS; Samantha Erwin[1], MS, PhD; Lauren E Charles[1,2], DVM, MS, PhD

[1]Artificial Intelligence and Data Analytics Division, Pacific Northwest National Laboratory, 902 Battelle Boulevard, Richland, WA, United States
[2]Paul G. Allen School for Global Health, Washington State University, Pullman, WA, United States
[3]Department of Statistics, North Carolina State University, Raleigh, NC, United States

Corresponding Author:
Lauren E Charles, DVM, MS, PhD
Artificial Intelligence and Data Analytics Division, Pacific Northwest National Laboratory, 902 Battelle Boulevard, Richland, WA, United States

Abstract

Background: Infectious diseases (IDs) have a significant detrimental impact on global health. Timely and accurate ID forecasting can result in more informed implementation of control measures and prevention policies.

Objective: To meet the operational decision-making needs of real-world circumstances, we aimed to build a standardized, reliable, and trustworthy ID forecasting pipeline and visualization dashboard that is generalizable across a wide range of modeling techniques, IDs, and global locations.

Methods: We forecasted 6 diverse, zoonotic diseases (brucellosis, campylobacteriosis, Middle East respiratory syndrome, Q fever, tick-borne encephalitis, and tularemia) across 4 continents and 8 countries. We included a wide range of statistical, machine learning, and deep learning models (n=9) and trained them on a multitude of features (average n=2326) within the One Health landscape, including demography, landscape, climate, and socioeconomic factors. The pipeline and dashboard were created in consideration of crucial operational metrics—prediction accuracy, computational efficiency, spatiotemporal generalizability, uncertainty quantification, and interpretability—which are essential to strategic data-driven decisions.

Results: While no single best model was suitable for all disease, region, and country combinations, our ensemble technique selects the best-performing model for each given scenario to achieve the closest prediction. For new or emerging diseases in a region, the ensemble model can predict how the disease may behave in the new region using a pretrained model from a similar region with a history of that disease. The data visualization dashboard provides a clean interface of important analytical metrics, such as ID temporal patterns, forecasts, prediction uncertainties, and model feature importance across all geographic locations and disease combinations.

Conclusions: As the need for real-time, operational ID forecasting capabilities increases, this standardized and automated platform for data collection, analysis, and reporting is a major step forward in enabling evidence-based public health decisions and policies for the prevention and mitigation of future ID outbreaks.

(JMIR Public Health Surveill 2025;11:e59971) doi:10.2196/59971

KEYWORDS

disease forecasting; machine learning; deep learning; epidemiology; One Health; decision-making; data visualization

Introduction

The frequency and magnitude of infectious disease (ID) events have seen a drastic incline in the past few decades, mainly attributed to climate change, urbanization, and globalization [1]. These events often involve the emergence of novel infectious agents or the re-emergence of a previously known pathogen. Timely and accurate prediction of such ID events is crucial for decision makers to decrease associated mortality, morbidity, and economic losses [2]. However, the complex and unpredictable nature of pathogen ecology makes forecasting the spatiotemporal dynamics of IDs a challenging task [3].

Digital information related to ID events is being generated and shared faster than ever before. Additionally, information associated with disease occurrence, such as meteorology, socioeconomics, demographics, land use, agriculture, social media, and internet trends, is often readily available [4]. To keep up with this unprecedented amount of data being generated, machine learning (ML)– and deep learning (DL)–based algorithms are being adopted by the research community. These methods have shown to be better at detecting cryptic patterns arising from interactions between multiple features, which are difficult, often impossible at times, to uncover with conventional prediction methods [5].

The time-series models have been previously used for forecasting ID events including brucellosis [6], Middle East respiratory syndrome [7], campylobacteriosis [8], and Q fever [5]. However, over the last decade, especially after the COVID-19 pandemic outbreak, considerable progress has been made in the field of ID surveillance, as the disease diagnosis, reporting, and intelligence-sharing infrastructure continue to grow on a global scale [9]. Dixion et al [5] compared the performance of various forecasting approaches across several diseases and countries and found that tree-based techniques had better predictive performance compared to statistical and DL techniques. Integrating diverse large-scale epidemiological data from multiple sources has further enhanced the accuracy and utility of disease prediction models [10]. Ensemble forecasting techniques, which merge predictions from multiple models, offer notable advantages over single-model approaches [11]. For example, Reich et al [12] used real-time multimodel ensembles for seasonal influenza in the United States, while Ma et al [13] demonstrated the effectiveness of integrating internet search data along with ensemble forecasting techniques to jointly forecast COVID-19 and influenza-like illnesses, underscoring the value of diverse data sources in public health forecasting.

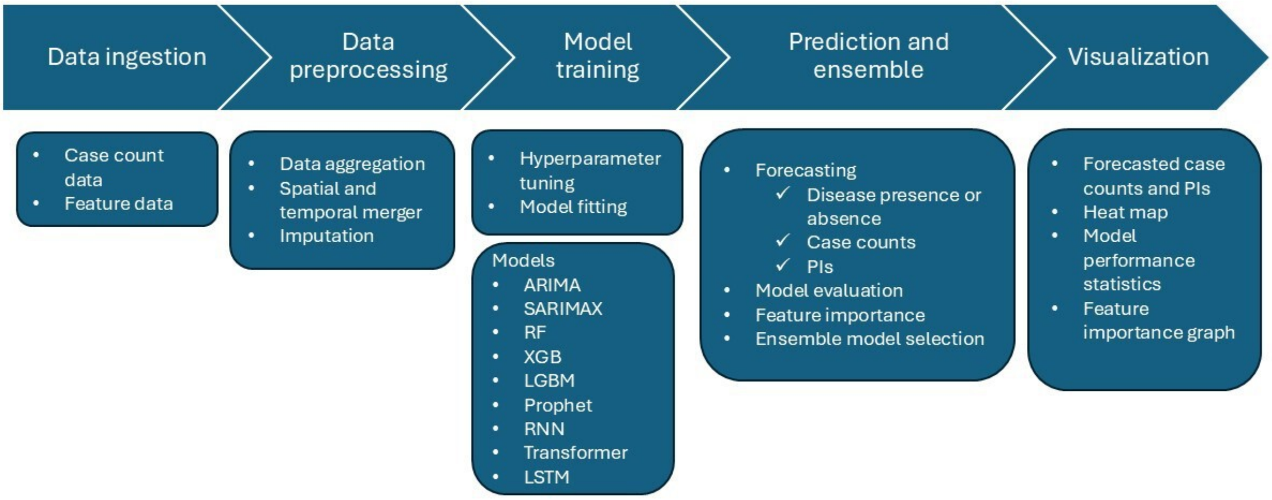
Despite several advantages, ML and DL techniques do come with a number of limitations, making them less desirable in the real-world operational setting. We conducted a scoping systematic review to determine the status and advances made in the field of ID prediction using ML and DL, which were published elsewhere [9]. One of our focuses in this systematic review was to identify if and how researchers were incorporating the necessary methods required to eventually enable the operational deployment of their ID forecasting models. Through this work, we identified multiple shortcomings in published modeling techniques that could hinder their ability to support operational biopreparedness and decision-making. Even though most of the studies focused on increasing the prediction accuracy of their models, which is critical for operational deployment, they did not address other important decision-making metrics such as computational efficiency, uncertainty quantification, interpretability, and generalizability. Modern ML and DL techniques are performance-driven, that is, they aim to generate better predictive or classification accuracy by minimizing errors [14]. As a result, other metrics that are crucial for operational decision-making are often overlooked by the scientific community. First, as the structural and functional complexities of the forecasting models evolve, the computational resources required to train such models also grow considerably. In addition, these complexities result in an expanded number of model hyperparameters that can be individually tuned using numerous possible techniques. This ever-increasing complexity gives rise to an even larger problem space with highly variable results. Second, since most of these techniques are nonparametric in nature, model uncertainties are not inherently estimated. Hence, model uncertainties are often overlooked even though these estimates are crucial for systematic and transparent decision-making [15]. Third, ML and DL models are considered black boxes, meaning that their internal logic and inner workings are mostly hidden from the end user [9]. Consequently, verifying and understanding the rationale behind the model forecasts are difficult and often neglected. With the above considerations in mind, we developed an ID forecasting pipeline that holistically focuses on these often neglected yet essential performance metrics required for operational decision support. By incorporating this pipeline into an interactive dashboard, we created a capability that allows users to easily visualize the ID forecasting results for use in operational planning and decision-making.

Methods

Overview

The graphical flowchart illustrating our forecasting pipeline including data ingest, preprocessing, model training, prediction, and visualization is presented in Figure 1.

Figure 1. Graphical flowchart illustrating the infectious disease forecasting pipeline. ARIMA: autoregressive integrated moving average; LGBM: light gradient boosting machine; LSTM: long short-term memory; PI: prediction interval; RF: random forest; RNN: recurrent neural network; SARIMAX: seasonal autoregressive integrated moving average with exogenous factors; XGB: extreme gradient boosting.



Case Count Data

The model outcome variable was case counts collected from EpiArchive [16] and Food and Agriculture Organization’s Emergency Prevention System-I [17]. We included the most consistent, available, and interpretable disease case counts from these sources. These data also included locations that reported only 0 disease cases. In instances where only the disease presence was reported without any information about the exact number of cases, the region, or the date of occurrence, such disease-location combinations were excluded from the analysis. We included all the diseases that spanned across multiple countries or regions and had a consistent reporting of disease events for our study period. Data were collected at the regional-level resolution for each country from January 2014 to December 2018, which contained case counts at either the weekly or monthly temporal resolution.

Feature Data

We collected feature data from a variety of open sources across the One Health landscape (Table 1). Broadly, the features included historic case counts (case count lags up to 6 months), country census data, socioeconomic factors, health data, agricultural trade data, landscape, and climate records. The feature information varied in spatial and temporal resolutions; data were typically available at monthly or yearly resolutions and covered either the country or regional level. A complete list of countries, regions, and the total number of features collected for each region is presented in Multimedia Appendix 1.

Table . The input feature data types, source, feature description by name, geographic location, geographic resolution, time period, and periodicity. The table was originally published by Dixon et al [5].

Data type	Individual features	Geographic location	Geographic resolution	Time period	Periodicity
Case counts [16]	Incidences of select human diseases	Countries of interest	Region level	2009-2018	Daily
Political borders [18]	Geopolitical borders (country and within the country)	Countries of interest	Region level	2018	Single instance
Climate [19,20]	Air temperature, humidity, precipitation, soil moisture, and wind speed	Global	Gridded 0.25 °×0.25 °, 1 °×1 °	2012 - 2018	Monthly
Gross domestic product [21]	Gross domestic product	Global	Country level	Varies	Yearly
Elevation [22]	Digital elevation map	Global	43,200×17,200 (30 arc seconds)	N/A[a]	N/A
Mortality [23]	Deaths by country, year, sex, age group, and cause of death	Global	Country level	2009 - 2018	Yearly
Municipal waste [24]	Municipal waste generation and treatment	Countries of interest	Country level	2009 - 2017	Yearly
Sociopolitical and physical data [25]	Socioeconomic and political attributes	Global	Varies by country; 1: 10-110 m	2019	Single instance
Population [26]	Population by age intervals by location	Global	Country level	2009 - 2015	Every 5 years
Population density [27]	Population density	Global	30 arc seconds	2009 - 2015	Every 5 years
Water potability and treatment [24]	Freshwater resources, available water, wastewater treatment plant capacity, surface water	Countries of interest	Country level	2009 - 2017	Yearly

[a] N/A: not applicable.

Data Preparation and Imputation

The harvested raw data contained information at varied spatial and temporal resolutions in different file formats. These raw data were harmonized to create the final datasets containing monthly case counts and input features for each region within the country. In situations where the feature data at the regional level were missing, the features aggregated by an average at the national level were used. Any features with more than 20% missing values across all regions and dates were not included as a predictor variable. The remaining missing values were imputed temporally using spline and forward-fill techniques for monthly and yearly feature data, respectively. Some features could not easily be imputed temporally through these techniques but were still sufficiently prevalent across regions to include in the models. In these instances, the majority of the missing data (less than 20% overall) came from only a few select regions. For these special cases, we performed geoimputation using the k-nearest neighbor method and restricted input to only regional feature data within the same country. We used distance-based weights for the features with 5 neighboring samples for imputation. The imputation techniques were implemented using the "pandas" and "skforecast" libraries in Python (Python Software Foundation) [26].

Model Choice

Overview

In our analysis, we included the following models discussed in detail below: statistical time-series models (ie, autoregressive integrated moving average [ARIMA], seasonal autoregressive integrated moving average with exogenous factors [SARIMAX], and Prophet), tree-based models (ie, random forest [RF], extreme gradient boosting [XGB], and light gradient boosting machine [LBGM]), and DL models (recurrent neural network [RNN], long short-term memory [LSTM], and transformer).

Statistical Time-Series Models

The ARIMA model is a traditional statistical technique used in time-series forecasting. These models use a linear, regression-type equation, in which the predictors are lags of the dependent variable or lags of the forecast errors. The SARIMAX is an extension of ARIMA that takes seasonal trends into account within time-series data. Because SARIMAX can also accommodate exogenous features, it is often preferable to traditional ARIMA. However, SARIMAX models may fail to converge if those exogenous features are highly correlated. To overcome this issue, the least absolute shrinkage and selection operator (LASSO) regression was used for feature selection and dimensionality reduction. These LASSO-selected features were later used as input variables for SARIMAX. The modeling techniques were implemented using "statsmodels" and "sklearn" libraries in Python [28,29].

Prophet is a decomposable time-series model that uses 3 main components to make its forecasts, namely, trend, seasonality, and holidays [30]. The trend component models the nonperiodic changes in the time series. Seasonality models the periodic, or seasonal, changes in the time series; this can be an important part of a disease forecasting model, as many diseases have seasonal trends. The Prophet model was implemented using the "darts" library in Python [31].

Tree-Based ML Models

Tree-based models make predictions by creating decision trees that divide the feature space using a series of binary decision thresholds. The RF is an extension of the decision tree method that uses an ensemble of decision trees to increase performance and reduce the risk of overfitting the training data. The XGB is an implementation of a decision tree that uses stochastic gradient boosting to sequentially improve prediction during training. Because of this methodology, XGB achieves superior accuracy in prediction tasks. LGBM is similar to XGB; however, it also applies a novel sampling method and feature bundling process that allows the method to get comparable performance to XGB but uses far less memory and trains much faster. These modeling techniques were implemented using "skforecast" library in Python [32].

DL Models

The RNN is built on DL architecture with a hidden state that memorizes sequential data. This added process provides the model with an understanding of temporal information in a time-series sequence. The LSTM is an extension of RNN with additional parameters that incorporate different scales of temporal information, which allow the model to effectively use both long- and short-term temporal data in its prediction. Hence, these models are routinely used in disease forecasting tasks.

Transformers [33] are another form of DL architecture that has been recently outperforming other models in many applications of ML. These models leverage a structure known as "attention," which allows them to model dependence in both short and long temporal lags. Unfortunately, these models are especially slow to train and often require very large amounts of data and computational power. The DL techniques were implemented as an encoder-record architecture using "darts" library in Python.

Model Training and Prediction

The model training involved hyperparameter tuning and model fitting, whereas model testing involved generating 1-step ahead forecasts (ie, monthly) and their prediction intervals (PIs). When creating predictions, all feature data and disease case counts up until the time of prediction were included in the model. The model training included hyperparameter tuning using a 5-fold time-series cross-validation split and model fitting. First, the models were trained using 2014 - 2016 data and tested on 2017 data. Subsequently, the model was retrained using 2014 - 2017 data and tested on 2018 data in a sliding window manner. A detailed description of model-specific hyperparameters and their search space used for model training is presented in Multimedia Appendix 1.

Model Evaluation

The forecasting models were evaluated in two ways: (1) F_1-score to evaluate the ability of a model to predict the presence or absence of a disease in each region (ie, did the region report 1 or more cases during the testing period) and (2) mean absolute error (MAE) for the subset of predictions where the disease was detected in a region (case counts greater than 0). The multimetric approach was adopted because of the high frequency of time-location pairs that had no disease present. Without separating these results, the MAE metric was largely driven by the results of the high-frequency disease and regions. Finally, we created ensemble models by selecting the best-performing technique based on the MAE of the testing dataset. Other metrics such as precision, recall, negative predictive value, true negatives, false positives, false negatives, and true positives were also estimated but not considered when building our ensemble model.

Uncertainty Quantification and Interpretability

The 95% PIs for our forecasted values were estimated to account for forecasting uncertainties. For statistical models (ARIMA, SARIMAX, and Prophet), PIs were readily obtained along with the model forecasts as a natural consequence of the model construction. However, PIs were not directly available for ML and DL techniques. Hence, we used Python packages that retroactively produced PIs using alternate techniques. The PIs for tree-based models (RF, XGB, and LGBM) were estimated by a bootstrapping technique using "skforecast" library, whereas for DL models (RNN, LSTM, and transformer), PIs were calculated by a nonparametric method known as quantile regression using the "Darts" library [31,32]. For each model forecast, we estimated the coverage probability of the 95% PIs (ie, instances where the PIs surrounded the true value) as a measure of calibration assessment for the prediction uncertainties.

Feature importance, based on Shapley Additive Explanation (SHAP) values for the best-performing models, was used as a generalized approach to interpretability for tree-based and DL models [34]. SARIMAX was the only statistical model that allowed for feature input and interpretability, which was provided by coefficient estimates from the model. However, the SARIMAX model was unable to incorporate the large number of features available for each region. To address this limitation, a feature selection approach was applied using LASSO regression to reduce the model's input variables.

The historic case count data are important and most frequently used predictive features in the ID prediction domain [9]. In our analysis, we used case count lags up to 6 months. To estimate the relative importance of these case count lags compared to other features, we calculated 2 metrics, namely, the feature ratio and mean reciprocal rank (MRR). A threshold of the top 10 features, as deemed by the models, was set to estimate the feature ratio and MRR. The feature ratio was the total number of lag case count features divided by the total number of features (up to the top 10 features). The MRR, on the other hand, also considers the position of the first relevant item in the ranked list, that is, case count lags. The MRR was defined as the mean of reciprocal ranks of case count lags across all features. The values range from 0 to 1, with a higher score signifying the greater relative feature importance rank of case count lags compared to other features for obtaining model forecasts.

Spatial Generalizability Using Transfer Learning

We used the transfer learning framework to test the models for their ability to produce accurate forecasts in regions that were fairly new to the disease. These regions were strategically picked by maximizing the global coverage and following the stratified random sampling framework. First, only the country-disease combinations where case count data were available were selected. This process also included the regions that reported 0 disease cases during our study period. Next, one region within the selected countries was randomly picked and denoted as the target region. Then, we selected a similar region to the target region to train the transfer learning models based on (1) the target and similar region being part of the same country and (2) the target and similar region containing comparable case counts over the last 6 months. The similar region pretrained model was then used to forecast the disease case counts of the target region, and performance was assessed.

Ethical Considerations

Data used in this study are open source and do not identify individual information, either directly or indirectly. Therefore, this research was exempted from ethical review.

Results

Overview

The summary statistics of the disease-location combinations aggregated at the country level are presented in Table 2, whereas a detailed breakdown of this summary for both test and train split is shown in Multimedia Appendix 1.

Table . Median (IQR) regional case counts per country and disease for the entire dataset from January 2014 to December 2018.

Disease	Country							
	Australia, median (IQR)	Germany, median (IQR)	Israel, median (IQR)	Japan, median (IQR)	Norway, median (IQR)	Saudi Arabia, median (IQR)	Sweden, median (IQR)	United States, median (IQR)
Brucellosis	3 (0-29)	11 (0-42)	30 (0-1072)	0 (0-3)	0 (0-2)	—[a]	1 (0-2)	4 (0-25)
Campylobacteriosis	7904 (0-43,205)	148,616 (50,644-197,616)	998 (0-7666)	—	306 (0-1456)	—	1103 (0-10,815)	1159 (0-25,327)
MERS[b]	0 (0-0)	—	—	0 (0-0)	—	3 (0-71)	—	—
Q fever	55 (0-974)	28 (0-769)	46 (0-103)	0 (0-1)	—	—	—	5 (0-26)
Tick-borne encephalitis	—	15 (0-897)	—	0 (0-2)	—	—	—	—
Tularemia	—	—	0 (0-0)	0 (0-1)	15 (0-69)	—	—	4 (0-89)

[a] Not available.

[b] MERS: Middle East respiratory syndrome.

There were 757 disease-region combinations ran through 9 different models in this study. As the first step, we calculated F_1-scores to determine the ability of the forecasting models to detect the presence or absence of a disease in a given region (Table 3). Overall, tree-based models (ie, XGB, RF, and LGBM) and Prophet consistently had better and comparable F_1-scores across the diseases. The DL models (ie, RNN, transformer, and LSTM) had significantly lower F_1-scores. Patterns were consistent across all diseases in each region. The additional evaluation metrics for model and disease combination including precision, recall, negative predictive value, true negatives, false positives, false negatives, and true positives are presented in Multimedia Appendix 1.

Table . F_1-score for all models forecasting the presence or absence of disease for each region in 2017 and 2018.

Disease	Model								
	ARIMA[a]	LGBM[b]	LSTM[c]	Prophet	RF[d]	RNN[e]	SARIMAX[f]	Transformer	XGB[g]
Brucellosis	0.65	0.67	0.55	0.69	0.67	0.56	0.66	0.59	0.67
Campylobacteriosis	0.93	0.96	0.96	0.97	0.96	0.95	0.91	0.94	0.96
MERS[h]	0.98	0.90	0.60	0.90	0.90	0.61	0.95	0.56	0.90
Q fever	0.78	0.78	0.64	0.79	0.78	0.67	0.74	0.65	0.78
Tick-borne encephalitis	0.68	0.98	0.59	0.97	0.98	0.55	0.95	0.54	0.98
Tularemia	0.54	0.73	0.49	0.71	0.73	0.50	0.72	0.55	0.73

[a] ARIMA: autoregressive integrated moving average.

[b] LGBM: light gradient boosting machine.

[c] LSTM: long short-term memory.

[d] RF: random forest.

[e] RNN: recurrent neural network.

[f] SARIMAX: seasonal autoregressive integrated moving average with exogenous factors.

[g] XGB: extreme gradient boosting.

[h] MERS: Middle East respiratory syndrome.

Next, we assessed the performance of each model for the subset of regions, where the total case counts were greater than 0 (ie, the disease was present) based on MAE. The model with the lowest MAE was inconsistent across diseases (Table 4).

Table . MAE[a] (95% CI) for each disease aggregated across all locations where the disease was present in 2017 and 2018.

Disease	Model								
	ARIMA[b], MAE (95% CI)	LGBM[c], MAE (95% CI)	LSTM[d], MAE (95% CI)	Prophet, MAE (95% CI)	RF[e], MAE (95% CI)	RNN[f], MAE (95% CI)	SARIMAX[g], MAE (95% CI)	Transformer, MAE (95% CI)	XGB[h], MAE (95% CI)
Brucellosis	0.8 (0.5 - 1.1)	0.8 (0.5 - 1.1)	0.84 (0.5 - 1.14)	0.9 (0.5 - 1.2)	0.8 (0.5 - 1.1)	0.9 (0.5 - 1.2)	1.1 (0.7 - 1.6)	0.8 (0.5 - 1.6)	0.8 (0.4 - 1.1)
Campylobacteriosis	237.9 (1.1 - 474.9)	43.8 (35.3 - 52.3)	45.2 (36.3 - 54.12)	43.4 (34.9 - 51.8)	39.8 (32.3 - 47.2)	46.6 (37.7 - 55.5)	243.7 (74.9 - 412.6)	50.9 (40.3 - 61.6)	43.1 (34.2 - 52.1)
Q fever	28.9 (25.4 - 83.1)	1.5 (0.9 - 2.1)	1.18 (0.8 - 1.6)	1.4 (0.9 - 1.9)	1.5 (0.8 - 2.1)	1.3 (0.9 - 1.8)	1.9 (1.2 - 2.7)	1.4 (0.9 - 2.0)	1.2 (0.9 - 1.5)
Tick-borne encephalitis	2.0 (0.6 - 3.3)	2.4 (0.7 - 4.1)	2.61 (0.7 - 4.53)	1.3 (0.5 - 2.2)	2.4 (0.6 - 4.1)	2.6 (0.7 - 4.6)	1.6 (0.6 - 2.6)	2.9 (0.7 - 5.2)	2.4 (0.6 - 4.1)
Tularemia	0.6 (0.4 - 0.7)	0.6 (0.5 - 0.7)	0.58 (0.5 - 0.71)	0.6 (0.5 - 0.7)	0.6 (0.5 - 0.7)	0.6 (0.4 - 0.7)	0.7 (0.5 - 0.9)	0.6 (0.5 - 0.7)	0.6 (0.5 - 0.7)

[a] MAE: mean absolute error.

[b] ARIMA: autoregressive integrated moving average.

[c] LGBM: light gradient boosting machine.

[d] LSTM: long short-term memory.

[e] RF: random forest.

[f] RNN: recurrent neural network.

[g] SARIMAX: seasonal autoregressive integrated moving average with exogenous factors.

[h] XGB: extreme gradient boosting.

The MAE of the best-performing model for each disease across all locations (with case counts greater than 0) is presented in Table 5, which we define as our "ensemble" model. Using this technique and when averaging across all regions, prediction years, and diseases, we observed that the ensemble model had the lowest MAE compared to other models (Table 6). While the difference does not appear to be a large amount, the difference is considerable, and the magnitude is visually minimized by the large error in the ARIMA and SARIMAX models. All the models included in our study had similar MAE except for ARIMA and SARIMAX since they had a few outlier predictions that skewed their MAE.

Table . Ensemble (best performing) models for each disease averaged across all regions where the disease was present in 2017 and 2018.

Disease	Best model	MAE[a] (95% CI)
Brucellosis	LGBM[b]	0.8 (0.5 - 1.1)
Campylobacteriosis	RF[c]	39.8 (32.3 - 47.2)
Q fever	LSTM[d]	1.2 (0.8 - 1.6)
Tick-borne encephalitis	Prophet	1.4 (0.5 - 2.2)
Tularemia	ARIMA[e]	0.6 (0.4 - 0.7)

[a] MAE: mean absolute error.

[b] LGBM: light gradient boosting machine.

[c] RF: random forest.

[d] LSTM: long short-term memory.

[e] ARIMA: autoregressive integrated moving average.

Table . MAE[a] (95% CIs) for each model aggregated across all diseases and regions where the disease was present in 2017 and 2018.

Model	MAE (95% CI)
ARIMA[b]	115.2 (5.2 - 225.2)
Ensemble	8.7 (6.5 - 23.9)
LGBM[c]	20.8 (16.5 - 25.1)
LSTM[d]	21.4 (16.9 - 25.9)
Prophet	20.6 (16.3 - 24.9)
RF[e]	19.0 (15.2 - 22.8)
RNN[f]	22.1 (17.6 - 26.7)
SARIMAX[g]	113.1 (34.6 - 191.6)
Transformer	24.2 (18.8 - 29.5)
XGB[h]	20.5 (16.0 - 25.0)

[a] MAE: mean absolute error.

[b] ARIMA: autoregressive integrated moving average.

[c] LGBM: light gradient boosting machine.

[d] LSTM: long short-term memory.

[e] RF: random forest.

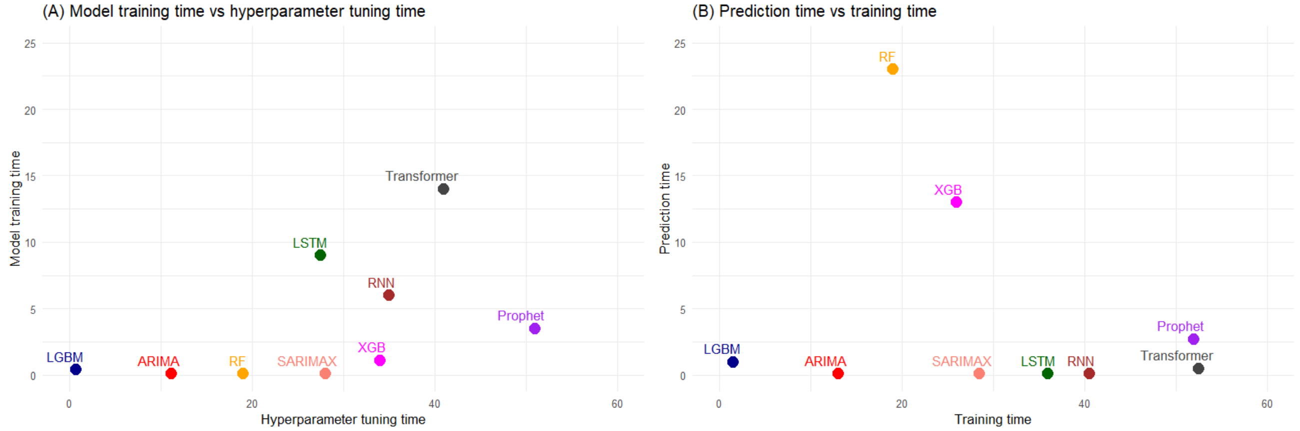
[f] RNN: recurrent neural network.

[g] SARIMAX: seasonal autoregressive integrated moving average with exogenous factors.

[h] XGB: extreme gradient boosting.

The average training time, which included hyperparameter tuning and model fitting, along with prediction time (ie, the time to generate 1-step ahead forecasts and their PIs) for each model averaged across all diseases, regions, and periods are presented in Figure 2. The LGBM followed by ARIMA were the most time-efficient models. Although RF and XGB had a training time comparable with other models, their average prediction time was considerably higher mainly because of the increased amount of time required by them to generate 95% PIs using the bootstrapping technique.

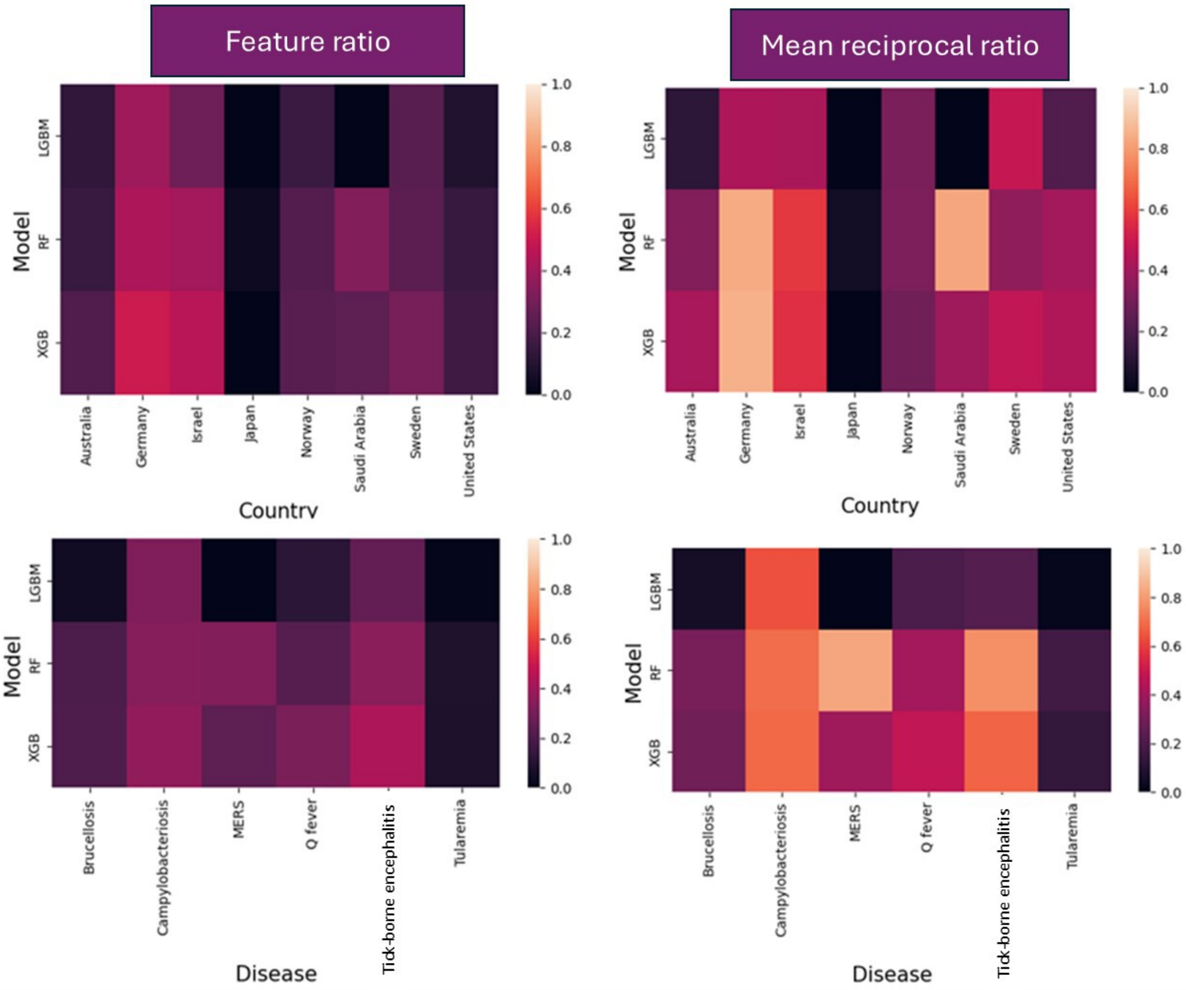
Figure 2. (A) Average hyperparameter tuning and model fitting time and (B) average training and prediction time for each model averaged across all diseases, locations, and periods. ARIMA: autoregressive integrated moving average; LGBM: light gradient boosting machine; LSTM: long short-term memory; RF: random forest; RNN: recurrent neural network; SARIMAX: seasonal autoregressive integrated moving average with exogenous factors; XGB: extreme gradient boosting.



Feature Importance

For the tree-based models (ie, LGBM, RF, and XGB), we selected up to 10 top features based on SHAP values that the model deemed important and calculated feature ratio and MRR to show how lag case count may be impacting our predictive performance (Figure 3). Overall, the feature ratios were low (<0.4) when aggregated across countries and diseases, suggesting that non–lag case count features were more important for forecasting compared to case lag data. Campylobacteriosis and tick-borne encephalitis had higher MRR compared to other diseases, indicating that lag case count features were ranked relatively higher among the top 10 features for these diseases. However, the low feature ratio and MRR for many diseases highlight how the additional feature variables can be valuable in disease forecasting, especially for less prevalent diseases. A complete breakdown of feature ratios and MRRs for each country-disease combination is presented in Multimedia Appendix 1.

Figure 3. Relative importance of case count lags to produce forecasts for tree-based models quantified as feature ratios and mean reciprocal ranks of the most important features (up to the top 10) aggregated across countries (top row) and diseases (bottom row). Higher scores signify the greater relative importance of case count lags compared to other features for obtaining model forecasts. LGBM: light gradient boosting machine; MERS: Middle East respiratory syndrome; RF: random forest; XGB: extreme gradient boosting.



The coverage probability of the forecasting models represented by the percent coverage of the true observations by 95% PIs is presented in Table 7. The LGBM had a 94.4% (n=14,707) coverage, indicating that the PIs were well calibrated, even better than traditional statistical models like ARIMA and SARIMAX. On the other hand, the DL models had less than 10% coverage, highlighting the fact that these models were not thoroughly calibrated. This observation was true when the coverage percentage was broken down across disease and country combinations in Multimedia Appendix 1.

Table . Coverage percentage of the 95% prediction intervals of the forecasting models[a].

Model	Coverage, n (%)
ARIMA[b]	14,582 (93.6)
LGBM[c]	14,707 (94.4)
LSTM[d]	1277 (8.2)
Prophet	13,086 (84)
RF[e]	13,756 (88.3)
RNN[f]	1246 (8)
SARIMAX[g]	13,959 (89.6)
Transformer	1262 (8.1)
XGB[h]	13,491 (86.6)

[a] The ideal coverage percentage is 95%.

[b] ARIMA: autoregressive integrated moving average.

[c] LGBM: light gradient boosting machine.

[d] LSTM: long short-term memory.

[e] RF: random forest.

[f] RNN: recurrent neural network.

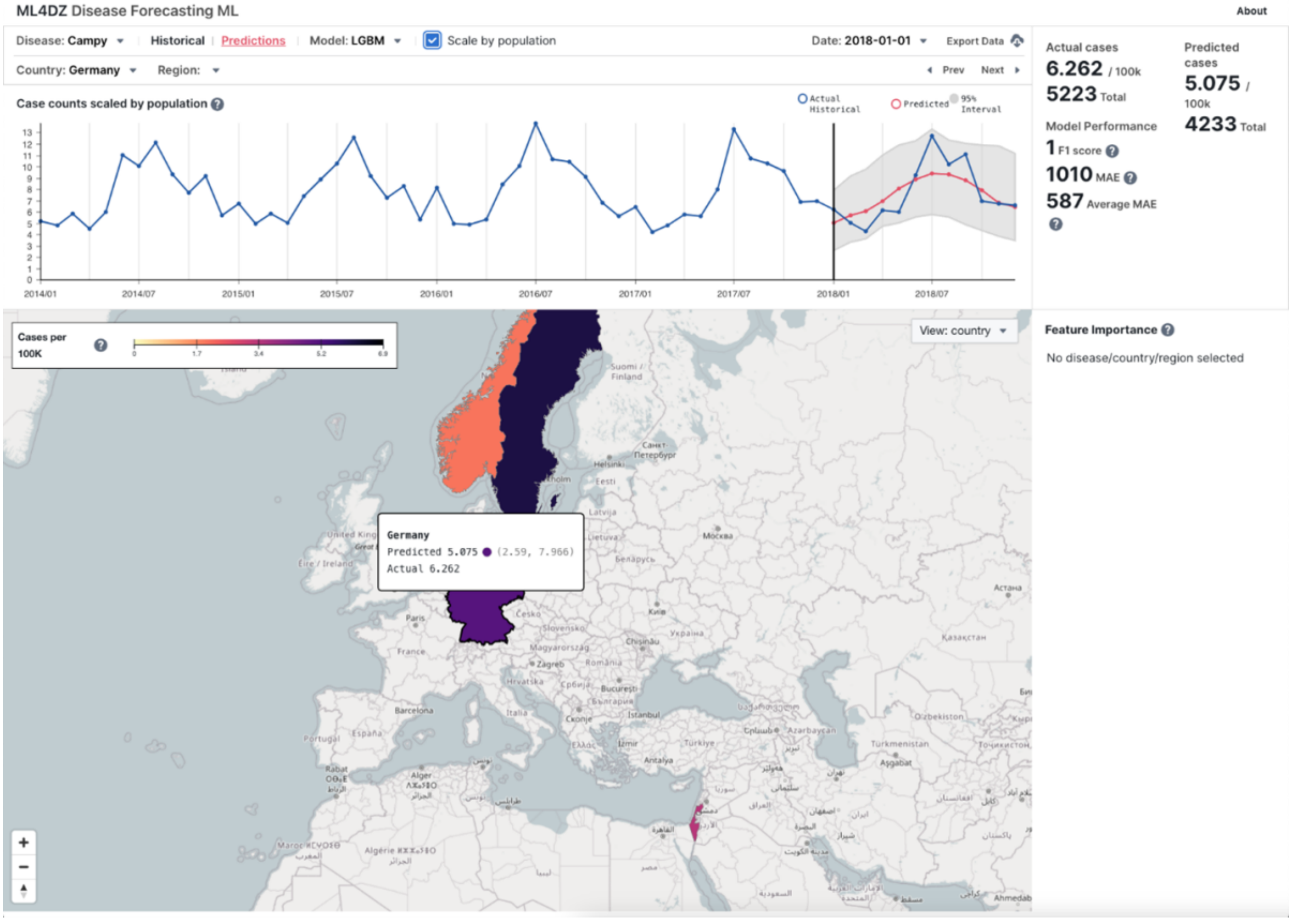
[g] SARIMAX: seasonal autoregressive integrated moving average with exogenous factors.

[h] XGB: extreme gradient boosting.

Results Visualization

We created an interactable dashboard to visualize the predictions made by forecasting models (Figure 4). The information presented in the user interface includes a time series of actual and predicted values with 95% PI, model performance metrics, and feature importance values for each disease and country or region. In Figure 5, campylobacteriosis in Schleswig-Holstein, Germany, for the year 2018 using the LGBM model is presented as an example, showing uncertainty quantification by PIs and feature importance as mean SHAP values. The lagging case count feature for the past 6 periods was the most important feature for the model, but variables related to global health spending and monetary export value of cattle also ranked high in feature importance.

Figure 4. Dashboard in prediction view with values based on case counts scaled by population. LGBM: light gradient boosting machine; MAE: mean absolute error; ML: machine learning.



ML4DZ Disease Forecasting ML

Disease: Campy | Historical | Predictions | Model: LGBM | ☑ Scale by population | Date: 2018-01-01

Country: Germany | Region:

Actual cases 6.262 / 100k 5223 Total

Predicted cases 5.075 / 100k 4233 Total

Model Performance

1 F1 score

1010 MAE

587 Average MAE

Feature Importance

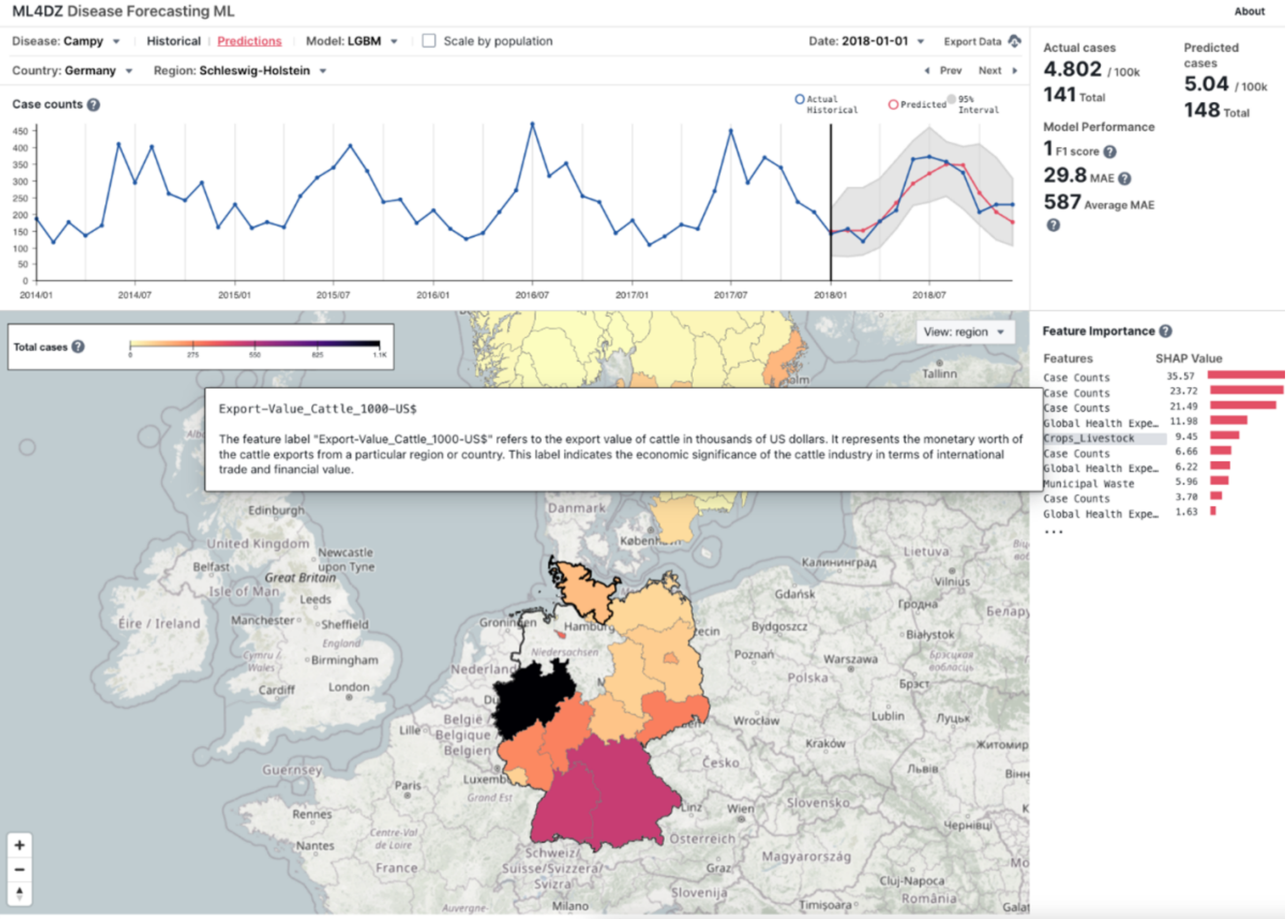
No disease/country/region selected

Germany

Predicted 5.075 ● (2.59, 7.966)

Actual 6.262

Figure 5. Dashboard with disease (“campy”=campylobacteriosis) and region (Schleswig-Holstein, Germany) selected in prediction view, which includes model performance statistics in summary and model feature importance list. Pop-out box provides a detailed explanation of the clicked and highlighted feature. LGBM: light gradient boosting machine; MAE: mean absolute error; ML: machine learning.



ML4DZ Disease Forecasting ML

Disease: Campy | Historical | Predictions | Model: LGBM | ☐ Scale by population | Date: 2018-01-01 | Export Data

Country: Germany | Region: Schleswig-Holstein

Actual cases 4.802 / 100k, 141 Total

Predicted cases 5.04 / 100k, 148 Total

Model Performance

1 F1 score

29.8 MAE

587 Average MAE

Export-Value_Cattle_1000-US$

The feature label "Export-Value_Cattle_1000-US$" refers to the export value of cattle in thousands of US dollars. It represents the monetary worth of the cattle exports from a particular region or country. This label indicates the economic significance of the cattle industry in terms of international trade and financial value.

Feature Importance

Features	SHAP Value
Case Counts	35.57
Case Counts	23.72
Case Counts	21.49
Global Health Expe…	11.98
Crops_Livestock	9.45
Case Counts	6.66
Global Health Expe…	6.22
Municipal Waste	5.96
Case Counts	3.70
Global Health Expe…	1.63
…	

Spatial Generalizability Using Transfer Learning

The spatial generalizability for the forecasting models was tested based on the performance of the pretrained ensemble model from the location with a comparable disease pattern (ie, similar region) to forecast the disease case counts of the selected region (ie, target region). The model tested (ensemble or best model) varied by location and included statistical (ARIMA: n=3) and tree-based models (LGBM: n=2 and XGB: n=1). In most cases, the MAE of a similar region was close to and often lower than the MAE of the target region (Table 8). However, campylobacteriosis forecasting in Kalmar, Sweden, was an exception, where the model MAE increased from 5.01 to 16.98 when Västmanland feature data were used for training. It is important to note that these are pairwise comparisons of a single metric calculated across a single region and do not include confidence bounds to determine if any differences are statistically significant. However, given that over half of the disease-region models produced an equal or lower MAE when testing in a new region, if a power analysis was conducted assuming a binary outcome and a continued trend, no amount of data would be sufficient to show that the transfer learning results would be significantly worse.

Table . The mean absolute error (MAE) values representing spatial generalizability of transfer learning models[a].

Disease	Country	Target region (cases in the past 6 months)	Ensemble model	Similar region (cases in the past 6 months)	MAE of baseline forecast	MAE of transfer learning forecast
Brucellosis	Israel	Jerusalem (34)	LGBM[b]	Akko (32)	5.41	2.3
Campylobacteriosis	Sweden	Kalmar (122)	XGB[c]	Västmanland (124)	5.01	16.98
MERS[d]	Saudi Arabia	Bisha (0)	ARIMA[e]	Eastern Province (0)	0.03	0.02
Q fever	Australia	New South Wales (37)	ARIMA	Queensland (48)	2.85	2.00
Tick-borne encephalitis	Germany	Saxony-Anhalt (3)	ARIMA	Brandenburg (3)	0.20	0.25
Tularemia	Japan	Ibaraki (0)	LGBM	Saga (0)	0.00	0.00

[a] For baseline forecasts, the models were trained and tested on target regions. For transfer learning forecasts, the models were trained on similar regions and tested on target regions.

[b] LGBM: light gradient boosting machine.

[c] XGB: extreme gradient boosting.

[d] MERS: Middle East respiratory syndrome.

[e] ARIMA: autoregressive integrated moving average.

Discussion

Principal Findings

As ML and DL techniques gain popularity in the ID forecasting domain, they are being extensively used across a wide range of pathogens with diverse ecology, geographic, and temporal scales. However, there is limited consensus among the forecasting community regarding the application and reporting of these results in a manner that lends themselves useful for operational decision-making in real-world circumstances [9]. In this study, we present a universal pipeline for analyzing data, generalizing the results, and automating reporting. We also address crucial operational metrics such as forecasting accuracy, computational efficiency, spatiotemporal generalizability, uncertainty quantification, and interpretability, which are essential in an operational scenario. This generalized and automated analytic pipeline is a major step toward addressing operational demands in the ID forecasting domain and to better inform public health and veterinary policies. We included 6 IDs with diverse transmission dynamics spanning 8 countries and 213 regions. Additionally, we trained our models using a broad range of data from demographic, geographic, climatic, and socioeconomic factors within the One Health landscape. This comprehensive approach enables our analysis to capture the intricate interplay of variables that drive infectious disease presence and transmission, more closely reflecting the complex realities observed in the real world.

Model Performance

We assessed forecasting model accuracies by their ability to detect the presence of a disease based on F_1-scores and by their ability to accurately forecast case counts when present using MAE. This 2-step approach prevented the inflation of model accuracy by the regions that did not encounter disease presence during our study period. Though we did not find a single best model suitable for all disease, region, and country combinations, the tree-based models were consistently better than all other models at detecting both disease presence and actual case counts with better F_1-scores and MAE values, respectively. Our ensemble technique, which selects the best-performing model for each disease-location pair, demonstrates superior forecasting performance by reducing prediction noise compared to single-model approaches, aligning with previous research [35-37]. While ideal ensemble methods would dynamically update to identify the best model at each time stamp, this demands extensive data, computational resources, and continuous validation. To address these challenges, we opted for a simpler approach by selecting the best model based on MAE for the testing dataset, minimizing resource and data requirements. However, ensemble techniques still require substantial computational power and analytic pipelines, posing challenges for low- and middle-income countries with limited resources. Investments in infrastructure, skill development, and data-sharing initiatives are crucial to fully leverage ML and DL techniques for ID forecasting in such settings.

Computational Efficiency

The computation time required for ID forecasting can broadly be split into model training and model prediction time. In our analysis, much of the training time was spent on hyperparameter tuning. Autoregressive and tree-based models took less time to tune than DL models. On the other hand, tree-based models (ie, RF and XGB) took considerably longer to produce predictions compared to the other models. This increase in time was mainly because of the time required to compute bootstrap PIs. The LGBM was an exception, as it used the least computational time for both training and testing compared to all the other models. Since our forecasting pipeline performed stepwise forecasting for each month with models retrained only once a year, more time was spent on predicting rather than model training. Hence, when building a forecasting pipeline where the models are retrained once a year and predictions are made every month, it is computationally optimal to choose models with lower prediction time rather than training time.

The accuracy of modeling techniques can drastically differ with the amount of time and computational resources spent on fine-tuning the model. Here, we chose the initial hyperparameter values for optimization based on their relative importance and time required for overall analysis with the goal of achieving comparable results. It is important to find a balance between model complexity and computational efficiency according to the individual operational needs and available resources.

Uncertainty Quantification

Since the effectiveness of health policies and operational decision-making is driven by the accuracy of forecasts, policy makers need to know the credibility of the models in the form of prediction uncertainties [38]. In traditional statistical techniques, such as ARIMA and SARIMAX, forecasting uncertainties are computed based on probability theory and a set of statistical assumptions and, therefore, are readily available. However, generating such PIs is rather complicated in ML and DL compared to traditional methods due to additional uncertainties associated with noise distributions, hyperparameters, overparameterization, and optimization that should be accounted for [39]. Therefore, most of the ID forecasting studies fail to report their model uncertainty despite their recent popularity [9]. In our study, we used bootstrapping and probabilistic methods to compute PIs around our point estimates of ML and DL methods. Such bounds are crucial to assessing future uncertainties, making operational decisions, planning policies for a range of possible outcomes, and comparing different forecast models thoroughly [40]. The LGBM was the best-calibrated model with almost 95% of true observations covered by 95% of PIs, followed by other tree-based and statistical models. The DL models had the least desirable coverage along with their inferior predictive performance and narrow prediction band. This may not be surprising, though as there is evidence in the literature that uncertainty estimates around DL predictions often fail to capture the true data distribution [41].

Spatiotemporal Generalizability

Overall, the predictive performance of transfer learning models was comparable to their base models as presented in Table 8. For example, using the Akko model to predict case counts of brucellosis in Jerusalem created an increase in predictive accuracy. This suggests the potential for spatial generalizability of our ensemble, that is, an ability to predict existing diseases in new regions with low error. However, campylobacteriosis forecasting in Kalmar, Sweden, was an exception, where the model MAE increased when Västmanland feature data were used for training. Since campylobacteriosis tends to have higher case counts, the difference in MAE cannot be directly comparable to the other diseases with much lower case counts and naturally lower MAE values. There are many possible reasons for the difference seen in the results. For example, the similar region model may not have contained the necessary similarities to the target region required to enable a lower prediction error, such as differences in regional management practices. The ensemble model (best-performing model for disease-location pair) was different between the test cases and included both statistical (only case counts) and tree-based (many features) models. More advanced transfer learning techniques that include multiple input features to identify similarity and emphasize model retraining by reallocating model weights between target and similar regions could be considered to improve spatial generalizability [2,42]. Regardless, these results are encouraging, and a more comprehensive study is warranted.

Model Interpretability

We included interpretability for each model in the form of feature importance, a critical metric that is often neglected in the ID forecasting domain [9]. Interpretability provides insight into the decision-making process of the prediction systems, which is crucial for implementing appropriate disease mitigation strategies. Our scoping systematic review of ID prediction models showed that historic case counts are the most commonly used input feature in the ID forecasting domain, while the other important predictors, such as climate, demographics, socioeconomics, and geography, are often neglected [9]. In this study, we estimated lower feature ratio and MRR for lag case count features across countries and diseases, indicating that the non–case count features were equally important, if not more, compared to case count lags. This study suggests that including data related to the full disease ecology is critical for obtaining accurate, reliable, and interpretable ID forecasts. For example, cattle export was one of the important informative non–case count features for forecasting campylobacteriosis in Schleswig-Holstein, Germany. A direct causal association between the disease and cattle exports cannot be made just by a feature importance plot. However, campylobacteriosis is the most commonly reported bacterial food-borne gastrointestinal infection in the European Union that is closely associated with the dairy industry [43]. Our results suggest that high dairy and other agriculture activities, including cattle trading in Schleswig-Holstein, could play an important role in the disease prevalence in the region and, if adjusted, could result in disease mitigation.

Data Visualization

To make the best-performing models and corresponding predictions accessible to decision makers, we developed an interactive data visualization dashboard. This dashboard provides the raw data, the model predictions with accompanying 95% PIs and feature importance, overall model performance, and an interactive global map. This dashboard allows a user to visualize forecasts for up to a year into the future in any country or region where the data exist and can be used to inform control measures to reduce the spread of an ID within a country or region.

Conclusions

Our study provides a generalized platform to analyze and report ID forecasts with an emphasis on analytical accuracy, computational efficiency, uncertainty quantification, interpretability, and generalizability. These 5 aspects are crucial in determining the forecasting approach optimal for each situation’s operational needs. While all the forecasting techniques come with their own strengths and weaknesses, choosing an optimal approach is usually a tradeoff between computational efficiency, model complexity, and forecasting accuracy. Recognizing and addressing such nuances will facilitate the use of ID forecasts in an operational environment for better preparedness and response during an ID emergency.

Acknowledgments

This work was funded by the Defense Threat Reduction Agency (project CB11029).

Data Availability

The datasets generated or analyzed during this study are publicly available with website links in the Methods section, including Table 1.

Authors' Contributions

LEC and KTP acquired the funding. RK, KTP, SD, SE, and LEC conceptualized and designed the study. RK, CH, SD, and SE analyzed the data. RK, CH, and KTP created visualizations. RK drafted the manuscript. RK, KTP, CH, and LEC reviewed and edited the manuscript. All authors agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and tables.
[DOCX File, 395 KB - publichealth_v11i1e59971_app1.docx]

References

1. Bloom RM, Buckeridge DL, Cheng KE. Finding leading indicators for disease outbreaks: filtering, cross-correlation, and caveats. J Am Med Inform Assoc 2007;14(1):76-85. [doi: 10.1197/jamia.M2178] [Medline: 17068353]
2. Keshavamurthy R, Charles LE. Predicting Kyasanur forest disease in resource-limited settings using event-based surveillance and transfer learning. Sci Rep 2023 Jul 8;13(1):11067. [doi: 10.1038/s41598-023-38074-0] [Medline: 37422454]
3. Woolhouse M. How to make predictions about future infectious disease risks. Philos Trans R Soc Lond B Biol Sci 2011 Jul 12;366(1573):2045-2054. [doi: 10.1098/rstb.2010.0387] [Medline: 21624924]
4. Bansal S, Chowell G, Simonsen L, Vespignani A, Viboud C. Big data for infectious disease surveillance and modeling. J Infect Dis 2016 Dec 1;214(Suppl 4):S375-S379. [doi: 10.1093/infdis/jiw400] [Medline: 28830113]
5. Dixon S, Keshavamurthy R, Farber DH, Stevens A, Pazdernik KT, Charles LE. A comparison of infectious disease forecasting methods across locations, diseases, and time. Pathogens 2022 Jan 29;11(2):185. [doi: 10.3390/pathogens11020185] [Medline: 35215129]
6. Shen L, Jiang C, Sun M, et al. Predicting the spatial-temporal distribution of human brucellosis in Europe based on convolutional long short-term memory network. Can J Infect Dis Med Microbiol 2022;2022:7658880. [doi: 10.1155/2022/7658880] [Medline: 35967090]
7. Da’ar OB, Ahmed AE. Underlying trend, seasonality, prediction, forecasting and the contribution of risk factors: an analysis of globally reported cases of Middle East Respiratory Syndrome Coronavirus. Epidemiol Infect 2018 Aug;146(11):1343-1349. [doi: 10.1017/S0950268818001541] [Medline: 29886854]
8. Adnan M, Gao X, Bai X, et al. Potential early identification of a large Campylobacter outbreak using alternative surveillance data sources: autoregressive modelling and spatiotemporal clustering. JMIR Public Health Surveill 2020 Sep 17;6(3):e18281. [doi: 10.2196/18281] [Medline: 32940617]
9. Keshavamurthy R, Dixon S, Pazdernik KT, Charles LE. Predicting infectious disease for biopreparedness and response: a systematic review of machine learning and deep learning approaches. One Health 2022 Dec;15:100439. [doi: 10.1016/j.onehlt.2022.100439] [Medline: 36277100]
10. Chae S, Kwon S, Lee D. Predicting infectious disease using deep learning and big data. Int J Environ Res Public Health 2018 Jul 27;15(8):1596. [doi: 10.3390/ijerph15081596] [Medline: 30060525]
11. Oidtman RJ, Omodei E, Kraemer MUG, et al. Trade-offs between individual and ensemble forecasts of an emerging infectious disease. Nat Commun 2021 Sep 10;12(1):5379. [doi: 10.1038/s41467-021-25695-0] [Medline: 34508077]
12. Reich NG, McGowan CJ, Yamana TK, et al. Accuracy of real-time multi-model ensemble forecasts for seasonal influenza in the U.S. PLOS Comput Biol 2019 Nov;15(11):e1007486. [doi: 10.1371/journal.pcbi.1007486] [Medline: 31756193]
13. Ma S, Ning S, Yang S. Joint COVID-19 and influenza-like illness forecasts in the United States using internet search information. Commun Med (Lond) 2023 Mar 24;3(1):39. [doi: 10.1038/s43856-023-00272-2] [Medline: 36964311]

14. Al-Jarrah OY, Yoo PD, Muhaidat S, Karagiannidis GK, Taha K. Efficient machine learning for big data: a review. Big Data Research 2015 Sep;2(3):87-93. [doi: 10.1016/j.bdr.2015.04.001]
15. Bhatt U, Antorán J, Zhang Y, et al. Uncertainty as a form of transparency: measuring, communicating, and using uncertainty. Presented at: Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society; May 19-21, 2021; Virtual Event, United States p. 401-413. [doi: 10.1145/3461702.3462571]
16. Generous N, Fairchild G, Khalsa H, Tasseff B, Arnold J. Epi Archive: automated data collection of notifiable disease data. OJPHI 2017 May 2;9(1). [doi: 10.5210/ojphi.v9i1.7615]
17. Welte VR, Vargas Terán M. Emergency Prevention System (EMPRES) for transboundary animal and plant pests and diseases. The EMPRES-livestock: an FAO initiative. Ann N Y Acad Sci 2004 Oct;1026:19-31. [doi: 10.1196/annals.1307.003] [Medline: 15604466]
18. GADM. URL: https://gadm.org/ [accessed 2024-02-24]
19. GES DISC. URL: https://disc.gsfc.nasa.gov/ [accessed 2024-02-24]
20. Earth Science Data Systems. NASA. 2024. URL: https://www.earthdata.nasa.gov/ [accessed 2024-02-24]
21. US Department of Agriculture. Economic Research Service—Home. URL: https://www.ers.usda.gov/ [accessed 2024-02-24]
22. DIVA-GIS | free, simple & effective. URL: https://diva-gis.org/ [accessed 2024-02-24]
23. World Health Organization (WHO). URL: https://www.who.int [accessed 2024-02-24]
24. OECD Statistics. URL: https://stats.oecd.org/ [accessed 2024-02-24]
25. Natural Earth. Free vector and raster map data at 1:10m, 1:50m, and 1:110m scales. URL: https://www.naturalearthdata.com/ [accessed 2024-02-24]
26. Population division. United Nations. URL: https://www.un.org/development/desa/pd/ [accessed 2024-02-24]
27. SEDAC: Socioeconomic Data and Applications Center. NASA. URL: https://sedac.ciesin.columbia.edu/ [accessed 2024-02-24]
28. Josef P, Skipper S, Kevin S, et al. Statsmodels: econometric and statistical modeling with python. 2010 Presented at: Proceedings of the 9th Python in Science Conference; Texas, US URL: https://www.statsmodels.org/ [accessed 2025-02-28]
29. scikit-learn: machine learning in Python. URL: https://scikit-learn.org/stable/ [accessed 2024-02-24]
30. Taylor SJ, Letham B. Forecasting at scale. PeerJ Preprints 2017 Sep;5:e3190v2. [doi: 10.7287/peerj.preprints.3190]
31. Darts: time series made easy in Python. Unit8. URL: https://unit8.com/resources/darts-time-series-made-easy-in-python/ [accessed 2024-02-24]
32. Amat Rodrigo J, Escobar Ortiz J. Skforecast. Zenodo. 2024. URL: https://zenodo.org/records/8382788 [accessed 2025-02-28]
33. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. Presented at: Advances in Neural Information Processing Systems; Dec 4-9, 2017; California, United States URL: https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html
34. Topical overviews. SHAP Latest Documentation. URL: https://shap.readthedocs.io/en/latest/overviews.html [accessed 2024-03-02]
35. Alfred R, Obit JH. The roles of machine learning methods in limiting the spread of deadly diseases: a systematic review. Heliyon 2021 Jun;7(6):e07371. [doi: 10.1016/j.heliyon.2021.e07371] [Medline: 34179541]
36. Khan SD, Alarabi L, Basalamah S. Toward smart lockdown: a novel approach for COVID-19 hotspots prediction using a deep hybrid neural network. Computers 2020;9(4):99. [doi: 10.3390/computers9040099]
37. Kasilingam D, Sathiya Prabhakaran SP, Rajendran DK, Rajagopal V, Santhosh Kumar T, Soundararaj A. Exploring the growth of COVID-19 cases using exponential modelling across 42 countries and predicting signs of early containment using machine learning. Transbound Emerg Dis 2021 May;68(3):1001-1018. [doi: 10.1111/tbed.13764] [Medline: 32749759]
38. Gilbert JA, Meyers LA, Galvani AP, Townsend JP. Probabilistic uncertainty analysis of epidemiological modeling to guide public health intervention policy. Epidemics 2014 Mar;6:37-45. [doi: 10.1016/j.epidem.2013.11.002] [Medline: 24593920]
39. Psaros AF, Meng X, Zou Z, Guo L, Karniadakis GE. Uncertainty quantification in scientific machine learning: methods, metrics, and comparisons. J Comput Phys 2023 Mar;477:111902. [doi: 10.1016/j.jcp.2022.111902]
40. Chatfield C. Prediction intervals for time-series forecasting. In: Armstrong JS, editor. Principles of Forecasting: A Handbook for Researchers and Practitioners: Springer; 2001:475-494. [doi: 10.1007/978-0-306-47630-3_21ISBN:978-0-306-47630-3]
41. Lakshminarayanan B, Pritzel A, Blundell C. Simple and scalable predictive uncertainty estimation using deep ensembles. Presented at: Advances in Neural Information Processing Systems; Dec 4-9, 2017; California, United States URL: https://proceedings.neurips.cc/paper_files/paper/2017/hash/9ef2ed4b7fd2c810847ffa5fa85bce38-Abstract.html [accessed 2025-02-28]
42. Roster K, Connaughton C, Rodrigues FA. Forecasting new diseases in low-data settings using transfer learning. Chaos Solitons Fractals 2022 Aug;161:112306. [doi: 10.1016/j.chaos.2022.112306] [Medline: 35765601]
43. Henningsen A, Henning C, Struve C, Mueller-Scheessel J. Economic impact of the mid-term review on agricultural production, farm income and farm survival: a quantitative analysis for local sub-regions of Schleswig-Holstein in Germany. Presented at: 2005 International Congress; Aug 23-27, 2005; Copenhagen, Denmark. [doi: 10.22004/ag.econ.24644]

Abbreviations

ARIMA: autoregressive integrated moving average
DL: deep learning
ID: infectious disease
LASSO: least absolute shrinkage and selection operator
LGBM: light gradient boosting machine
LSTM: long short-term memory
MAE: mean absolute error
ML: machine learning
MRR: mean reciprocal rank
PI: prediction interval
RF: random forest
RNN: recurrent neural network
SARIMAX: seasonal autoregressive integrated moving average with exogenous factors
SHAP: Shapley Additive Explanation
XGB: extreme gradient boosting

Edited by KS Sahu; submitted 27.04.24; peer-reviewed by S Ma, Y Hua; revised version received 10.12.24; accepted 24.12.24; published 21.03.25.

Please cite as:
Keshavamurthy R, Pazdernik KT, Ham C, Dixon S, Erwin S, Charles LE
Meeting Global Health Needs via Infectious Disease Forecasting: Development of a Reliable Data-Driven Framework
JMIR Public Health Surveill 2025;11:e59971
URL: https://publichealth.jmir.org/2025/1/e59971
doi:10.2196/59971

Vaccine Hesitancy and Associated Factors Among Caregivers of Children With Special Health Care Needs in the COVID-19 Era in China: Cross-Sectional Study

Mingyan Li[1*], MD; Changxuan Sun[2*], MSc; Chai Ji[1], MD; Meiying Gao[1], MSc; Xia Wang[1], MSc; Dan Yao[1], MD; Junxia Guo[1], MSc; Lidan Sun[1], MA; Abdul Rafay[3], HS; Antonita Shereen George[4], MBBS; Sanduni Hasara Samararathna Samararathna Muhandiramge[4], MBBS; Guannan Bai[1], PhD

[1]Department of Child Health Care, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, China

[2]The Fourth Affiliated Hospital of Soochow University, Medical Center of Soochow University, Suzhou Dushu Lake Hospital, Suzhou, China

[3]School of Medicine, Imperial College London, London, United Kingdom

[4]First Affiliated Hospital of Zhejiang University, Zhejiang University School of Medicine, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Guannan Bai, PhD
Department of Child Health Care, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, China

Abstract

Background: Immunization is a cost-effective way to prevent infectious diseases in children, but parental hesitancy leads to low vaccination rates, leaving children at risk. Caregivers of children with special health care needs are more hesitant about vaccines than those of healthy children.

Objective: The aim of the study is to investigate the changes in caregivers' vaccination hesitation of children with special health care needs before, during, and after the COVID-19 pandemic in China and to identify associated factors for caregivers' attitudes toward National Immunization Program (NIP) and non-NIP vaccines.

Methods: We included 7770 caregivers of children with special health care needs (median age 7.0, IQR 2.4-24.1 months) who visited the Vaccination Consultation Clinic at Children's Hospital, Zhejiang University School of Medicine (Hangzhou, China) from May 2017 to May 2023. General and clinical information was extracted from the immunization evaluation system for children with special health care needs and medical records. We compared the differences in caregivers' willingness and hesitation for vaccinating their children across the 3 stages of the COVID-19 pandemic using chi-square tests. Multinomial logistic regression models were used to identify independent variables that were associated with caregivers' willingness and hesitation toward NIP and non-NIP vaccines.

Results: There is a statistically significant difference in caregivers' vaccine hesitancy before, during, and after the COVID-19 pandemic ($P<.05$). During the COVID-19 pandemic, the percentages of choosing NIP, alternative non-NIP, and non-NIP vaccines are highest (n=1428, 26%, n=3148, 57.4%, and n=3442, 62.7%, respectively) than those at other 2 stages. In comparison, caregivers' hesitation toward NIP and non-NIP vaccines is lowest (n=911, 16.6% and n=2045, 37.3%, respectively). Despite the stages of the COVID-19 pandemic, multiple factors, including children's age and sex, parents' educational level, comorbidities, and history of allergy, were significantly associated with caregivers' attitude toward NIP and non-NIP vaccines ($P<.05$). The profiles of risk factors for hesitancy toward NIP and non-NIP vaccines are different, as indicated by the results from the logistic regression models.

Conclusions: This study demonstrated that caregivers' willingness to vaccinate their children with special health care needs with NIP and non-NIP vaccines was highest during the COVID-19 pandemic in China, and their hesitancy was lowest. Additionally, we have identified multiple factors associated with caregivers' willingness and hesitancy to vaccinate their children. These findings provide evidence-based support for developing personalized health education strategies.

(JMIR Public Health Surveill 2025;11:e67487) doi:10.2196/67487

KEYWORDS

COVID-19; caregivers; children with special health care needs; vaccination hesitancy; decision-making

Introduction

Immunization is widely recognized as an exceptionally cost-effective strategy for preventing and eliminating infectious diseases in children. In China, a dual-tier vaccination program is implemented. The government provides vaccines under the National Immunization Program (NIP) at no cost across the nation, whereas non-NIP vaccines incur some costs [1]. It is mandatory for all children to receive NIP vaccines according to governmental guidelines. The NIP vaccines in China, which include 14 vaccines targeting 15 diseases, have achieved a coverage rate exceeding 95% among children younger than 6 years of age at the national level [2,3]. In contrast, non-NIP vaccines, which complement the NIP vaccines and are crucial for comprehensive disease prevention and control, exhibit relatively low coverage rates. These rates vary significantly across different regions and socioeconomic groups primarily because non-NIP vaccines are voluntary, self-funded, and not required for school enrollment. Zhou et al [4] observed that the coverage of non-NIP vaccines was lowest among left-behind families and highest among local urban families. Generally, the coverage rates for most non-NIP vaccines are below 50%, with specific rates ranging from 1.8% for the third dose of the rotavirus vaccine to 67.1% for the first dose of the varicella vaccine [5].

Numerous factors can affect vaccine coverage rates. Studies demonstrate a direct correlation between caregivers' attitudes toward vaccines and vaccination rates [6]. A significant number of caregivers express hesitancy regarding vaccinating their children often due to concerns about safety, perceived inconvenience, financial costs [7], and a lack of trust in scientific evidence [8]. This hesitancy results in insufficient vaccination coverage and leaves children vulnerable to vaccine-preventable diseases [9]. Notably, caregivers of children with special health care needs are more hesitant about vaccines than those of healthy children [10]. Children with special health care needs refers to a population of children who either have or are at an elevated risk for chronic physical, developmental, behavioral, or emotional conditions. In China, about 2.12 million newborns face birth defects or immune issues each year [11]. Recently, there has been a focus on vaccinating children with special health care needs, with researchers studying their vaccination rates [12,13] and factors associated with their vaccine hesitancy [14-16]. Expert groups from countries such as the United States, Australia [17], and China [18] have developed guidelines and consensus to improve vaccination rates for children with special health care needs.

The global COVID-19 pandemic has heightened vaccine awareness [19]. China approved emergency use of inactivated COVID-19 vaccines for children aged 3 to 17 years in July 2021 [20]. Media focus on vaccination during the pandemic has shifted public and caregivers' attitudes toward vaccines [21,22]. A survey across 6 countries showed that the pandemic has encouraged families who previously avoided vaccines to consider influenza vaccination [23]. A meta-analysis showed an increased willingness among caregivers to vaccinate their children against seasonal flu during the COVID-19 pandemic [24]. Caregivers of children with special health care needs may be more anxious than the general population [25]. The heightened public awareness of vaccines may lead to potential changes in vaccination attitudes among caregivers of children with special health care needs. However, there is a notable lack of research addressing vaccine hesitancy within this specific population.

In 2016, the Health Commission of Zhejiang province, China, set up the Vaccination Consultation Clinic at the Children's Hospital, Zhejiang University School of Medicine [26]. Specialists there have been advising on vaccinations for children with special health care needs. Therefore, we conducted this retrospective study to investigate the attitudes toward vaccines among caregivers of children with special health care needs and evaluate the changes in caregivers' hesitancy toward vaccines before, during, and after the COVID-19 pandemic. In addition, we also aimed to identify associated factors, including the stages of the COVID-19 pandemic, for caregivers' attitudes toward NIP and non-NIP vaccines.

Methods

Study Design and Data Collection

Since 2016, the children's hospital where this study was conducted has established the sole vaccination consultation clinic dedicated to providing assessments and tailored vaccination recommendations for children with special health care needs. The majority of these children were referred by physicians from vaccination sites throughout the province. This hospital is a grade A tertiary children's hospital located in Zhejiang province, primarily serving pediatric patients within the province. In 2019, it was designated as a National Clinical Research Center for Child Health, thereby broadening its patient base. Based on this clinic, we conducted a retrospective analysis of caregivers' vaccination willingness monitoring data from May 18, 2017, to May 31, 2023. Data collection used paper questionnaires until March 31, 2021, followed by an electronic monitoring system from April 1, 2021, onward. Trained research assistants transferred data collected from paper questionnaires into an electronic database managed by EpiData software (The EpiData Association). A double data entry process was implemented to ensure consistency and accuracy.

Ethical Considerations

This study was conducted according to the Declaration of Helsinki [27] and was approved by the medical ethics committee of the Children's Hospital Zhejiang University School of Medicine (2025-IRB-0005-P-01). This study involved secondary analysis of existing anonymized data that posed no privacy risks to individuals. The need for informed consent was waived by the ethics committee of Children's Hospital, Zhejiang University School of Medicine. No financial compensation was provided as the study did not involve direct participant interaction.

Study Population

Data were collected from children with special health care needs who visited the Vaccination Consultation Clinic at Children's Hospital, Zhejiang University School of Medicine between May 18, 2017, and May 31, 2023. For the purposes of this study, children with special health care needs are defined as children with a medical condition that has persisted or is anticipated to persist for 12 months or longer, meeting at least one of the following criteria: (1) a persistent need for prescribed medications, (2) a persistent need for medical care that is higher than the average for children at the same age, (3) a persistent need for special treatments, (4) limitations on activities that are available for the majority of children at the same age, and (5) the presence of a persistent behavioral or developmental condition that requires treatment or counseling [26,28]. In total, 7770 children with special health care needs were included in this study.

Measurements

Information on demographics, family history of diseases, history of adverse events following immunization, parental education levels, parental attitudes toward vaccination, and reasons for being referred to our Vaccination Consultation Clinic were extracted from the immunization evaluation system for children with special health care needs. Concurrently, medical records, such as the specific diagnosis of health conditions, were obtained from the hospital's electronic medical records system.

We used 2 questions to measure caregivers' hesitation toward NIP and non-NIP vaccines. The first question is "What is your attitude toward NIP vaccines?" There are three options: (1) I would like to choose NIP vaccines for my child; (2) if there are alternatives to the NIP vaccine (ie, alternative non-NIP vaccines), I am willing to pay and choose the alternative non-NIP vaccines for my child; and (3) I am not sure and hesitating to choose the NIP vaccines. If caregivers chose the option (3), we defined it as "hesitation toward NIP vaccines." The second question is "What is your attitude toward the non-NIP vaccines?" There are two options: (1) I would like to choose non-NIP vaccines for my child and (2) I am not sure and hesitating. If caregivers chose the option (2), we defined it as "hesitation toward non-NIP vaccines."

Based on the timeline of the COVID-19 pandemic in Zhejiang province, China, we defined stage I, II, and III. More specifically, stage I was "before the COVID-19 pandemic," referred the time period from May 18, 2017, when the first questionnaire was filled to January 22, 2020; stage II was "during the COVID-19 pandemic," referred from January 23, 2020, to December 18, 2022, when all the COVID-19 measures were eased in Zhejiang province; and stage III as from December 19, 2022, to May 31, 2023.

Statistical Analysis

First, we conducted a descriptive analysis to present the general and clinical characteristics of the study population. Normality was tested for the continuous variable, that is, children's age in our study. The distribution of age was not normal, so we calculated the median and IQR. Categorical variables were presented as numbers and percentages. In addition, we compared the difference in the general characteristics of the study population across 3 stages of the COVID-19 pandemic. Regarding the nonnormally distributed variable, we used the Kruskal-Wallis rank test, while for categorical variables, chi-square tests were applied. Furthermore, we described the spectrum of disease for all children enrolled in this study across the 3 phases of the COVID-19 pandemic. In addition, we described the disease spectrum of children with special health care needs in the overall population and the subgroup of each stage of the COVID-19 pandemic. Second, we described the percentages of caregivers' choices for vaccines at 3 stages of the COVID-19 pandemic and compared the differences in the percentages of choices across 3 stages using chi-square tests. Finally, multinomial logistic regression models were used to identify independent variables that were associated with caregivers' hesitation toward NIP and non-NIP vaccines. In addition, we conducted separate logistic regression analyses for different stages of the COVID-19 pandemic.

All statistical analyses were conducted using the R software (version 4.3.1; R Foundation for Statistical Computing). The statistical significance was indicated when $P<.05$ (2 tails).

Results

General and Clinical Characteristics of the Study Population

Table 1 presents the general characteristics of the study population. In total, 4418 (56.9%) children included in the analysis were male, and 3534 (45.5%) were aged less than 6 months. A total of 6051 (77.9%) questionnaires were filled in by mothers. Mothers of 3061 (39.4%) children had an educational level of college or above, while fathers of 3004 (38.7%) children had an educational level of college or above. There were statistically significant differences in children's age, caregiver who filled in the questionnaire, and maternal and paternal educational level across the 3 COVID-19 pandemic stages (all $P<.001$).

Table . General characteristics of children and their caregivers (N=7770).

Items	Overall	Stage I[a]	Stage II[b]	Stage III[c]	*P* value
Child's sex, n (%)					.50
Male	4418 (56.9)	1001 (56.8)	3108 (56.6)	309 (59.3)	
Female	3352 (43.1)	761 (43.2)	2379 (43.4)	212 (40.7)	
Age of the child (months), median (IQR)	7.0 (2.4-24.1)	6.4 (2.3-19.6)	7.6 (2.5-25.7)	5.3 (2.0-18.2)	<.001
Age group of children (months), n (%)					<.001
0 - 6	3534 (45.5)	845 (48)	2416 (44)	273 (52.4)	
7 - 12	1228 (15.8)	306 (17.4)	840 (15.3)	82 (16.7)	
13 - 24	1053 (13.6)	235 (13.3)	755 (13.8)	63 (12.1)	
25 - 72	1376 (17.7)	305 (17.3)	1001 (18.2)	70 (13.3)	
≥72	579 (7.4)	71 (4)	475 (8.7)	33 (6.3)	
Caregiver who filled in the questionnaire, n (%)					<.001
Mother	6051 (77.9)	1484 (84.2)	4208 (76.7)	359 (68.9)	
Father	1419 (18.3)	242 (13.7)	1049 (19.1)	128 (24.6)	
Others	281 (3.6)	33 (1.9)	214 (3.9)	34 (6.5)	
Missing	19 (0.2)	3 (0.2)	16 (0.3)	0 (0)	
Maternal educational level, n (%)					<.001
Middle school or below	1253 (16.1)	316 (17.9)	883 (16.1)	54 (10.4)	
High school or equivalent	1233 (15.9)	317 (18)	855 (15.6)	61 (11.7)	
Two-year college or associate degree	1948 (25.1)	463 (26.3)	1378 (25.1)	107 (20.5)	
Bachelor degree or above	3061 (39.4)	618 (35.1)	2209 (40.3)	234 (44.9)	
Missing	275 (3.5)	48 (3)	162 (3)	65 (12.5)	
Paternal educational level, n (%)					<.001
Middle school or below	1205 (15.5)	302 (17.1)	850 (15.5)	53 (10.2)	
High school or equivalent	1401 (18)	358 (20.3)	978 (17.8)	65 (12.5)	
Two-year college or associate degree	1796 (23.1)	435 (24.7)	1244 (22.7)	117 (22.5)	
Bachelor degree or above	3004 (38.7)	611 (34.7)	2170 (39.6)	223 (42.8)	
Missing	364 (4.7)	56 (3.2)	245 (4.5)	63 (12.1)	

[a] Stage Ⅰ was defined as the time period from May 18, 2017, to January 22, 2020.

[b] Stage II was defined as the time period from January 23, 2020, to December 18, 2022.

[c] Stage III was defined as the time period from December 19, 2022, to May 31, 2023.

Table 2 demonstrates the clinical characteristics of the study population. A total of 2729 (35.1%) children had comorbidity, 4739 (61%) children had a history of allergies, and 330 (4.2%) children had a history of adverse events following immunization. Regarding the reasons to visit the vaccination consultation clinic, 5245 (67.5%) caregivers reported that they were referred by doctors at the vaccination site, and 2198 (28.3%) caregivers came to consult because they did not know whether children with special health care needs can receive vaccines. In total, 336 (4.3%) caregivers came to consult because they did not know whether certain medications applied to the child may affect vaccination, and 130 (1.7%) caregivers due to adverse events in previous vaccinations. In addition, 288 (3.7%) caregivers were referred by the specialists because they were not sure about the vaccination of children with special health care needs, and 376 (4.8%) caregivers came to consult because their children were mandatory to be vaccinated in order to go to school.

Table . Clinical characteristics of the study population (N=7770).

Items	Overall, n (%)	Stage I[a], n (%)	Stage II[b], n (%)	Stage III[c], n (%)
Having comorbidity				
Yes	2729 (35.1)	592 (66.4)	1915 (34.9)	222 (42.6)
No	5041 (64.9)	1170 (33.6)	3572 (65.1)	299 (57.4)
History of allergy				
Yes	4739 (61)	446 (25.3)	3772 (68.7)	521 (100)
No	3026 (38.9)	1316 (74.7)	1710 (31.2)	0 (0)
Missing	5 (0.1)	0 (0)	5 (0.1)	0 (0)
History of AEFI [d]				
Yes	330 (4.2)	81 (5)	227 (4.1)	22 (4.2)
No	7358 (94.7)	1681 (95.4)	5178 (94.4)	499 (95.8)
Missing	82 (1.1)	0 (0)	82 (1.5)	0 (0)
Reasons for visiting the vaccination consultation clinic				
Referred by doctors at the vaccination site	5245 (67.5)	1407 (79.8)	3513 (64)	325 (62.4)
Caregivers came to consult because they did not know whether children with special health care needs can receive vaccines	2198 (28.3)	289 (16.4)	1716 (31.3)	193 (37)
Caregivers came to consult because they did not know whether certain medications applied to the child may affect vaccination	336 (4.3)	8 (0.4)	275 (5)	53 (10.2)
Caregivers came to consult because of adverse events in previous vaccinations	130 (1.7)	5 (0.3)	106 (1.9)	19 (3.6)
Referred by the specialist because they were not sure about vaccination of children with special health care needs	288 (3.7)	15 (0.8)	264 (4.8)	9 (1.7)
Required by school entry	376 (4.8)	13 (0.7)	300 (5.5)	63 (12.1)
Missing	86 (1.1)	25 (1.4)	58 (1.1)	3 (0.6)

[a] Stage Ⅰ was defined as the time period from May 18, 2017, to January 22, 2020.

[b] Stage II was defined as the time period from January 23, 2020, to December 18, 2022.

[c] Stage III was defined as the time period from December 19, 2022, to May 31, 2023.

[d] AEFI: adverse events following immunization.

Disease Spectrum of the Study Population

Multimedia Appendix 1 shows the spectrum of disease in the overall study population and at each stage of the COVID-19 pandemic. The 3 most common diseases of those children with special health care needs were circulatory system diseases, nervous system diseases, and neonatal diseases in the overall population, and this pattern remains the same at each stage.

Caregivers’ Attitude Toward NIP Vaccines and Non-NIP Vaccines

Figure 1 shows caregivers’ choice for NIP vaccines or alternative non-NIP vaccines and their hesitancy at each stage of the COVID-19 pandemic. Figure 2 shows caregivers’ choice of non-NIP vaccines and their hesitation at 3 stages of the COVID-19 pandemic. During the COVID-19 pandemic (ie, stage II), the percentages of choosing alternative non-NIP or non-NIP vaccines are highest (n=3148, 57.4% and n=3442, 67.7%, respectively) than those at the other 2 stages, while the percentages of caregivers' hesitation toward NIP and non-NIP vaccines are lowest (n=911, 16.6% and n=2045, 37.3%, respectively). The percentages of caregivers' attitudes toward NIP vaccines significantly differ across 3 stages of the COVID-19 pandemic (*P*<.001; Table S1 in Multimedia Appendix 2). The percentages of caregivers' attitudes toward non-NIP vaccines significantly differ across 3 stages of the COVID-19 pandemic (*P*<.001; Table S2 in Multimedia Appendix 2).

Figure 1. Caregivers' choices for NIP and alternative non-NIP vaccines and their hesitancy at 3 stages of the COVID-19 pandemic. NIP: National Immunization Program.

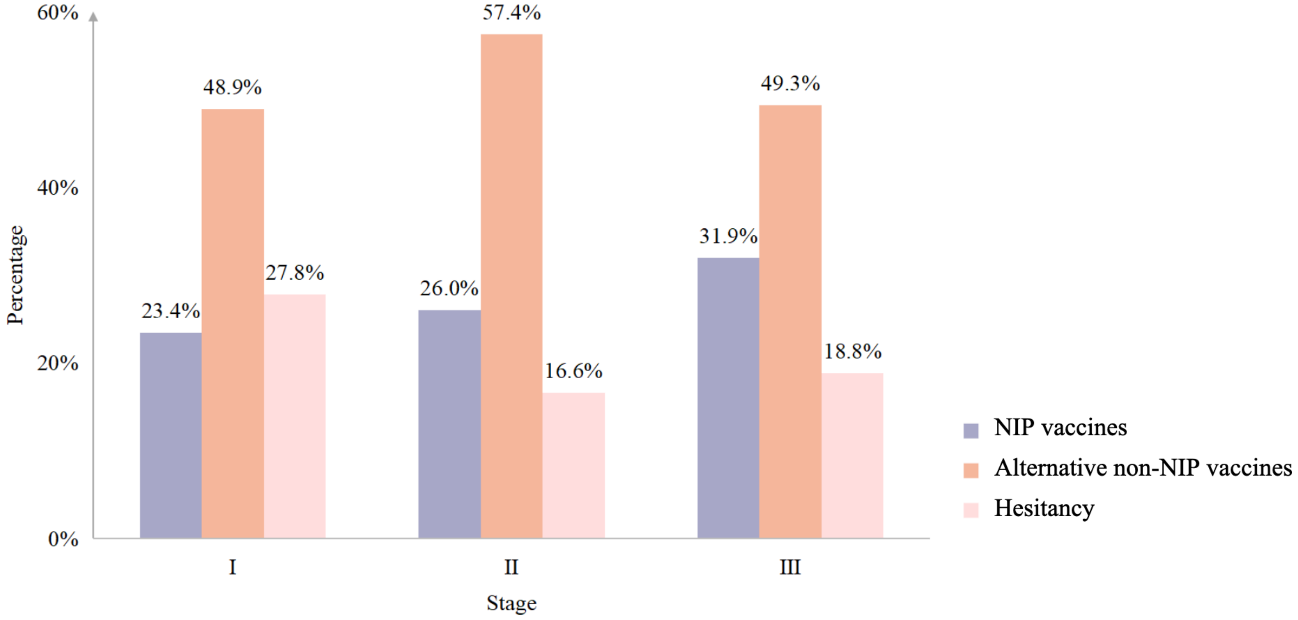
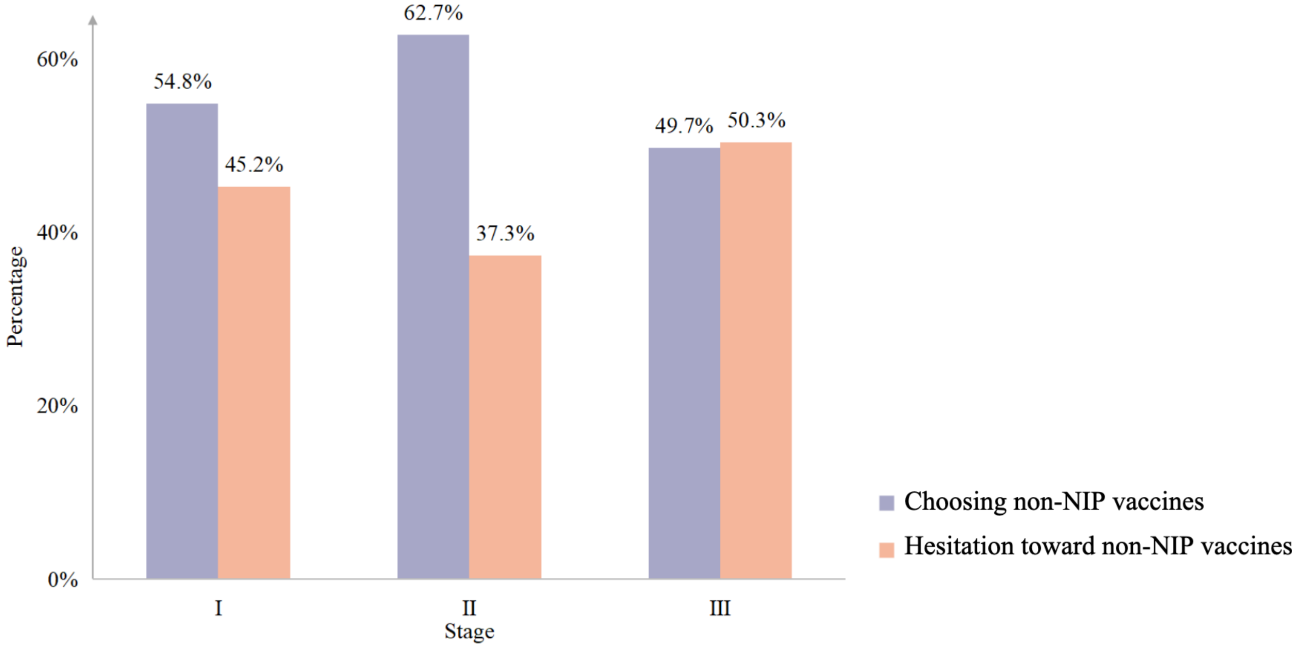


Figure 2. Caregivers' choices for non-NIP vaccines and their hesitancy at 3 stages of the COVID-19 pandemic. NIP: National Immunization Program.



Associated Factors for Caregivers' Choices of NIP Vaccines and Alternative Non-NIP Vaccines and Hesitations Toward NIP Vaccines

Table 3 presents the results of multinomial logistic regression regarding the associated factors for caregivers' choices of alternative non-NIP vaccines and hesitations toward NIP vaccines using caregivers' choices for NIP vaccines as a reference.

Table . Associated factors for caregivers' hesitation toward NIP[a] vaccines using multinomial logistic regression analysis.

Variables	Choosing alternative non-NIP vaccines, OR[b] (95% CI)	Hesitation toward NIP vaccines, OR (95% CI)
Stage of COVID-19 pandemic		
I	Reference	Reference
II	1.3 (1.2-1.6)[c]	0.6 (0.5-0.8)[c]
III	0.9 (0.7-1.2)	0.6 (0.4-0.8)[d]
Children's sex		
Male	Reference	Reference
Female	0.8 (0.8-1.0)[d]	0.9 (0.8-1.1)
Age groups of children (months)		
0 - 6	Reference	Reference
7 - 12	1.1 (0.9-1.3)	0.6 (0.5-0.7)[c]
13 - 24	0.8 (0.7-1.0)	0.4 (0.3-0.5)[c]
25 - 72	0.7 (0.6-0.8)[c]	0.4 (0.3-0.5)[c]
≥72	0.4 (0.3-0.5)[c]	0.5 (0.4-0.6)[c]
Caregivers who filled the questionnaire		
Mother	Reference	Reference
Father	0.8 (0.7-0.9)[c]	1.1 (0.9-1.3)
Others	1.3 (0.9-1.8)	2.4 (1.6-3.4)[c]
Maternal educational level		
Middle school or below	Reference	Reference
High school or equivalent	1.2 (1.0-1.5)	1.1 (0.8-1.4)
Two-year college	1.5 (1.2-1.8)[d]	1.2 (0.9-1.5)
Bachelor degree or above	2.3 (1.8-3.0)[c]	1.3 (1.0-1.8)
Paternal educational level		
Middle school or below	Reference	Reference
High school or equivalent	1.1 (0.9-1.3)	1.0 (0.8-1.3)
Two-year college	1.5 (1.2-1.9)[d]	1.0 (0.7-1.3)
Bachelor degree or above	1.4 (1.1-1.8)[d]	0.9 (0.7-1.2)
Having comorbidity		
No	Reference	Reference
Yes	1.0 (0.9-1.2)	1.2 (1.0-1.4)[e]
History of allergy		
No	Reference	Reference
Yes	0.6 (0.5-0.7)[c]	0.7 (0.6-0.8)[c]
History of AEFI[f]		
No	Reference	Reference
Yes	1.2 (0.9-1.5)	0.8 (0.5-1.1)

[a] NIP: National Immunization Program.

[b] OR: odds ratio.

[c] *P*<.001.

[d] *P*<.01.

[e] *P*<.05.

[f] AEFI: adverse events following immunization.

Choosing Alternative Non-NIP Vaccines Versus NIP Vaccines

Stage II of the COVID-19 pandemic (*P*<.001), maternal educational level of 2-year college (*P*=.002) and of bachelor degree or above (*P*<.001), as well as paternal educational level of 2-year college (*P*=.001) and bachelor degree or above (*P*=.010) were significantly associated with higher odds to choose alternative non-NIP vaccines. In contrast, the child being female (*P*=.005), the child's age of 25 - 72 months (*P*<.001) and older than 72 months (*P*<.001), fathers filling in the questionnaire (*P*<.001), and having a history of allergy (*P*<.001) were significantly associated with lower odds of choosing alternative non-NIP vaccines.

Hesitation Toward NIP Vaccines Versus Choosing NIP Vaccines

Caregivers other than parents filling in the questionnaire (*P*<.001) and having comorbidity (*P*=.02) were significantly associated with higher odds of hesitations toward NIP vaccines. Stage II (*P*<.001) and III (*P*=.003) of the COVID-19 pandemic, the child age older than 6 months (*P*<.001) and having a history of allergy (*P*<.001) were significantly associated with lower odds of hesitations toward NIP vaccines.

Associated Factors for Caregivers' Choices of Non-NIP Vaccines and Hesitations Toward Non-NIP Vaccines

Table 4 presents the results of multinomial logistic regression regarding the associated factors for caregivers' hesitations toward non-NIP vaccines using the willingness of choosing non-NIP vaccines as a reference. The child age older than 72 months (*P*<.001), father (*P*<.001) and caregivers other than parents (*P*=.006) filling in the questionnaire, having comorbidity (*P*=.03), and having a history of allergy (*P*<.001) were significantly associated with higher odds of hesitation toward non-NIP vaccines. Stage II of the COVID-19 pandemic (*P*<.001), the child age of 7-12 months (*P*=.01), 13-24 months (*P*=.01), 25 - 72 months (*P*=.001), maternal educational level of bachelor degree or above (*P*=.001) and paternal educational level of 2-year college (*P*=.002), and bachelor degree or above (*P*=.003) were significantly associated with lower odds of hesitation toward the non-NIP vaccines.

Table . Associated factors for caregivers' hesitation toward non-NIP[a] vaccines using multinomial logistic regression analysis.

Predictors	Hesitation toward non-NIP vaccines, OR[b] (95% CI)
Stage of COVID-19 pandemic	
I	Reference
II	0.6 (0.5-0.7)[c]
III	1.0 (0.8-1.3)
Children's sex	
Male	Reference
Female	1.1 (1.0-1.2)
Age groups of children (months)	
0 - 6	Reference
7 - 12	0.8 (0.7-1.0)[d]
13 - 24	0.8 (0.7-1.0)[d]
25 - 72	0.8 (0.7-0.9)[c]
≥72	1.6 (1.3-1.9)[c]
Caregivers who filled the questionnaire	
Mother	Reference
Father	1.6 (1.4-1.9)[c]
Others	1.5 (1.1-1.9)[e]
Maternal educational level	
Middle school or below	Reference
High school or equivalent	0.9 (0.8-1.1)
Two-year college	0.8 (0.7-1.0)
Bachelor degree and above	0.7 (0.6-0.8)[c]
Paternal educational level	
Middle school or below	Reference
High school or equivalent	1.0 (0.8-1.2)
Two-year college	0.7 (0.6-0.9)[e]
Bachelor degree and above	0.7 (0.6-0.9)[e]
Having comorbidity	
No	Reference
Yes	1.1 (1.0-1.2)[d]
History of allergy	
No	Reference
Yes	1.3 (1.2-1.4)[c]
History of AEFI[f]	
No	Reference
Yes	0.8 (0.7-1.1)

[a] NIP: National Immunization Program.

[b] OR: odds ratio.

[c] *P*<.001.

[d] *P*<.05.

[e] $P<.01$.

[f] AEFI: adverse events following immunization.

In addition, we presented the results of the separate logistic regression analysis for different stages of the COVID-19 pandemic in Tables S3-S8 in Multimedia Appendix 2.

Discussion

Principal Findings

This study used a large-scale, retrospective dataset to examine shifts in caregivers' hesitancy of choosing NIP and non-NIP vaccines for children with special health care needs across the pre-, during, and postpandemic periods of COVID-19. Furthermore, we identified factors associated with vaccine hesitancy, including the different phases of the COVID-19 pandemic. To the best of our knowledge, this study has contributed to provide real-world evidence and comprehensive data on vaccine hesitancy among caregivers of children with special health care needs, addressing a gap in the existing literature.

Our findings indicate that, irrespective of the stage of the COVID-19 pandemic, the proportion of individuals opting for alternative non-NIP vaccines surpassed those choosing NIP vaccines. This trend may be attributed to the unique characteristics of the population under study, that is, children with special health care needs. These children differ from the healthy, general pediatric population, as they are managing certain health conditions. Consequently, their caregivers may have greater caution when selecting vaccines. The alternative non-NIP vaccines available are either inactivated vaccines, as opposed to live-attenuated NIP vaccines, or vaccines with a higher number of valences. It is understandable that caregivers opt for inactivated vaccines over live-attenuated vaccines to mitigate the potential side effects associated with live pathogens. Furthermore, parents may prefer to administer vaccines with multiple valences in a single injection to their child.

Our study revealed that caregivers' acceptance of NIP vaccines increased across the pre-, during, and postpandemic periods of COVID-19. Wang et al [29] also found that there has been an increasing acceptance of routine childhood vaccination and COVID-19 vaccination from 2020 to 2021 via a repeated cross-sectional survey. We found that the proportion of caregivers of children with special health care needs willing to vaccinate their children with alternative non-NIP vaccines was highest during the pandemic. Concurrently, we observe the lowest proportion of caregivers exhibiting hesitation toward NIP vaccines at the same stage. These findings align with the results of a multinomial logistic analysis, which indicated that this stage (ie, during the pandemic) is associated with significantly higher odds of caregivers' willingness to select alternative non-NIP vaccines. Additionally, stages II and III are significantly associated with lower odds of caregivers' hesitation toward NIP vaccines. Zhang et al [30] reported that the coverage rate of non-NIP vaccines increased by 25.8% and 34.7%, respectively, in 2020 and 2021 when it was the nonpharmaceutical intervention period compared to the prepandemic period (ie, in 2019). Our finding also echoed with previous conclusion that the COVID-19 pandemic has caused unprecedented impacts including parental trust in vaccines [31]. During the COVID-19 pandemic, both the government and the mass media have substantially intensified their advocacy and promotion of the COVID-19 vaccine. They have also enhanced risk communication to raise awareness about the benefits and risks of vaccination and to promote the scientific belief that the benefits outweigh the risks during the COVID-19 pandemic [32]. This heightened publicity has concurrently enhanced public awareness regarding vaccines in general, potentially fostering a more positive attitude and greater acceptance of vaccines among the population. Consequently, caregivers have become more inclined to vaccinate their children during the pandemic; they are even willing to pay for the self-funded, alternative NIP vaccines.

Regarding caregivers' attitudes toward non-NIP vaccines, our study indicates that caregivers' willingness to vaccinate their children with special health care needs was at its peak, and their hesitancy was at its lowest during the pandemic. This observation aligns with the results obtained from the multinomial regression analysis. The increased willingness can likely be attributed to the extensive promotion of vaccines, particularly the COVID-19 vaccine, by governments and social media platforms. However, following the COVID-19 pandemic, our study observed a decline in the proportion of caregivers opting for non-NIP vaccines and an increase in their hesitancy toward these vaccines. In this study, stage III corresponds to the 6-month period following the relaxation of COVID-19 measures in China, during which the Chinese public exhibited a varied and mixed attitude toward vaccines. This period was characterized by both trust and mistrust in vaccines, exacerbated by reports of incidents such as "fake vaccines" or vaccine scandals that were disseminated through social media. There has been a scarcity of literature regarding the impact of the COVID-19 pandemic on non-NIP vaccination intention for children. Wang et al [33] found that in a community sample, approximately 70% of parents did not change their intention for self-paid, non-NIP vaccines for their children, and about 20% of parents increased their intention, while about 10% decreased the intention after the COVID-19 pandemic. In this study, a short period of stage III resulted in a limited sample size for data collection, potentially leading to findings that may be attributable to chance. Therefore, we recommend conducting future studies with an increased sample size by extending the study period during the "postpandemic" phase to validate or refute our findings.

Our study indicated a sex disparity in caregivers' selection of alternative, self-funded non-NIP vaccines. Caregivers of female children were less likely to choose these alternative non-NIP vaccines compared to caregivers of male children. This pattern remained the same before and during the pandemic but was reversed after the pandemic when caregivers of female children were more likely to choose the alternative non-NIP compared to caregivers of male children. Although global immunization coverage does not show significant differences between male and female children, certain countries and communities may experience sex-related barriers to immunization due to prevailing social and cultural norms [34]. The outbreak of the COVID-19 pandemic might have increased the intention of caregivers of female children to choose paid alternative non-NIP vaccines for them.

Regarding the age of children, our study indicated that caregivers of older children demonstrated a lower propensity to vaccinate their children with alternative non-NIP vaccines or exhibited reduced vaccine hesitancy compared to caregivers of very young children (ie, 0‐6 months). This pattern remains the same in the separate logistic analysis in each stage. Notably, our findings revealed that the likelihood of vaccine hesitancy toward non-NIP vaccines among caregivers of children older than 72 months was 1.58 times greater than that among caregivers of children aged 0‐6 months. A potential experiential factor may account for this finding. Non-NIP vaccines are not mandatory for children, and advocacy or public awareness efforts are less robust compared to those for NIP vaccines. Advocacy predominantly targets young children, typically younger than 6 years of age. Consequently, caregivers of older children may lack awareness regarding the appropriate timing and types of non-NIP vaccines beneficial for their children. Therefore, our study suggests that it is important and necessary to educate caregivers of children older than 6 years.

In this study, our findings indicate that fathers, in comparison to mothers, exhibited a lower propensity to choose for alternative non-NIP vaccines and demonstrated a greater likelihood of hesitancy in selecting non-NIP vaccines for their children. Furthermore, the analysis revealed that caregivers other than parents showed a higher probability of hesitancy in choosing non-NIP vaccines for children when compared to mothers. This finding is noteworthy and may be attributed to the cultural norm within Chinese families, where mothers predominantly assume the role of primary caregiver and are chiefly responsible for making medical decisions concerning their children. While it is important to acknowledge the contributions of fathers and other caregivers in child-rearing, their involvement and roles may not be as prominent or comparable to those of mothers.

Our study demonstrates that a relatively high maternal and paternal educational level, such as a 2-year college and bachelor or above degree, is associated with the higher intention for choosing alternative non-NIP vaccines and less hesitation to both NIP and non-NIP vaccines, which is consistent with previous studies. For instance, a study in Shanghai showed that higher income, higher education, and greater access to vaccines made respondents more willing to immunize their children with non-NIP [35]. Economic status remained a significant factor in determining the acceptability of non-NIP vaccines [36]. Parents' high education is usually linked to high health literacy, which greatly affects their ability to use health information to make health decisions for their children [37].

In this study, caregivers of children with comorbidities exhibit increased hesitancy regarding NIP and non-NIP vaccines. This hesitancy is understandable, as the existing health conditions of these children already present significant concerns, prompting caregivers to exercise greater caution when considering vaccination decisions. We also found that caregivers of children with allergy had increased hesitancy regarding non-NIP vaccines, which may stem from concerns about an elevated risk of adverse effects following vaccination in children with a history of allergies.

The vaccination of children with special health care needs is unique, as their health issues may lead to delayed or missed vaccinations, leaving them vulnerable to infections. We analyzed how China's macroprevention and control strategies for COVID-19 may affect the vaccination intentions of this group using a large sample. Our findings indicate that during the COVID-19 pandemic, caregivers exhibited a marked increase in their willingness to receive vaccinations, accompanied by a significant decrease in vaccine hesitancy, compared to levels observed both prior to and following the pandemic. This empirical conclusion aligns with observations made by physicians in vaccination clinics. The onset of the pandemic appears to have heightened parental awareness regarding vaccines, thereby enhancing vaccination rates. This study revealed that a significant proportion of parents continue to exhibit hesitancy toward vaccines, highlighting the persistence of a knowledge gap and the potential for enhancement in health education initiatives. Furthermore, the research identified several factors associated with vaccine hesitancy, including the child's age and sex, the educational level of the parents, and the presence of an allergy history. These findings aim to furnish evidence-based support for the formulation of personalized and targeted health education strategies in future endeavors. The findings of this study further indicate the necessity for social mobilization from a macrolevel perspective to collaboratively enhance vaccine awareness and advance health equity.

Strengths and Limitations

To the best of our knowledge, this study is the first to use a large sample size (N=7770) to examine attitudes, specifically hesitancy, toward vaccinations among caregivers of children with special health care needs before, during, and after the COVID-19 pandemic. Additionally, our study has documented a comprehensive set of variables, which allows us to apply multivariate regression analysis to identify associated factors for caregivers' attitudes toward NIP and non-NIP vaccines in China. We consider it to be among the few studies of its kind conducted globally.

However, several limitations merit consideration. First, this study used a cross-sectional design, which restricts our ability to infer causality and confines our analysis to assessing associations. Second, we did not use a standardized or validated instrument, such as the Vaccine Hesitancy Scale, to measure vaccine hesitancy. Instead, we used 2 self-developed questions as a proxy for assessing vaccine hesitancy. Third, our study exclusively assessed caregivers' willingness or attitudes toward vaccinating their children with NIP or non-NIP vaccines. It is important to acknowledge the potential discrepancy between the respondents' self-reported willingness to vaccinate and their actual vaccination behavior, as survey responses may not always accurately reflect real-world actions [38].

Conclusions

This study demonstrated that caregivers' willingness to vaccinate their children with special health care needs with NIP and non-NIP vaccines was highest and their hesitancy was lowest during the COVID-19 pandemic in China. Additionally, we have identified multiple factors associated with caregivers' willingness and hesitancy of vaccinating their children. These findings provide evidence-based support for developing personalized health education strategies and further indicate the necessity for social mobilization at a macrolevel to collaboratively enhance vaccine awareness and health literacy regarding vaccinating children with special health care needs.

Acknowledgments

This study was financially supported by the Medical Science and Technology Project of Zhejiang province (2020385296). The funders had no role in the study design; data collection, analysis, or interpretation; in the report's writing; or in the decision to submit the paper for publication.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to safety and privacy concerns but are available from CJ (6198011@zju.edu.cn) on reasonable request.

Authors' Contributions

ML, CS, CJ, and GB were involved in the study conceptualization and study design. CS and MG carried out the data analyses. ML and GB wrote the first draft of the manuscript. XW, DY, JG, and LS contributed to the data collection, data entering, and data cleaning. AR, ASG, and SHSSM contributed to searching the literature, visualizing the results, and editing the manuscript. CJ and GB supervised the project. All authors provided critical revisions of the paper for important intellectual content, contributed to the interpretation of the data, and approved the final version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The spectrum of disease in the overall study population and at each stage of the COVID-19 pandemic. (A) The spectrum of disease in the overall study population, (B) the spectrum of disease at stage I of the COVID-19 pandemic, (C) the spectrum of disease at stage II of the COVID-19 pandemic, and (D) the spectrum of disease at stage III of the COVID-19 pandemic.

[TIFF File, 6291 KB - publichealth_v11i1e67487_app1.tiff]

Multimedia Appendix 2

Additional tables.

[DOCX File, 49 KB - publichealth_v11i1e67487_app2.docx]

References

1. National Immunization Program vaccine children's immunization procedures and instructions. National Health Commission of the People's Republic of China. 2021. URL: http://www.nhc.gov.cn/jkj/s3581/202103/590a8c7915054aa682a8d2ae8199e222/files/7559f978e6be4ec585a9a1fe0d9224f6.pdf [accessed 2025-03-13]
2. Childhood Immunization Schedule for National Immunization Program Vaccines—China (version 2021). China CDC Wkly 2021 Dec 24;3(52):1101-1108. [doi: 10.46234/ccdcw2021.270] [Medline: 35186365]
3. Chen S, Rodewald LE, Du AH, Tang S. Advancing the National Immunization Program in an era of achieving universal vaccine coverage in China and beyond. Infect Dis Poverty 2024 Mar 13;13(1):25. [doi: 10.1186/s40249-024-01192-6] [Medline: 38475849]
4. Zhou Y, Li D, Cao Y, et al. Immunization coverage, knowledge, satisfaction, and associated factors of non-National Immunization Program vaccines among migrant and left-behind families in China: evidence from Zhejiang and Henan provinces. Infect Dis Poverty 2023 Oct 13;12(1):93. [doi: 10.1186/s40249-023-01145-5] [Medline: 37833775]
5. Zhang H, Lai X, Mak J, et al. Coverage and equity of childhood vaccines in China. JAMA Netw Open 2022 Dec 1;5(12):e2246005. [doi: 10.1001/jamanetworkopen.2022.46005]
6. Baumgaertner B, Ridenhour BJ, Justwan F, Carlisle JE, Miller CR. Risk of disease and willingness to vaccinate in the United States: a population-based survey. PLoS Med 2020 Oct;17(10):e1003354. [doi: 10.1371/journal.pmed.1003354] [Medline: 33057373]
7. Doherty M, Schmidt-Ott R, Santos JI, et al. Vaccination of special populations: protecting the vulnerable. Vaccine (Auckl) 2016 Dec 20;34(52):6681-6690. [doi: 10.1016/j.vaccine.2016.11.015] [Medline: 27876197]

8. Shahani R, Chu J, Rufai OH, et al. Understanding the role of psychosocial factors in Pakistani parents' hesitancy to vaccinate their kids: the mediating role of knowledge and mistrust of science about the COVID-19 vaccine. Vaccines (Basel) 2022 Aug 5;10(8):1260. [doi: 10.3390/vaccines10081260] [Medline: 36016148]
9. Xu B, Zhang Y, Zhou C, Wang Q, Luan R. Vaccination coverage and immunization safety among children with special health status in Chongqing, China. Hum Vaccin Immunother 2022 Dec 30;18(7):2138466. [doi: 10.1080/21645515.2022.2138466] [Medline: 36469710]
10. McPherson M, Arango P, Fox H, et al. A new definition of children with special health care needs. Pediatrics 1998 Jul;102(1 Pt 1):137-140. [doi: 10.1542/peds.102.1.137] [Medline: 9714637]
11. Wang L, Hao YH, Bai Y. Public health and clinical care integration to improve immunization of children with special health care needs. Zhonghua Yu Fang Yi Xue Za Zhi 2022 Jun 6;56(6):673-678. [doi: 10.3760/cma.j.cn112150-20220104-00010] [Medline: 35785850]
12. de Souza RP, Ribeiro ALR, de Menezes SAF, Machado LFA. Incidence of respiratory syncytial virus infection in children with congenital heart disease undergoing immunoprophylaxis with palivizumab in Pará state, north region of Brazil. BMC Pediatr 2019 Aug 28;19(1):299. [doi: 10.1186/s12887-019-1681-6] [Medline: 31462289]
13. Yang L, Peng J, Deng J, et al. Vaccination status of children with epilepsy or cerebral palsy in Hunan rural area and a relative KAP survey of vaccinators. Front Pediatr 2019;7:84. [doi: 10.3389/fped.2019.00084] [Medline: 30984716]
14. Talbird SE, Carrico J, La EM, et al. Impact of routine childhood immunization in reducing vaccine-preventable diseases in the United States. Pediatrics 2022 Sep 1;150(3):e2021056013. [doi: 10.1542/peds.2021-056013] [Medline: 35821599]
15. Luan L, Zhang Z, Xu J, et al. Evaluation of vaccination status of children with special health care needs in Suzhou. Hum Vaccin Immunother 2020;19(2):2254965. [doi: 10.1080/21645515.2023.2254965]
16. Laforgia N, Di Mauro A, Bianchi FP, et al. Are pre-terms born timely and right immunized? Results of an Italian cohort study. Hum Vaccin Immunother 2018 Jun 3;14(6):1398-1402. [doi: 10.1080/21645515.2018.1428509] [Medline: 29351055]
17. Cossarizza A, Chang HD, Radbruch A, et al. Guidelines for the use of flow cytometry and cell sorting in immunological studies (third edition). Eur J Immunol 2021 Dec;51(12):2708-3145. [doi: 10.1002/eji.202170126] [Medline: 34910301]
18. Notice of the National Health Commission on issuing the National Immunization Program Vaccine Children's Immunization Procedures and Instructions (2021 edition). National Health Commission of People's Republic of China. URL: http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=590a8c7915054aa682a8d2ae8199e222 [accessed 2024-10-07]
19. Olusanya OA, Bednarczyk RA, Davis RL, Shaban-Nejad A. Addressing parental vaccine hesitancy and other barriers to childhood/adolescent vaccination uptake during the coronavirus (COVID-19) pandemic. Front Immunol 2021;12:663074. [doi: 10.3389/fimmu.2021.663074] [Medline: 33815424]
20. Qin C, Wang R, Tao L, Liu M, Liu J. Association between risk perception and acceptance for a booster dose of COVID-19 vaccine to children among child caregivers in China. Front Public Health 2022;10:834572. [doi: 10.3389/fpubh.2022.834572] [Medline: 35372197]
21. Lee M, You M. Direct and indirect associations of media use with COVID-19 vaccine hesitancy in South Korea: cross-sectional web-based survey. J Med Internet Res 2022 Jan 6;24(1):e32329. [doi: 10.2196/32329] [Medline: 34870605]
22. Zhou X, Song S, Zhang Y, Hou Z. Deep learning analysis of COVID-19 vaccine hesitancy and confidence expressed on twitter in 6 high-income countries: longitudinal observational study. J Med Internet Res 2023 Nov 6;25:e49753. [doi: 10.2196/49753] [Medline: 37930788]
23. Goldman RD, McGregor S, Marneni SR, et al. Willingness to vaccinate children against influenza after the coronavirus disease 2019 pandemic. J Pediatr 2021 Jan;228:87-93. [doi: 10.1016/j.jpeds.2020.08.005]
24. Wang Z, Chen S, Fang Y. Parental willingness and associated factors of pediatric vaccination in the era of COVID-19 Pandemic: a systematic review and meta-analysis. Vaccines (Basel) 2022 Sep 2;10(9):1453. [doi: 10.3390/vaccines10091453] [Medline: 36146530]
25. Geweniger A, Haddad A, Barth M, et al. Mental health of children with and without special healthcare needs and of their caregivers during COVID-19: a cross-sectional study. BMJ Paediatr Open 2022 Jun;6(1):e001509. [doi: 10.1136/bmjpo-2022-001509] [Medline: 36053611]
26. Li M, Ji C, Wang B, et al. Incomplete vaccination among children with special health care needs in Zhejiang, China: analysis of retrospective data. Front Pediatr 2019;7:173. [doi: 10.3389/fped.2019.00173] [Medline: 31119119]
27. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
28. Bethell CD, Read D, Stein RE, Blumberg SJ, Wells N, Newacheck PW. Identifying children with special health care needs: development and evaluation of a short screening instrument. Ambul Pediatr 2002;2(1):38-48. [doi: 10.1367/1539-4409(2002)0022.0.co;2] [Medline: 11888437]
29. Wang Q, Xiu S, Yang L, et al. Changes in parental attitudes toward COVID-19 vaccination and routine childhood vaccination during the COVID-19 pandemic: repeated cross-sectional survey study. JMIR Public Health Surveill 2022 May 13;8(5):e33235. [doi: 10.2196/33235] [Medline: 35486516]
30. Zhang C, Hu W, Li Y, Lv Y, Zhang S. Impact of the COVID-19 pandemic on routine vaccination services in Shaanxi province, Northwest China: non-pharmaceutical intervention period and mass COVID-19 vaccination period. Hum Vaccin Immunother 2023 Aug 1;19(2):2251826. [doi: 10.1080/21645515.2023.2251826] [Medline: 37665634]

31. Saxena S, Skirrow H, Bedford H. Routine vaccination during covid-19 pandemic response. BMJ 2020 Jun 16;369:m2392. [doi: 10.1136/bmj.m2392] [Medline: 32546575]
32. Wu J, Li Q, Silver Tarimo C, et al. COVID-19 vaccine hesitancy among Chinese population: a large-scale national study. Front Immunol 2021;12:781161. [doi: 10.3389/fimmu.2021.781161] [Medline: 34912346]
33. Wang S, Xu J, Zhu J. Impact of the COVID-19 pandemic on self-paid vaccination intentions for children: a cross-sectional study in China. BMJ Open 2024 Aug;14(8):e083056. [doi: 10.1136/bmjopen-2023-083056]
34. Gender and immunisation. Gavi. URL: https://www.gavi.org/our-alliance/strategy/gender-and-immunisation [accessed 2024-10-08]
35. Wu L, Huang Z, Guo X, Liu J, Sun X. Measuring parents’ acceptance of non-national immunization program vaccines for children and its influencing factors during the COVID-19 pandemic in Shanghai, China. Hum Vaccin Immunother 2022 Nov 30;18(5):2069427. [doi: 10.1080/21645515.2022.2069427] [Medline: 35727599]
36. Zhang X, Syeda ZI, Jing Z, et al. Rural-urban disparity in category II vaccination among children under five years of age: evidence from a survey in Shandong, China. Int J Equity Health 2018 Jun 22;17(1):87. [doi: 10.1186/s12939-018-0802-4] [Medline: 29929503]
37. Morrison AK, Glick A, Yin HS. Health literacy: implications for child health. Pediatr Rev 2019 Jun;40(6):263-277. [doi: 10.1542/pir.2018-0027] [Medline: 31152099]
38. Zhang V, Zhu P, Wagner AL. Spillover of vaccine hesitancy into adult COVID-19 and influenza: the role of race, religion, and political affiliation in the United States. Int J Environ Res Public Health 2023 Feb 15;20(4):3376. [doi: 10.3390/ijerph20043376] [Medline: 36834071]

Abbreviations

NIP: National Immunization Program

Edited by A Mavragani; submitted 13.10.24; peer-reviewed by E Taherifard, Y Fang; revised version received 26.01.25; accepted 11.02.25; published 26.03.25.

Please cite as:
Li M, Sun C, Ji C, Gao M, Wang X, Yao D, Guo J, Sun L, Rafay A, George AS, Muhandiramge SHSS, Bai G
Vaccine Hesitancy and Associated Factors Among Caregivers of Children With Special Health Care Needs in the COVID-19 Era in China: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e67487
URL: https://publichealth.jmir.org/2025/1/e67487
doi:10.2196/67487

Original Paper

Predictors of COVID-19 Vaccination Intention and Behavior Among Young People in a European Union Country With Low COVID-19 Vaccination Rates: Cross-Sectional Study

Sara Atanasova[1], PhD; Tanja Kamin[1], PhD; Nina Perger[1], PhD

[1]Faculty of Social Sciences, University of Ljubljana, Ljubljana, Slovenia

Corresponding Author:
Sara Atanasova, PhD
Faculty of Social Sciences
University of Ljubljana
Kardeljeva pl. 5
Ljubljana, 1000
Slovenia
Phone: 386 1 5805 167
Email: sara.atanasova@fdv.uni-lj.si

Abstract

Background: Vaccination against COVID-19 is a critical measure for managing the pandemic and achieving herd immunity. In 2021, Slovenia had a significantly lower COVID-19 vaccination rate compared to the average rate in the European Union, with individuals aged younger than 37 years showing the highest hesitancy. Previous studies primarily explored vaccination willingness before vaccines were available to young people, leaving a gap in understanding the factors influencing vaccination behavior and differences within the population of young people.

Objective: This study aimed to investigate a wide set of predictors influencing COVID-19 vaccination intention and behavior among young people in Slovenia. Specifically, we aimed to compare vaccinated and unvaccinated young people, further categorizing the unvaccinated group into those who were hesitant, those who intended to vaccinate in the near future, and those who refused vaccination.

Methods: An integrated model, based on the health belief model and theory of planned behavior, was developed, and it included additional contextual factors (such as trust in science, trust in vaccines, conspiracy theory tendencies, etc) and health-related and sociodemographic characteristics. Data were collected in August 2021 via the online access survey panel JazVem (Valicon), targeting individuals aged 15-30 years in Slovenia. Quotas ensured that the sample (n=507) was quasi-representative according to age, gender, education, and region. Bivariate analyses and multinomial logistic regression were performed to explore the determinants of vaccination intention and behavior.

Results: Among respondents, 45.8% (232/507) were vaccinated, 30.0% (152/507) refused vaccination, 12.4% (63/507) were hesitant, and 11.8% (60/507) intended to undergo vaccination in the near future. Vaccinated individuals were predominantly aged 23-26 years, had higher education, and reported above-average material status. Refusers were more common among the youngest (15-18 years) and oldest (27-30 years) groups, had lower education, and showed higher conspiracy theory tendencies. Multinomial regression analysis revealed that unvaccinated respondents who perceived greater COVID-19–related health consequences were more likely to delay vaccination (adjusted odds ratio [aOR] 2.0, 95% CI 1.2-3.3) or exhibit hesitancy (aOR 1.9, 95% CI 1.1-3.2) compared with vaccinated respondents. Subjective norms were less influential among hesitant individuals (aOR 0.4, 95% CI 0.2-0.7) and refusers (aOR 0.3, 95% CI 0.2-0.7) than among vaccinated individuals. Self-efficacy in managing health problems was less evident among those who delayed vaccination to the near future (aOR 0.5, 95% CI 0.3-0.9) than among vaccinated individuals.

Conclusions: This study underscores the complexity of vaccination intentions and behaviors among young people, emphasizing the necessity for public health strategies promoting vaccination to be tailored to the specific reasons for nonvaccination within different subgroups. Interventions aimed at addressing vaccine hesitancy and delays should particularly focus on individuals with lower education and material disadvantages. By fostering trust and enhancing self-efficacy, these interventions could more effectively promote vaccine uptake.

(JMIR Public Health Surveill 2025;11:e64653) doi:10.2196/64653

KEYWORDS

vaccine uptake; young people; COVID-19 vaccine; health belief model; theory of planned behavior

Introduction

The rapid development of COVID-19 vaccines is considered as a pandemic success story, and vaccination against COVID-19 is considered as one of the most effective preventive measures [1,2]. The herd immunity threshold for SARS-CoV-2 has been assumed to be between 50% and 67%, and the administration of SARS-CoV-2 vaccines was suggested to be an important strategy to reach this threshold [3]. It has been suggested that around 75% to 90% of the population needs to be vaccinated against COVID-19 to successfully control the pandemic [4]. Slovenia lagged far behind this goal; it was among the countries with the lowest vaccination rate against COVID-19 in the European Union (EU) since vaccination against COVID-19 became available to citizens [5]. By the end of 2021, the 27 EU member states had vaccinated an average of 77% of their population aged 18 years or older. However, Slovenia reported a vaccination completion rate of only 65.6% among adults [6].

Individuals showing the most hesitancy regarding COVID-19 vaccination in Slovenia were those younger than 37 years [7]. This is consistent with the findings of previous studies globally that among adults, younger age groups had a lower willingness to be vaccinated against COVID-19 [8-15]. Vaccination against COVID-19 among those most likely to get infected and to transmit the virus, such as young people in school settings, was identified as a crucial strategy for reducing ongoing transmission and preventing outbreaks [16]. This approach was considered essential for establishing herd immunity; protecting elderly people and immunocompromised individuals from severe outcomes, hospitalization, and death [17]; and ensuring the continued functioning of the workforce [16].

The vaccination behavior of young people is complex, as highlighted by a systematic review of qualitative studies on adolescents' understanding and attitudes toward vaccines [18]. It has been suggested that health literacy plays a crucial role in adolescents' comprehension of the benefits and risks of vaccination. However, other factors, such as peer influence, parental guidance, trust in health care providers, and misconceptions about the vaccine's safety and efficacy, also significantly influence vaccine uptake among adolescents [18]. Given the unprecedented speed at which COVID-19 vaccines were developed compared to historical vaccine development and testing timelines, a higher degree of hesitancy toward COVID-19 vaccination was somewhat anticipated. A study examining the factors that discourage individuals from participating in vaccine trial registries or clinical trials identified primary concerns related to the safety of novel vaccines and a lack of trust in those involved in vaccine development [19].

Similar findings were obtained for vaccination behavior related to COVID-19, with attitude toward vaccination, trust in health staff and scientists, time of information, and conspiracy beliefs about COVID-19 being identified as the best predictors of the intention of vaccination among young adults in Spain at the beginning of the vaccination campaign [20]. Chaufan [21], for example, questioned the prevailing risk-benefit analysis with regard to younger people. If vaccine uptake significantly reduces health risks in the general population, the younger population has a much lower risk of COVID-19–related complications. Given their lower risk of COVID-19–related health complications, the reluctance to vaccinate in this population group, which is often seen as unreasonable, may seem understandable [22]. Their perceived benefits of vaccination seem to be more collectively beneficial, while the risks associated with vaccination are perceived to be individual. This vaccination was new, and there was no broad-based data available on the efficacy and immunogenicity of the vaccination in the young population at the beginning of the vaccination campaign against COVID-19 [23].

Another scoping review suggested that the primary factors of young people's acceptance of the COVID-19 vaccine include the desire to protect themselves and close family members or friends, fear of infection, professional recommendations, and employer obligations [24]. On the other hand, the primary hesitancy factors include concerns about the safety of the vaccine and its side effects, effectiveness, and efficacy, as well as a lack of trust in the pharmaceutical industry and government, conspiracies, and favoring natural immunity [24]. The same study also suggested the need for additional research into COVID-19 vaccine–related decision-making dynamics for specific adolescent and youth population age ranges to better understand how vaccination-related behavior is influenced by environmental or social factors as well as personal health and susceptibility [24].

The significant affective distress experienced by the younger population due to COVID-19–related protective measures, particularly isolation and social distancing [25], may substantially influence both their willingness to receive the vaccination, as suggested by a study [26], and their actual uptake of COVID-19 vaccination. This is especially pertinent if vaccination is perceived as a means to achieve greater mobility and fewer restrictions on in-person social interactions. Negative emotions (affective responses) associated with recommended protective behaviors may impact core beliefs (cognitive responses) and undermine evidence-based reasoning, thereby diminishing the perceived importance of these behaviors in decision-making processes (behavioral responses) [25]. These findings highlight that vaccination intention and behavior depend on several contextual factors. Most studies conducted prior to our research primarily focused on investigating the willingness to receive COVID-19 vaccination at a time when vaccines were not yet available to young people. Furthermore, studies involving the youth have often relied on opportunistic samples, predominantly drawn from student populations [27-30]. Such approaches have resulted in narrow and unrepresentative samples that fail to adequately reflect the broader category of young people. Our study aims to address these limitations.

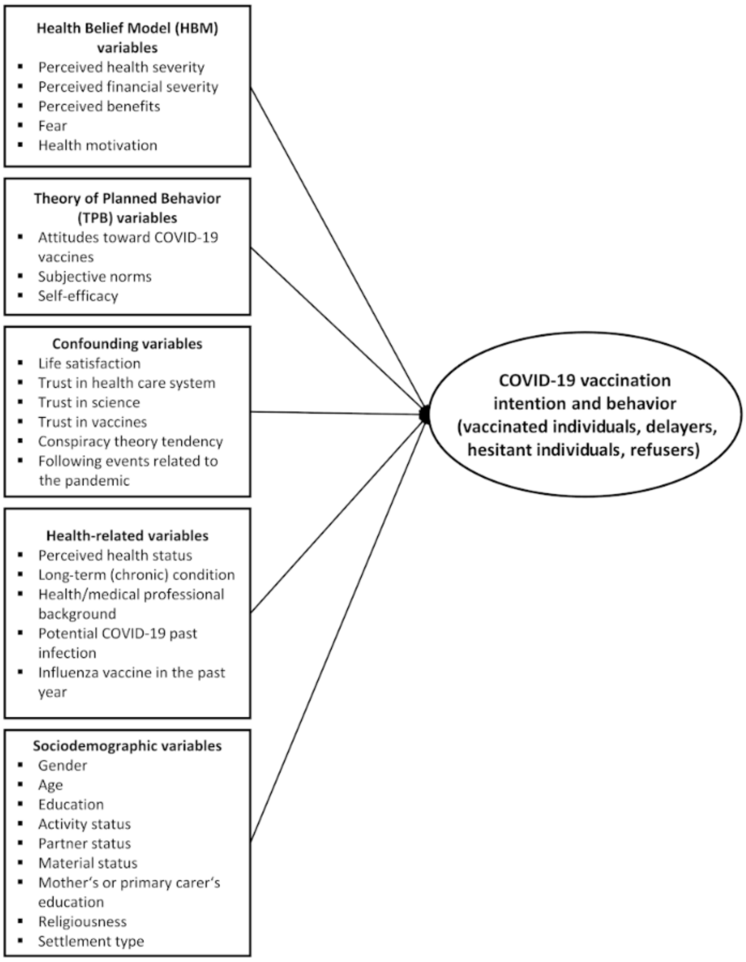
The main aim of this study is to investigate a wide set of predictors influencing COVID-19 vaccination intention and behavior among young people in Slovenia. Specifically, based on a quasi-representative sample of young people in Slovenia, this study aims to gain a deeper understanding of the cognitive, affective, and behavioral determinants of vaccine uptake within this population. It examines factors related to COVID-19 vaccination intention and behavior, including sociodemographic and health-related variables, as well as other predictors identified in previous studies and theoretical frameworks as significant in influencing vaccination behavior [14,24]. The study adopts and further develops the model by Shmueli [31], incorporating factors from 2 behavioral models (health belief model [HBM] and theory of planned behavior [TPB]) along with additional social factors such as trust in science and tendency toward conspiracy theories. These factors have been shown to be strong predictors of the willingness to receive the COVID-19 vaccine in the adult population [32]. Building on this integrated model, the study contributes by comparing vaccinated and unvaccinated young people and further categorizing the unvaccinated group into those who are hesitant, those who intend to vaccinate, and those who refuse vaccination. Our goal is to enhance the body of research that explores the decision-making processes within distinct groups of young people. The study seeks to better understand and inform public health strategies for increasing vaccination rates for COVID-19 and future viral outbreaks among young people.

Methods

Framework of the Study

Based on the study by Schmueli [31], we developed an integrated model that integrates the constructs of 2 prominent health behavior theories (HBM and TPB) and includes a number of other contextual factors (eg, trust in science, trust in vaccines, conspiracy theory tendencies, etc) and health-related and sociodemographic characteristics of the studied population to identify the predictors of COVID-19 vaccination intention and behavior among young people in Slovenia (Figure 1). The HBM, which has generated a great deal of research interest in various health-related behaviors to date, is based on a set of core beliefs (cognitive factors) regarding risk susceptibility, risk severity, benefits, and barriers [33]. According to the widely tested TPB [34,35], the underlying behavioral intention as a direct antecedent of behavior is determined by the attitude toward the behavior, subjective norms, and perceived behavioral control. Both models aim to predict behavior, and in our case, we target COVID-19 vaccination–related behavior.

Figure 1. Framework of the predictors of COVID-19 vaccination intention and behavior.



Study Design

The data for this cross-sectional study were collected using an online survey questionnaire administered by the online survey panel JazVem (Valicon, a Slovenian marketing research company) [36]. Existing members of the JazVem online survey panel were contacted via email to participate in the survey. Members voluntarily sign up with the JazVem online survey panel, agree with the terms and privacy conditions of the online survey panel, sign participation consent, provide baseline information, and regularly receive emails inviting them to take part in different research studies. For participation in JazVem surveys, panelists receive small incentives in the form of points, which can be redeemed for modest rewards. Online panels operate under rigorous recruitment and quality assurance protocols to ensure participants can enroll only once, avoid overexposure to surveys, and maintain high levels of engagement. More information about the study design can be found in the CHERRIES (Checklist for Reporting Results of Internet e-Surveys) in Multimedia Appendix 1.

Data Collection

The online questionnaire used in this study was developed by the authors and included 79 questions in the Slovene language, which took around 15 to 20 minutes to answer. The questionnaire covered questions about the perception of COVID-19–related protective measures, questions regarding behavior for vaccination against COVID-19, questions related to HBM and TPB constructs, questions about trust in science and vaccines, questions about the respondent’s health, and questions related to sociodemographic characteristics. The questionnaire and measurement instruments were evaluated for content validity by 8 experts from the fields of public health, health communication, sociology, communication studies, statistics, social science methodology, and psychology.

Data were collected between August 11, 2021, and August 17, 2021. Eligible participants were aged 15 to 30 years and resided in Slovenia. Quotas were used to achieve a quasi-representative sample of the general population in Slovenia according to age, gender, education, and region based on data from the Statistical Office of the Republic of Slovenia. Based on quota sampling, 1197 JazVem panelists were invited to participate in the online survey. Among these, 568 clicked on the link for the online survey and 555 viewed the introduction page of the survey. Of the 568 respondents, 525 fully completed the survey, with a completion rate of 92.4% (525/568). After the data screening and cleaning procedures, the final sample comprised 507 respondents. The data were weighted to match the general population distribution according to age, gender, education, and region. More information about the data collection can be found in the CHERRIES in Multimedia Appendix 1.

Measurement Instruments

Vaccine Uptake

The dependent variable vaccine uptake was measured with the following three questions: (1) “Please tell us about your decision regarding the COVID-19 vaccination. Have you been vaccinated against COVID-19?” (2) “Do you intend to get vaccinated against COVID-19 in the next 30 days?” and (3) “Do you intend to get vaccinated against COVID-19 in the next 6 months?” The classification and naming of unvaccinated respondents was guided by the Stages of Change model [37]. This model posits several behavioral stages preceding the desired recommended behavior, which in our case was vaccination behavior: (1) Precontemplation: individuals in this stage are not considering vaccination; (2) Contemplation: individuals in this stage are considering vaccinating in the next 6 months, and according to the theory, they are considering the recommended behavior but still weighing the pros and cons; and (3) Preparation: individuals in this stage intend to get vaccinated in the next 30 days, and according to the theory, they may have taken some initial steps, such as seeking information or planning an appointment. Based on the respondents’ answers, we computed vaccine uptake variables and divided the respondents into 4 categories as follows: respondents who reported being vaccinated were categorized as “vaccinated respondents;” respondents who reported not being vaccinated but intended to get vaccinated in the next 30 days were categorized as “delayers;” respondents who reported not being vaccinated but intended to get vaccinated in the next 6 months were categorized as “hesitant respondents;” and respondents who reported not being vaccinated and who did not intend to get vaccinated in the next 30 days or 6 months were categorized as “refusers.”

The independent variables (ie, predictors of vaccine uptake among young people) were arranged into 5 groups: HBM variables, TPB variables, confounding variables, health-related variables, and sociodemographic variables.

HBM Variables

Perceived health severity was measured with an item adapted from Shmueli [31], with respondents asked to assess on a 5-point Likert scale (1 [insignificant consequences] to 5 [very serious consequences]) the possible health consequences in the case of COVID-19 infection.

Perceived financial severity was measured with an item adapted from Shmueli [31], with respondents asked to assess on a 5-point Likert scale (1 [insignificant consequences] to 5 [very serious consequences]) the possible financial consequences in the case of COVID-19 infection.

Perceived benefits were measured using 3 items from Chu and Liu [38], with 1 item assessing the perceived individual benefits of COVID-19 vaccines (eg, “COVID-19 vaccines are effective in preventing COVID-19”) and 2 items assessing the perceived community benefits of COVID-19 vaccines (eg, “Having myself vaccinated against COVID-19 is beneficial for the health of others in my community”). Exploratory factor analysis (EFA) indicated that the items represent a single factor, explaining 79.1% of the variance. The Cronbach alpha (α=.95) demonstrated outstanding internal consistency.

Fear was measured with an item from Chu and Liu [38], with respondents asked to assess on a 5-point scale (1 [not at all] to 5 [very much]) how afraid they feel when they think about COVID-19.

Health motivation was measured with 2 items from Shmueli [31]. These items (“I make sure to eat healthy and diverse food every day” and “I ensure that I exercise and work out regularly”) were summed in an index and demonstrated acceptable internal consistency (α=.77; r=0.6; P<.001; r_{SB}=0.77).

The HBM also includes susceptibility, perceived barriers, and cues to action, which were measured in our study but only in the subsample of unvaccinated respondents. Because the purpose of this study is to compare different groups of young people, including those who accept vaccination, we did not include these variables in the analysis.

TPB Variables

Attitudes toward COVID-19 vaccines were measured using 6 items adapted from Chu and Liu [38]. Respondents were asked to evaluate how they personally feel about COVID-19 vaccination using 5-point scales with various descriptors (negative to positive, undesirable to desirable, bad to good, harmful to beneficial, foolish to wise, and pointless to meaningful). The EFA indicated that the items represent a single factor, explaining 85.7% of the variance, and the Cronbach alpha (α=.97) demonstrated outstanding internal consistency.

Subjective norms were measured using an adapted scale comprising 3 items from Chu and Liu [38]. Respondents were asked to indicate their level of agreement on a 5-point Likert scale (1 [strongly disagree] to 5 [strongly agree]) with statements concerning their similar or significant others (eg, “Most people who are similar to me have been or will be vaccinated against COVID-19”). The EFA showed that the items represent a single factor, explaining 71.5% of the variance, and the Cronbach alpha (α=.88) demonstrated good internal consistency.

To measure self-efficacy, 3 items adapted from the General Self-Efficacy Scale of Schwarzer and Jerusalem [39] were used to ask respondents to rate their agreement on a 5-point scale (1 [strongly disagree] to 5 [strongly agree]) with statements related to managing their health (eg, “If unexpected events related to my health occur, I know how to handle them effectively”). The EFA revealed that the items represent a single factor, explaining 65.3% of the variance. The Cronbach alpha (α=.85) indicated good internal consistency.

Confounding Variables

Life satisfaction was measured using a single-item question asking respondents to assess their current satisfaction with life on a scale from 0 (extremely dissatisfied) to 10 (extremely satisfied).

Similarly, trust in the health care system was measured using a single-item question asking respondents to rate their trust in the Slovenian health care system on a scale from 0 (don’t trust at all) to 10 (completely trust).

Trust in science was measured using the Trust in Science and Scientists Inventory [40]. However, since the EFA did not confirm the unidimensionality of the original scale, the measure was slightly adjusted to include 9 items. These items asked respondents to assess their agreement on a 5-point Likert scale (1 [strongly disagree] to 5 [strongly agree]) with statements related to the value of and trust in science (eg, “I believe that the work of scientists contributes to a better life for everyone”). The EFA revealed that the items represent a single factor, explaining 55.1% of the variance, and the Cronbach alpha (α=.92) demonstrated outstanding internal consistency.

Trust in vaccines was measured using 2 items that asked respondents to indicate their agreement on a 5-point Likert scale (1 [strongly disagree] to 5 [strongly agree]) with statements related to the safety and efficiency of COVID-19 vaccines (eg, “I trust that the vaccines available to us against the coronavirus that causes COVID-19 are safe”). The 2 items were summed in an index and demonstrated outstanding internal consistency (α=.92; r=0.8; P<.001; r_{SB}=0.92).

Conspiracy theory tendency was measured using 3 items adapted from the Slovenian public opinion research #Novanormalnost [41]. Respondents were asked to assess their agreement with statements suggesting COVID-19 is a part of a conspiracy (eg, “Vaccination against the coronavirus that causes COVID-19 is an attempt to control the population”), using a 5-point Likert scale (1 [strongly disagree] to 5 [strongly agree]). The EFA revealed that the 3 items represent a single factor, explaining 70.3% of the variance, and the Cronbach alpha (α=.87) showed good internal consistency.

Following events related to the pandemic was measured using a single-item question that asked respondents to assess the extent to which they follow pandemic-related events on a 5-point Likert scale (1 [I don’t follow events at all] to 5 [I follow events very often]).

Health-Related Variables

Perceived health status was measured using a single-item question on a 5-point Likert scale (1 [very bad] to 5 [excellent]), asking respondents to evaluate their current physical and mental health. The online questionnaire also included a question asking respondents to indicate whether they have a long-term (chronic) condition.

Additionally, single-item questions were used to determine whether respondents have a health or medical professional background and whether they received an influenza vaccine in the past year.

Potential COVID-19 past infection was measured by asking respondents to report their perceived probability of having been infected with COVID-19 in the past (probably infected or probably not infected).

Sociodemographic Variables

The online questionnaire included questions asking respondents to indicate their gender and age, with age transformed from a numeric to a categorical variable (15-18 years, 19-22 years, 23-26 years, and 27-30 years) to examine differences across specific age groups. Respondents were also asked to report their current education level (elementary school or less; 2-, 3-, 4-, or 5-year high school; and college, university, or higher education) and their activity status (high school student, student, employed, and unemployed).

Single-item questions were used to determine whether respondents have a partner, consider themselves religious, and live in urban or rural settlements. Respondents were also asked to assess their mother’s or primary carer’s education level (elementary school or less; 2-, 3-, 4-, or 5-year high school; and college, university, or higher education) and their material status on a scale from 1 (significantly below average) to 5 (significantly above average). Due to a low number of observations in categories 1 (significantly below average) and 5 (significantly above average), the material status variable was recoded according to 3 categories: below average, average, and above average.

Statistical Analysis

Before the analysis, the data were weighted using the random iterative weighting method according to gender, age group, education, and region. The weighting process was conducted by Valicon, the JazVem online panel survey provider. The prevalence of vaccination-related behavior was measured as the percentage of vaccinated respondents, delayers, hesitant respondents, and refusers in the total study sample. A series of EFAs were conducted to determine the internal structure of the selected measurement instruments. Factors were extracted using principal axis factoring with oblimin rotation. We performed EFA using the following criteria: (1) eigenvalue greater than 1, (2) scree test, (3) items loading on the same factor (≥0.04), (4) no cross-loading, and (5) conceptual interpretability of factors. The quality of the measurement instrument was also assessed using reliability analysis, and the Cronbach α coefficient was computed, which ranges between 0 and 1.0. The rough guidelines are that a value of .9 or higher indicates outstanding internal consistency, a value of .8 or higher indicates good internal consistency, a value between .7 and .8 indicates acceptable reliability, and a value between .6 and .7 indicates questionable reliability [42]. For 2-item scales, reliability was estimated using the Spearman-Brown coefficient (r_{SB}), which was calculated based on the interitem Pearson correlation (r). Bivariate analyses were performed to compare groups according to sociodemographic characteristics, health-related variables, confounding variables, HBM variables, and TPB variables, using chi-square tests for categorical variables and ANOVA along with appropriate post hoc tests (Games-Howell) for continuous variables.

Multinomial logistic regression analysis was used to estimate the selected determinants of vaccine uptake with “vaccinated” as the reference group for comparisons. The selection of variables for the multinomial regression model was guided by their statistical significance in bivariate analyses and their theoretical relevance to the research question, ensuring the inclusion of predictors with stronger explanatory power while minimizing the risk of overfitting. Variables that did not show statistically significant associations with the outcome in bivariate analyses were excluded. While the inclusion of all variables was initially considered, challenges with model convergence underscored the importance of a parsimonious model that prioritizes interpretability and robustness. The final model, therefore, balances these considerations to provide reliable and meaningful insights. Confounding variables were included in the multinomial regression model alongside the primary independent variables (HBM and TPB variables, and sociodemographic and health-related variables). By incorporating these confounders into the model, we aimed to account for their potential influence on the outcome variable and to isolate the association between the primary independent variables and the dependent variable. This approach ensures that the estimated effects of the primary predictors are adjusted for the potential bias introduced by the confounding variables. We report adjusted odds ratios (aORs) and 95% CIs. Data were analyzed using IBM SPSS version 27 software.

Ethical Considerations

Data collection was conducted by the online survey panel provider Valicon, who is a member of the European Society for Opinion and Market Research (ESOMAR), and the data collection and research adhere to the International Chamber of Commerce and ESOMAR international code and professional standards of social research practice. We (the authors) had no access to the respondents’ information and were provided with an anonymized dataset that did not contain any identifiable personal information. The study was conducted in accordance with the Code of Ethics for Researchers of the University of Ljubljana [43] and the World Medical Association’s Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects [44]. The study received approval from the Ethics Committee of the Faculty of Social Sciences, University of Ljubljana (number: 801-2024-025/TD).

Results

Participant Characteristics

The results showed that the majority of respondents (232/507, 45.8%) had been vaccinated against COVID-19, 30.0% (152/507) refused vaccination, 12.4% (63/507) were hesitant, and 11.8% (60/507) reported that they intended to be vaccinated against COVID-19 in the near future (ie, delayers). The respondents were on average 23.4 years old (SD 4.4 years), comprising 51.6% (262/507) men and 48.4% (245/507) women (Table 1). The majority had completed high school (290/507, 57.3%), were employed (194/507, 38.3%), were students (184/507, 36.2%), and had a partner (258/507, 50.9%). On average, respondents reported the perceived material status as average (mean 2.1, SD 0.64). Most participants identified as religious (316/507, 62.3%) and resided in urban areas (260/507, 51.3%).

Table 1. Sociodemographic participant characteristics in the total sample and according to COVID-19 vaccination–related behavior groups.

Variable	Total sample (N=507), n (%)	Vaccinated individuals (n=232, 45.8%), n (%)	Delayers (n=60, 11.8%), n (%)	Hesitant individuals (n=63, 12.4%), n (%)	Refusers (n=152, 30.0%), n (%)	Chi-square (*df*)	*P* value
Gender						2.8 (3)	.43
Male	262 (51.6)	127 (48.5)	33 (12.6)	29 (11.1)	73 (27.9)		
Female	245 (48.4)	105 (42.9)	27 (11.0)	34 (13.9)	79 (32.2)		
Age (years) groups						39.7 (9)	<.001
15-18	102 (20.2)	28 (27.5)	18 (17.6)	17 (16.7)	39 (38.2)		
19-22	123 (24.3)	66 (53.7)	18 (14.6)	14 (11.4)	25 (20.3)		
23-26	133 (26.3)	81 (60.9)	6 (4.5)	11 (8.3)	35 (26,3)		
27-30	148 (29.2)	57 (38.5)	17 (11.7)	20 (13.5)	54 (36.5)		
Education						25.6 (6)	<.001
Elementary school or less	69 (13.6)	15 (21.7)	12 (17.4)	14 (20.3)	28 (40.6)		
2-, 3-, 4-, or 5-year high school	290 (57.3)	131 (45.2)	35 (12.1)	34 (11.7)	90 (31.0)		
College, university, or higher education	147 (29.1)	85 (57.8)	12 (8.2)	15 (10.2)	35 (23.8)		
Activity status						34.9 (9)	<.001
High school student	106 (20.9)	31 (29.2)	18 (17.0)	14 (13.2)	43 (40.6)		
Student	184 (36.2)	109 (59.2)	22 (12.0)	16 (8.7)	37 (20.1)		
Employed	194 (38.3)	85 (43.8)	16 (8.2)	28 (14.4)	65 (33.5)		
Unemployed	23 (4.5)	7 (30.4)	4 (17.4)	4 (17.4)	8 (34.8)		
Partner status						7.4 (3)	.06
Do not have a partner	249 (49.1)	109 (43.8)	39 (15.7)	32 (12.9)	69 (27.7)		
Have a partner	258 (50.9)	123 (47.7)	21 (8.1)	31 (12.0)	83 (32.2)		
Material status (n=494)						24.5 (6)	<.001
Below average	81 (16.4)	26 (32.1)	9 (11.1)	11 (13.6)	35 (43.2)		
Average	288 (58.3)	123 (42.7)	37 (12.8)	36 (12.5)	92 (31.9)		
Above average	125 (25.3)	78 (62.4)	8 (6.4)	15 (12.0)	24 (19.2)		
Mother’s or primary carer’s education (n=488)						17.4 (6)	.008
Elementary school or less	40 (8.2)	13 (32.5)	5 (12.5)	6 (15.0)	16 (40.0)		
2-, 3-, 4-, or 5-year high school	242 (49.4)	99 (40.9)	23 (9.5)	31 (12.8)	89 (36.8)		
College, university, or higher education	206 (42.2)	112 (54.4)	26 (12.6)	24 (11.7)	44 (21.4)		
Religiousness						7.5 (3)	.06
Nonreligious	191 (37.7)	100 (52.4)	15 (7.9)	22 (11.5)	54 (28.3)		
Religious	316 (62.3)	132 (41.8)	45 (14.2)	41 (13.0)	98 (31.0)		
Settlement type						13.1 (3)	.004
Urban	260 (51.3)	137 (52.7)	32 (12.3)	28 (10.8)	63 (24.2)		
Rural	247 (48.7)	95 (38.3)	28 (11.3)	35 (14.1)	90 (36.3)		

The majority of respondents reported having a very good (254/507, 50.2%) or good (150/507, 29.6%) health status (mean 3.7, SD 0.8). Additionally, 67.9% (344/507) of respondents indicated that they were probably not infected by coronavirus in the past (Table 2).

Table 2. Health-related participant characteristics in the total sample and according to COVID-19 vaccination–related behavior groups.

Variable	Total sample (N=507), n (%)	Vaccinated individuals (n=232, 45.8%), n (%)	Delayers (n=60, 11.8%), n (%)	Hesitant individuals (n=63, 12.4%), n (%)	Refusers (n=152, 30.0%), n (%)	Chi-square (*df*)	*P* value
Perceived health status						20.9 (12)	.05
Very bad	6 (1.2)	2 (33.3)	1 (16.7)	0 (0.0)	3 (50.0)		
Bad	32 (6.3)	18 (56.3)	5 (15.6)	4 (12.5)	5 (15.6)		
Good	150 (29.6)	63 (42.0)	27 (18.0)	21 (14.0)	39 (26.0)		
Very good	254 (50.2)	118 (46.5)	20 (7.9)	34 (13.4)	82 (32.3)		
Excellent	64 (12.6)	30 (46.9)	6 (9.4)	3 (4.7)	25 (39.1)		
Long-term (chronic) condition						4.3 (3)	.23
No	399 (78.7)	181 (45.3)	43 (10.8)	48 (12.0)	128 (32.0)		
Yes	108 (21.3)	51 (47.2)	17 (15.7)	15 (13.9)	25 (23.1)		
Potential COVID-19 past infection						14.4 (3)	.002
Probably not infected	344 (67.9)	171 (49.7)	46 (13.4)	36 (10.5)	91 (26.5)		
Probably infected	163 (32.1)	60 (36.8)	14 (8.6)	27 (16.6)	62 (38.0)		
Health or medical professional background						5.3 (3)	.15
No	428 (84.4)	192 (44.9)	46 (10.7)	55 (12.9)	135 (31.5)		
Yes	79 (15.6)	40 (50.0)	14 (17.5)	8 (10.0)	18 (22.5)		
Influenza vaccine in the past year						19.1 (3)	<.001
No	457 (90.1)	195 (42.7)	55 (12.0)	60 (13.1)	147 (32.2)		
Yes	50 (9.9)	37 (74.0)	5 (10.0)	3 (6.0)	5 (10.0)		

Bivariate Analyses and Comparisons Among Vaccination-Related Behavior Groups

Bivariate analyses and comparisons among vaccination-related behavior groups (Table 1) showed that vaccinated individuals were predominantly aged 23 to 26 years (81/133, 60.9%) and were more likely to have a higher education (85/147, 57.8%) and above-average material status (78/125, 62.4%). In comparison, refusers were more common among younger (15-18 years: 39/102, 38.2%) and older age groups (27-30 years: 54/148, 36.5%), had lower levels of education (elementary or less: 28/69, 40.6%), reported below-average material status (35/81, 43.2%), and were more likely from rural areas (90/247, 36.3%; Table 1). Additionally, refusers more often reported a potential COVID-19 past infection (62/163, 38.0%; Table 2). Results also showed that delayers shared characteristics with both vaccinated individuals and refusers, whereas hesitant respondents shared more characteristics with refusers, including rural residence, a lower education level, and a higher rate of potential COVID-19 past infection (Tables 1 and 2).

Table 3 presents participant characteristics across vaccination behavior groups based on confounding, HBM, and TPB variables. Vaccinated individuals reported significantly higher trust in the health care system (mean 6.5, SD 2.2), science (mean 3.9, SD 0.6), and COVID-19 vaccines (mean 3.9, SD 0.9) compared to refusers (Table 3). Conspiracy belief tendencies were on average the highest among refusers (mean 3.6, SD 1.0) and on average the lowest among vaccinated individuals (mean 1.9, SD 1.0). Pandemic-related events were followed more closely by vaccinated individuals (mean 3.1, SD 1.0) and delayers (mean 2.9, SD 0.9) than hesitant individuals (mean 2.7, SD 1.0) and refusers (mean 2.5, SD 1.0). Delayers perceived greater health (mean 2.8, SD 0.9) and financial (mean 2.7, SD 1.1) consequences from COVID-19 compared to refusers. Vaccinated individuals perceived the greatest benefit of vaccination (mean 4.0, SD 0.9), while refusers perceived the lowest benefit (mean 1.8, SD 0.8). Fear of COVID-19 was the lowest among refusers (mean 1.9, SD 1.0) and the highest among delayers (mean 3.0, SD 1.2). According to the TPB model, vaccinated individuals had the highest average scores for positive attitudes toward COVID-19 vaccines (mean 4.1, SD 1.1) and subjective norms (mean 3.8, SD 0.9), which gradually decreased among other groups and were the lowest on average among refusers (Table 2). Self-efficacy in coping with health problems was the highest among vaccinated individuals (mean 3.8, SD 0.7) and refusers (mean 3.7, SD 0.8) but was the lowest among delayers (mean 3.4, SD 0.7). Health motivation did not differ significantly between groups (Table 3).

Table 3. Descriptive statistics of confounding variables, health belief model variables, and theory of planned behavior variables in the total sample and according to COVID-19 vaccination–related behavior groups.

Variable	Total sample (N=507), mean (SD)	Vaccinated individuals (n=232, 45.8%), mean (SD)	Delayers (n=60, 11.8%), mean (SD)	Hesitant individuals (n=63, 12.4%), mean (SD)	Refusers (n=152, 30.0%), mean (SD)	ANOVA[a] *F* (*df*)	*P* value
Confounding variables							
Life satisfaction (scale 0-10)	6.6 (2.2)	6.8 (2.0)	6.1 (2.1)	6.8 (1.9)	6.5 (2.5)	2.1 (3, 502)	.10
Trust in the health care system (scale 0-10)	5.8 (2.4)	6.5 (2.2)[b]	5.8 (1.6)[b,c]	5.8 (2.4)[d]	4.7 (2.5)[b,c,d]	19.8 (3, 502)	<.001
Trust in science (scale 1-5)	3.5 (0.8)	3.9 (0.6)[b,c,d]	3.6 (0.6)[b,c]	3.4 (0.67)[b,d]	3.1 (0.8)[b,c,d]	46.1 (3, 502)	<.001
Trust in vaccines (scale 1-5)	3.0 (1.3)	3.9 (0.9)[b,c,d,e]	3.1 (0.8)[b,c,d,e]	2.5 (0.9)[b,c,d,e]	1.6 (0.8)[b,c,d,e]	226.6 (3, 502)	<.001
Conspiracy theory tendency (scale 1-5)	2.7 (1.2)	1.9 (1.0)[b,c,d,e]	2.6 (0.8)[b,c,d,e]	3.1 (0.8)[b,c,d,e]	3.6 (1.0)[b,c,d,e]	98.3 (3, 502)	<.001
Following events related to the pandemic (scale 1-5)	2.9 (1.0)	3.1 (1.0)[b,d]	2.9 (0.9)[c]	2.7 (1.0)[b,d]	2.5 (1.0)[b,c]	12.6 (3, 502)	<.001
Health belief model variables							
Perceived health severity (scale 1-5)	2.3 (1.0)	2.4 (0.9)[b]	2.8 (0.9)[b,c]	2.7 (0.9)[d]	1.9 (0.8)[b,c,d]	23.3 (3, 502)	<.001
Perceived financial severity (scale 1-5)	2.3 (1.1)	2.4 (1.2)[b]	2.7 (1.1)[c]	2.3 (1.0)	2.0 (1.00)[b,c]	6.1 (3, 502)	<.001
Perceived benefits (scale 1-5)	3.0 (1.3)	4.0 (0.9)[b,c,d,e]	3.1 (0.7)[b,c,d,e]	2.5 (0.9)[b,c,d,e]	1.8 (0.8)[b,c,d,e]	221.8 (3, 502)	<.001
Fear (scale 1-5)	2.6 (1.2)	2.8 (1.1)[b]	3.0 (1.2)[c]	2.9 (1.0)[d]	1.9 (1.0)[b,c,d]	31.0 (3, 502)	<.001
Health motivation (scale 1-5)	3.5 (0.9)	3.5 (0.9)	3.5 (0.9)	3.6 (0.8)	3.6 (1.0)	0.2 (3, 502)	.87
Theory of planned behavior variables							
Attitudes (scale 1-5)	3.0 (1.5)	4.1 (1.1)[b,c,d,e]	3.0 (1.0)[b,c,d,e]	2.4 (0.9)[b,c,d,e]	1.4 (0.7)[b,c,d,e]	260.2 (3, 502)	<.001
Subjective norms (scale 1-5)	3.0 (1.1)	3.8 (0.9)[b,c,d,e]	3.1 (0.6)[b,c,d,e]	2.5 (0.9)[b,c,d,e]	1.9 (0.8)[b,c,d,e]	159.5 (3, 502)	<.001
Self-efficacy (scale 1-5)	3.7 (0.7)	3.8 (0.7)[b]	3.4 (0.7)[b,c]	3.6 (0.7)	3.7 (0.8)[c]	4.8 (3, 502)	<.001

[a] Games-Howell post hoc test.

[b] Category has a statistically different mean value (*P*<.05) of the corresponding variable in comparison to the mean value in the other category with the same superscript.

[c] Category has a statistically different mean value (*P*<.05) of the corresponding variable in comparison to the mean value in the other category with the same superscript.

[d] Category has a statistically different mean value (*P*<.05) of the corresponding variable in comparison to the mean value in the other category with the same superscript.

[e] Category has a statistically different mean value (*P*<.05) of the corresponding variable in comparison to the mean value in the other category with the same superscript.

Predictors of Vaccine Uptake

The results of multinomial logistic regression models with predictors of vaccine uptake, using “vaccinated respondents” as the reference group for comparisons, are shown in Table 4. Variables included in the multinomial regression model were selected based on their statistical significance in bivariate analyses and their theoretical relevance to the research question (more details are provided in the Statistical Analysis subsection in the Methods section). Multinomial regression analysis (Table 4) showed that young adults aged 23 to 26 years (compared with those aged 27 to 30 years) were less likely to delay vaccination (aOR 0.1, 95% CI 0.031-0.5). Young adults who had completed elementary school or less (compared to respondents with higher levels of education) were more likely to delay vaccination (aOR 9.8, 95% CI 1.6-61.1) or were hesitant toward vaccination (aOR 11.7, 95% CI 1.5-89.7). Results also showed that young people who were high school students (compared with unemployed respondents) were less likely to be hesitant (aOR 0.0081, 95% CI 0.0002-0.4; Table 4). Similarly, respondents who were students (aOR 0.1, 95% CI 0.0077-0.9) or employed (aOR 0.1, 95% CI 0.0059-0.6; compared with unemployed respondents) were also less likely to refuse vaccination. Participants whose mothers or primary carers had completed elementary school or less were significantly more likely to refuse vaccination (aOR 6.6, 95% CI 1.3-34.4) compared to those with university-educated mothers or primary carers. Similarly, those whose mothers or primary carers finished high school were more likely to refuse vaccination (aOR 2.8, 95% CI 1.0-8.0) compared to the same university-educated group. The results also showed that respondents who felt that COVID-19 could have greater health consequences for them were more likely to delay vaccination (aOR 2.0, 95% CI 1.2-3.3) or be hesitant toward it (aOR 1.9, 95% CI 1.1-3.2). In addition, more negative attitudes toward COVID-19 vaccines were more likely among delayers, hesitant respondents, and refusers than among vaccinated respondents (Table 4). Compared with vaccinated young people, subjective norms were significantly less likely to be present among hesitant respondents (aOR 0.4, 95% CI 0.2-0.7) and refusers (aOR 0.3, 95% CI 0.2-0.7). Self-efficacy to cope with health problems was less present among respondents who delayed vaccination (aOR 0.5, 95% CI 0.3-0.9) than among vaccinated respondents (Table 4). For the data in Table 4, the R^2 was 0.7 (Cox and Snell) and 0.8 (Negelkerke), and the model χ^2 was 610.2 (*df*=81; *P*<.001).

Table 4. Predictors of vaccine uptake by multinomial logistic regression (vaccinated respondents as the reference group).

Predictor variable	Delayers, aOR[a] (95% CI)	Hesitant individuals, aOR (95% CI)	Refusers, aOR (95% CI)
Sociodemographic variables			
Age group (reference: 27-30 years)			
15-18 years	2.9 (0.2-39.4)	22.2 (0.8-60.2)	3.13 (0.1-124.9)
19-22 years	0.6 (0.12-2.6)	0.8 (0.2-3.6)	0.2 (0.035-1.2)
23-26 years	0.1[b] (0.031-0.5)	0.3 (0.1-1.1)	0.4 (0.1-1.5)
Education (reference: college, university, or higher education)			
Elementary school or less	9.8[c] (1.6-61.1)	11.7[b] (1.5-89.7)	6.8 (0.7-65.3)
2-, 3-, 4-, or 5-year high school	1.9 (0.7-5.5)	2.5 (0.9-7.4)	4.0[c] (1.3-12.6)
Activity status (reference: unemployed)			
High-school student	0.2 (0.007-4.1)	0.0081[b] (0.0002-0.4)	0.1 (0.00088-3.0)
Student	0.7 (0.1-6.0)	0.1 (0.015-1.2)	0.1[c] (0.0077-0.9)
Employed	0.4 (0.1-3.1)	0.2 (0.023-1.4)	0.1[c] (0.0059-0.6)
Material status (reference: above average)			
Below average	1.8 (0.4-7.5)	1.3 (0.3-5.6)	3.3 (0.7-17.3)
Average	1.9 (0.7-5.5)	0.6 (0.2-1.9)	1.3 (0.4-4.4)
Mother’s or primary carer’s education (reference: college, university, or higher education)			
Elementary school or less	2.2 (0.5-9.8)	4.2 (0.9-19.7)	6.6[c] (1.3-34.4)
2-, 3-, 4-, or 5-year high school	0.9 (0.4-2.0)	1.4 (0.5-3.5)	2.8[c] (1.0-8.0)
Settlement type (reference: rural)			
Urban	0.9 (0.4-2.0)	0.9 (0.4-2.1)	0.9 (0.3-2.1)
Health-related variables			
Potential COVID-19 past infection (reference: probably infected)			
Probably not infected	1.2 (0.5-2.9)	0.6 (0.2-1.4)	0.9 (0.3-2.6)
Influenza vaccine in the past year (reference: yes)			
No	1.4 (0.4-5.0)	1.9 (0.4-10.3)	1.5 (0.2-13.0)
Cofounding variables			
Trust in the health care system	1.1 (0.9-1.4)	1.2 (1.0-1.5)	1.0 (0.8-1.3)
Trust in science	2.2 (0.9-4.9)	1.2 (0.5-2.6)	2.1 (0.9-5.0)
Trust in vaccines	0.7 (0.3-1.3)	0.7 (0.4-1.5)	0.5 (0.2-1.1)
Conspiracy theory tendency	0.9 (0.5-1.6)	1.5 (0.8-2.6)	1.3 (0.7-2.3)
Following events related to the pandemic	0.9 (0.6-1.4)	0.8 (0.5-1.2)	0.8 (0.5-1.3)
Health belief model variables			
Perceived health severity	2.0[b] (1.2-3.3)	1.9[b] (1.1-3.2)	0.8 (0.4-1.4)
Perceived financial severity	1.3 (0.9-1.8)	0.8 (0.5-1.2)	0.8 (0.5-1.3)
Perceived benefits	0.7 (0.3-1.3)	0.5 (0.2-1.1)	0.6 (0.3-1.4)
Fear	0.9 (0.6-1.3)	1.2 (0.8-1.8)	0.7 (0.5-1.1)
Theory of planned behavior variables			
Attitudes	0.5[c] (0.3-0.9)	0.5[b] (0.3-0.9)	0.2[d] (0.1-0.4)
Subjective norms	0.7 (0.4-1.3)	0.4[b] (0.2-0.7)	0.3[d] (0.2-0.7)

Predictor variable	Delayers, aOR[a] (95% CI)	Hesitant individuals, aOR (95% CI)	Refusers, aOR (95% CI)
Self-efficacy	0.5[c] (0.3-0.9)	0.9 (0.5-1.7)	1.0 (0.5-2.0)

[a] aOR: adjusted odds ratio.

[b] $P<.01$.

[c] $P<.05$.

[d] $P<.001$.

Discussion

Principal Findings

The aim of this study was to gain a deeper understanding of the determinants influencing vaccination intention and behavior among young people in Slovenia, an EU country where the adult COVID-19 vaccination rate significantly lagged the EU average. We applied an integrated model to analyze and compare vaccinated and unvaccinated young people, further categorizing the unvaccinated people into 3 groups: those hesitant to vaccinate, those willing but delaying vaccination, and those refusing vaccination against COVID-19. Our goal was to enhance the body of research exploring the decision-making processes within these distinct groups of young people regarding COVID-19 vaccination.

In our study, a substantial proportion of respondents self-reported being vaccinated against COVID-19. This rate was slightly higher than the vaccination rate of the general population (aged 18 years or older) in Slovenia, which was approximately 43% for the first dose by the end of August 2021 [5]. Compared to the vaccinated general population, this is a favorable percentage, as previous studies have indicated that younger adults are among the groups with lower willingness to be vaccinated against COVID-19 [8-15]. This finding may be partially explained by policy interventions in Slovenia during the time of data collection. Notably, the introduction of the so called “COVID-19 pass” (certificate of recovery, vaccination, or testing) was designed to facilitate safer and freer mobility of citizens within the EU. Several studies have shown that the implementation of COVID-19 passes improved both vaccination uptake and intention [45].

Our study demonstrated that unvaccinated individuals are not a homogeneous group. Not all respondents who were unvaccinated at the time of data collection had explicitly refused vaccination against COVID-19. Specifically, almost half of these respondents indicated that they had considered getting vaccinated but had not yet done so. Some of these were hesitant, while others expressed a willingness to be vaccinated but had not yet taken action. Only half of the unvaccinated respondents explicitly refused vaccination against COVID-19.

In comparison to the findings of other studies [46], our study did not find gender to be a significant factor influencing vaccination behavior. However, the influence of other sociodemographic and health-related variables on vaccine uptake among young people aligns with previous research [46]. Key predictors of vaccination were age, education level, activity status, material status, mother’s or primary carer’s education, settlement type, potential COVID-19 past infection, and influenza vaccination in the past year. Vaccinated individuals were predominantly aged 23 to 26 years, were more likely to have higher education, and had above-average material status. In contrast, vaccine refusers were more common among the youngest (15-18 years) and oldest (27-30 years) age groups, had lower education levels (elementary or less), had mothers or primary caregivers with lower education, reported below-average material status, were more likely to reside in rural areas, and more frequently reported a potential past COVID-19 infection. These findings suggest that the factors significantly predicting vaccination behavior are closely related to the primary socioeconomic and cultural resources available to an individual [47]. Delayers, while distinct from both vaccinated individuals and refusers, shared some sociodemographic and health-related characteristics with each group. Hesitant respondents, on the other hand, shared more characteristics with refusers, including rural residence, lower education levels, and higher rates of potential past COVID-19 infection.

The factors from the HBM and TPB models that significantly differentiated the vaccinated respondents from all other groups were greater trust in science, trust in vaccines, trust in the health care system, and positive attitudes toward COVID-19 vaccination. These characteristics appeared to be strongly associated with higher levels of education, as suggested by previous studies [48,49]. Perceived health consequences of COVID-19, negative attitudes toward vaccines, and weaker levels of subjective norms were associated with a greater likelihood of delaying or refusing vaccination. Among unvaccinated respondents, perceived self-efficacy in coping with health problems emerged as a significant factor, particularly for those who were delaying vaccination, as this group reported the lowest levels of self-efficacy. Interestingly, despite exhibiting heightened fear of COVID-19 and perceiving the health and financial burdens of contracting the disease as greater than those reported by other groups, delayers may experience affective and cognitive conflict [25]. This conflict likely stems from the simultaneous presence of fear of contracting the disease and negative attitudes toward COVID-19 vaccines. Combined with low confidence in their ability to manage health issues, this internal conflict may explain their expressed willingness to vaccinate but delay in taking action.

The findings of our study indicated that hesitant respondents shared many similarities with delayers, including the same identified conflict between fear of COVID-19 and negative attitudes toward vaccines. However, hesitant respondents exhibited higher self-efficacy in coping with health problems and reported lower levels of subjective norms. Vaccination refusers, in contrast, reported lower levels of subjective norms, were more likely to have mothers or primary carers with lower levels of education (elementary school or less), and were more likely to hold negative attitudes toward COVID-19 vaccination. Notably, they exhibited a concerning combination of increased self-efficacy in coping with health problems and the highest scores for conspiracy theory tendencies. This profile aligns with a group previously described in the literature as dysfunctionally empowered [50].

Strengths and Limitations

The strength of this study lies in its comprehensive approach to examining the factors influencing COVID-19 vaccine uptake among young people in Slovenia. It addresses a critical research gap by focusing on the specific population group of young people who have frequently demonstrated hesitant attitudes toward vaccination [13]. By comparing vaccinated and unvaccinated individuals and further categorizing the unvaccinated individuals based on their willingness to be vaccinated, the study provides a detailed analysis of the factors underlying vaccination intention and hesitancy. These insights could inform the development of more targeted and effective public health interventions and strategies to address vaccine hesitancy and improve vaccination rates among the youth.

This study also has some limitations. First, data collection through an online access survey panel may have resulted in overrepresentation of certain demographic groups among young adults who have internet access and are more digitally literate. This could exclude other groups and affect the representativeness of the results. Second, the survey relied on self-reported behavior, which is subject to bias. Third, the cross-sectional design of the study precludes the identification of cause-and-effect relationships. Fourth, variables, such as susceptibility, perceived barriers, and cues to action, from the HBM were measured only in the subsample of unvaccinated respondents. As these variables were not included in the analysis, the study may not fully capture all factors influencing vaccine decision-making across the entire sample. Fifth, the multinomial regression analysis results showed wide CIs, indicating greater uncertainty in the estimated effects of the predictor variables on vaccination-related behavior groups. This may result from small sample sizes, overly complex models, sparse data, or other factors. This should be carefully considered when interpreting our findings. Future research using similar methodological frameworks and a similar design should address these issues to enhance the robustness and interpretability of the results. Finally, the data are limited to the social, cultural, and political contexts of Slovenia, where specific COVID-19–related protective measures and crisis communication may have significantly influenced the vaccination behavior of the youth [25].

Implications

Our findings suggest that young people exhibit different attitudes toward vaccination against COVID-19. Consequently, public health authorities should implement diverse strategies tailored to address these attitudes and fears, and the underlying factors influencing them. This recommendation extends beyond COVID-19 and is equally relevant to other diseases, particularly those with lower immunization coverage and vaccination rates [51].

According to the stages of change model [37], which provides a framework for understanding how individuals progress toward a desired behavioral change, such as COVID-19 vaccination, we can categorize vaccine refusers as being in the first stage of behavior change, known as precontemplation. Individuals in this stage are often unaware of the problem, are unmotivated to address it, or may even deny its relevance to themselves. Effective strategies for this group include raising awareness of the issue, providing clear information about the benefits of vaccination, and addressing misconceptions with targeted communication tailored to their specific concerns about vaccine safety and health risks. The groups of delayers and hesitant individuals appear to be either in the contemplation or preparation stage. Those in the contemplation stage are considering vaccination but have not yet made a definitive decision. Strategies for this group should focus on helping them weigh the perceived benefits of vaccination, such as protection against COVID-19 and its potential health and financial consequences, against the perceived costs, such as potential vaccine side effects. Delayers, on the other hand, seem to align more with the preparation stage. They are actively planning to get vaccinated but remain the most uncertain group when it comes to making health-related decisions. To support this group, strategies should include encouragement to finalize their decision, such as prompting them to schedule a vaccination appointment and offering reliable information on the effectiveness, safety, and availability of vaccines.

Conclusions

This cross-sectional study provides survey data on the vaccination behavior (vaccination uptake and refusal) and intention (delayed vaccination and hesitancy) of young people regarding COVID-19 vaccination. It also examines the sociodemographic, health-related, and behavioral predictors of these intentions, integrating constructs from 2 health behavior theories: HBM and TPB.

The findings of our study highlight that young people exhibit diverse behavioral and volitional positions regarding COVID-19 vaccination. Several sociodemographic, health-related, and behavioral factors, such as educational level, material status, subjective norms, perceived self-efficacy, attitudes toward vaccines, fear of contracting the disease, and conspiracy theory tendencies, significantly influence these positions. With regard to vaccination as the desired behavior, individuals are positioned in different stages of behavior change. From a normative perspective, these stages should be carefully considered when designing and implementing measures to promote vaccination for both COVID-19 and other diseases among young people. Greater efforts should be made to target individuals with lower levels of education, those with material disadvantages, and those residing in rural areas.

The findings of this study can guide health policy makers and professionals in optimizing young people's willingness to vaccinate and can provide valuable insights for addressing other vaccinations, future infectious disease outbreaks, and the implementation of effective preventive measures. Interventions should be tailored to address the primary reasons for nonvaccination, with priority target groups based on their potential for behavior change. Determined refusers may not be the most effective group to focus on initially. Instead, delayers, who demonstrate the greatest willingness to vaccinate among unvaccinated individuals, require targeted strategies, such as those aimed at increasing their self-efficacy in managing health-related decisions.

Acknowledgments

This work was financially supported by the Slovenian Research and Innovation Agency under the research program Social Psychology and Sociology of Everyday Life (research core funding number: P5-0183). The funder played no role in the study design; the collection, analysis, and interpretation of the data; or the decision to submit the article for publication. The funder accepts no responsibility for the content.

Data Availability

The data underlying this article will be shared by the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CHERRIES (Checklist for Reporting Results of Internet E-Surveys).
[PDF File (Adobe PDF File), 145 KB - publichealth_v11i1e64653_app1.pdf]

References

1. Kokić Z, Kon P, Djurković-Djaković O. Effectiveness of vaccination in preventing COVID-19: a community study comparing four vaccines. Vaccines (Basel) 2023 Feb 24;11(3):544 [FREE Full text] [doi: 10.3390/vaccines11030544] [Medline: 36992128]
2. Rahmani K, Shavaleh R, Forouhi M, Disfani HF, Kamandi M, Oskooi RK, et al. The effectiveness of COVID-19 vaccines in reducing the incidence, hospitalization, and mortality from COVID-19: A systematic review and meta-analysis. Front Public Health 2022 Aug 26;10:873596 [FREE Full text] [doi: 10.3389/fpubh.2022.873596] [Medline: 36091533]
3. Omer SB, Yildirim I, Forman HP. Herd immunity and implications for SARS-CoV-2 control. JAMA 2020 Nov 24;324(20):2095-2096. [doi: 10.1001/jama.2020.20892] [Medline: 33074293]
4. Anderson RM, Vegvari C, Truscott J, Collyer BS. Challenges in creating herd immunity to SARS-CoV-2 infection by mass vaccination. The Lancet 2020 Nov;396(10263):1614-1616. [doi: 10.1016/s0140-6736(20)32318-7]
5. COVID-19 Vaccine Tracker. European Center for Disease Control and Prevention. URL: https://vaccinetracker.ecdc.europa.eu/public/ [accessed 2024-06-20]
6. OECD/European Union. Vaccination against COVID-19. In: Health at a Glance: Europe 2022. Paris, France: OECD Publishing; 2022:198-199.
7. Namera cepljenja stagnira in ne preseže 60%. Prebivalstvo je razdeljeno po starosti, oceni lastne rizičnosti ter odnosu do virusa in cepljenja nasploh. Valicon. URL: https://tinyurl.com/yc5hwf9v [accessed 2024-05-25]
8. Aw J, Seng JJB, Seah SSY, Low LL. COVID-19 vaccine hesitancy-a scoping review of literature in high-income countries. Vaccines (Basel) 2021 Aug 13;9(8):900 [FREE Full text] [doi: 10.3390/vaccines9080900] [Medline: 34452026]
9. Freeman D, Loe BS, Yu L, Freeman J, Chadwick A, Vaccari C, et al. Effects of different types of written vaccination information on COVID-19 vaccine hesitancy in the UK (OCEANS-III): a single-blind, parallel-group, randomised controlled trial. The Lancet Public Health 2021 Jun;6(6):e416-e427. [doi: 10.1016/s2468-2667(21)00096-7]
10. Kelly BJ, Southwell BG, McCormack LA, Bann CM, MacDonald PDM, Frasier AM, et al. Predictors of willingness to get a COVID-19 vaccine in the U.S. BMC Infect Dis 2021 Apr 12;21(1):338 [FREE Full text] [doi: 10.1186/s12879-021-06023-9] [Medline: 33845781]
11. Latkin CA, Dayton L, Yi G, Colon B, Kong X. Mask usage, social distancing, racial, and gender correlates of COVID-19 vaccine intentions among adults in the US. PLoS One 2021 Feb 16;16(2):e0246970 [FREE Full text] [doi: 10.1371/journal.pone.0246970] [Medline: 33592035]
12. Prickett KC, Habibi H, Carr PA. COVID-19 vaccine hesitancy and acceptance in a cohort of diverse New Zealanders. Lancet Reg Health West Pac 2021 Sep;14:100241 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100241] [Medline: 34377999]
13. Schwarzinger M, Watson V, Arwidson P, Alla F, Luchini S. COVID-19 vaccine hesitancy in a representative working-age population in France: a survey experiment based on vaccine characteristics. Lancet Public Health 2021 Apr;6(4):e210-e221 [FREE Full text] [doi: 10.1016/S2468-2667(21)00012-8] [Medline: 33556325]

14. Vollmann M, Salewski C. To get vaccinated, or not to get vaccinated, that is the question: illness representations about COVID-19 and perceptions about COVID-19 vaccination as predictors of COVID-19 vaccination willingness among young adults in The Netherlands. Vaccines (Basel) 2021 Aug 24;9(9):941 [FREE Full text] [doi: 10.3390/vaccines9090941] [Medline: 34579178]
15. Carter J, Rutherford S, Borkoles E. COVID-19 vaccine uptake among younger women in rural Australia. Vaccines (Basel) 2021 Dec 27;10(1):26 [FREE Full text] [doi: 10.3390/vaccines10010026] [Medline: 35062687]
16. Russell FM, Greenwood B. Who should be prioritised for COVID-19 vaccination? Hum Vaccin Immunother 2021 May 04;17(5):1317-1321 [FREE Full text] [doi: 10.1080/21645515.2020.1827882] [Medline: 33141000]
17. Ford C, MacKay M, Thaivalappil A, McWhirter J, Papadopoulos A. COVID-19 vaccine communications on Instagram and vaccine uptake in young adults: a content assessment and public engagement analysis. Emerging Adulthood 2023 Dec 21;12(2):224-235. [doi: 10.1177/21676968231222439]
18. Mitchell H, Lim R, Gill PK, Dhanoa J, Dubé È, Bettinger JA. What do adolescents think about vaccines? Systematic review of qualitative studies. PLOS Glob Public Health 2022;2(9):e0001109 [FREE Full text] [doi: 10.1371/journal.pgph.0001109] [Medline: 36962668]
19. Themistocleous S, Argyropoulos CD, Vogazianos P, Shiamakkides G, Noula E, Nearchou A, et al. Perspectives of European patient advocacy groups on volunteer registries and vaccine trials: VACCELERATE survey study. JMIR Public Health Surveill 2024 Apr 04;10:e47241 [FREE Full text] [doi: 10.2196/47241] [Medline: 38573762]
20. Mateo-Canedo C, Sanabria-Mazo JP, Comendador L, Rojas JS, Carmona M, Crespo-Puig N, et al. Predictive factors of hesitancy to vaccination against SARS-CoV-2 virus in young adults in Spain: Results from the PSY-COVID study. Vaccine X 2023 Aug;14:100301 [FREE Full text] [doi: 10.1016/j.jvacx.2023.100301] [Medline: 37091731]
21. Chaufan C. Is Covid-19 "vaccine uptake" in postsecondary education a "problem"? A critical policy inquiry. Health (London) 2024 Nov 15;28(6):831-857 [FREE Full text] [doi: 10.1177/13634593231204169] [Medline: 37968946]
22. Kelsall J, Sorell T. Two kinds of vaccine hesitancy. Social Epistemology 2024 Apr 15;39(1):40-55. [doi: 10.1080/02691728.2024.2330466]
23. Gao P, Kang L, Liu J, Liu M. Immunogenicity, effectiveness, and safety of COVID-19 vaccines among children and adolescents aged 2-18 years: an updated systematic review and meta-analysis. World J Pediatr 2023 Nov;19(11):1041-1054 [FREE Full text] [doi: 10.1007/s12519-022-00680-9] [Medline: 36723827]
24. Blahut R, Flint A, Orlando E, DesChatelets J, Khowaja A. A scoping review on the decision-making dynamics for accepting or refusing the COVID-19 vaccination among adolescent and youth populations. BMC Public Health 2023 Apr 28;23(1):784 [FREE Full text] [doi: 10.1186/s12889-023-15717-5] [Medline: 37118794]
25. Kamin T, Perger N, Debevec L, Tivadar B. Alone in a time of pandemic: solo-living women coping with physical isolation. Qual Health Res 2021 Jan 19;31(2):203-217. [doi: 10.1177/1049732320971603] [Medline: 33213283]
26. Nilsson S, Mattson J, Berghammer M, Brorsson A, Forsner M, Jenholt Nolbris M, et al. To be or not to be vaccinated against COVID-19 - The adolescents' perspective - A mixed-methods study in Sweden. Vaccine X 2021 Dec;9:100117 [FREE Full text] [doi: 10.1016/j.jvacx.2021.100117] [Medline: 34693273]
27. Crescenzi-Lanna L, Valente R, Cataldi S, Martire F. Predictors of young people's anti-vaccine attitudes in the context of the COVID-19 pandemic. Public Underst Sci 2024 Jan 28;33(1):73-87 [FREE Full text] [doi: 10.1177/09636625231179830] [Medline: 37377105]
28. Zhou Y, Li R, Shen L. Targeting COVID-19 vaccine-hesitancy in college students: An audience-centered approach. J Am Coll Health 2024 Dec 28;72(9):3526-3535. [doi: 10.1080/07448481.2023.2180988] [Medline: 36853986]
29. Huang HY, Gerend MA. The role of trust, vaccine information exposure, and Health Belief Model variables in COVID-19 vaccination intentions: Evidence from an HBCU sample. J Health Psychol 2024 May 05;29(6):621-632. [doi: 10.1177/13591053241227388] [Medline: 38312019]
30. Small ML, Lennon RP, Dziak JJ, Smith RA, Sommerville G, Bharti N. College students' COVID-19 vaccine beliefs and intentions: Implications for interventions. J Am Coll Health 2024 May 27;72(4):989-994. [doi: 10.1080/07448481.2022.2065205] [Medline: 35622961]
31. Shmueli L. Predicting intention to receive COVID-19 vaccine among the general population using the health belief model and the theory of planned behavior model. BMC Public Health 2021 Apr 26;21(1):804 [FREE Full text] [doi: 10.1186/s12889-021-10816-7] [Medline: 33902501]
32. Milošević Đorđević J, Mari S, Vdović M, Milošević A. Links between conspiracy beliefs, vaccine knowledge, and trust: Anti-vaccine behavior of Serbian adults. Soc Sci Med 2021 May;277:113930 [FREE Full text] [doi: 10.1016/j.socscimed.2021.113930] [Medline: 33873008]
33. Prestwich A, Kenworthy J, Conner M. Health Behavior Change: Theories, Methods and Interventions. London, England: Routledge; 2024.
34. McEachan RRC, Conner M, Taylor NJ, Lawton RJ. Prospective prediction of health-related behaviours with the Theory of Planned Behaviour: a meta-analysis. Health Psychology Review 2011 Sep;5(2):97-144. [doi: 10.1080/17437199.2010.521684]

35. McEachan R, Taylor N, Harrison R, Lawton R, Gardner P, Conner M. Meta-analysis of the reasoned action approach (RAA) to understanding health behaviors. Ann Behav Med 2016 Aug 11;50(4):592-612 [FREE Full text] [doi: 10.1007/s12160-016-9798-4] [Medline: 27169555]
36. JazVem. URL: https://www.jazvem.si/ [accessed 2024-11-25]
37. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [doi: 10.1037//0022-006x.51.3.390] [Medline: 6863699]
38. Chu H, Liu S. Integrating health behavior theories to predict American's intention to receive a COVID-19 vaccine. Patient Educ Couns 2021 Aug;104(8):1878-1886 [FREE Full text] [doi: 10.1016/j.pec.2021.02.031] [Medline: 33632632]
39. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Weinman J, Wright S, Johnston M, editors. Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs. Windsor, UK: NFER-Nelson; 1995.
40. Nadelson L, Jorcyk C, Yang D, Jarratt Smith M, Matson S, Cornell K, et al. I just don't trust them: the development and validation of an assessment instrument to measure trust in science and scientists. School Sci & Mathematics 2014 Jan 19;114(2):76-86. [doi: 10.1111/ssm.12051]
41. Raziskava #Novanormalnost. Valicon. URL: https://www.valicon.net/wp-content/uploads/2020/12/Novanormalnost-15.pdf [accessed 2024-11-25]
42. DeVellis R, Thorpe C. Scale Development: Theory and Applications. Thousand Oaks, CA: SAGE Publications, Inc; 2021.
43. Code of ethics for researchers at the University of Ljubljana. University of Ljubljana. URL: https://www.uni-lj.si/assets/Sluzba-za-raziskovalno-dejavnost/Etika-in-integriteta/EN/Code-of-ethics-for-researchers.pdf [accessed 2025-02-03]
44. World Medical Association. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. Bull World Health Organ 2001;79(4):373-374 [FREE Full text] [Medline: 11357217]
45. Natalia YA, Delporte M, De Witte D, Beutels P, Dewatripont M, Molenberghs G. Assessing the impact of COVID-19 passes and mandates on disease transmission, vaccination intention, and uptake: a scoping review. BMC Public Health 2023 Nov 17;23(1):2279 [FREE Full text] [doi: 10.1186/s12889-023-17203-4] [Medline: 37978472]
46. Biswas MR, Alzubaidi MS, Shah U, Abd-Alrazaq AA, Shah Z. A scoping review to find out worldwide COVID-19 vaccine hesitancy and its underlying determinants. Vaccines (Basel) 2021 Oct 25;9(11):1243 [FREE Full text] [doi: 10.3390/vaccines9111243] [Medline: 34835174]
47. Kamin T, Anker T. Cultural capital and strategic social marketing orientations. Journal of Social Marketing 2014;4(2):94-110. [doi: 10.1108/JSOCM-08-2013-0057]
48. Stead M, Jessop C, Angus K, Bedford H, Ussher M, Ford A, et al. National survey of attitudes towards and intentions to vaccinate against COVID-19: implications for communications. BMJ Open 2021 Oct 28;11(10):e055085 [FREE Full text] [doi: 10.1136/bmjopen-2021-055085] [Medline: 34711602]
49. Miller JD, Ackerman MS, Laspra B, Polino C, Huffaker JS. Public attitude toward Covid-19 vaccination: The influence of education, partisanship, biological literacy, and coronavirus understanding. FASEB J 2022 Jul 03;36(7):e22382 [FREE Full text] [doi: 10.1096/fj.202200730] [Medline: 35657606]
50. Petrič G, Atanasova S, Kamin T. Impact of social processes in online health communities on patient empowerment in relationship with the physician: emergence of functional and dysfunctional empowerment. J Med Internet Res 2017 Mar 13;19(3):e74 [FREE Full text] [doi: 10.2196/jmir.7002] [Medline: 28288953]
51. Immunization coverage. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/immunization-coverage [accessed 2024-05-25]

Abbreviations

aOR: adjusted odds ratio
EFA: exploratory factor analysis
ESOMAR: European Society for Opinion and Market Research
EU: European Union
HBM: health belief model
TPB: theory of planned behavior

Edited by C Argyropoulos; submitted 23.07.24; peer-reviewed by T Ricks, N Kaur, M Sarhan; comments to author 18.10.24; revised version received 11.12.24; accepted 19.01.25; published 21.02.25.

Please cite as:
Atanasova S, Kamin T, Perger N
Predictors of COVID-19 Vaccination Intention and Behavior Among Young People in a European Union Country With Low COVID-19 Vaccination Rates: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e64653
URL: https://publichealth.jmir.org/2025/1/e64653
doi:10.2196/64653
PMID:39983109

Net Reproduction Number as a Real-Time Metric of Population Reproducibility

Chiara Achangwa[1], PhD; Changhee Han[2], MS; Jun-Sik Lim[3], DVM, MPH; Seonghui Cho[1], BBA; Sangbum Choi[4], PhD; Sukhyun Ryu[1], MD, PhD

[1]Department of Preventive Medicine, College of Medicine, The Catholic University of Korea, R6117, Omibus Park, 222 Banpo-daero, Seoul, Republic of Korea

[2]Computer Information System, Georgia State University, Atlanta, GA, United States

[3]IHAP, Université de Toulouse, INRAE, ENVT, Toulouse, France

[4]Department of Statistics, Korea University, Seoul, Republic of Korea

Corresponding Author:
Sukhyun Ryu, MD, PhD
Department of Preventive Medicine, College of Medicine, The Catholic University of Korea, R6117, Omibus Park, 222 Banpo-daero, Seoul, Republic of Korea

Abstract

Abstract: The total fertility rate (TFR) is a biased estimate of the population reproductive potential when there is a sex-ratio imbalance at birth, and it does not account for the mortality of women of childbearing age. This study aimed to estimate the reproduction rate (Rt), which adjusts for the sex-ratio imbalance and the mortality of women of childbearing age, and to assess the differences in the timing of when the population reached the replacement level of the TFR and Rt. We first estimated the Rt using the probability of survival in women and the number of female births. Then, using a time-series analysis, we compared the time series of the TFR and Rt in the Korean population between 1975 and 2022. We found the Rt showed a below replacement level of the population a year earlier than the TFR. However, the estimate of the time-series analysis of Rt was not significantly different from the estimates of the TFR. Our finding suggests that the Rt can provide timely information on the adjusted population reproductive potential and is easier for the public to interpret compared to TFR.

(JMIR Public Health Surveill 2025;11:e63603) doi:10.2196/63603

KEYWORDS

fertility rate; reproducibility; reproduction rate; population control; Korea; sex ratio; imbalance; mortality; woman; female; childbearing age; reproductive age; giving birth; assessment; time series; Korean; impact analysis; birth control; reproduction

Introduction

Between 1962 and 1993, Korea implemented a successful family planning policy. In 1993, this policy was discontinued; in 2004, childbirth-promoting policies were implemented (Multimedia Appendix 1). In 2023, Korea had the lowest fertility rate (0.8) worldwide [1]. The total fertility rate (TFR), the average offspring number per childbearing-age females (15 - 49 years), is a common metric to assess population change potential. However, this cohort-based measure is biased when there is male-to-female sex-ratio imbalance at birth [2,3]. The TFR does not account for mortality rates among childbearing-age women, possibly affecting population reproducibility [2,4]. These limitations reduce TFR’s ability to accurately reflect a country’s population replacement dynamics. Therefore, the net reproduction rate (Rt), the number of daughters a woman of childbearing age would produce under prevailing fertility and mortality conditions, is better. Like other real-time epidemiological metrics (eg, the effective reproduction number in infectious disease modeling) [5], the Rt can be calculated and updated regularly with new population data; it can provide timely insights into population sustainability. The Rt is easier for public understanding, as a population is below the replacement level when the Rt < 1 [6], in contrast to the TFR, with a threshold of 2.1. Despite this, no previous studies have evaluated the population reproducibility using the Rt in Korea.

This study assessed the differences in the timing of reaching population replacement level of the TFR and Rt and the estimated difference of the time series of the TFR and Rt by two major population control policies.

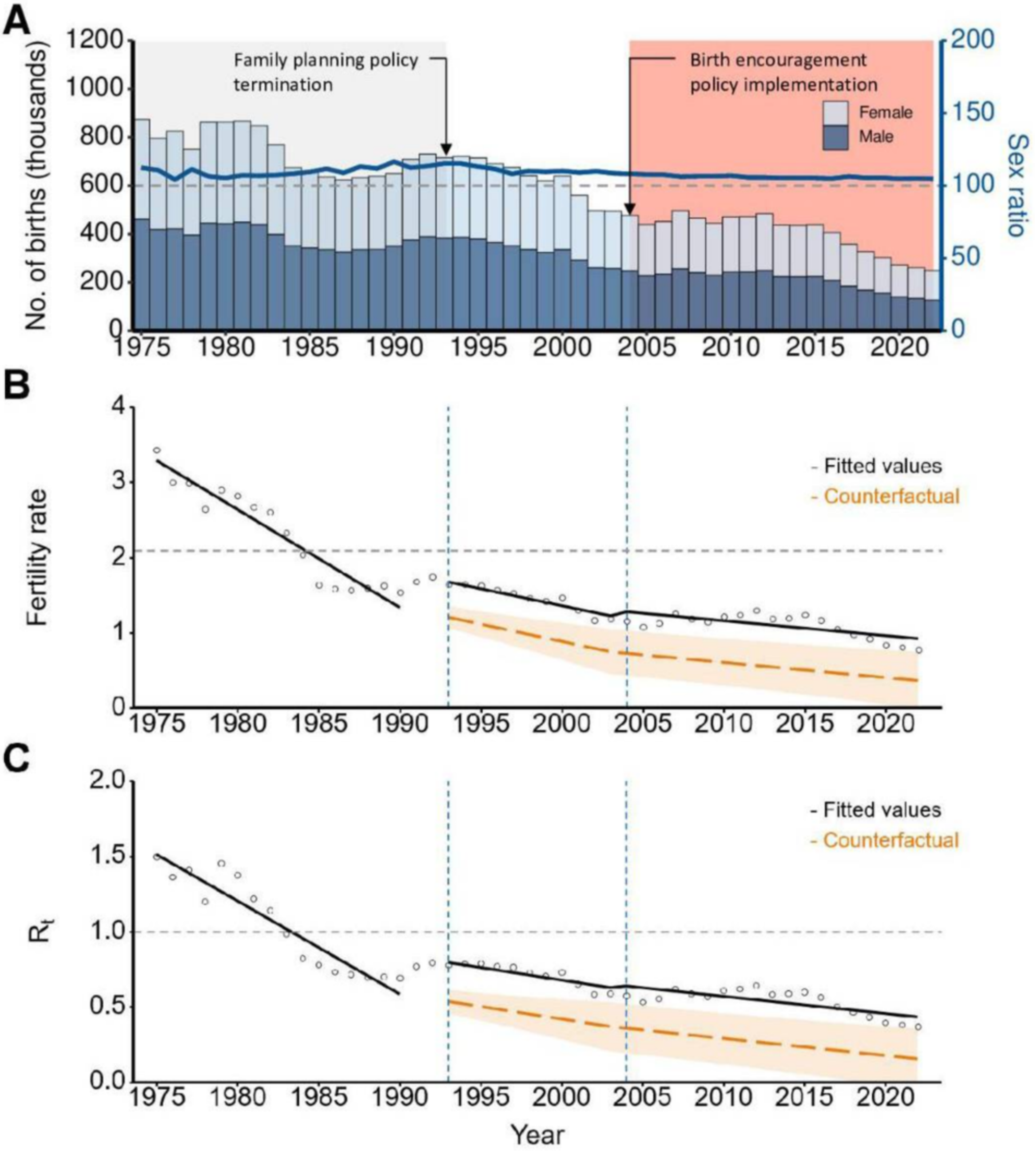
Methods

We collected the annual number of live births, number of women, mortality rate of women, and male-to-female ratio of women of childbearing age between 1975 and 2022 through the Korean National Statistic Agency [7] to calculate the TFR and Rt (Multimedia Appendix 2). To identify the different estimates of policy impact (1975-1993: family planning policy; 1993-2004: childbirth encouragement policy), we conducted an interrupted time series (ITS) with segmented regression to examine the time trend and its level change in the TFR and Rt. We also conducted a cross-correlation analysis to evaluate the temporal relationship between the TFR and Rt. Then, we compared the estimates of Rt multiplied by 2.1 (TFR threshold level) with the TFR estimates along with 95% CIs. All analyses were conducted using R software (version 4.4.0; R Foundation for Statistical Computing).

Results

The number of live birth number decreased from 874,030 in 1975 to 249,186 in 2022. Similarly, the male-to-female sex ratio decreased from 112 in 1975 to 105 in 2022 (Figure 1A). The TFR remained below 2.1 since 1984 (TFR=2.04) and decreased further to 0.78 in 2022 (Figure 1B). The Rt remained below 1 since 1983 (Rt=0.98) and decreased to 0.4 in 2022 (Figure 1C).

Figure 1. The annual number of live births, total fertility rate (TFR), and reproduction rate (R_t) in South Korea, 1975-2022. (**A**) The bar-colored blue and sky-blue indicate the annual number of male and female births, respectively. The blue solid line indicates the yearly sex ratio of male to female births. (**B**) The interrupted time series model with the TFR. The interruption (dashed vertical line) was set to 1993 (when the family planning policy was discontinued) taking into account for the transition period of the policy and 2004 (when the birth encouragement policy was implemented) to identify the changes in the TFR level and slope. The horizontal dashed line indicates a critical threshold of the TFR at 2.1. The dashed orange line indicates the annual TFR based on a counterfactual scenario without changing the birth control policy, and the orange shaded area indicates 95% confidence intervals (CIs) of the TFR. (**C**) The interrupted time series model with the estimated R_t; the critical threshold of R_t=1. An $R_t < 1$ indicates that the population's reproductive performance falls below the replacement level. The dashed orange line indicates the annual R_t based on a counterfactual scenario without changing the birth control policy, and the orange shaded area indicates 95% CIs of R_t.



The ITS showed an immediate increase in the mean TFR (55%) and Rt (26%) and an increased slope change of 9% in the TFR and 4% in Rt following the family planning policy discontinuation (Figure 1B and 1C and Table 1). After the birth encouragement policy introduction, the slope of the TFR (3%) and Rt (1%) increased, with no significant level change. When the Rt was multiplied by 2.1, the estimates were within the 95% CI of the TFR estimate (Table 1). A high correlation between the TFR and Rt at lag 0 indicated no temporal differences (Multimedia Appendix 3).

Table . Estimates from the interrupted time-series analysis using the total fertility rate and reproduction rate in South Korea, 1975-2022.

	Mean total fertility rate, % (95% CIs)[a]	Mean reproduction rate, % (95% CIs)[a]	Mean reproduction rate, (95% CIs) multiplied by 2.1[b]
Immediate level change following the discontinuation of family planning	54.9 (33.1 to 67.2)	25.9 (18.3 to 33.8)	54.4 (38.4 to 71.0)
Post-intervention slope change following the discontinuation of family planning	8.9 (7.6 to 10.4)	4.2 (3.3 to 5.5)	8.8 (6.9 to 11.6)
Immediate level change following the birth-encouragement policy implementation	3.2 (1.3 to 4.8)	1.2 (0.1 to 2.0)	2.5 (0.2 to 4.2)
Post-intervention slope change following the birth-encouragement policy implementation	5.8 (-5.5 to 16.4)	3.3 (-4.6 to 11.2)	6.9 (-9.7 to 23.5)

[a] Estimates of the mean and 95% confidence intervals (CIs) from the interrupted time series with a segmented regression model to examine the time trend and its level change.

[b] Estimates of the reproduction rate were multiplied by 2.1 (threshold level of total fertility rate) along with 95% CIs.

Discussion

The threshold level of the population replacement was captured a year earlier through the Rt compared to the TFR. This is likely due to sex-ratio imbalances in Korea.

The trend levels and slope changes of the TFR and Rt increased following the birth control policy discontinuation [8]. These significant level changes were likely affected by previous birth control policies [9]. However, after the child encouragement policy implementation in 2004, the TFR and Rt were far below the population replacement threshold, consistent with a previous study that reported no positive effect of child encouragement policies on the fertility rate [10], likely due to sociocultural factors influencing fertility behavior (eg, changing gender roles and economic pressures) [10]. Our study could be applied to other countries experiencing similar socioeconomic and cultural dynamics, particularly those with comparable fertility patterns and sex-ratio imbalances [4].

This study had limitations. Sensitivity analyses were not included in the parameter estimation models. The ITS models were interrupted in 1993 to reflect the discontinuation of the family planning policy, accounting for the policy transition period. The ITS may not fully capture the nonlinear trends after 2015. We did not consider the qualitative characteristics of each policy.

The Rt can be used as a useful and timely metric of population reproducibility, particularly in the presence of sex-ratio imbalance at birth. Furthermore, the Rt threshold of 1 may be easier for public interpretation compared to the TFR, as the public became familiar with the Rt parameter during the COVID-19 pandemic.

Acknowledgments

Ethics approval was waived by the institutional review board at the Catholic University of Korea (2024-1205-0001). The authors attest that there was no use of generative artificial intelligence technology in the generation of text, figures, or other informational content in this manuscript. This work was supported by the Gorvernment-wide Research and Development to Advance Infectious Disease Prevention and Control, Republic of Korea (grant number RS-2023-KH140322), and the Basic Science Research Program through the National Research Foundation of Korea by the Ministry of Education (grant numbers NRF-2020R1I1A3066471 and 2022R1A2C1008514).

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Authors' Contributions

SR conceived the study and designed the statistical methods. CA, CH, LSJ, and SC collected and assimilated the data. CA, CH, LSJ, and SC performed the data analysis. CA and SR wrote the first manuscript draft. CA, SC, and SR critically reviewed and edited the manuscript. All authors helped to interpret the results, critically revised the manuscript, and provided final approval of the version for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Key population policies in South Korea between 1970 and 2022.
[DOCX File, 20 KB - publichealth_v11i1e63603_app1.docx]

Multimedia Appendix 2
Description of total fertility rate, reproduction rate, and time-series analysis.
[DOCX File, 22 KB - publichealth_v11i1e63603_app2.docx]

Multimedia Appendix 3
Yearly lagged cross-correlation coefficients between total fertility rate and net reproduction number in South Korea (1975–2022).
[DOCX File, 108 KB - publichealth_v11i1e63603_app3.docx]

References

1. World population dashboard. United Nations Population Fund.: New York, United States: UNFPA; 2023. URL: https://www.unfpa.org/data/world-population-dashboard
2. Gietel-Basten S, Scherbov S. Is half the world's population really below "replacement-rate"? PLoS ONE 2019;14(12):e0224985. [doi: 10.1371/journal.pone.0224985] [Medline: 31790416]
3. Chao F, Gerland P, Cook AR, Guilmoto CZ, Alkema L. Projecting sex imbalances at birth at global, regional and national levels from 2021 to 2100: scenario-based Bayesian probabilistic projections of the sex ratio at birth and missing female births based on 3.26 billion birth records. BMJ Glob Health 2021 Aug;6(8):e005516. [doi: 10.1136/bmjgh-2021-005516] [Medline: 34341019]
4. Aitken RJ. The changing tide of human fertility. Hum Reprod 2022 Apr 1;37(4):629-638. [doi: 10.1093/humrep/deac011] [Medline: 35079808]
5. Han C, Seo H, Cho S, Chiara A, Ryu S. Impact of travel restrictions for travellers from China on the internal spread of SARS-CoV-2 in South Korea. J Travel Med 2023 Sep 5;30(5):taad047. [doi: 10.1093/jtm/taad047] [Medline: 37040325]
6. Lotka AJ. Analytical Theory of Biological Populations 1998.
7. Korean Statistical Information Service Monthly, quarterly, and annual population trends.: Daejeon, South Korea: Statistics Korea; 2023. URL: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B8000G&conn_path=I2
8. Oh YR. How does Korea's family planning policy promote familism? Korea Social Policy Review 2020;27(1):213-248. [doi: 10.17000/kspr.27.1.202003.213]
9. Jang Y, Kim N, Lee S, Jin D. Korea's Population Policy: History and Future 2010.
10. Jeong K, Yoon J, Cho HJ, Kim S, Jang J. The relationship between changes in the korean fertility rate and policies to encourage fertility. BMC Public Health 2022 Dec 8;22(1):2298. [doi: 10.1186/s12889-022-14722-4] [Medline: 36482452]

Abbreviations

ITS: interrupted time series
TFR: total fertility rate

Edited by A Mavragani; submitted 24.06.24; peer-reviewed by A Aktuna, Y Ko; revised version received 22.11.24; accepted 04.12.24; published 12.02.25.

Please cite as:
Achangwa C, Han C, Lim JS, Cho S, Choi S, Ryu S
Net Reproduction Number as a Real-Time Metric of Population Reproducibility
JMIR Public Health Surveill 2025;11:e63603
URL: https://publichealth.jmir.org/2025/1/e63603
doi:10.2196/63603

Dissemination and Implementation Approach to Increasing Access to Local Pre-Exposure Prophylaxis (PrEP) Resources With Black Cisgender Women: Intervention Study With Vlogs Shared on Social Media

Mandy J Hill[1], DrPH, MPH, BS; Laurenia Mangum[2], PhD, MPH, LMSW; Sandra J Coker[1], MD; Tristen Sutton[3], BA; Diane M Santa Maria[4], DrPH, MSN

[1]Department of Emergency Medicine, McGovern Medical School, UTHealth Houston, 6431 Fannin St, Houston, TX, United States

[2]Jane Addams College of Social Work, University of Illinois Chicago, Chicago, IL, United States

[3]Tristen Sutton Consulting, Houston, TX, United States

[4]Cizik School of Nursing, UTHealth Houston, Houston, TX, United States

Corresponding Author:
Mandy J Hill, DrPH, MPH, BS
Department of Emergency Medicine, McGovern Medical School, UTHealth Houston, 6431 Fannin St, Houston, TX, United States

Abstract

Background: Black cisgender women account for only 2% of pre-exposure prophylaxis (PrEP)-eligible people in the United States who use PrEP to prevent HIV. Owing to the low PrEP use, Black cisgender women continue to contract HIV more frequently than women from every other racial group. Intervention efforts that can bridge the link between knowing that PrEP prevents HIV and support with access to PrEP are necessary for Black cisgender women.

Objective: The purposes of the vlogs through the campaign were to share information about ways to prevent HIV using PrEP and fact-based education and provide access to PrEP resources with active links to local PrEP providers at local community health centers.

Methods: In Phase 1, the study team formerly piloted full-length video blog posts (ie, vlogs; 10 - 12 min each) with 26 women during an emergency department visit. Using the findings from Phase 1, Phase 2 involved a prospective 6-month social media marketing campaign. The study team led a Texas-Development CFAR-funded pilot grant to disseminate brief vlog snippets (30 s) of excerpts from full-length vlogs with a larger group of Black women in Harris County. Community members, who were aged 18 - 55 years, usually consume content that is often viewed by Black cisgender women (ie, health and beauty) and reside in neighborhoods (based on zip code) in Harris County where most residents are Black or African American. They were shown a series of brief vlog snippets on their social media pages, along with a brief message about PrEP and an active hyperlink to local PrEP resources. The study team assessed implementation outcomes, including the feasibility and acceptability, appropriateness of vlogs, access to PrEP resources at local clinics, and clinical outcomes such as increased PrEP awareness among Black cisgender women.

Results: Within 6 months, the campaign reached 110.8k unique individuals (the number of unique accounts that have seen your content at least once) who identified as women. When stratified by age, video plays (the number of times a video starts playing) at 50% of the vlogs (n=30,877) were most common among women aged 18 - 24 years (n=12,017) and least common among women aged 45 - 54 years (n=658). Key performance indicators showed that 1,098,629 impressions (the number of times a user saw the vlog) and 1,002,244 total video plays resulted in 15,952 link clicks to local PrEP resources.

Conclusions: The campaign demonstrated the feasibility and acceptability of this approach with Black cisgender women and illustrated preliminary effectiveness at supporting access to local PrEP resources with Black cisgender women. Further dissemination and implementation of this approach is necessary to fully assess whether vlog viewership and clicks on links to PrEP resources can meaningfully empower Black cisgender women to access PrEP and help them to assess whether PrEP is personally a useful HIV prevention option.

(JMIR Public Health Surveill 2025;11:e67367) doi:10.2196/67367

KEYWORDS

PrEP; cisgender Black women; social media campaign; PrEP access; HIV prevention; vlogging; dissemination and implementation; pre-exposure prophylaxis; dissemination; implementation; HIV; prevention; human immunodeficiency virus; cisgender; social media; marketing; campaign; education; sexually transmitted diseases; STDs; vlog

Introduction

Background

Pre-exposure prophylaxis (PrEP) describes a medication that is proven to be 99% effective when taken as prescribed to reduce the risk of getting HIV by taking a pill or injection among people who do not have HIV but may be vulnerable to being introduced to or "exposed" to it. The marginal use of PrEP hinders progress toward decreasing the incidence of HIV cases among Black cisgender women. Black women account for nearly 60% of new HIV cases among women in the United States while comprising less than 15% of the female population [1,2] and also accounting for 67% of new HIV diagnoses among all women in the southern part of the United States [3]. Additionally, Black women in Houston are 18.4 times more likely to be living with HIV [4] than other women in the United States. Optimal progress toward ending the HIV epidemic (EHE) goals requires strategies that will interrupt transmission pathways among women in HIV hotspot locations. PrEP, when taken as prescribed as an oral daily or bimonthly injection medication, can reduce the risk of HIV through sex by 99% [5,6]. Interim findings of the PURPOSE 1 trial presented at the AIDS 2024 Conference confirm that lenacapavir, a long-acting injectable PrEP technology that can be taken twice a year, offered 100% protection against HIV transmission among cisgender women enrolled [7]. However, less than 2% of eligible Black cisgender women in the United States take PrEP [8].

The critical gap between the availability of PrEP and use by Black cisgender women, arguably the most HIV-vulnerable demographic of women in the United States, requires the attention of researchers and interventionists. The identified barriers to PrEP uptake with Black women include limited PrEP awareness, low perceived HIV risk, concerns about adverse effects, concerns about costs, limited marketing, and distrust in the health care system [9]. Real-life competing priorities (ie, housing instability and food insecurity) trump HIV prevention for many Black women [10]. Social and structural factors, including racism, sexism, and classism, are often barriers to PrEP access, especially among young Black women [10,11]. Black women have cited medical racism as a key reason for distrusting the medical system and, as a result, not choosing to use PrEP [10-13]. These barriers, in addition to legacy of medical mistrust [13-15], the dehumanization of Black women [16], and stigmatizing experiences within health care settings [15-18], create a seemingly impermeable barrier to the possibility of increasing PrEP use among Black women. However, public messaging of PrEP as an effective HIV prevention message via innovation by culturally immersed thought leaders, researchers, and interventionists who are of the community can be believed to be true, real, and relatable for Black cisgender women.

To date, there is a paucity of research focused on behavioral interventions that can overcome the complex barriers that suppress PrEP uptake among Black cisgender women. As such, progress towards EHE has been slow within and among sexual networks of Black cisgender women and there has been no significant decrease in HIV incidence or any meaningful increase in PrEP uptake within the last decade [8], which is a significant unmet population health need.

The time that it takes between intervention development and integration of innovative, evidence-based interventions (EBIs) into routine clinical practice can worsen the structural and systemic barriers that impede PrEP use among Black cisgender women. This process can take over a decade, leaving the science outdated, EBIs being ineffective, and improvement in clinical outcomes unaddressed. The measures of acceptability and feasibility in this study are poised to inform the adoption, scale-up, and sustainability of EBIs intending to increase PrEP initiation from the perspective of Black cisgender women. The acceptability of EBIs typically refers to the perception among stakeholders that a given treatment, service, practice, or innovation is agreeable, or palatable, which can change over time [19]. In this research, we are interested in Black cisgender women's acceptability of HIV prevention vlogs as a pathway for both information dissemination and linkage to PrEP resources at local clinics. The research team designed the vlogs with an intentional focus on race, gender, and culture to meet the needs of Black cisgender women. Local clinics already use webpages to provide information about PrEP to community members. In this case, vlogs could drive traffic to the clinic webpages. Lastly, the feasibility measure will indicate how well vlogs can be used to motivate access to PrEP. Findings of feasibility and acceptability measures related to the influence of vlog use, as a health communication strategy, through a social media campaign on information access to PrEP knowledge and access to PrEP will produce a meaningful addition to the paucity of research on behavioral interventions that overcome complex barriers to PrEP access and use with Black cisgender women, and may inform pathways to facilitate PrEP uptake with Black women in the future.

Some progress has been made with identifying facilitators to PrEP uptake with Black women, including women's empowerment and advocacy, a need for increased PrEP-specific education, and the positive influence of PrEP-engaged women's testimonials [9]. In a study (2017 - 2019) where 26% of cisgender women (145/565) in NYC initiated PrEP, Latina women (29.7%, 79/565) and Black women (26.1%, 47/565) were more likely to initiate PrEP when compared to White women (16.3%, 14/565) [10]. Findings demonstrated that PrEP initiation was related to PrEP awareness, low income, unstable housing, receipt of navigation services, and reports of noninjection substance use and a recent sexual relationship with an HIV-positive partner [10]. Although some research has been done, there is a need for more focus on health communication with tailoring to the population. Black women shared a limited awareness of PrEP exacerbated by the lack of Black women-specific marketing [9]. In a qualitative study assessing the permanent perceptions Black cisgender women made based on external influences such as media messaging, researchers determined that an internal belief that PrEP is not for heterosexual people was developed because some Black cisgender women felt that commercials and marketing for PrEP were geared toward male same gender–loving couples and people of trans experience [20]. Opportunities to support Black women-specific social marketing could increase awareness and knowledge regarding the potential sexual health benefits of PrEP [9]. Additional research outcomes suggest that Black women want to know that the out-of-pocket costs for PrEP will be affordable, that they will have insurance coverage, that they will have positive interactions with informed and culturally competent clinical staff, access to a discreet and convenient clinic, and that PrEP will be safe for them [10]. Emerging research has demonstrated a growing interest in social media–based sexual health interventions for Black women [21].

Goal of This Study

This project addressed social and structural determinants of HIV for Black cisgender women by using an intersectional framework that brought together health communication through the lens of culture, race, and gender. Study findings here will contribute to EHE goals by potentially increasing PrEP access among HIV-naïve Black women and informing the Project ROLE intervention, which is currently being developed and will be fully tested to assess the efficacy of a new vlog series among Black cisgender women that will address PrEP knowledge, uptake, persistence, and adherence from the perspectives of both Black women and health care providers of Black women through community health centers and on social media through a NIMH-funded R34 mechanism (1R34MH136826-01). Findings of the Texas Development Center for AIDS Research pilot study substantiated the feasibility of social media as a plausible source of information to Black cisgender women for PrEP access and a pathway to increase uptake for Black cisgender women in Harris County. The purpose of the study was to better understand Black cisgender women's behaviors when engaging in social media-based HIV prevention interventions designed to provide direct access to web-based sexual health resources related to PrEP. Our research question was, "Can a culturally tailored social media-based vlog HIV prevention intervention motivate Black cisgender women to seek online access to PrEP resources?"

Methods

Prior Work

With the awareness of literary findings on the facilitators and barriers to PrEP uptake among Black cisgender women, the research presented here is an extension of a research concept piloted in an emergency department, whereby vlogs, which can be described as videos of blog posts, were used to debunk culturally believed myths in the Black community on how HIV and sexually transmitted infections are transmitted [22-27].

Behavioral Framework of the Prior Study

The premise of the health communication message of the vlogs stemmed from findings of a qualitative research study where Black cisgender women enrolled in an emergency department setting described a minimized value for sex [28]. Women who used the word "just" to describe sexual acts demonstrated a higher prevalence of experience with condomless sex, a previous diagnosis of a sexually transmitted disease (STI), and abuse. Misperceptions about what is safe or risky needs to be proven or refuted in order to promote sexual health and inform healthy sexual decision-making. The goal of the vlogs in the prior study that was a response to the research study (2015 ‐ 2016) was to resonate with Black women and create an opportunity to replace common misconceptions with the correct information needed to inform healthy sexual decision-making. Based on the protocol of the prior study approved by the University of Texas Health Science Center at Houston–Committee for the Protection of Human Subjects (CPHS), the proposed long-term programming stemming from the pilot study was to explore the utility of social media in reaching marginalized populations.

Overview of Methods of the Prior Study

Investigators completed a pilot randomized controlled trial of an HIV prevention intervention of vlogs comparing 2 educational strategies, a storytelling strategy versus an interactive gaming strategy, to assess and compare changes in knowledge pre- and post-intervention. In total, 26 Black cisgender women aged 18 ‐ 45 years who reported heterosexual sex within the last 3 months were enrolled [23-27]. The team used actors who reflected the target audience on the premises of race, gender, and colloquialism as the communication method in an effort to enhance cultural relevance with the intervention's content. The vlog strategy was innovative, relatable, and motivating for Black women.

Study Design

A test of whether vlogs could motivate online access to PrEP resources among Black cisgender women using 2 aims through a 90-day social media campaign was conducted. The findings of Aim 1, which include the development and testing of the vlog intervention, have been detailed in other studies [10].

During Aim 2, the social media campaign's performance was evaluated by assessing whether vlog viewership translated to linking Black women in Harris County to local PrEP resources. We assessed the primary outcome, engagement rate by reach, and secondary outcomes, post interactions, and visualizations (ie, views initiated, viewership, and completion rates) and evaluated the social media campaign's feasibility and acceptability at linking Black cisgender women to local PrEP resources. The social media consultant developed a report with metrics, analytics, and key performance indicators (including engagement with the campaign; vlog reach to Black women, ie, visualization; community engagement; and quantity of vlog material consumed ie, time) that measure study outcomes to quantify how well vlogs performed at motivating independent access to PrEP information as measured by clicks on hyperlinks to local PrEP resources.

Study Population

The social media campaign aimed to engage adult Black cisgender women aged 18 - 45 years who self-reported on their social media page profiles that they were assigned female at birth.

Eligible participants included cisgender women in Harris County who were (1) assigned female sex at birth and currently identified as female, (2) between 18 and 45 years of age, (3) fluent in English, and (4) had a phone or internet access.

Study Procedures

Development of the Marketing Strategy

A certified Social Media Advertisement expert designed a marketing campaign for implementation across social media websites, including Facebook and Instagram, using an algorithm to target individuals self-identifying as female, aged 18 - 45 years, usually consume content that is often viewed by Black cisgender women (ie, health and beauty), and reside in neighborhoods (based on zip code) in Harris County where most residents are Black or African American. These women were shown a series of brief vlog snippets on their social media pages, along with a brief message about PrEP and an active hyperlink to local PrEP resources. This strategy included a bi-monthly evaluation with an adjustment approach to reach a balanced sample and sustain active engagement with the target community. The study team reviewed the algorithm and made revisions as necessary before launching the campaign. In 2022, Facebook changed its demographic collection procedures and does not allow researchers to create inclusion criteria and exclude by race. Therefore, it is probable that non-Black women viewed the material.

Development of the Study's Landing Page

The social media ads provided active links to a landing page that housed access to local PrEP resources. The landing page was housed on the academic health center's website and was co-developed and managed by a web developer and the PI of the grant. The landing page's content included a link to a brief survey, described as a pre-screening tool, where individuals could describe brief demographic factors (eg, race and gender). Below the link to the brief survey, a question, "Did you Know?" is followed by a few sentences describing the status of HIV in the South with Black women and how the Texas D-CFAR pilot project aims to address the current health inequity. A table titled, "Pre-Exposure Prophylaxis (PrEP) Providers in Houston" is provided with the name of each local PrEP-providing agency, their physical address, and active links to their website and social media pages for ease of access to those who visit the page. The bottom of the landing page contains a link to the Texas-Development CFAR pilot protocol page, which provides more details about the study, inclusive of a project summary, background information, objectives, funding and collaborators, as well as a bio of project leadership [24].

Development of 60-Second Vlog Trailers

A study team member with robust vlogging experience on social media created 10 vlog trailers with brief excerpts from each vlog for the social media campaign. The trailers highlighted excerpts on PrEP within engaging aspects of vlogs with intentions to garner the immediate interest of Black women. The study team reviewed all 10 trailers, recommended edits, and collectively chose the top 5. When the trailers were refined and finalized, the social media ad expert launched the campaign with all 5 trailers.

Development of Post Content to Accompany Vlog Trailers

Each social media post included written content with at least 280 characters, including the hyperlink to local PrEP resources on the study's landing page. An example of a post was, "What does it mean to be at risk for HIV? Check out the vlog and click on the link to learn more about PrEP" (83 characters). We included hashtags like #PrEPforHER, #OWNYourHealth, and #PreventHIV, when character limits allowed.

Launching the Campaign

The social media ad expert launched the campaign on March 1, 2023, and concluded on September 8, 2023. Viewers of the vlog trailers were identified by an algorithm led by the social media ad expert who mobilized clicks on advertisements through social media sites using a strategy with established effectiveness. The social media ad expert tracked whether viewers clicked on referral sources. His algorithm determined how often the trailers were shown to social media audiences. Through careful monitoring on a weekly basis, the trailers with the least activity were removed from the campaign in an iterative fashion. The highest performing trailers (n=2) remained for the duration of the campaign. There was no participant enrollment process for individuals to view the social media campaign, and only publicly available data associated with paid advertisements were used to collect data. Individuals were not engaged on an individual basis. The social media campaign was publicly available.

Access to Local PrEP Clinics Through the Social Media Campaign

The project links cisgender women in an EHE priority jurisdiction to local PrEP resources, providing them with access to local PrEP clinic information through a landing page from a social media campaign, and introduced PrEP as an effective HIV prevention option through highly trafficked social media channels as a means of encouraging PrEP initiation with Black cisgender women in Harris County using vlogging for sexual health. Harris County residents who viewed the vlogs on their social media page had access to more information about PrEP by clicking on hyperlinks to local PrEP resources that accompany the vlog trailers. Hyperlinks were connected to websites and social media pages of local PrEP providers for review and following.

Ethical Considerations

The University of Texas Health Science Center at Houston–CPHS determined this study did not meet the regulatory definition of human subjects' research. Thus, no further review was required.

Data Analysis

A report developed with Facebook Insights and Google Analytics was used to evaluate the performance of the social media campaign. The analytic strategy collected data based on age, gender, race, and interest in health and beauty. Descriptive statistics assessed the self-identified gender and age of the profiles of individuals who engaged with the campaign. The study team also evaluated key performance indicators that included impressions (the number of times a user saw the vlog), video plays (the number of times a video starts playing), and the number of clicks to local PrEP resources (the total number of times users have clicked on a specific link). The study team also evaluated the economic cost of implementing a sexual health–focused social media campaign to increase PrEP access to Black cisgender women.

Results

Intervention Characteristics

Social media platforms that were leveraged for the campaign included Facebook and Instagram. Key performance indicators of the social media campaign stratified by age revealed that young people aged 18 - 24 years engaged most with the campaign (Table 1). Video plays at 50% of the vlogs (n=30,877) were most common among women aged 18 - 24 years (n=12,017) and least common among women aged 45 - 54 years (n=658). Video plays at 50% means that individuals started playing the video and watched 50% of the video. Table 2 shows an overview of women engaging with the social media campaign by age. Across cohorts, women aged 18 - 24 years were mostly represented, followed by a decrease in engagement with social media campaign in each subsequent cohort, with women aged 45 - 54 years being the least represented.

Table . Key performance indicators of the social media campaign, stratified by age (March-September 2023).

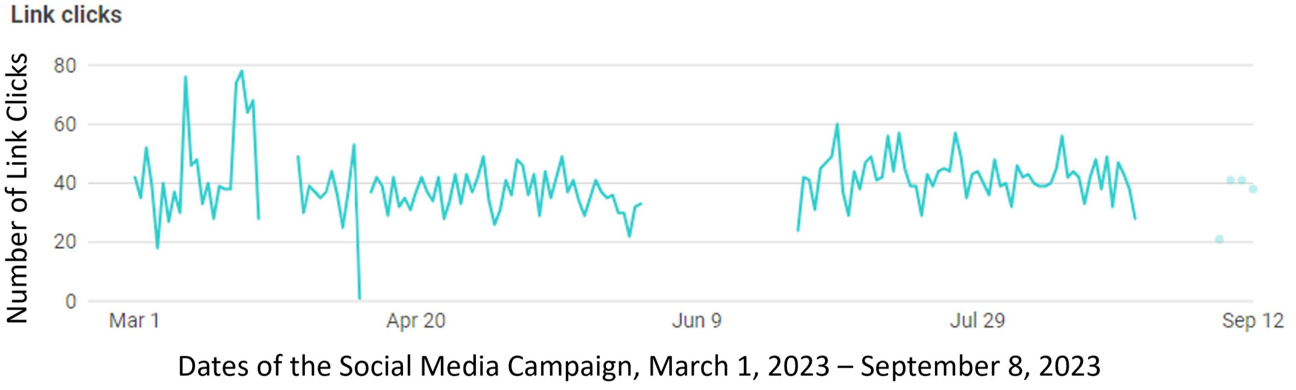
Age (years)	Link clicks	Reach	Impressions	Cost per link click (USD)	Amount spent (USD)	Video plays at 25%	Video plays at 50%	Video plays at 75%	Video plays at 95%	Video plays at 100%	Video plays
18 - 24	2330	39,968	491,656	$1.27	$2967.69	15,752	12,017	10,192	7072	6413	465,262
25 - 34	1834	38,816	336,524	$1.27	$2337.52	16,741	10,129	7494	5652	4913	302,026
35 - 44	1662	29,376	250,782	$1.28	$2124.98	14,534	8073	5204	4292	3567	217,758
45 - 54	126	2592	19,667	$1.43	$180.76	1255	658	413	335	286	17,198
Total	5952	110,752	1,098,629	$1.28	$7610.94	48,282	30,877	23,303	17,351	15,179	1,002,244

Table . Age distribution of women engaging with the social media campaign over 6 months.

Age group (years)	Total count (approximate)
13-17	0
18-24	2250
25-34	1800
35-44	1600
45-54	100
55-64	0
65+	0

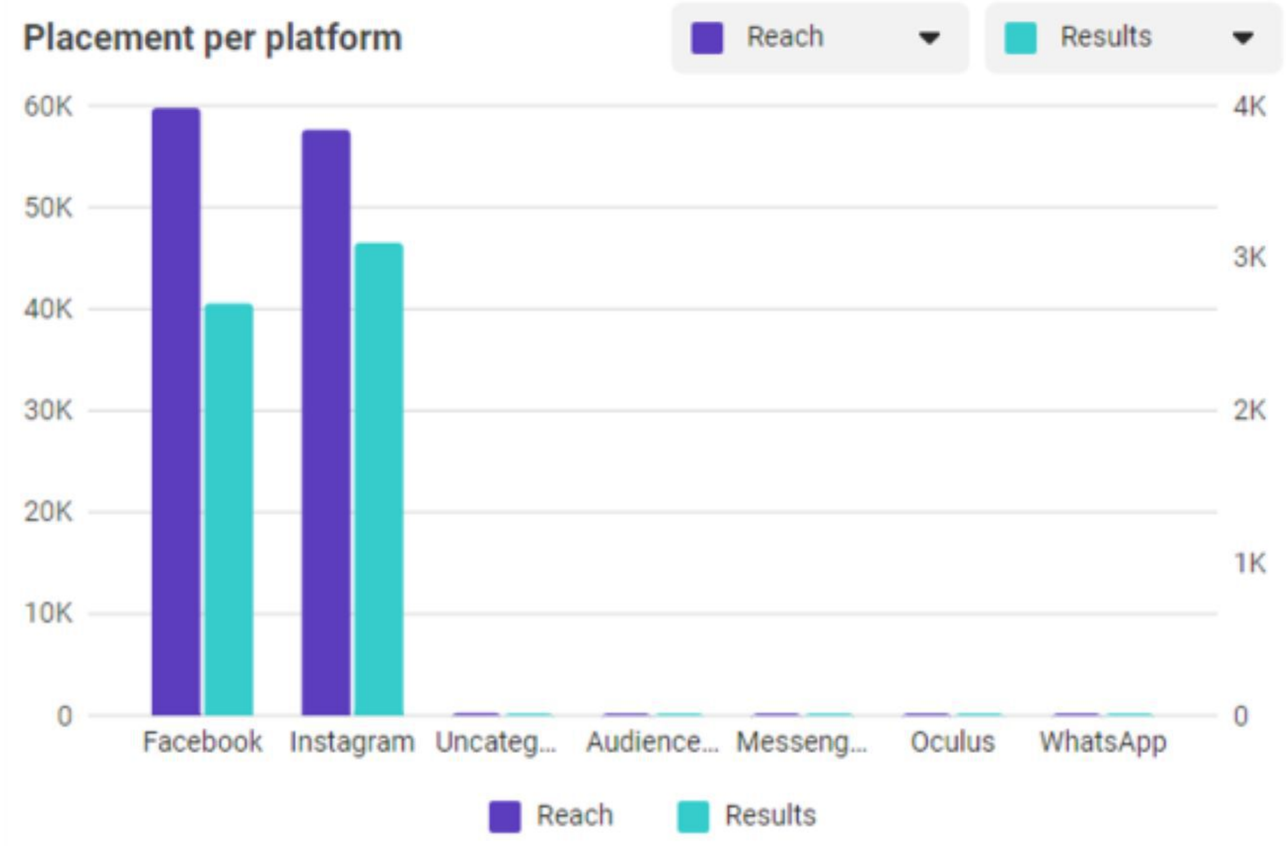
A line graph was used to assess the trend in clicks of the campaign over time (Figure 1). The trend reveals a steady link click rate of around 40 across the breadth of the 6-month campaign, largely within a 20 - 60 link click rate.

Figure 1. Summary of the trend in link clicks over the breadth of the social media campaign.



The final report showed 1,098,629 impressions and 1,002,244 total video plays and resulted in 15,952 link clicks to local PrEP resources. The campaign reached more individuals on Instagram (59,776) than on Facebook (57,623), but there were more clicks on Facebook (3101) than on Instagram (2703) (Figure 2).

Figure 2. Social media campaign performance based on placement per platform with a mobile device only. All platforms apart from Facebook and Instagram have a value of 0.



The evaluation of the economic cost of the social media campaign revealed that the cost was $1.28 per link click. The total amount spent was $7410.94 across the 6-month campaign. When stratified by social media platform, the campaign performed slightly better on Instagram than on Facebook, costing $3992.85 versus $3411.55, respectively (Table 3).

Table . Key performance indicators: by overall campaign and by platform (March-September 2023).

Key performance indicators and measures	Value
Overall social media campaign	
Reach (post engagement), n	110,752
Total impressions, n	1,098,629
Link clicks, n	5952
Cost (USD)	$7410.94
Cost per click (USD)	$1.28
Platform	
Facebook	
Reach, n	57,623
Link clicks, n	3101
Cost (USD)	$3411.55
Instagram	
Reach, n	59,776
Link clicks, n	2703
Cost (USD)	$3,992.85

Discussion

Study Findings and Comparison With Previous Works

The social media campaign over 6 months reached close to 6000 women. The campaign primarily used Facebook and Instagram. The highest engagement was observed among individuals aged 18 - 24 years, particularly women, who accounted for the most video plays (n=12,017). Engagement was the lowest among women aged 45 - 54 years (n=658). The distribution of engagement showed a consistent decline in participation across older age groups, with women aged 18 - 24 years being the most represented and women aged 45 - 54 years the least. The final report indicated 1,098,629 impressions and 1,002,244 total video plays, leading to 15,952 link clicks to local PrEP resources. The campaign reached more individuals on Instagram (59,776) than on Facebook (57,623), but Facebook generated more clicks (3101) compared to Instagram (2703).

When exploring the distribution of women who engaged with the social media campaign, it was evident that younger women were more likely to be represented. This finding supports existing research demonstrating that mobile health apps or electronic sexual and reproductive health interventions are effective in engaging young Black women [29,30]. Additionally, the findings align with previous research demonstrating that generational differences exist in the ways in which women access health information [28]. Many people access health information through their smartphones as the first step in gathering valid and correct information. This may be different from older generations, such as older millennials, generation X, or boomers, who developed sexual and reproductive health practices before the invention and proliferation of the internet and, as such, sought out professionals such as physicians, nurses, and other health care providers as trusted sources of sexual and reproductive health information. While seeking out medical professionals is a trusted and effective route, mobile health and e-health apps now remove the barrier of access to and utilization of health care and providers, due to the nature of needing insurance or being insured, in addition to costs related to co-pays, transportation, and time to take off work [29]. Additionally, these approaches to HIV prevention remove the burden of actively participating in one's sexual health wellness and accommodate diverse lifestyles [29,31]. Moreover, the findings of this study demonstrate that women aged 18 - 35 years engaged in the social media campaign. Given that newer cases of STIs such as chlamydia, gonorrhea, and syphilis are the highest among women aged 18 - 25 years [19] and new HIV diagnoses are reported among women aged 25 - 35 years [19], it is advantageous to explore the sustainability and scale-up of such an intervention to address the current public health needs. The decrease in engagement as age increased could indicate the need for tailored approaches in future campaigns, perhaps targeting specific age groups with content that resonates more with their interests and preferences [32].

Silva et al [29] sampled male and female youth aged 16 - 49 years on social media (including Twitter, subsequently rebranded X, and Facebook), along with other mass media platforms, in a study supporting intergenerational communication on sexual reproductive health and family planning with youth. While findings show strong potential, findings also showed that young people still need supportive guidance and approval from trusted adults in their region of the world (eg, sub-Saharan Africa) [29]. Another study assessed sexual health promotion campaigns on Twitter and Facebook and found significant potential with health promotion initiatives based on the number of interactions with users [29]. Although social media has been explored for health promotion and sexual health generally, the focus on PrEP engagement appears to be missing. This reveals a meaningful gap and unmet need that this study was poised to fill.

Although minimal, a greater number of women used Instagram as opposed to Facebook. However, link clicks on Facebook were higher than that on Instagram. It would be advantageous to retain both social media sites for future programming, as each site tends to garner a different viewership demographic [33]. Despite the differences in user preferences, both platforms were deemed acceptable to women engaged in the social media campaign over 6 months.

Compared to the over 1 million impressions and over 1 million total video plays, clicks to the local PrEP resources were close to 16,000, demonstrating that while women are interested in the campaign and videos, there is less engagement with PrEP. This may point to a critical point in women's decision-making in terms of their self-assessment of HIV risk or vulnerability and the need to try a new HIV prevention strategy. The novelty of the innovation reinforces the fact that such social media campaigns for Black women are scant, and the scale-up of these interventions would make engaging in HIV prevention practices commonplace and more widely adopted. This study, however, is the first of its kind to actually embed local PrEP resources within the vlogs to drive traffic to the clinics. This is an ideal strategy as it addresses the retention of participants. Findings from this study suggest that social media campaigns as an intervention strategy to increase PrEP access are likely a compatible complement to EHE goals and outreach efforts to enhance access to PrEP providers among populations who are vulnerable to new HIV diagnoses. Video plays were over 1 million views, surpassing the expectations of the investigative team while solidifying the feasibility of this approach at reaching cisgender women through social media with PrEP messaging.

Limitations

There are limitations to our study. First, there were 2 breaks in the campaign, stalled by communication gaps within the team and logistical challenges within the institution occurring between March 23, 2023 and March 30, 2023 and May 30, 2023 and June 27, 2023. Although this occurred, our findings remain robust, and if the study team had not had breaks in the campaign, the results could have been similar or higher for reach, clicks, vlog views, and accessing local PrEP clinic resources. Second, in regard to the racial profile of the study population, the ability of any social media campaign on Facebook or Instagram to market advertisements by race was removed. In lieu of this sudden and unpredicted change, the study team and consultant advertisement expert used 2 proxies in lieu of race, which were content, specifically health and beauty, and zip codes in areas where a significant majority were Black. As an example, zip code 77033 was one of several zip codes used. According to the United States Postal Service, the majority of residents in 77033, specifically 75.4% as of January 13, 2025, are Black or African American. Individuals who resided in zip code 77033 often viewed health and beauty content and described themselves as female; they were shown the vlogs during the social media campaign. Based on the expertise of our consultant, the team was confident that this approach would garner viewership from a majority Black female population. However, it is likely that the vlogs were viewed by a minority of non-Black women.

Future Research

Future studies examining implementation outcomes such as acceptability could engage the users in focus groups or in-depth qualitative interviews to determine reasons for staying on the site or linking to local PrEP clinics. Given that many clinics operate on limited budgets, it may be advantageous to invest in a formal feedback session on their webpage for opportunities to garner usage of the site further and enhance the user experience and ability to get the information needed quickly. Future studies could also investigate the appropriateness of such an intervention. Additionally, the utilization of a hybrid type 2 design to examine the scale-up and testing of both the intervention effectiveness and implementation strategy would allow for the simultaneous assessment of HIV prevention via vlogs with multiple sites.

Conclusions

The campaign demonstrated the feasibility, acceptability, and appropriateness of this approach with Black cisgender women and illustrated preliminary effectiveness at supporting access to local PrEP resources with Black cisgender women. Further dissemination and implementation of this approach is necessary to fully assess whether vlog viewership and clicks on links to PrEP resources can meaningfully empower Black cisgender women to access PrEP and help them to assess whether PrEP is personally a useful HIV prevention option.

Acknowledgments

We would like to acknowledge the Texas Development Center for AIDS Research (5P30AI161943-04; PI: Giordano, Rice, and Santa Maria) for funding this pilot research. We would also like to thank Carmen Avalos for all of her administrative support. We would like to acknowledge Sarah Sapp, MPH for her research coordination on this project. As the Debunking Myths study served as the foundation for this work, we would like to thank the Texas Emergency Medicine Research Center for funding that investigator-initiated study in 2021, which served as a springboard for this work, that led to the current funded study that will extend this work.

Conflicts of Interest

MH is the recipient of an Investigator Research Award from Merck Pharmaceuticals Project grant # 100099, which funded a qualitative study in Houston and Austin, Texas, with an objective to explore perspectives and beliefs related to PrEP readiness among Black cis women. A small percentage of MH’s effort toward the conduct of the research was supported as a part of the funding provided. The primary outcome of this study is different but related to the outcome of the project presented here in this manuscript and overlaps in the geographic location and population of focus.

References

1. Ojikutu BO, Mayer K. HIV prevention among black women in the US-time for multimodal integrated strategies. JAMA Netw Open 2021 Apr 1;4(4):e215356. [doi: 10.1001/jamanetworkopen.2021.5356] [Medline: 33835180]
2. HIV surveillance in urban and nonurban areas through 2017. : National Center for HIV/AIDS VH, STD, and TB Prevention; 2019 Nov 8 URL: https://stacks.cdc.gov/view/cdc/149300 [accessed 2025-03-17]
3. D’Amore J, Hung O, Chiang W, Goldfrank L. The epidemiology of the homeless population and its impact on an urban emergency department. Acad Emerg Med 2001 Nov;8(11):1051-1055. [doi: 10.1111/j.1553-2712.2001.tb01114.x] [Medline: 11691667]
4. AIDSVu. Houston/harris county - rates of persons living with HIV, 2018. 2021. URL: https://aidsvu.org/local-data/united-states/south/texas/houston [accessed 2021-07-02]
5. Aaron E, Blum C, Seidman D, et al. Optimizing delivery of HIV preexposure prophylaxis for women in the United States. AIDS Patient Care STDS 2018 Jan;32(1):16-23. [doi: 10.1089/apc.2017.0201] [Medline: 29323558]
6. Chou R, Evans C, Hoverman A, et al. Preexposure prophylaxis for the prevention of HIV infection: evidence report and systematic review for the US Preventive Services Task Force. J Am Med Assoc 2019;321(22):2214-2230. [doi: 10.1001/jama.2019.2591]
7. Lewin S. Long-acting technologies. Presented at: 25th International AIDS Conference; Jul 29, 2024; Munich, Germany.
8. Conley C, Johnson R, Bond K, Brem S, Salas J, Randolph S. US Black cisgender women and pre-exposure prophylaxis for human immunodeficiency virus prevention: a scoping review. Womens Health (Lond) 2022;18:17455057221103098. [doi: 10.1177/17455057221103098] [Medline: 35699104]

9. Willie TC, Knight D, Baral SD, et al. Where's the "Everyday Black Woman"? An intersectional qualitative analysis of Black Women's decision-making regarding HIV pre-exposure prophylaxis (PrEP) in Mississippi. BMC Public Health 2022 Aug 23;22(1):1604. [doi: 10.1186/s12889-022-13999-9] [Medline: 35999528]
10. Cheek JB, Feldman MB, Montero N, Gambone GF, Hoffman S, Blackstock OJ. Pre-exposure prophylaxis (PrEP) initiation among Black and Latina cisgender women receiving HIV prevention care coordination services in New York City. AIDS Behav 2022 Oct;26(10):3174-3184. [doi: 10.1007/s10461-022-03661-1] [Medline: 35362904]
11. Mayer KH, Agwu A, Malebranche D. Barriers to the wider use of pre-exposure prophylaxis in the United States: a narrative review. Adv Ther 2020 May;37(5):1778-1811. [doi: 10.1007/s12325-020-01295-0] [Medline: 32232664]
12. D'Angelo AB, Davis Ewart LN, Koken J, Bimbi D, Brown JT, Grov C. Barriers and facilitators to pre-exposure prophylaxis uptake among Black women: a qualitative analysis guided by a socioecological model. J Assoc Nurses AIDS Care 2021;32(4):481-494. [doi: 10.1097/JNC.0000000000000241] [Medline: 34171885]
13. Nydegger LA, Dickson-Gomez J, Ko Ko T. A longitudinal, qualitative exploration of perceived HIV risk, healthcare experiences, and social support as facilitators and barriers to PrEP adoption among Black women. AIDS Behav 2021 Feb;25(2):582-591. [doi: 10.1007/s10461-020-03015-9] [Medline: 32886220]
14. Irie WC, Mahone A, Johnson B, et al. "Just the Stigma Associated with PrEP Makes You Feel Like It'S HIV Itself": exploring PrEP stigma, skepticism, and medical mistrust among Black cisgender women in urban and rural counties in the U.S. Deep South. Arch Sex Behav 2024 Mar;53(3):1187-1195. [doi: 10.1007/s10508-023-02769-2] [Medline: 38195827]
15. Heads AM, Hill MJ, Suchting R, Yammine L, Gilmore-Thomas A. Predictors of anticipated PrEP stigma among women with self-reported problematic substance use: implications for engaging women in the PrEP care continuum. Arch Sex Behav 2021 Oct;50(7):2955-2964. [doi: 10.1007/s10508-021-02031-7] [Medline: 34561793]
16. Thomas T, Hill M. Reversing the dehumanization of Black women. In: Black Women and Public Health: Regenerative History, Practice, and Planning: Suny Press; 2022. [doi: 10.2307/jj.18253045.6]
17. Chandler R, Guillaume D, Wells J, Hernandez N. Let Me PREP You to PREP Me: amplifying the voices of Black women and their providers to consider PREP as an HIV prevention option. Int J Environ Res Public Health 2022 Jan 27;19(3):1414. [doi: 10.3390/ijerph19031414] [Medline: 35162438]
18. Drossman DA, Ruddy J. Improving patient-provider relationships to improve health care. Clin Gastroenterol Hepatol 2020 Jun;18(7):1417-1426. [doi: 10.1016/j.cgh.2019.12.007] [Medline: 31843593]
19. Centers for Disease Control and Prevention. Sexually transmitted infections surveillance, 2023. National overview of STIS in 2023 | STI statistics. CDC. 2024.
20. Hill MJ, Sapp S, McCants S, et al. An examination of perceptions among black women on their awareness of and access to pre-exposure prophylaxis (PrEP). Int J Environ Res Public Health 2024 Aug 16;21(8):1084. [doi: 10.3390/ijerph21081084] [Medline: 39200693]
21. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
22. Hill M, Santa Maria D, Coker S, Sapp S, Sutton T. Leveraging video logs to enhance HIV prep access to black women in harris county. Presented at: American Public Health Association (APHA) 2023 Annual Meeting and Expo; Nov 12-15, 2023; Atlanta, GA.
23. Hill M, Sapp S, Sutton T, Coker S, Santa Maria D. Leveraging video logs to support HIV prep access to black women in harris county. Presented at: Inter-CFAR Women and HIV Symposium; Dec 7, 2023; San Francisco, CA.
24. Hill MJ, Coker S. Novel use of video logs to deliver educational interventions to Black women for disease prevention. West J Emerg Med 2022 Feb 28;23(2):211-221. [doi: 10.5811/westjem.2021.12.54012] [Medline: 35302455]
25. Coker S, Hill M. Debunking myths on HIV risks: a targeted strategy to challenge perceptions of low HIV/STI risks among African American women during an emergency department visit. Presented at: National Medical Association Annual Meeting, Emergency Medicine Section; 2018; Orlando, FL.
26. Coker S, Hill M, Stotts A. Leveraging the emergency department visit to debunk myths on HIV risks among substance-using African American women. Presented at: The College on Problems of Drug and Alcohol Dependence Annual Meeting; 2019; San Antonio, TX.
27. Hill M, Coker S, Hudson T. Leveraging the emergency department visit to debunk myths on HIV and STI risks among african american women. Presented at: The National Conference for HIV, HCV, STI, and LGBTQ Health; 2020; Virtual URL: https://www.scitechnol.com/peer-review/just-an-indicator-of-minimized-value-of-the-sexual-act-nHxD.php?article_id=7768 [accessed 2025-03-12]
28. Hill M, Andrews S, Granado M, Nielsen E, Grimes R. Just': an indicator of minimized value of the sexual act. HIV AIDS Res J 2018;1(2):1-9.
29. Silva M, Loll D, Ezouatchi R, et al. Evaluating a youth-designed sexual and reproductive health mass and social media campaign in Côte d'Ivoire: triangulation of three independent evaluations. Sex Reprod Health Matters 2023 Dec;31(1):2248748. [doi: 10.1080/26410397.2023.2248748] [Medline: 37728549]

30. Tesema N, Guillaume D, Francis S, Paul S, Chandler R. Mobile phone apps for HIV prevention among college-aged Black women in Atlanta: mixed methods study and user-centered prototype. JMIR Form Res 2023 Feb 23;7:e37987. [doi: 10.2196/37987] [Medline: 36821362]
31. Texas Emergency Medicine Research Center. TX-DCFAR Pilot Project: McGovern Medical School at UTHealth Houston; 2022. URL: https://www.cdc.gov/sti-statistics/annual/summary.html#:~:text=Findings&text=In%202023%2C%20over%202.4%20million,1.6%20million%20cases%20of%20chlamydia [accessed 2025-03-12]
32. Chandler R, Hernandez N, Guillaume D, Grandoit S, Branch-Ellis D, Lightfoot M. A community-engaged approach to creating a mobile HIV prevention app for Black women: focus group study to determine preferences via prototype demos. JMIR Mhealth Uhealth 2020 Jul 24;8(7):e18437. [doi: 10.2196/18437] [Medline: 32706723]
33. Alhabash S, Ma M. A tale of four platforms: motivations and uses of Facebook, Twitter, Instagram, and Snapchat among college students? Social Media + Society 2017 Jan;3(1). [doi: 10.1177/2056305117691544]

Abbreviations

CPHS: University of Texas Health Science Center at Houston's Committee for the Protection of Human Subjects
EBI: evidence-based intervention
EHE: ending the HIV epidemic
PrEP: pre-exposure prophylaxis
STI: sexually transmitted disease

Edited by A Mavragani, T Sanchez; submitted 09.10.24; peer-reviewed by D Dinh, M Pyra; revised version received 13.01.25; accepted 15.01.25; published 28.03.25.

Please cite as:
Hill MJ, Mangum L, Coker SJ, Sutton T, Santa Maria DM
Dissemination and Implementation Approach to Increasing Access to Local Pre-Exposure Prophylaxis (PrEP) Resources With Black Cisgender Women: Intervention Study With Vlogs Shared on Social Media
JMIR Public Health Surveill 2025;11:e67367
URL: https://publichealth.jmir.org/2025/1/e67367
doi:10.2196/67367

Digital Surveillance of Mental Health Care Services in Saudi Arabia: Cross-Sectional Study of National e-Referral System Data

Abdullah A Alharbi[1], PhD; Nawfal A Aljerian[2,3], SBEM; Meshary S Binhotan[4,5], PhD; Hani A Alghamdi[6], PhD; Ali K Alsultan[2], EMBA; Mohammed S Arafat[2], EMBA; Abdulrahman Aldhabib[2], MBBS; Yasser A Alaska[7], MBBS; Eid B Alwahbi[2], MSc; Mohammed A Muaddi[1], PhD; Ahmad Y Alqassim[1], PhD; Ronnie D Horner[8], PhD

[1]Department of Family and Community Medicine, Faculty of Medicine, Jazan University, Jazan, Saudi Arabia
[2]Medical Referrals Centre, Ministry of Health, Riyadh, Saudi Arabia
[3]Department of Emergency Medicine, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia
[4]Department of Emergency Medical Services, College of Applied Medical Sciences, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia
[5]King Abdullah International Medical Research Center, Riyadh, Saudi Arabia
[6]Department of Family and Community Medicine, College of Medicine, King Saud University, Riyadh, Saudi Arabia
[7]Department of Emergency Medicine, King Saud University, Riyadh, Saudi Arabia
[8]Department of Health Services Research and Administration, College of Public Health, University of Nebraska Medical Center, Omaha, NE, United States

Corresponding Author:
Mohammed A Muaddi, PhD
Department of Family and Community Medicine, Faculty of Medicine, Jazan University, Jazan, Saudi Arabia

Abstract

Background: Mental illness affects an estimated 25% of the global population, with treatment gaps persisting worldwide. The COVID-19 pandemic has exacerbated these challenges, leading to a significant increase in mental health issues globally. In Saudi Arabia, the lifetime prevalence of mental disorders is estimated at 34.2%, yet 86.1% of those with a 12-month mental disorder report no service use. To address these challenges, digital health solutions, particularly electronic referral (e-referral) systems, have emerged as powerful tools to improve care coordination and access. Saudi Arabia has pioneered the nationwide Saudi Medical Appointments and Referrals Centre (SMARC), a centralized e-referral system using artificial intelligence and predictive analytics.

Objectives: This study aims to analyze Saudi Arabia’s novel nationwide e-referral system for mental health services, using SMARC platform data to examine referral patterns, and service accessibility. This study also aims to demonstrate how digital health technology can inform and improve mental health care delivery and policy making.

Methods: This retrospective, cross-sectional study used secondary data from SMARC on 10,033 psychiatric e-referrals in Saudi Arabia during 2020‐2021. Referrals were assessed by patient sociodemographic variables, geographic data, and e-referral characteristics including date, type, bed type, and reason for e-referral. Descriptive statistical analyses identified referral patterns, while regression modeling determined predictors of external referrals to other regions.

Results: Analysis of 10,033 psychiatric e-referrals revealed that 58.99% (n=5918) were for patients aged 18‐44 years, 63.93% (n=6414) were for men, and 87.10% (n=8739) were for Saudi nationals. The Western Business Unit generated 45.17% (n=4532) of all e-referral requests. Emergency cases were the most common type of referral overall, followed by routine inpatient and routine outpatient department referrals. However, in the Northern Business Unit, routine inpatient referrals were most frequent. Two-thirds of requests were for ward beds, while critical beds were rarely requested. “Unavailable subspecialty” was the primary reason for referrals across all regions. The utilization of the mental health e-referral system varied across regions, with the Northern Border and Albaha regions showing the highest rates, while Madinah, Eastern, and Riyadh regions demonstrated lower use. Temporal analysis showed almost similar monthly patterns in 2020 and 2021. There was an overall increase in referrals in 2021 compared with 2020.

Conclusions: This pioneering study of mental health e-referrals in Saudi Arabia demonstrates how digital health transformation, particularly through an e-referral system, has significantly enhanced access to mental health services nationwide in Saudi Arabia. The success of this digital initiative demonstrates how digital health solutions can transform health care access, particularly in mental health services, offering a valuable model for other health care systems.

(JMIR Public Health Surveill 2025;11:e64257) doi:10.2196/64257

KEYWORDS

digital health; mental health; health policy; epidemiology; Saudi Arabia; SMARC; health care transformation; e-referral; Saudi Medical Appointments and Referrals Centre

Introduction

Mental illness remains a serious global health concern, with the World Health Organization estimating that it affects 25% of the population worldwide [1-3]. Significant treatment gaps persist globally, as around 70% of those needing mental health management lack access to care [4-8]. This high rate of untreated patients partly stems from specialist shortages and geographic barriers distancing patients from care [9,10], contributing to a major treatment gap. Furthermore, stigma and lack of mental health literacy continue to be significant barriers to seeking care, particularly in Middle Eastern countries [11].

The global landscape of mental health has evolved significantly in recent years, with the COVID-19 pandemic exacerbating existing challenges and creating new ones. Recent studies have shown that these gaps have widened during the COVID-19 pandemic, with a significant increase in mental health issues reported globally [12,13]. A meta-analysis estimated that the pandemic led to an additional 53.2 million cases of major depressive disorder and 76.2 million cases of anxiety disorders globally in 2020 alone [14]. These findings underscore the urgent need for scalable and accessible mental health interventions worldwide.

In response to these challenges, digital health solutions have emerged as powerful tools to improve coordination and access to mental health care. Electronic referral (e-referral) systems, in particular, have shown promise in reducing wait times, improving care coordination, and increasing access to specialist care [15,16]. The Kingdom of Saudi Arabia (KSA) has been an early pioneer in leveraging digital health through the nationwide Saudi Medical Appointments and Referrals Centre (SMARC). This centralized e-referral system was revamped and relaunched in 2019 to unify and streamline referrals across the country's 13 administrative regions. The advanced SMARC platform uses artificial intelligence and predictive analytics to coordinate referrals between public and private hospitals, addressing previous fragmentation. As of 2019, the majority of governmental and private Saudi health care facilities are connected to SMARC, enabling efficient nationwide referral management—a major milestone in the KSA's health sector digital transformation under Saudi vision 2030 [17,18].

Recent studies have highlighted significant challenges and opportunities to improve mental health care in Saudi Arabia. Estimates show that 34.2% of the general Saudi population experience a mental disorder during their lifetime [19]. Regarding depressive disorders, there was about a 60% increase in number of new cases from 1.42 million cases in 2011 to 2.28 million cases in 2021. The incidence rate also rose by around 23% during the same period, going from 4,913.71 to 6,039.77 per 100,000 people [20]. Despite this high prevalence, access to mental health care remains limited, with a national survey revealing that about 84% (597/711) of participants with a 12-month mental disorder reported no service use [21]. Barriers to care include low perceived need for help, cultural stigma, and attitudinal obstacles [22]. These findings emphasize the critical role that e-referral systems such as SMARC could play in addressing these challenges.

SMARC facilitates analysis of mental health patterns at a national scale, providing a unique opportunity for epidemiological and health services research on mental illnesses [1,2]. However, no large-scale studies have been conducted to date in the KSA, making such analysis a high priority [23]. This innovative study is the first to examine Saudi Arabia's adoption of an e-referral system for mental health services, establishing digital monitoring of mental health care referrals. By analyzing national referral trends using SMARC platform data, it provides crucial insights into care patterns and resource distribution. The research showcases Saudi Arabia's innovative use of technology to evaluate and enhance mental health service provision and accessibility. Our findings, derived from SMARC data, illuminate potential gaps in the mental health care system and could inform evidence-based policies and interventions. This aligns with the Saudi's digital-driven health care transformation, offering valuable insights to optimize mental health care delivery throughout Saudi Arabia. Ultimately, this study contributes to the broader goal of improving mental health outcomes by leveraging technology to guide strategic health care decisions and resource allocation.

Methods

Study Design and Setting

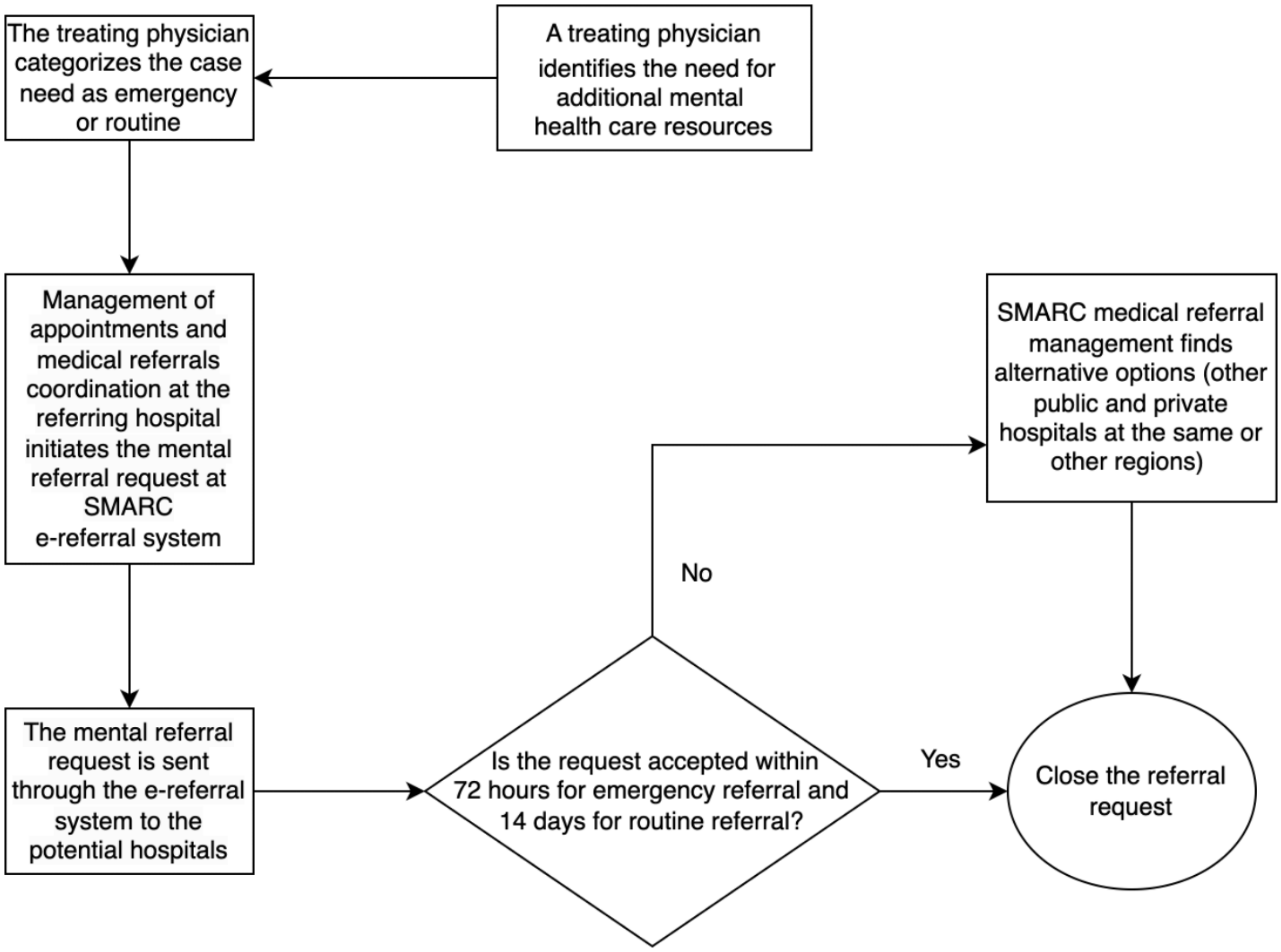
This descriptive study of psychiatric health care delivery used data by the SMARC system between January 2020 and December 2021. The SMARC system is concerned with all e-referrals across the KSA. Unlike conventional referral systems, this system primarily serves e-referral between secondary, tertiary, and specialized health care providers. The SMARC system routinely collects an array of data concerning each e-referral request.

A descriptive approach was chosen for this study due to several key factors. First, this is the first comprehensive analysis of the Saudi national e-referral system for mental health services, necessitating an exploratory approach to establish baseline patterns and trends. Second, the descriptive nature of our study provides a crucial foundation for generating hypotheses that can guide future in-depth research. Third, given the novelty of the data source and the lack of prior comprehensive analyses in this context, a descriptive approach allows for a broad understanding of the current landscape of mental health referrals in Saudi Arabia. Finally, our descriptive findings offer immediate, actionable insights for policy makers and health care planners, which is particularly valuable given the urgency of addressing mental health needs in the region.

To initiate a referral request, a treating physician has to first identify the need and type of the referral, emergency or routine (both inpatient and outpatient). The Office of Coordination and Eligibility for Treatment, which is allocated in every hospital, digitally uploads the referral request at SMARC e-referral system. The Office of Coordination and Eligibility for Treatment can identify 3 potential receiving hospitals or can leave the task to the e-referral system.

The referral request is sent, through the e-referral system, to the potential receiving hospitals for acceptance. The e-referral system has built-in time frames for requests to be accepted, depending on the referral type. For emergency requests, the system allows hospitals 72 hours to accept the request, while routine requests have 14 days. If the initially identified hospitals reject the request within the allotted time frame, further hospitals will receive the request for acceptance consideration. Once the allotted time frame elapses with no acceptance, the request will automatically be escalated to SMARC medical referral management team to secure an acceptance. The team will look at alternative options including other governmental and private hospitals, both at the same region and other regions if needed. During the process of referral initiation and acceptance, all patients continue receiving the necessary health care management at the referring hospital, to ensure stability, until being transferred to the receiving hospital. Figure 1 describes the process of mental referral request initiation and acceptance.

Figure 1. Flowchart depicting the process of initiating and accepting psychiatric e-referral requests through the SMARC in Saudi Arabia from January 2020 to December 2021. This cross-sectional study analyzed data on 10,033 psychiatric e-referrals across all regions of Saudi Arabia during this 2-year period. e-Referral: electronic referral; SMARC: Saudi Medical Appointments and Referrals Centre.



Study Variables

All e-referral data provided by the SMARC system were used in this analysis, totaling 10,033 referrals. The variables included sociodemographic variables such as age, sex, and nationality of the patient with the e-referral request, as well as the patient's administrative area of residence and business unit (BU) from which the request originated. BU refers to a geographical and administrative area designated for health care service delivery. The health care system comprises 5 BUs, each managed by a government-owned holding company. The units are distributed based on geographical coverage: Central BU includes Riyadh and AL Qassim; Western BU covers Makkah, Madinah, and Albaha; Eastern BU includes the Eastern region; Northern BU consists of Al-jouf, Northern Border, Tabuk, and Hail; and Southern BU includes Aseer, Jazan, and Najran. Each BU is tasked with managing and providing health care services within its respective regions.

Other variables pertain to the e-referral request itself, including the date, type, bed type, and reason. The type of e-referral includes routine outpatient department (OPD), routine inpatient, and emergency. Routine referrals (both OPD and inpatient) are those that do not require prompt intervention and are supposed to be accepted within 14 days. An emergency e-referral is a case requiring medical or surgical intervention within 72 hours to avoid mortality or morbidity. The bed type refers to OPD, ward bed, or critical bed. A critical bed indicates a case needing admission to the intensive care unit. Reasons for e-referral include unavailable specialty, unavailable physician, unavailable machine (meaning no tools), and unavailable bed (meaning no available bed at the hospital).

Statistical Analysis

Descriptive statistics were calculated for all variables, including frequencies and percentages. To explore potential regional differences, we conducted cross-tabulations of each variable across the 5 BUs and performed χ^2 tests to assess associations. STATA (version 16; StataCorp LLC) was used for all analyses. Furthermore, we generated color-coded maps using ArcGIS software (ArcGIS 10.0; Environmental Systems Research Institute, Inc) to visually represent the distribution of e-referrals requests across different administrative areas.

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and approved by the institutional review board of the Ministry of Health (MOH) (IRB log No. 23 - 77-E; date: September 20, 2023). The requirement for individual informed consent was waived due to the retrospective nature of the study and the use of deidentified data. Patient privacy and confidentiality were strictly maintained throughout the research process. All data obtained from the SMARC system were anonymized prior to analysis. Access to the dataset was restricted to authorized research team members and stored on secure, password-protected servers. The study did not involve any direct contact with patients or any intervention that could affect their care. We confirm that no images or information that could lead to the identification of individual participants is included in this manuscript or any supplementary materials.

Results

Sociodemographic Characteristics of Psychiatry Patients With e-Referral Requests

Of the 10,033 patients with a psychiatry-related e-referrals during the study period, 58.99% (n=5918) were aged 18 - 44 years, followed by those aged 45 - 64 years (n=1869, 18.63%). Patients younger than 18 years or older than 65 years comprised relatively few requests. Nearly two-thirds of requests were for men (6414/10,033, 63.93%) compared with women (3619/10,033, 36.07%). The majority (8739/10,033, 87.1%) were Saudi nationals. Approximately 45% (n=4532) of all e-referral requests originated from the Western BU, while the other regions each generated around one-third or less than this proportion (Table 1).

Table . Sociodemographic characteristics of psychiatry patients with e-referral requests in Saudi Arabia from January 2020 to December 2021[a].

Characteristics	Participants (N=10,033), n (%)
Age (years)	
<11	844 (8.41)
12 - 17	617 (6.15)
18 - 44	5918 (58.99)
45 - 64	1869 (18.63)
>65	785 (7.82)
Sex	
Male	6414 (63.93)
Female	3619 (36.07)
Nationality	
Non-Saudi	1294 (12.90)
Saudi	8739 (87.10)
Business unit	
Central	1128 (11.24)
Eastern	1220 (12.16)
Western	4532 (45.17)
Northern	1603 (15.98)
Southern	1550 (15.45)
Year	
2020	4179 (41.65)
2021	5854 (58.35)

[a] This cross-sectional study analyzed data on 10,033 psychiatric e-referral requests across the 5 business units (Central, Eastern, Western, Northern, and Southern) in Saudi Arabia between January 2020 and December 2021. Data used were the data routinely collected by the Saudi Medical Appointments and Referrals Centre (SMARC) system. The SMARC system is concerned with all e-referral requests in Saudi Arabia.

Sociodemographic Characteristics of Psychiatry Patients With e-Referral Requests According to BUs

Table 2 shows the sociodemographic characteristics of patients with a psychiatry-related e-referral request for each of the 5 BUs. In general, the patients being referred for psychiatric care in each BU were similar in characteristics to those identified for the population as a whole. These associations within regions were all significant at the $P<.001$ level.

Table . Sociodemographic profile of psychiatric e-referral patients across different business units in Saudi Arabia from January 2020 to December 2021[a].

Characteristics	Central (1128/10,033, 11.24%), n (%)	Eastern (1220/10,033, 12.16%), n (%)	Western (4532/10,033, 45.17%), n (%)	Northern (1603/10,033, 15.98%), n (%)	Southern (1550/10,033, 15.45%), n (%)
Age (years) (*P* <.001)					
<11	72 (6.38)	130 (10.66)	341 (7.52)	121 (7.55)	180 (11.61)
12 - 17	71 (6.29)	105 (8.61)	260 (5.74)	73 (4.55)	108 (6.97)
18 - 44	704 (62.41)	643 (52.70)	2878 (63.50)	829 (51.72)	864 (55.74)
45 - 64	230 (20.39)	253 (20.74)	799 (17.63)	334 (20.84)	253 (16.32)
>65	51 (4.25)	89 (7.30)	254 (5.60)	246 (15.35)	145 (9.35)
Sex (*P*<.001)					
Males	694 (61.52)	739 (60.57)	2998 (66.15)	998 (62.26)	988 (63.74)
Females	434 (38.48)	481 (39.43)	1534 (33.85)	605 (37.74)	562 (36.26)
Nationality (*P*<.001)					
Non-Saudi	107 (9.49)	106 (8.69)	658 (14.52)	189 (11.79)	234 (15.10)
Saudi	1021 (90.51)	1114 (91.31)	3874 (85.48)	1414 (88.21)	1316 (84.90)

[a] This cross-sectional study analyzed data on 10,033 psychiatric e-referral requests across the 5 business units (Central, Eastern, Western, Northern, and Southern) in Saudi Arabia between January 2020 and December 2021. Data used were the data routinely collected by the Saudi Medical Appointments and Referrals Centre (SMARC) system. The SMARC system is concerned with all e-referral requests in Saudi Arabia. Correlation between variables was examined through a χ^2 test, revealing a highly significant difference at a significance level of $P \leq .001$.

e-Referrals Characteristics of Psychiatry Patients by BU

Table 3 presents the results of psychiatry e-referral requests overall and for each of the 5 BUs. Out of the 3 types of psychiatry e-referrals , the most common type was for emergency cases and the least reported type was routine OPD e-referral. Emergency psychiatric e-referral was also the most common type for each BU with the exception of the Northern BUs, where routine inpatient e-referral was the highest. With regard to bed types, two-thirds of all requests were for ward beds; very few requests were for critical beds. Ward beds were also the most commonly requested type of bed across BUs. “Unavailable subspecialty” was the most commonly reported reason for the psychiatry-related e-referrals overall, whereas the “unavailability of a machine” was the least reported reason. Among each BUs, a similar pattern was observed for the reason for the e-referral.

Table . Characteristics of psychiatric e-referrals[a] across different business units in Saudi Arabia from January 2020 to December 2021[b].

Characteristic	Central (1128/10,033, 11.24%), n (%)	Eastern (1220/10,033, 12.16%), n (%)	Western (4532/10,033, 45.17%), n (%)	Northern (1603/10,033, 15.98%), n (%)	Southern (1550/10,033, 15.45%), n (%)	Total (N=10,033), n (%)
e-Referral types (*P*<.001)						
Routine OPD[c]	411 (36.44)	447 (36.64)	1297 (28.62)	326 (20.34)	437 (28.19)	2918 (29.08)
Routine inpatient	192 (17.02)	325 (26.64)	1473 (32.50)	675 (42.11)	410 (26.45)	3075 (30.65)
Emergency	525 (46.54)	448 (36.72)	1762 (38.88)	602 (37.55)	703 (45.35)	4040 (40.27)
Bed type (*P*<.001)						
OPD	409 (36.26)	447 (36.64)	1292 (28.51)	324 (20.21)	437 (28.19)	2918 (29.08)
Ward	709 (62.85)	750 (61.48)	3090 (68.18)	1224 (76.36)	1064 (68.65)	6828 (68.06)
Critical	10 (0.89)	23 (1.89)	150 (3.31)	55 (3.43)	49 (3.16)	287 (2.86)
Reason of e-referral (*P*<.001)						
Unavailable subspecialty	767 (68)	776 (63.61)	3197 (70.54)	1046 (65.25)	1107 (71.42)	6893 (68.70)
Unavailable physician	249 (22.07)	394 (32.30)	944 (20.83)	473 (29.51)	292 (18.84)	2352 (23.44)
Unavailable machine	26 (2.30)	34 (2.79)	97 (2.14)	61 (3.81)	96 (6.19)	314 (3.13)
Unavailable bed	86 (7.62)	16 (1.31)	294 (6.49)	23 (1.43)	55 (3.15)	474 (4.72)

[a] e-referral: electronic referral.

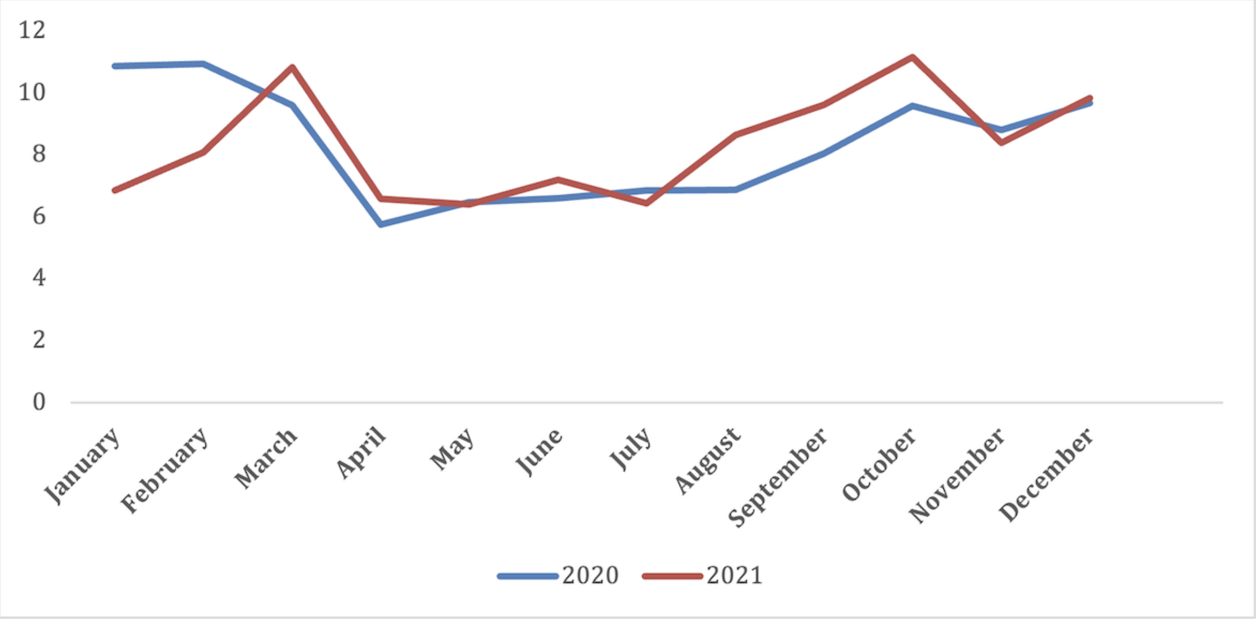
[b] This cross-sectional study analyzed data on 10,033 psychiatric e-referrals across the 5 business units (Central, Eastern, Western, Northern, and Southern) in Saudi Arabia between January 2020 and December 2021. Data used were the data routinely collected by the Saudi Medical Appointments and Referrals Centre (SMARC) system. The SMARC system is concerned with all e-referral in Saudi Arabia. . Correlation between variables was examined through a χ^2 test, revealing a highly significant difference at a significance level of $P\leq.001$.

[c] OPD: outpatient department.

Trends and Patterns of Psychiatry e-Referral Requests

Figure 2 shows the trend of psychiatry-related e-referral requests across months separately for 2020 and 2021. For both years, the pattern was similar across the months between March and December, with slightly higher peaks for 2021. However, there is a clear discrepancy or divergence between years in the pattern of requests between January and March.

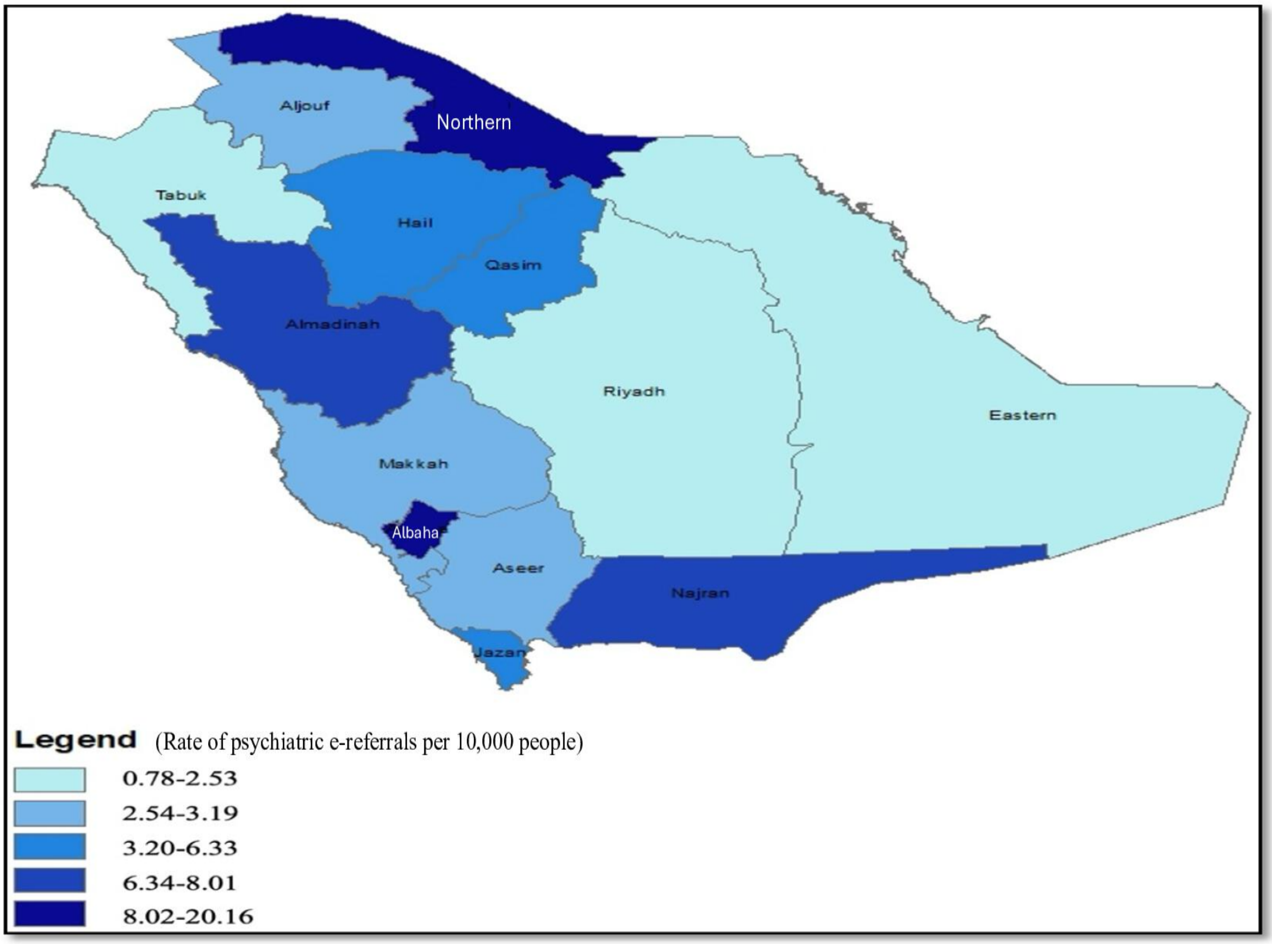
Figure 2. Monthly trend of psychiatric e-referrals in Saudi Arabia from January 2020 to December 2021, based on data routinely collected by the Saudi Medical Appointments and Referrals Centre (SMARC). This cross-sectional study analyzed data on 10,033 psychiatric e-referrals processed through the SMARC system across all regions of Saudi Arabia during this 2-year period. The line graph depicts the pattern of psychiatric e-referrals per month. e-Referral: electronic referral.



Geographic Patterns of Psychiatric e-Referral Requests Across Different Administrative Areas in Saudi Arabia

Figure 3 shows the administrative regions' rates of e-referral requests per 10,000 people. The Northern Border and Albaha regions had the highest rates and fell into the highest quantile. In contrast, the Madinah, Eastern, and Riyadh regions had the lowest rates of e-referral requests per 10,000 people. See Table S1 in Multimedia Appendix 1 for more details on the rates by region.

Figure 3. usGeographic distribution of psychiatric e-referral requests per 10,000 people across administrative regions of Saudi Arabia from January 2020 to December 2021, based on data from a cross-sectional study of 10,033 e-referrals. The map, developed using ArcGIS, depicts the rate of psychiatric e-referrals per 10,000 people in each region, with darker shades of blue indicating higher rates (upper quantiles) and lighter shades indicating lower rates (lower quantiles). The map highlights regional variations in the utilization of the Saudi Medical Appointments and Referrals Centre for psychiatric referrals during the study period. e-Referral: electronic referral.



Discussion

Principal Findings

This national study primarily aims to analyze psychiatric e-referral patterns across Saudi Arabia, while also providing an overview of the sociodemographic and clinical characteristics of mental health referrals across different BUs under the new health care transformation, emphasizing regional needs. The analysis is enabled by Saudi Arabia’s investments in advanced digital health systems, the nationwide e-referral platform SMARC, allowing examination of mental health referral trends at an unprecedented scale. Our findings indicate that e-referrals predominantly involved Saudi nationals, males, and working-age adults (aged 18 - 44 years), with most referrals from the Western BU and an increase in 2021 compared with 2020. Moreover, among these mental health referrals, emergency referrals were the most common e-referral type, ward beds being the most requested, and the essential cause was the lack of specialized providers.

Males were referred more frequently than females in the e-referral data. More than half of referrals were for males, highlighting a considerable gender imbalance that warrants further investigation into the underlying causes. The higher male referral rate may stem from men delaying or avoiding mental health treatment, leading to increased severity when finally referred, as untreated conditions may worsen over time [24,25]. This is supported by our findings where more emergency referrals were initiated for males compared with females, indicating the potential delay in seeking help (Table S2 in Multimedia Appendix 1). This delay in help-seeking among men has been linked with a higher risk of suicide [26], potentially explaining the higher referral rate. Broader literature supports this trend, reporting higher psychiatric emergency referrals and admission rates to psychiatric units for males compared with females [27,28].

The gender variation in referrals could also be attributed to differences in the prevalence of mental disorders. Males are more frequently associated with drug and substance use, disruptive disorders, schizophrenia, and a higher risk of suicide [29-33], often requiring more intensive resources. Conversely, females tend to experience higher rates of depression, anxiety, and somatic complaints [6,34], which can often be managed without referral. Studies in Saudi Arabia reflect similar patterns, with substance use and disruptive disorders more common in males, and anxiety and depressive disorders more prevalent among females [19,35].

These differences may have influenced the referral rates due to the availability of the required health care resources for each condition, since a referral is initiated when the needed resources such as psychiatry subspecialty are not available at the patient's initial point of care. Nevertheless, these findings do not assess the general mental health care access in Saudi Arabia. Despite limited data of mental health care access, two-thirds of patients with depression who accessed primary health care were females [36], suggesting that both males and females can access mental health care services. However, our data lack detailed diagnostic information, necessitating further research to confirm these speculations and develop evidence-based recommendations.

The peak referral rate among 18‐ to 44-year-olds indicates a need for youth-focused mental health services and preventive strategies. Previous research found that this age group has a high mental illness burden [37,38]. However, responsibilities, such as work-life balance, finances, and caregiving [39,40], may contribute to increased mental health issues as reflected in higher referral requests among this age group.

Monthly referral patterns showed a bimodal distribution with peaks in March and October (Figure 1). This pattern suggests that there may be specific factors driving increased demand for mental health services during these months. These trends are likely due to the COVID-19 pandemic, , as pandemic waves escalated in Saudi Arabia during these months [41]. Isolation, anxiety, and stress likely increased emotional health needs during the pandemic, while lockdowns reduced access to in-person mental health services [42]. Social distancing, job losses, and fear of contracting the virus may have increased depression and anxiety [43]. During pandemic peak periods, mental health facilities may have had to limit in-person appointments or close, delaying referral processes [12].

Regional variation in referrals warrants more research. The Western region had the highest proportion of referrals, possibly due to its higher population density and mental health resources versus other regions. The Western region has more psychiatric manpower and hospital beds per capita compared with the Central region. For example, the Western region has a slightly higher proportion of psychiatrists (2.88 per 100,000 people vs 2.68 per 100,000 people) and psychiatric beds (1.32 per 10,000 people vs 1.03 per 10,000 people) than the Central region (Table S1 in Multimedia Appendix 1). This greater capacity may enable more internal referrals. Distinguishing internal versus external referrals by region could elucidate the observed patterns.

Saudi Arabia National Health Care System

Saudi Arabia's health care system serves 34 million people across 13 regions through a mix of public and private facilities. Aligning with Vision 2030 reforms, the system aims to improve quality, efficiency, and value by integrating the public and private sectors [44,45]. This involves establishing 5 new BUs to manage the regions alongside national insurance companies, overseen by the MOH and new insurance centers.

Fortunately, under Vision 2030, the MOH is pioneering a new health care paradigm focused on enhancing social, mental, and physical well-being through a patient-centered model of care (MOC) [46]. Structured around 6 systems, including mental health within chronic care [47], the MOC seeks to tailor health care delivery to individual needs. The MOH is developing a national mental health strategy following the MOC's principles of patient-centered and integrated services starting at the community level [23]. This addresses care gaps and establishes accessible, quality services, supported by the new "National Committee for Mental Health Promotion" [48].

Applying e-Referral Findings to Enhance Mental Health Systems

International and national studies have highlighted the difficulty in measuring quality via health care utilization rates and the significant variation in service quality among countries, cities, hospitals, and even among BUs [44,49,50]. In our study, mental health referrals varied significantly across BUs by demographic and clinical characteristics. Referral rates for those younger than 18 years peaked in the Eastern BU and then Southern BU, ages 18‐44 years in the Western BU, and older adults in the Northern BU. Increasing child psychiatrists, youth mental health programs, school initiatives, and young adult services are suggested for these BUs [51]. Female referrals were slightly higher in the Eastern BU, mirroring the overall elevated referral rate for females found in a prior study examining total referral patterns across all health conditions, not just mental health specifically [44]. It is recommended to tailor gender-specific services for women and develop outreach programs to encourage support-seeking, especially for mental health issues associated with women such as pre- and postpartum mental issues [52,53] across all BU regions, particularly those with high referrals. These findings provide actionable evidence for decentralizing and customizing mental health services for local needs.

The Southern BU had double or triple the referrals due to unavailable machines compared with other BUs, aligning with a previous study showing that the Southern BU had the highest referral rate for all health care scopes [44]. However, in this study, the Central BU had the highest referrals due to unavailable beds. Referrals from unavailable services or psychiatrists were similar across BUs. OPD referrals were highest in the Central and Eastern BUs, while routine inpatient referrals were highest in the Northern BU, and emergency conditions referrals were highest in the Central and Southern BUs. To enhance the distribution of mental health resources, a system could be implemented to standardize essential services and psychiatrists across BUs [54]. Infrastructure investments, increased psychiatric beds, and expanded outpatient and inpatient services may be especially beneficial for BUs with high e-referrals rates.

The World Health Organization advocates integrating mental health into primary care to enhance access, reduce stigma, promote social integration, and build health care capacity for mental disorders [55]. Aligning with this, Saudi Arabia has expanded mental health training for primary care physicians and general practitioners to provide community-based services [56,57]. General practitioners in primary care settings typically address minor mental health disorders through counseling, while more complex cases require specialist care [23]. Ongoing training is needed to enhance care quality by mental health professionals. Nonetheless, primary care plays a key role in early detection of mental health issues and appropriate e-referral to specialized care based on condition severity and complexity [56].

Policy Implications

Our analysis revealed regional differences in Saudi Arabia's mental health e-referral patterns. We propose several policies to improve e-referral rates based on our findings. First, invest in preventive measures, including reducing stigma through awareness campaigns, to decrease the prevalence and severity of mental health issues. This can be achieved through education, counseling or support access, physical activity promotion, stress management programs, community networks, early intervention for at-risk groups, and encouraging early help-seeking before severity worsens, potentially decreasing late-stage e-referral [58-61]. Second, strengthening primary care providers' mental health training could enable earlier identification and management [62]. Integrated care models with mental health professionals also facilitate connections across levels, aligning with Saudi Arabia's new health clusters unifying primary, hospital, and specialty care [63,64]. Third, expand telehealth and e-mental health tools to increase access to mental health services [65,66]. Fourth, grow the qualified mental health workforce through education and training investments to expand capacity [67]. Overall, multifaceted strategies focused on training, integration, technology, capacity building, and stigma reduction could enhance Saudi Arabia's mental health e-referral system and outcomes.

It is important to note that this study was conducted during the peak of the COVID-19 pandemic, which may have influenced the observed referral patterns. While this timing introduces potential biases, it also provides a unique opportunity to understand mental health service utilization during a global crisis. Our findings serve as a crucial baseline for future comparisons with postpandemic periods. This research not only contributes to national policy making but also adds to the limited global literature on mental health service evaluations during extraordinary circumstances. Future studies should aim to compare these findings with postpandemic data to fully understand the long-term impacts of the crisis on mental health referral patterns and service delivery. Such comparisons will be invaluable in developing resilient mental health systems capable of adapting to both routine and crisis scenarios.

Limitations

This study's limitations stem from its reliance on secondary e-referral data, which lack detailed clinical information and comprehensive mental health infrastructure data across regions. This restricted our ability to conduct in-depth analyses of referral trend drivers and fully evaluate care access gaps. The study's timing during the peak of the COVID-19 pandemic may have introduced biases in referral patterns, potentially limiting the generalizability of findings to noncrisis periods. However, this unique context also offers valuable insights into mental health service needs and utilization during global emergencies. We acknowledge that e-referral data alone cannot capture the full complexity of mental health issues in Saudi Arabia, and the scarcity of comparable research in Saudi Arabia limited contextualization of our findings. To address these limitations, we propose future research directions including longitudinal studies incorporating detailed diagnostic information, cultural factors, and resource availability data. Future qualitative research could explore whether regional differences exist in mental health attitudes and help-seeking behaviors. Furthermore, integrating socioeconomic data may provide further insights into factors influencing health care utilization. Collaborative research involving religious scholars would provide valuable insights. Postpandemic research is also crucial to determine how referral patterns have evolved. These comprehensive approaches would not only address current study limitations but also offer a more nuanced understanding of mental health in Saudi Arabia, informing targeted interventions and policy development to improve mental health care delivery and access nationwide.

Conclusions

This nationwide study provides the first in-depth analysis specific to mental health e-referral patterns in Saudi Arabia, enabled by the large study sample of more than 10,033 referrals. Mental e-referrals predominantly involved Saudi nationals, males, and working-age adults (aged 18-44 years), with most referrals from the Western regions and a notable increase in 2021 compared with 2020. Emergency referrals were most common, with ward beds being the most requested due to a lack of mental health specialty services. Regional variations in e-referral patterns were observed. Given the widespread challenges in mental health care globally, these findings emphasize the importance to expand community-based services and implement preventive measures—challenges encountered by many health care systems worldwide. The SMARC e-referral system's ability to facilitate large-scale epidemiologic analysis, identify care patterns, and inform policy demonstrates the value of digital health platforms in improving coordination and transforming mental health care delivery. This study provides an exemplar that could serve as a model for other nations seeking to address mental health system challenges. This research not only informs national mental health policies but also contributes to the global understanding of mental health service delivery during extraordinary circumstances, paving the way for more resilient and adaptable health care systems.

Acknowledgments

The authors would like to thank all personnel involved in maintaining the Saudi Medical Appointments and Referrals Centre database. The authors attest that there was no use of generative artificial intelligence (AI) technology in the generation of text, figures, or other informational content of this manuscript. However, language editing features of AI were used in some parts of the manuscript to detect and correct grammatical and language mistakes.

Data Availability

The datasets analyzed during this study are not publicly available due to their sensitive nature as national governmental health data. However, aggregated or anonymized data may be available from the corresponding author on reasonable request, subject to approval from the relevant governmental authorities.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Analysis of psychiatric e-referral data from the Saudi Medical Appointments and Referrals Centre (SMARC) system, January 2020 to December 2021. Table S1 presents the proportion of 10,033 psychiatric e-referral requests by administrative area and business units in Saudi Arabia. It shows the number and rate of requests per 10,000 people across business units and regions, allowing comparison of e-referral utilization geographically. Table S2 displays the distribution of 10,033 psychiatric e-referral reasons categorized by sex. It enables comparison of referral causes between men and women. This cross-sectional study analyzed e-referral data routinely collected in the SMARC system. Data are presented as frequency (n [%]).
[DOCX File, 19 KB - publichealth_v11i1e64257_app1.docx]

References

1. Steel Z, Marnane C, Iranpour C, et al. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. Int J Epidemiol 2014 Apr;43(2):476-493. [doi: 10.1093/ije/dyu038] [Medline: 24648481]
2. Alghadeer SM, Alhossan AM, Al-Arifi MN, et al. Prevalence of mental disorders among patients attending primary health care centers in the capital of Saudi Arabia. Neurosciences (Riyadh) 2018 Jul;23(3):238-243. [doi: 10.17712/nsj.2018.3.20180058]
3. WHO Mental Health Gap Action Programme (mhGAP): World Health Organization; 2012.
4. Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. Lancet 2013 Nov;382(9904):1575-1586. [doi: 10.1016/S0140-6736(13)61611-6]
5. Murray CJL, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec;380(9859):2197-2223. [doi: 10.1016/S0140-6736(12)61690-0]
6. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150. [doi: 10.1016/S2215-0366(21)00395-3] [Medline: 35026139]
7. Nishi D, Ishikawa H, Kawakami N. Prevalence of mental disorders and mental health service use in Japan. Psychiatry Clin Neurosci 2019 Aug;73(8):458-465. [doi: 10.1111/pcn.12894] [Medline: 31141260]
8. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866. [Medline: 15640922]
9. Wainberg ML, Scorza P, Shultz JM, et al. Challenges and opportunities in global mental health: a research-to-practice perspective. Curr Psychiatry Rep 2017 May;19(5):1-10. [doi: 10.1007/s11920-017-0780-z] [Medline: 28425023]
10. Ibrahim NK. Epidemiology of mental health problems in the middle east. In: Handbook of Healthcare in the Arab World: Springer Cham; 2021:133-149. [doi: 10.1007/978-3-030-36811-1_12]
11. Elshamy F, Hamadeh A, Billings J, Alyafei A. Mental illness and help-seeking behaviours among Middle Eastern cultures: a systematic review and meta-synthesis of qualitative data. PLoS One 2023;18(10):e0293525. [doi: 10.1371/journal.pone.0293525] [Medline: 37883515]
12. Duden GS, Gersdorf S, Stengler K. Global impact of the COVID-19 pandemic on mental health services: a systematic review. J Psychiatr Res 2022 Oct;154:354-377. [doi: 10.1016/j.jpsychires.2022.08.013] [Medline: 36055116]
13. Toro-Devia O, Leyton G. COVID-19 pandemic and mental healthcare: impact on health insurance with guaranteed universal access in Chile. Front Public Health 2022;10:1005033. [doi: 10.3389/fpubh.2022.1005033] [Medline: 36684988]
14. Santomauro DF, Mantilla Herrera AM, Shadid J, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet 2021 Nov;398(10312):1700-1712. [doi: 10.1016/S0140-6736(21)02143-7]
15. Iorfino F, Piper SE, Prodan A, et al. Using digital technologies to facilitate care coordination between youth mental health services: a guide for implementation. Front Health Serv 2021;1:745456. [doi: 10.3389/frhs.2021.745456] [Medline: 36926493]
16. Azamar-Alonso A, Costa AP, Huebner LA, Tarride JE. Electronic referral systems in health care: a scoping review. Clinicoecon Outcomes Res 2019;11:325-333. [doi: 10.2147/CEOR.S195597] [Medline: 31190925]

17. Aljerian NA, Alharbi AA, Alghamdi HA, et al. External vs internal e-referrals: results from a nationwide epidemiological study utilizing secondary collected data. Risk Manag Healthc Policy 2024;17:739-751. [doi: 10.2147/RMHP.S453042] [Medline: 38562249]
18. Alnassar A, Aljerian N, Alhosaini A, et al. Trends of referrals throughout the Kingdom, a retrospective analysis of the Saudi Medical Appointments and Referrals Centre Registry, Saudi Arabia. IJIRMS 2022;7(11):563-578. [doi: 10.23958/ijirms/vol07-i11/1517]
19. Altwaijri YA, Al-Subaie AS, Al-Habeeb A, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the Saudi National Mental Health Survey. Int J Methods Psychiatr Res 2020 Sep;29(3):e1836. [doi: 10.1002/mpr.1836] [Medline: 33245595]
20. Saudi Arabia. Institute for Health Metrics and Evaluation. URL: https://www.healthdata.org/research-analysis/health-by-location/profiles/saudi-arabia [accessed 2024-12-30]
21. Alangari AS, Knox SS, Kristjansson AL, et al. Barriers to mental health treatment in the Saudi National Mental Health Survey. Int J Environ Res Public Health 2020 May 30;17(11):3877. [doi: 10.3390/ijerph17113877] [Medline: 32486182]
22. Alattar N, Felton A, Stickley T. Mental health and stigma in Saudi Arabia: a scoping review. Mental Health Rev J 2021 May 27;26(2):180-196. [doi: 10.1108/MHRJ-08-2020-0055]
23. Al-Subaie AS, Al-Habeeb A, Altwaijri YA. Overview of the Saudi National Mental Health Survey. Int J Methods Psychiatr Res 2020 Sep;29(3):e1835. [doi: 10.1002/mpr.1835] [Medline: 33245612]
24. Picco L, Abdin E, Pang S, et al. Association between recognition and help-seeking preferences and stigma towards people with mental illness. Epidemiol Psychiatr Sci 2018 Feb;27(1):84-93. [doi: 10.1017/S2045796016000998] [Medline: 27927259]
25. Holzinger A, Floris F, Schomerus G, Carta MG, Angermeyer MC. Gender differences in public beliefs and attitudes about mental disorder in western countries: a systematic review of population studies. Epidemiol Psychiatr Sci 2012 Mar;21(1):73-85. [doi: 10.1017/s2045796011000552] [Medline: 22670415]
26. Sagar-Ouriaghli I, Godfrey E, Bridge L, Meade L, Brown JSL. Improving mental health service utilization among men: a systematic review and synthesis of behavior change techniques within interventions targeting help-seeking. Am J Mens Health 2019;13(3):1557988319857009. [doi: 10.1177/1557988319857009] [Medline: 31184251]
27. Singh G, Chaudhury S, Saldanha D, Singh V, Marella S, Vhora R. Psychiatric emergency referrals in a tertiary care hospital. Med J DY Patil Vidyapeeth 2018;11(4):312. [doi: 10.4103/MJDRDYPU.MJDRDYPU_180_17]
28. Aflalo M, Soucy N, Xue X, Colacone A, Jourdenais E, Boivin JF. Characteristics and needs of psychiatric patients with prolonged hospital stay. Can J Psychiatry 2015 Apr;60(4):181-188. [doi: 10.1177/070674371506000405]
29. Degenhardt L, Bharat C, Glantz MD, et al. The epidemiology of drug use disorders cross-nationally: findings from the WHO's World Mental Health Surveys. Int J Drug Policy 2019 Sep;71:103-112. [doi: 10.1016/j.drugpo.2019.03.002] [Medline: 31255918]
30. Fayyad J, Sampson NA, Hwang I, et al. The descriptive epidemiology of DSM-IV Adult ADHD in the World Health Organization World Mental Health Surveys. Atten Defic Hyperact Disord 2017 Mar;9(1):47-65. [doi: 10.1007/s12402-016-0208-3] [Medline: 27866355]
31. Scott KM, Lim CCW, Hwang I, et al. The cross-national epidemiology of DSM-IV intermittent explosive disorder. Psychol Med 2016 Nov;46(15):3161-3172. [doi: 10.1017/S0033291716001859] [Medline: 27572872]
32. Suanrueang P, Peltzer K, Suen MW, Lin HF, Er TK. Trends and gender differences in mental disorders in hospitalized patients in Thailand. Inquiry 2022;59:469580221092827. [doi: 10.1177/00469580221092827] [Medline: 35420043]
33. Schneider B, Lukaschek K, Baumert J, Meisinger C, Erazo N, Ladwig KH. Living alone, obesity, and smoking increase risk for suicide independently of depressive mood findings from the population-based MONICA/KORA Augsburg cohort study. J Affect Disord 2014 Jan;152-154:416-421. [doi: 10.1016/j.jad.2013.10.007] [Medline: 24238953]
34. Hamdan A. Mental health needs of Arab women. Health Care Women Int 2009 Jul;30(7):595-613. [doi: 10.1080/07399330902928808] [Medline: 19492205]
35. Altwaijri Y, Kazdin AE, Al-Subaie A, et al. Lifetime prevalence and treatment of mental disorders in Saudi youth and adolescents. Sci Rep 2023 Apr 15;13(1):6186. [doi: 10.1038/s41598-023-33005-5] [Medline: 37061556]
36. Becker SM. Detection of somatization and depression in primary care in Saudi Arabia. Soc Psychiatry Psychiatr Epidemiol 2004 Dec;39(12):962-966. [doi: 10.1007/s00127-004-0835-4] [Medline: 15583903]
37. Askari MS, Mauro PM, Kaur N, Keyes KM. Age, period, and cohort trends in perceived mental health treatment need and differences by mental health severity in the United States, 2008-2019. Community Ment Health J 2023 May;59(4):631-640. [doi: 10.1007/s10597-022-01044-3] [Medline: 36305995]
38. Höglund P, Hakelind C, Nordin S. Severity and prevalence of various types of mental ill-health in a general adult population: age and sex differences. BMC Psychiatry 2020 May 11;20(1):209. [doi: 10.1186/s12888-020-02557-5] [Medline: 32393209]
39. Finegold D, Mohrman S, Spreitzer GM. Age effects on the predictors of technical workers' commitment and willingness to turnover. J Organ Behav 2002 Aug;23(5):655-674. [doi: 10.1002/job.159]
40. Darcy C, McCarthy A, Hill J, Grady G. Work-life balance: one size fits all? An exploratory analysis of the differential effects of career stage. Eur Manag J 2012 Apr;30(2):111-120. [doi: 10.1016/j.emj.2011.11.001]
41. Salam AA, Al-Khraif RM, Elsegaey I. COVID-19 in Saudi Arabia: an overview. Front Public Health 2021;9:736942. [doi: 10.3389/fpubh.2021.736942] [Medline: 35186861]

42. Jeffers A, Meehan AA, Barker J, et al. Impact of social isolation during the COVID-19 pandemic on mental health, substance use, and homelessness: qualitative interviews with behavioral health providers. Int J Environ Res Public Health 2022 Sep 25;19(19):12120. [doi: 10.3390/ijerph191912120] [Medline: 36231422]
43. Spencer-Laitt D, Eustis EH, Barlow DH, Farchione TJ. The impact of COVID-19 related social distancing on mental health outcomes: a transdiagnostic account. Int J Environ Res Public Health 2022 May 28;19(11):6596. [doi: 10.3390/ijerph19116596] [Medline: 35682179]
44. Aljerian NA, Alharbi AA, AlOmar RS, et al. Showcasing the Saudi e-referral system experience: the epidemiology and pattern of referrals utilising nationwide secondary data. Front Med (Lausanne) 2024;11:1348442. [doi: 10.3389/fmed.2024.1348442] [Medline: 38994343]
45. Alharbi AA, Alqassim AY, Muaddi MA, Alghamdi SS. Regional differences in COVID-19 mortality rates in the Kingdom of Saudi Arabia: a simulation of the new model of care. Cureus 2021 Dec;13(12):e20797. [doi: 10.7759/cureus.20797] [Medline: 34987945]
46. Chowdhury S, Mok D, Leenen L. Transformation of health care and the new model of care in Saudi Arabia: Kingdom’s Vision 2030. J Med Life 2021;14(3):347-354. [doi: 10.25122/jml-2021-0070] [Medline: 34377200]
47. Health sector transformation strategy. Kingdom of Saudi Arabia - Ministry of Health. URL: https://www.moh.gov.sa/en/Ministry/vro/Documents/Healthcare-Transformation-Strategy.pdf [accessed 2024-12-30]
48. National Center for Mental Health Promotion. URL: https://ncmh.org.sa/view/16/3th [accessed 2024-11-30]
49. Corallo AN, Croxford R, Goodman DC, Bryan EL, Srivastava D, Stukel TA. A systematic review of medical practice variation in OECD countries. Health Policy 2014 Jan;114(1):5-14. [doi: 10.1016/j.healthpol.2013.08.002] [Medline: 24054709]
50. Alharbi AA, Alqassim AY, Gosadi IM, et al. Regional differences in COVID-19 ICU admission rates in the Kingdom of Saudi Arabia: a simulation of the new model of care under vision 2030. J Infect Public Health 2021 Jun;14(6):717-723. [doi: 10.1016/j.jiph.2021.04.012] [Medline: 34020211]
51. Burns J, Birrell E. Enhancing early engagement with mental health services by young people. Psychol Res Behav Manag 2014;7:303-312. [doi: 10.2147/PRBM.S49151] [Medline: 25473320]
52. Chrzan-Dętkoś M, Walczak-Kozłowska T, Lipowska M. The need for additional mental health support for women in the postpartum period in the times of epidemic crisis. BMC Pregnancy Childbirth 2021 Feb 8;21(1):114. [doi: 10.1186/s12884-021-03544-8] [Medline: 33557768]
53. Waqas A, Koukab A, Meraj H, et al. Screening programs for common maternal mental health disorders among perinatal women: report of the systematic review of evidence. BMC Psychiatry 2022 Jan 24;22(1):54. [doi: 10.1186/s12888-022-03694-9] [Medline: 35073867]
54. Conceição MA, Monteiro MM, Kasraian D, et al. The effect of transport infrastructure, congestion and reliability on mental wellbeing: a systematic review of empirical studies. Transp Rev 2023 Mar 4;43(2):264-302. [doi: 10.1080/01441647.2022.2100943]
55. Greenhalgh T. WHO/WONCA report—integrating mental health in primary care: a global perspective. London J Prim Care (Abingdon) 2009;2(1):81-82. [doi: 10.1080/17571472.2009.11493254] [Medline: 26042178]
56. Al-Habeeb A, Helmi B, Qureshi N. Mental and social health atlas: an update, Ministry of Health, Saudi Arabia, 2015. Int Neuropsychiatric Dis J 2015;6(3):1-20. [doi: 10.9734/INDJ/2016/23002]
57. Koenig HG, Al Zaben F, Sehlo MG, et al. Mental health care in Saudi Arabia: past, present and future. Open J Psychiatry 2014;04(02):113-130. [doi: 10.4236/ojpsych.2014.42016]
58. Wahlbeck K. Public mental health: the time is ripe for translation of evidence into practice. World Psychiatry 2015 Feb;14(1):36-42. [doi: 10.1002/wps.20178] [Medline: 25655149]
59. Arango C, Díaz-Caneja CM, McGorry PD, et al. Preventive strategies for mental health. Lancet Psychiatry 2018 Jul;5(7):591-604. [doi: 10.1016/S2215-0366(18)30057-9] [Medline: 29773478]
60. Castillo EG, Ijadi-Maghsoodi R, Shadravan S, et al. Community interventions to promote mental health and social equity. Curr Psychiatry Rep 2019 Mar 29;21(5):35. [doi: 10.1007/s11920-019-1017-0] [Medline: 30927093]
61. Costello EJ. Early detection and prevention of mental health problems: developmental epidemiology and systems of support. J Clin Child Adolesc Psychol 2016 Nov;45(6):710-717. [doi: 10.1080/15374416.2016.1236728]
62. Funk M, Saraceno B, Drew N, Faydi E. Integrating mental health into primary healthcare. Ment Health Fam Med 2008 Mar;5(1):5-8. [Medline: 22477840]
63. Khan K. Saudi Arabia launches health clusters. International Trade Administration. 2022 Jul 25. URL: https://www.trade.gov/market-intelligence/saudi-arabia-launches-health-clusters [accessed 2024-12-30]
64. Alharbi AA, Muhayya M, Alkhudairy R, et al. The pattern of emergency department length of stay in Saudi Arabia: an epidemiological nationwide analyses of secondary surveillance data. Front Public Health 2023;11:1265707. [doi: 10.3389/fpubh.2023.1265707] [Medline: 38162606]
65. Mulvaney-Day N, Dean D, Miller K, Camacho-Cook J. Trends in use of telehealth for behavioral health care during the COVID-19 pandemic: considerations for payers and employers. Am J Health Promot 2022;36(7):1237-1241. [doi: 10.1177/08901171221112488e]

66. Lui JHL, Marcus DK, Barry CT. Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. Prof Psychol Res Pract 2017;48(3):199-210. [doi: 10.1037/pro0000122]
67. Kakuma R, Minas H, van Ginneken N, et al. Human resources for mental health care: current situation and strategies for action. Lancet 2011 Nov;378(9803):1654-1663. [doi: 10.1016/S0140-6736(11)61093-3]

Abbreviations

BU: business unit
e-referral: electronic referral
KSA: Kingdom of Saudi Arabia
MOC: model of care
MOH: Ministry of Health
OPD: outpatient department
SMARC: Saudi Medical Appointments and Referrals Centre

Edited by A Mavragani; submitted 12.07.24; peer-reviewed by F Verbeke, M Sengupta; revised version received 21.10.24; accepted 28.10.24; published 24.01.25.

Please cite as:
Alharbi AA, Aljerian NA, Binhotan MS, Alghamdi HA, Alsultan AK, Arafat MS, Aldhabib A, Alaska YA, Alwahbi EB, Muaddi MA, Alqassim AY, Horner RD
Digital Surveillance of Mental Health Care Services in Saudi Arabia: Cross-Sectional Study of National e-Referral System Data
JMIR Public Health Surveill 2025;11:e64257
URL: https://publichealth.jmir.org/2025/1/e64257
doi:10.2196/64257

Assessing the Impact of Distance Traveled and Birth Volumes of Hospital Maternity Units on Newborn Outcomes: Population-Based Cohort Study

Anna Cantarutti[1*], PhD; Riccardo Boracchini[1*], MSc; Roberto Bellù[2], MD; Raffaella Ronco[1], MSc; Federico Rea[1], PhD; Anna Locatelli[3], MD; Rinaldo Zanini[4], MD; Giovanni Corrao[1], PhD

[1]National Centre for Healthcare Research & Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy
[2]Neonatal Intensive Care Unit, Woman and Child Health Department, Azienda Ospedaliera della Provincia di Lecco, Lecco, Italy
[3]Department of Obstetrics, IRCCS San Gerardo dei Tintori Foundation, University of Milano-Bicocca, Monza, Italy
[4]General Management, IRCCS Istituto Giannina Gaslini, Genova, Italy
[*]these authors contributed equally

Corresponding Author:
Anna Cantarutti, PhD
National Centre for Healthcare Research & Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy

Abstract

Background: The centralization of childbirth and newborn care in large maternity units has become increasingly prevalent in Europe. While this trend offers potential benefits such as specialized care and improved outcomes, it can also lead to longer travel and waiting times, especially for women in rural areas.

Objective: This study aimed to evaluate the association between hospital maternity unit (HMU) volumes, road travel distance (RTD) to the hospital, and other neonatal outcomes.

Methods: We conducted a population-based cohort study including all live births in hospitals without intensive care units between 2016 and 2019 in the Lombardy region, Italy. Given the hierarchical structure of our data (births nested within hospitals), we employed log-binomial regression models with random intercepts to estimate relative risks and 95% CIs for evaluating the association between HMU volumes (≥1500 births/year) and RTD (<5 km) with the risk of being transferred and/or death after birth (primary outcome). Secondary outcomes included a low Apgar score at 5 minutes and low adherence to antenatal care (ANC). We controlled for several potential confounders including adherence to the ANC pathway for the primary and low Apgar outcomes. To explore the influence of HMU volumes on the primary outcome, we identified the fractional polynomial model that best described this relationship.

Results: Of 65,083 live births, 71% (n=45,955) occurred in low-volume hospitals (<1000 births/year), 21% (n=13,560) involved long-distance travel (>15 km), 1% (n=735) were transferred and/or died after birth, 0.5% (n=305) had a low Apgar score at 5 minutes, and 64% (n=41,317) completely adhered to ANC. The risk of transfer and/or death increased as HMU volume decreased, ranging from 1% for hospitals with 1000 - 1500 births/year to a 3.6-fold high risk for hospitals with <500 births/year (compared to high-volume hospitals). Travel distance did not affect the primary outcome. Neither HMU volume nor RTD were associated with low Apgar scores. Conversely, the risk of complete adherence to ANC decreased with lower HMU volumes but increased with shorter RTD. Additionally, high-volume hospitals demonstrated a decreasing trend in the frequency of the primary outcome, with transfer and/or death rates ranging from 2% to 0.5% and flattening to 0.5% in hospitals, with activity volume ≥1000 mean births/year.

Conclusions: Our findings showed an excess risk of neonatal transfer and/or death for live births in HMUs with low activity volumes without an intensive care unit. In contrast, RTD primarily affected adherence to ANC. Moreover, data suggest that 1000 births/year could be an optimal cutoff for maternity hospitals to ensure an appropriate standard of care at delivery.

(JMIR Public Health Surveill 2025;11:e58944) doi:10.2196/58944

KEYWORDS

maternity level; road travel distance; delivery; regionalization; neonatal outcome; childbirth; newborn; cohort; birth; population-based; maternal; maternal outcomes; neonatal; European; health system; health care system; perinatal care; antenatal; mortality; neonatal mortality; perinatal

Introduction

Driven by a commitment to improving maternal and neonatal outcomes, several European health care systems, including the Italian National Health Service (NHS) have embraced the regionalization of perinatal care in large maternity units since the 1980s. Perinatal regionalization aims to optimize access to quality care by organizing maternity and neonatal services into distinct levels based on the complexity of care required. This approach involves (1) ensuring that pregnant women and newborns are directed to the appropriate level of care based on their needs, (2) implementing early screening and assessment mechanisms to identify women at risk, and (3) facilitating timely and appropriate transfers between levels of care when necessary. By implementing these strategies, perinatal regionalization can improve outcomes for both mothers and infants while ensuring the efficient use of health care resources [1-4].

While perinatal regionalization offers potential benefits, its implementation remains debated. Robust, real-world evidence is needed to inform decision-makers about both the advantages of centralization (eg, reduced travel time, specialized care) and the potential disadvantages (eg, increased travel distance for some women). Moreover, only a few studies have comprehensively examined the combined effects of hospital maternity unit (HMU) volume and road travel distance (RTD) on maternal and neonatal outcomes [5-7]. However, the centralization of HMUs inevitably leads to the closure of some units and increases the distance (and travel time) to the hospital for some mothers.

Regionalization of perinatal care in HMUs with at least 1000 births/year has been undertaken since 2010 in Italy, although some maternity units continue to record fewer than 500 births/year. Despite the heterogenous geographical conformation of Italy, the distances between maternity units are relatively modest, compared to those reported by North American studies [8]. This finding combined with the ongoing efforts to regionalize HMU volumes to units with at least 1000 births/year, makes Italy an ideal setting to investigate the relationship between HMU volume, travel distance, and perinatal outcomes.

In this study, we conducted a population-based cohort study in Lombardy, the largest and most populous region of Italy, to investigate the singular and combined effects of HMU volumes and RTD on specific outcomes in births occurring in hospitals without intensive care units (ICUs). Additionally, we studied the influence of perinatal factors like antenatal care (ANC) and sociodemographic features of the mothers.

Methods

Data Source and Study Cohort

The study cohort consisted of all live infants born in Lombardy from January 1, 2016, to December 31, 2019. Lombardy is a region in Italy that accounts for approximately 16% of the country's population, comprising nearly 10 million inhabitants. The health care utilization of all residents of Lombardy is covered by the government-funded NHS, which employs an automated system of databases to collect a variety of information. This system includes demographic and administrative data for all beneficiaries of the Regional Health Service (approximately covering the entire resident population), such as residence municipalities. Additional databases include the hospital discharges registry, which records all patients discharged from public or private hospitals; the outpatient drug prescription registry that reports all dispensations of NHS-reimbursable drugs; and the specialist visits and diagnostic exams registry including a specific automated system that collects data from the regional Department of Mental Health, accredited by the NHS and focused on outpatient specialist mental health care. Lastly, the Certificates of Delivery Assistance provide detailed information about pregnancy, childbirth, and fetal presentation at delivery. A unique deidentification code is systematically used for all databases; as a result, linking these records enables the creation of a large birth cohort and establishing relevant traits and care pathways for mothers and newborns.

The criteria for selecting the study cohort almost completely overlapped with those previously reported by our group [9]. Briefly, using the Certificates of Delivery Assistance database, we identified all live births in Lombardy between 2016 and 2019 and available identification codes to women who met the inclusion criteria: (1) were beneficiaries of the NHS and had been residents of Lombardy for at least one year before pregnancy, (2) were aged 15 to 55 years at delivery, and (3) delivered between 22 to 42 weeks of gestation, based on the first day of the last menstrual period ascertained via maternal reports or ultrasonography. All births recorded in maternity hospitals equipped with an ICU were excluded. Further, records with incomplete data were excluded because baseline covariates such as sociodemographic and gestational information may be missing for some women, and limiting analyses to the subset of women for whom complete data were available would not result in a significant loss of information (Figure 1).

Figure 1. Flowchart of inclusion and exclusion criteria. NHS: National Health Service.

The results were reported in agreement with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement guidelines [10].

Categorization of HMU and RTD

The HMU volumes for each hospital were obtained based on the mean number of births recorded between 2016 and 2019. Categories were defined based on the activity volume of the hospital in which the birth occurred (ie, volumes ≤500 births/year, 500 - 1000 births/year, 1000-1500 births/year, and >1500 births/year).

The RTD between the hospital and the mother's residence municipalities was calculated according to the distance matrix (ie, an open-source distance matrix database provided by the Italian Institute of Statistics) [11]. Births were classified based on the mother having traveled <5 km, 5 - 15 km, 15 - 21.5 km, or ≥21.5 km to the HMU.

Outcomes

The primary outcome of interest was transfer to a different hospital or death of the infant during the same birth hospitalization. The secondary outcomes included (1) a low Apgar score (<7) at 5 minutes after birth [12] and (2) complete adherence of the mothers to ANC during pregnancy. Maternal ANC adherence was evaluated based on the promptness and appropriateness of the number and timing of ANC interventions, including prenatal visits, ultrasound examinations, and laboratory tests, in relation to the length of pregnancy. In particular, the ANC assessment included (1) appropriateness of prenatal visits (ie, at least 5 visits during pregnancy or 2-4 visits for women with a gestational duration of <27 weeks); (2) promptness of prenatal visits (ie, at least 1 of the 4 visits within the 12th week of gestation); (3) appropriateness of ultrasound examinations (ie, ≥2 examinations during pregnancy, with at least 1 within the 12th week of gestation); and (4) appropriateness of laboratory tests, defined as completion of recommended tests for each trimester [13]. Women were identified as treatment-adherent if they followed all 4 of the recommendations.

Definitions of the Covariates

Several baseline maternal characteristics were considered. Sociodemographic features included nationality, marital status, employment, and educational attainment (ie, low: ≤5 years, medium: 6 - 13 years, and high: ≥14 years of study corresponding to primary education or none, lower or upper secondary education, and at least a bachelor's degree, respectively). Clinical information was obtained from the inpatient hospital registry, the outpatient drug prescriptions registry, and the specialist visits and diagnostic exams registry. Maternal comorbidities, including psychiatric comorbidities, medical comorbidities, obstetric characteristics, and health care utilization measures (ie, number of hospitalizations and distinct prescription drugs used as general markers of comorbidity) [14] were evaluated. Concomitant medications (ie, nonsteroidal anti-inflammatory drugs, antidepressants, antiepileptics, psycholeptics, antihypertensive drugs, and antidiabetic drugs) were measured from 1 year before the last menstrual period through the 22nd week of gestation. Information regarding the ANC of mothers was measured during pregnancy.

Ethical Considerations

According to the Italian Medicines Agency guidelines [15], retrospective studies that do not involve direct contact with patients do not need written consent to process personal data when they are used for research purposes. Thus, no ethics board review or approval was required.

To preserve privacy, each identification code was automatically anonymized, with only the Regional Health Authority having access to identifying information, which may be released upon request from judicial authorities to guarantee data deidentification.

Statistical Analysis

The baseline maternal characteristics and distribution of the ANC received by mothers during pregnancy were described overall and stratified by the primary outcome. The *t* test, χ^2 test, or Fisher exact test were used as appropriate. Moreover, the distribution of outcomes was expressed as absolute and relative percentages across HMU volumes and RTD categories. The Cochran-Armitage trend test was implemented for trend analysis.

As our data had a multilevel structure—with births (level 1) nested within hospitals (level 2)—log-binomial regression models with random intercepts were used to estimate relative risks (RRs) with 95% CIs for associations between HMU volumes (reference: ≥1500 births/year) and RTD (reference: RTD<5 km) and other outcomes of interest. The interaction between HMU volumes and RTD was also assessed. Models were adjusted for a disease score that was calculated based on the primary outcome, including all covariates listed above. For analyses in which ANC adherence was the outcome of interest, we excluded variables related to the ANC pathway from the disease score calculations.

A stratified analysis was performed by HMU volume, categorized as low (<1000 births/year) and high (≥1000 births/year) HMU volumes. Further, RTD was categorized as short (<15 km) and long RTD (≥15 km). Finally, to examine the influence of HMU volumes on the primary outcome (ie, transfer to a different hospital or death), a fractional polynomial model (FPM) was applied (first, second, or third degree), which described the relationship through a stepwise selection [16]. FPMs are particularly useful to preserve the continuous nature of the covariates in regression models while accounting for potential nonlinear relationships. Hospitals were the statistical units for this analysis.

All analyses were performed using the SAS (version 9.4; SAS Institute). *P* values <.05 were considered statistically significant.

Results

Among the 175,366 live births recorded in Lombardy between January 1, 2016, and December 31, 2019, 30,730 births were excluded because the mothers had not received health care from the NHS for at least 1 year before pregnancy, while 9 were excluded because the mothers were younger than 15 or older than 55 years. Furthermore, 216 births were excluded because the gestational age at delivery was outside the range of 22 to 42 weeks, 75,395 births were excluded because they occurred in hospitals with an ICU, and 3933 births were excluded because complete sets of information about the mothers were not available (Figure 1). The final cohort consisted of 65,083 live births, of which 735 (1.13%) newborns were either transferred to another hospital (n=731) or died (n=4), 305 (0.47%) had a low Apgar score at 5 minutes after birth, and 41,317 (63.5%) fully adhered to ANC.

Mothers of newborns included in this cohort had a mean age of 32 (SD 5.4) years and mean gestational duration of 39 (SD 1.4) weeks. Most mothers had moderate educational attainment, were employed and married at delivery, and were of Italian nationality, and 55% (35,799/65,083) had already given birth. Maternal comorbidities were generally less prevalent and did not significantly differ across groups, except for psychoses and substance dependence, which were more common among mothers whose newborns experienced hospital transfer or death. Additionally, mothers in the transfer or death group showed higher rates of antiepileptic and antidiabetic drug use (Table 1).

Table . Sociodemographic, clinical, and pregnancy characteristics of mothers in Lombardy from 2016 - 2019.

Cohort characteristics	Live births (N=65,083)	Individuals not transferred and/or experienced death (n=64,348)	Individuals transferred and/or experienced death (n=735)	*P* value[a]
Sociodemographic characteristics[b]				
Age (years), mean (SD)	32.27 (5.37)	32.27 (5.36)	32.17 (5.49)	.64
Educational attainment, years of study[c], n (%)				<.001
≥14	19,036 (29.25)	18,859 (29.31)	177 (24.08)	
6 - 13	29,720 (45.66)	29,385 (45.67)	335 (45.58)	
≤5	16,327 (25.09)	16,104 (25.03)	223 (30.34)	
Employed, n (%)				<.001
No	22,832 (35.08)	41,835 (65.01)	416 (56.60)	
Yes	42,251 (64.92)	22,513 (34.99)	319 (43.40)	
Married, n (%)				.99
No	24,695 (37.94)	39,932 (62.06)	456 (62.04)	
Yes	40,388 (62.06)	24,4416 (37.94)	279 (37.96)	
Italian nationality, n (%)				.01
No	18,109 (27.82)	46,474 (72.22)	500 (68.03)	
Yes	46,974 (72.18)	17,874 (27.78)	235 (31.97)	
Maternal comorbidities[d], n (%)				
Depression and anxiety	394 (0.61)	387 (0.60)	7 (0.95)	.22
Preeclampsia	44 (0.07)	43 (0.07)	1 (0.14)	.39
Hypertension	90 (0.14)	87 (0.14)	3 (0.41)	.08
Diabetes	219 (0.34)	214 (0.33)	5 (0.68)	.10
Obesity or overweight	70 (0.11)	69 (0.11)	1 (0.14)	.55
Psychoses[e]	120 (0.18)	116 (0.18)	4 (0.54)	.048
Neuropathic, nonneuropathic, and other pain	140 (0.22)	139 (0.22)	1 (0.14)	>.99
Substance dependence	21 (0.03)	19 (0.03)	2 (0.27)	.02
Concomitant medication, n (%)				
Antidepressants	1395 (2.14)	1375 (2.14)	20 (2.72)	.28
Psycholeptics	159 (0.24)	155 (0.24)	4 (0.54)	.11
Antiepileptics	426 (0.65)	416 (0.65)	10 (1.36)	.02
Antihypertensive drugs	824 (1.27)	811 (1.26)	13 (1.77)	.22
Antidiabetic drugs	751 (1.15)	729 (1.13)	22 (2.99)	<.001
NSAIDs[f]	2440 (3.75)	2406 (3.74)	34 (4.63)	.21
Health care utilization, n (%)				
Hospitalization	8033 (12.34)	7935 (12.33)	98 (13.33)	.41
Number of distinct prescription drugs (≥1)	48,386 (74.35)	47,827 (74.33)	559 (76.05)	.29
Pregnancy characteristics[b], n (%)				
Gestational duration (weeks), mean (SD)	38.99 (1.44)	39.01 (1.41)	37.33 (3)	<.001
Multiple pregnancy	877 (1.35)	846 (1.31)	31 (4.22)	<.001
Parity				<.001

Cohort characteristics	Live births (N=65,083)	Individuals not transferred and/or experienced death (n=64,348)	Individuals transferred and/or experienced death (n=735)	*P* value[a]
Nulliparous	29,284 (44.99)	28,888 (44.89)	384 (52.24)	
Others	35,799 (55.01)	35,460 (55.11)	351 (47.76)	
Secondary outcomes, n (%)				
Low Apgar score at 5 minutes	305 (0.47)	205 (0.32)	100 (13.61)	<.001
Complete adherence to ANC	41,317 (63.48)	40,947 (63.63)	370 (50.34)	<.001
Antenatal maternal care[g]				
Appropriateness of gynecological visits	58,341 (89.64)	57,709 (89.68)	632 (85.99)	.001
Promptness of gynecological visits	61,338 (94.25)	60,650 (94.25)	688 (93.61)	.45
Appropriateness of ultrasound examinations	60,997 (93.72)	60,317 (93.74)	680 (95.52)	.18
Appropriateness of laboratory tests, n (%)				<.001
Complete adherence	46,718 (71.78)	46,298 (71.95)	420 (57.14)	
Partial adherence	16,622 (25.54)	16,353 (25.41)	269 (36.60)	
Nonadherence	1743 (2.68)	1697 (2.64)	46 (6.26)	

[a] *t* test, Chi-square, or Fisher exact test, as appropriate.

[b] Data related to the current pregnancy.

[c] ≥14: at least bachelor's degree; 6 - 13: lower or upper secondary education; ≤5: primary education or none.

[d] Measured from one year before pregnancy through 22 weeks of gestation.

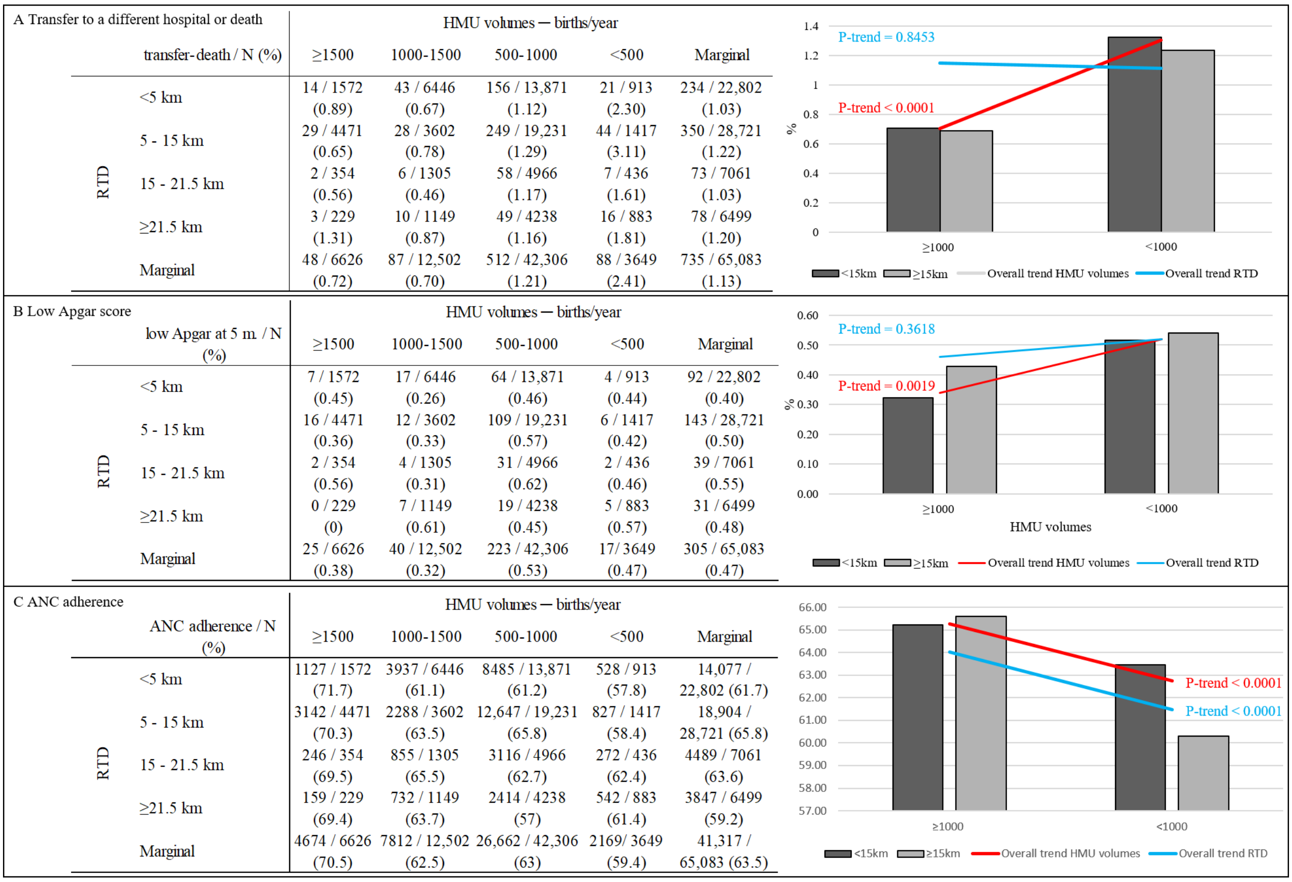
[e] Included diagnoses of migraine/headache, epilepsy, bipolar disorder, personality disorder, other psychiatric disorders, psychosis or schizophrenia, and sleep disorder.

[f] NSAID: nonsteroidal anti-inflammatory drug.

[g] ANC: antenatal care.

Figure 2 illustrates the distribution of outcomes by HMU volume and RTD. The rate of the primary outcome (transfer or death) significantly increased with decreasing HMU volume (*P* trend<.001). Hospitals with fewer than 500 births/year had a rate of 2.41%, compared with 0.72% in hospitals with 1500 births/year (Figure 2A). No significant differences were observed for RTD (*P* trend=.84). However, the lowest rate of the primary outcome was consistently observed in births with both high HMU volumes (≥1000 births/year) and short RTD (<15 km), representing the "best scenario" compared with all other combinations of volume and distance. Although too uncommon for detailed analysis across all categories, low Apgar scores were more prevalent in low HMUs with longer RTD (*P* trend<.002) (Figure 2B). Complete adherence to ANC decreased with lower HMU volume and increased with shorter RTD (*P* trend<.001, Figure 2C).

Figure 2. Distribution of maternal and neonatal outcomes based on HMU volume and RTD in Lombardy (N=65,083). (A) Transfer to a different hospital or death; (B) low Apgar score; (C) ANC adherence. ANC: antenatal care; HMU: hospital maternity unit; RTD: road travel distance.



A Transfer to a different hospital or death

	transfer-death / N (%)	≥1500	1000-1500	500-1000	<500	Marginal
		HMU volumes — births/year				
RTD	<5 km	14 / 1572 (0.89)	43 / 6446 (0.67)	156 / 13,871 (1.12)	21 / 913 (2.30)	234 / 22,802 (1.03)
	5 - 15 km	29 / 4471 (0.65)	28 / 3602 (0.78)	249 / 19,231 (1.29)	44 / 1417 (3.11)	350 / 28,721 (1.22)
	15 - 21.5 km	2 / 354 (0.56)	6 / 1305 (0.46)	58 / 4966 (1.17)	7 / 436 (1.61)	73 / 7061 (1.03)
	≥21.5 km	3 / 229 (1.31)	10 / 1149 (0.87)	49 / 4238 (1.16)	16 / 883 (1.81)	78 / 6499 (1.20)
	Marginal	48 / 6626 (0.72)	87 / 12,502 (0.70)	512 / 42,306 (1.21)	88 / 3649 (2.41)	735 / 65,083 (1.13)

B Low Apgar score

	low Apgar at 5 m. / N (%)	≥1500	1000-1500	500-1000	<500	Marginal
		HMU volumes — births/year				
RTD	<5 km	7 / 1572 (0.45)	17 / 6446 (0.26)	64 / 13,871 (0.46)	4 / 913 (0.44)	92 / 22,802 (0.40)
	5 - 15 km	16 / 4471 (0.36)	12 / 3602 (0.33)	109 / 19,231 (0.57)	6 / 1417 (0.42)	143 / 28,721 (0.50)
	15 - 21.5 km	2 / 354 (0.56)	4 / 1305 (0.31)	31 / 4966 (0.62)	2 / 436 (0.46)	39 / 7061 (0.55)
	≥21.5 km	0 / 229 (0)	7 / 1149 (0.61)	19 / 4238 (0.45)	5 / 883 (0.57)	31 / 6499 (0.48)
	Marginal	25 / 6626 (0.38)	40 / 12,502 (0.32)	223 / 42,306 (0.53)	17/ 3649 (0.47)	305 / 65,083 (0.47)

C ANC adherence

	ANC adherence / N (%)	≥1500	1000-1500	500-1000	<500	Marginal
		HMU volumes — births/year				
RTD	<5 km	1127 / 1572 (71.7)	3937 / 6446 (61.1)	8485 / 13,871 (61.2)	528 / 913 (57.8)	14,077 / 22,802 (61.7)
	5 - 15 km	3142 / 4471 (70.3)	2288 / 3602 (63.5)	12,647 / 19,231 (65.8)	827 / 1417 (58.4)	18,904 / 28,721 (65.8)
	15 - 21.5 km	246 / 354 (69.5)	855 / 1305 (65.5)	3116 / 4966 (62.7)	272 / 436 (62.4)	4489 / 7061 (63.6)
	≥21.5 km	159 / 229 (69.4)	732 / 1149 (63.7)	2414 / 4238 (57)	542 / 883 (61.4)	3847 / 6499 (59.2)
	Marginal	4674 / 6626 (70.5)	7812 / 12,502 (62.5)	26,662 / 42,306 (63)	2169/ 3649 (59.4)	41,317 / 65,083 (63.5)

The association between low HMU volume and the primary outcome was confirmed by log-binomial regression models (Figure 3). Newborns born in maternity units with fewer than 500 births/year highlighted had up to a 3-fold higher risk of being transferred or dying compared with those delivered in hospitals with 1500 or more births/year (adjusted RR 1.69, 95% CI 0.65-4.41) and newborns born in hospitals with an HMU volume between 500 - 1000 and <500 (adjusted RR 3.57, 95% CI 1.26-10.13), respectively (Figure 3A). In contrast, no statistically significant associations were observed when evaluating the RTD and its interaction with volumes while considering the low Apgar score as a secondary outcome (Figure 3B). Conversely, the RR of adhering to ANC recommendations decreased with lower HMU volume. Women who delivered in maternity units with 1000 - 1500 births/year had a 9% (95% CI 7% - 11%) lower adherence rate, whereas those in units with <500 births/year had a 14% (95% CI 11% - 17%) lower adherence rate compared to high-volume hospitals. Additionally, longer RTD was associated with a 3% (95% CI 1% - 6%) lower adherence rate (Figure 3C).

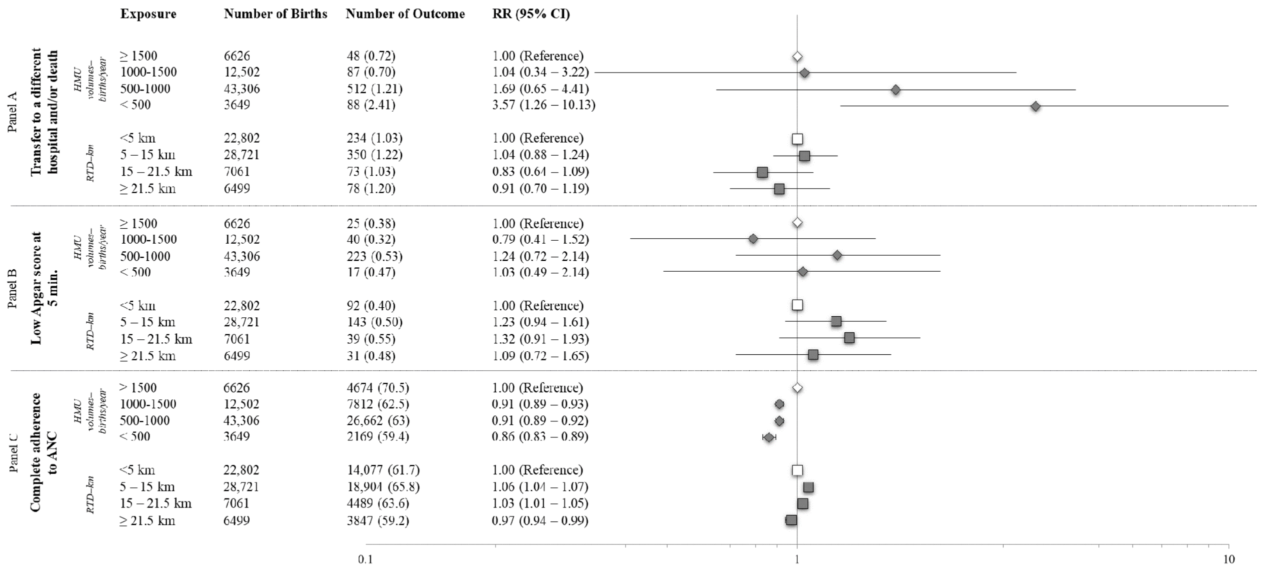
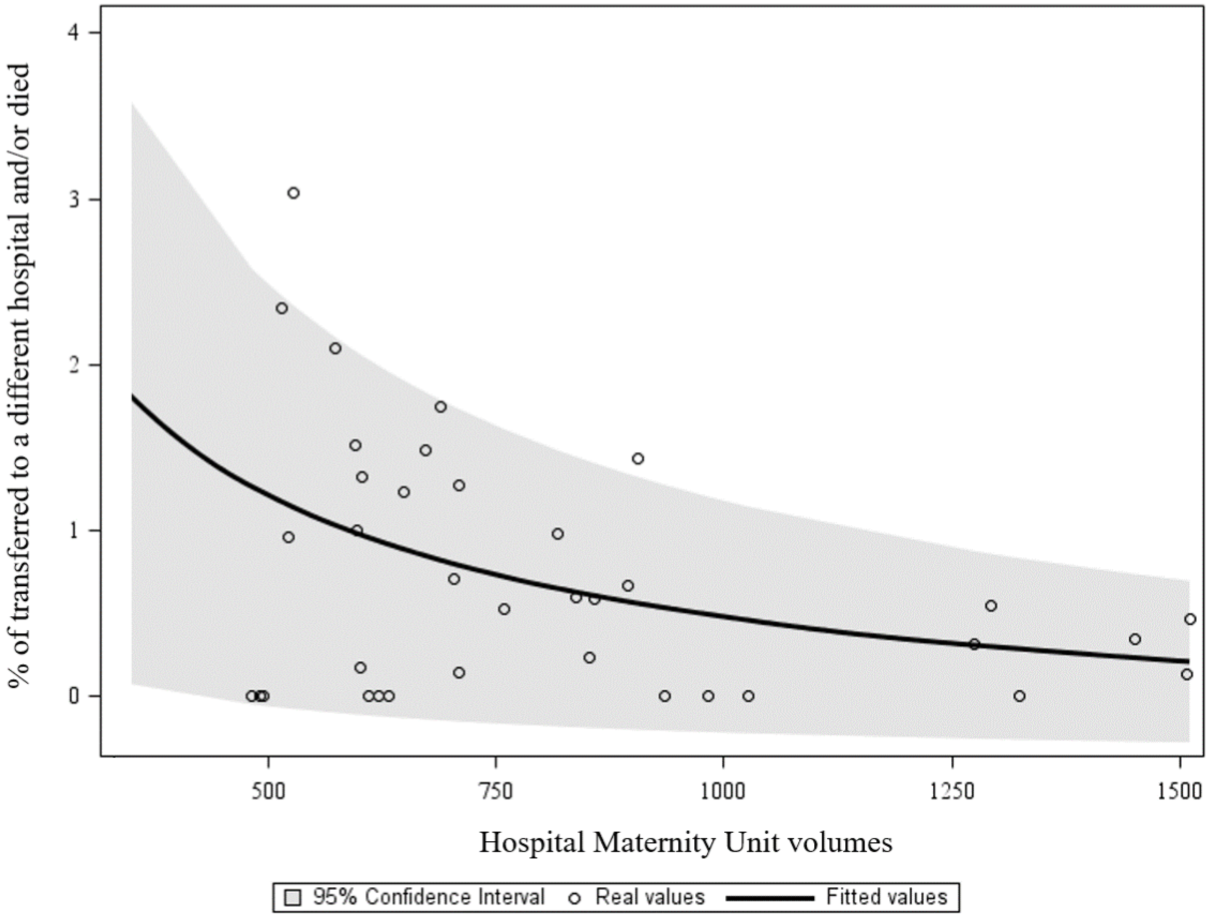


Figure 3. Association between HMU volumes, RTD, and primary and secondary outcomes in Lombardy in 2016 - 2019 (N=65,083). (A) transfer to a different hospital or death; (B) low Apgar score; (C) ANC adherence. ANC: antenatal care; HMU: hospital maternity unit; RTD: road travel distance.

These results were confirmed through a stratified analysis by hospital birth volume (high: ≥1000 births/year, low: <1000 births/year) (data not shown). Figure 4 shows the distribution of the percentage of newborns who were transferred to a different hospital or died as a function of HMU volumes. The FPM identified was 1x+ln (x)*1x, where x represents the mean number of births in each hospital in 2019. Each point on the graph represents a hospital included in our cohort. The solid black line demonstrates a clear decreasing trend in the frequency of the primary outcome (ie, percentage of individuals transferred to a different hospital and/or died), ranging from 2% in hospitals with lower activity volumes to 0.5% in those with higher activity volumes. There is also a specific trend of the curve flattening at an HMU volume of 1000 births/year or greater, where the frequency of the primary outcome was around 0.5%.

Figure 4. Distribution of the percentage of newborns who were transferred to a different hospital or died as a function of HMU volumes in Lombardy in 2016 - 2019.



Discussion

Our findings showed an excess risk of both neonatal transfer and death for live births delivered in HMUs with low activity volumes, comparable to services without ICUs. In contrast, the distance between the mothers' residences and hospital locations did not affect the primary outcome. Neither HMU volume nor RTD were significantly associated with low Apgar scores. However, mothers referred to hospitals with higher activity volumes and shorter travel distances demonstrated better adherence to ANC pathways, suggesting that being assisted by a lower-volume and distant hospital is associated with poorer ANC and correlates with a higher risk of neonatal problems. Finally, the data show a flattening trend in the prevalence of transfer to a different hospital or death in hospitals with an activity volume of 1000 births/year or more. This finding suggests this cutoff as the acceptable number of births for each maternity hospital to ensure an appropriate standard of care at delivery.

Maternity care reconfiguration is a complex and controversial issue in health care. It involves consolidating or closing small, less specialized maternity units and centralizing childbirth services at larger, more specialized hospitals. There are a number of reasons supporting the regionalization of HMUs. These include (1) declining birth rates, (2) improved safety and outcomes, and (3) saving costs.

In many countries, including Italy, the birth rate has declined for several decades. This has reduced the number of births—mainly at already small hospitals—increasing the risk of inefficient operation, low occupancy rates, and increasing maternal and neonatal adverse outcomes. However, regionalization of perinatal care could be the right solution to improve outcomes for mothers and newborns. For example, in Portugal and many other countries, regionalization involved the closure of a massive number of HMUs with fewer than 1000 births/year. This has consistently reduced several maternal and neonatal adverse outcomes such as the stillbirth rate, which ranged from 4 per 1000 births in 2000 to 2.2 per 1000 births in 2021 [17]. In contrast, in Italy, despite a decreased number of births, the stillbirth rate has remained consistent over the years, ranging from 2.4 to 2.2 per 1000 births during the same period. However, several studies have shown that women who give birth at large hospitals with high volumes of deliveries tend to have better outcomes, including lower rates of complications and infant mortality [18]. Our results are consistent with findings from other studies that reported significantly lower rates of stillbirths and neonatal mortality in both rural and urban regions after the closure of low-volume HMUs [19].

While data indicate lower neonatal mortality rates in high-volume birth hospitals at an early age, closing low-volume units requires extreme caution due to potential disadvantages, and is the subject of ongoing debate. Studies have raised concerns about increased unplanned out-of-hospital births and high neonatal mortality and stillbirth rates following closures [20,21]. Furthermore, studies in vast rural areas with limited access to perinatal care have revealed high rates of adverse birth outcomes and increased stress and anxiety among pregnant women [22,23]. However, our results did not show an association between long distances traveled to access perinatal care and the analyzed neonatal outcomes, except among hospitals with a low volume of activity and for full adherence to ANC.

Moreover, centralizing childbirth services can help to reduce health care costs by eliminating duplication of services and improving the efficient use of resources. However, there are several potential drawbacks to maternity care reconfiguration including increased travel times for women (if women must travel farther and for a longer time to give birth, this can increase their risk of complications), which may make receiving adequate prenatal care more difficult. Unfortunately, in this study, we were only able to calculate the distance to access perinatal care, which is a proxy for travel time rather than the time travel itself. Moreover, the closure of small maternity units can leave some women with limited access to childbirth services, especially in rural or geographically disadvantaged areas. It must be considered that access to ANC is often guaranteed in nonhospital settings (eg, public clinics) and that low-level hospitals can offer adequate ANC while organizing a structured referral to large hospitals for delivery. High-performing hospitals and well-equipped neonatal ICUs can also manage critical situations without the need to transfer newborns.

The decision to regionalize perinatal care units is complex and must be made on a case-by-case basis, considering all relevant factors. There is no one-size-fits-all solution that would be suitable for all communities, and health care providers and policy makers should consider several factors before making decisions about maternity care reconfiguration. These include the specific needs of the community, availability of transportation, and access to specialized care.

A strength of our study is the large, population-based sample reporting real-world evidence on the associations between HMU volumes and RTD, and the transfer to a different hospital or death (identified through a robust definition), low Apgar score at 5 minutes, and complete adherence to ANC. Moreover, we considered the individual prenatal care pathway for each mother, strengthening our analysis by addressing a potential confounding factor.

A limitation of our study is the lack of data on the actual time to travel to HMUs, as only data on RTD were available; we acknowledge that factors such as traffic congestion, mode of transportation, and individual driving habits could influence actual travel times. In addition, due to privacy concerns, we lack data on the exact location of the mothers' residence; therefore, we used the mothers' residence municipality for the calculation of the RTD. This misclassification might have reduced the difference in outcomes between RTD categories, thus underestimating the effect of RTD on outcomes. However, we believe that the RTD data used in our analysis provides a valuable approximation of accessibility, particularly when considered in conjunction with other factors such as HMU volumes and neonatal outcomes. Moreover, we were not able to consider several other factors, such as smoking, alcohol, illicit drug use, and social factors related to migration status that may affect access to both ANC and the HMU and influence the outcomes of interest. Additionally, the exclusion of mothers who experienced a stillbirth or those with missing information for sociodemographic or clinical characteristics likely affected less healthy women. However, missing data were infrequent (at most 0.4% for employment status). Nevertheless, outcome prevalence did not vary between complete and incomplete cases, supporting the hypothesis of data "missing at random" (data not shown). In our previous study, we showed that multiple imputations did not change the complete-case analysis results [24]. Finally, our results are derived from Lombardy, a specific Italian region, which may impact the generalizability of results. However, we believe that our study's key insights can apply to a broader range of settings, particularly those with similar health care systems and demographic characteristics. However, cultural beliefs and preferences regarding childbirth and health care can affect the decision-making process for mothers and health care providers. While our study provides valuable insights into the relationship between HMU volume, RTD, and neonatal outcomes in Lombardy, it is important to consider these factors when applying the findings to other regions. Several studies have reported that regionalized care enhances patient outcomes by concentrating specialized expertise at high-volume centers and fostering collaboration among health care providers within a defined region [25]. Moreover, studies have consistently demonstrated the benefits of regionalized care for high-risk infants, particularly those born very or late preterm and very-low-birth weight (<32 weeks of gestation) [26,27]. Further research may be needed to assess the generalizability of our results in diverse settings.

In conclusion, our results support the requirement outlined in the Italian Ministry of Health's decree 70/2015, which mandates at least 1000 deliveries per year in HMUs as a standard. It is necessary to emphasize that maternal-newborn separation caused by the transfer of the newborn can be incredibly detrimental to their development and bonding, making neonatal transfer a major concern [28]. Therefore, it is important to have a comprehensive plan to mitigate any potential negative impacts of maternity care reorganization. This plan should include measures to ensure that women have access to appropriate prenatal care and that they are able to give birth in a safe and supportive environment, regardless of their location, as advocated by the World Health Organization.

Acknowledgments

RZ and GC contributed equally to this work. The project was funded by a grant from the Italian Ministry of Education, University and Research 'Fondo d'Ateneo per la Ricerca' portion, year 2023 (grant number 2023-ATE-0565), and a grant from the Italian Ministry of Health ("Metodi per il monitoraggio e la valutazione dell'assistenza sanitaria nell'ambito del Nuovo Sistema di Garanzia, con particolare riferimento alle cure integrate e al confronto dei risultati pre e post pandemia da CoViD-19" project, grant number J59H06000160001). The funding agencies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

GC received research support from the European Community, the Italian Agency of Drug, and the Italian Ministry for University and Research.

Data Availability

Data may be obtained from a third party and are not publicly available.

Conflicts of Interest

GC took part in various projects funded by pharmaceutical companies (ie, Novartis, GSK, Roche, AMGEN, and BMS). He also received honoraria as a member of the advisory board from Roche.
The other authors have no conflicts of interest to declare. We affirm that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of any organization or company.

References

1. Rashidian A, Omidvari AH, Vali Y, et al. The effectiveness of regionalization of perinatal care services--a systematic review. Public Health (Fairfax) 2014 Oct;128(10):872-885. [doi: 10.1016/j.puhe.2014.08.005] [Medline: 25369352]
2. Barfield WD, Papile LA, Baley JE, American Academy of Pediatrics Committee on Fetus and Newborn, et al. Levels of neonatal care. Pediatrics 2012 Sep 1;130(3):587-597. [doi: 10.1542/peds.2012-1999] [Medline: 23641169]
3. Promoting effective perinatal care (PEPC) in the European region. World Health Organization. 2000 Nov. URL: https://iris.who.int/bitstream/handle/10665/108550/WHO-EURO-2000-4052-43811-61678-eng.pdf?sequence=1&isAllowed=y
4. Six success stories in making pregnancy safer. World Health Organization. 2010. URL: https://iris.who.int/bitstream/handle/10665/374844/WHO-EURO-2010-8879-48651-72242-eng.pdf?sequence=1&isAllowed=y [accessed 2022-02-20]
5. Aubrey-Bassler FK, Cullen RM, Simms A, et al. Population-based cohort study of hospital delivery volume, geographic accessibility, and obstetric outcomes. Int J Gynaecol Obstet 2019 Jul;146(1):95-102. [doi: 10.1002/ijgo.12832] [Medline: 31032903]
6. Darling EK, Lawford KMO, Wilson K, Kryzanauskas M, Bourgeault IL. Distance from home birth to emergency obstetric services and neonatal outcomes: a cohort study. J Midwif Womens Health 2019 Mar;64(2):170-178. [doi: 10.1111/jmwh.12896] [Medline: 30325580]
7. Pfniss I, Gold D, Holter M, et al. Birth during off-hours: impact of time of birth, staff´s seniority, and unit volume on maternal adverse outcomes-a population-based cross-sectional study of 87 065 deliveries. Birth 2023 Jun;50(2):449-460. [doi: 10.1111/birt.12663] [Medline: 35789033]
8. Handley SC, Lorch SA. Regionalization of neonatal care: benefits, barriers, and beyond. J Perinatol 2022 Jun;42(6):835-838. [doi: 10.1038/s41372-022-01404-7] [Medline: 35461330]
9. Lupattelli A, Corrao G, Gatti C, Rea F, Trinh NTH, Cantarutti A. Antidepressant continuation and adherence in pregnancy, and risk of antenatal hospitalization for unipolar major depressive and/or anxiety disorders. J Affect Disord 2023 Oct;339:502-510. [doi: 10.1016/j.jad.2023.07.066] [Medline: 37459974]
10. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
11. Matrici di contiguità, distanza e pendolarismo. Instituto Nazionale di Statistica. 2004 May 9. URL: https://www.istat.it/it/archivio/157423 [accessed 2024-02-20]
12. Casey BM, McIntire DD, Leveno KJ. The continuing value of the Apgar score for the assessment of newborn infants. N Engl J Med 2001 Feb 15;344(7):467-471. [doi: 10.1056/NEJM200102153440701] [Medline: 11172187]
13. Corrao G, Cantarutti A, Locatelli A, et al. Association between adherence with recommended antenatal care in low-risk, uncomplicated pregnancy, and maternal and neonatal adverse outcomes: evidence from Italy. Int J Environ Res Public Health 2020 Dec 29;18(1):173. [doi: 10.3390/ijerph18010173] [Medline: 33383661]

14. Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Glynn RJ. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. Am J Epidemiol 2001 Nov 1;154(9):854-864. [doi: 10.1093/aje/154.9.854] [Medline: 11682368]
15. Determinazione 20 marzo 2008 [Article in Italian]. Gazzetta Ufficiale. URL: https://www.gazzettaufficiale.it/eli/id/2008/03/31/08A02109/sg [accessed 2024-02-20]
16. Royston P, Sauerbrei W. Stability of multivariable fractional polynomial models with selection of variables and transformations: a bootstrap investigation. Stat Med 2003 Feb 28;22(4):639-659. [doi: 10.1002/sim.1310] [Medline: 12590419]
17. Stillbirth rate (per 1,000 total births). World Bank Group. URL: https://genderdata.worldbank.org/indicators/sh-dyn-stlb/?year=2000 [accessed 2024-02-20]
18. Walther F, Kuester D, Bieber A, Malzahn J, Rüdiger M, Schmitt J. Are birth outcomes in low risk birth cohorts related to hospital birth volumes? A systematic review. BMC Pregnancy Childbirth 2021 Jul 27;21(1):531. [doi: 10.1186/s12884-021-03988-y] [Medline: 34315416]
19. Allen VM, Jilwah N, Joseph KS, et al. The influence of hospital closures in Nova Scotia on perinatal outcomes. J Obstet Gynaecol Can 2004 Dec;26(12):1077-1085. [doi: 10.1016/s1701-2163(16)30435-2] [Medline: 15607044]
20. Lorch SA, Srinivas SK, Ahlberg C, Small DS. The impact of obstetric unit closures on maternal and infant pregnancy outcomes. Health Serv Res 2013 Apr;48(2 Pt 1):455-475. [doi: 10.1111/j.1475-6773.2012.01455.x] [Medline: 22881056]
21. Kozhimannil KB, Hung P, Henning-Smith C, Casey MM, Prasad S. Association between loss of hospital-based obstetric services and birth outcomes in rural counties in the United States. JAMA 2018 Mar 27;319(12):1239-1247. [doi: 10.1001/jama.2018.1830] [Medline: 29522161]
22. Grzybowski S, Stoll K, Kornelsen J. Distance matters: a population based study examining access to maternity services for rural women. BMC Health Serv Res 2011 Jun 10;11:147. [doi: 10.1186/1472-6963-11-147] [Medline: 21663676]
23. Kornelsen J, Stoll K, Grzybowski S. Stress and anxiety associated with lack of access to maternity services for rural parturient women. Aust J Rural Health 2011 Feb;19(1):9-14. [doi: 10.1111/j.1440-1584.2010.01170.x] [Medline: 21265919]
24. Cantarutti A, Merlino L, Giaquinto C, Corrao G. Use of antidepressant medication in pregnancy and adverse neonatal outcomes: a population-based investigation. Pharmacoepidemiol Drug Saf 2017 Sep;26(9):1100-1108. [doi: 10.1002/pds.4242] [Medline: 28594145]
25. Lorch SA, Myers S, Carr B. The regionalization of pediatric health care. Pediatrics 2010;126:1182. [doi: 10.1542/peds.2010-1119] [Medline: 21041285]
26. Lasswell SM, Barfield WD, Rochat RW, Blackmon L. Perinatal regionalization for very low-birth-weight and very preterm infants. JAMA 2010 Sep 1;304(9):992. [doi: 10.1001/jama.2010.1226] [Medline: 20810377]
27. Salazar EG, Passarella M, Formanowski B, Phibbs CS, Lorch SA, Handley SC. The impact of volume and neonatal level of care on outcomes of moderate and late preterm infants. J Perinatol 2024 Oct;44(10):1409-1415. [doi: 10.1038/s41372-024-01901-x] [Medline: 38413758]
28. Bergman NJ. Birth practices: maternal-neonate separation as a source of toxic stress. Birth Defects Res 2019 Sep 1;111(15):1087-1109. [doi: 10.1002/bdr2.1530] [Medline: 31157520]

Abbreviations

ANC: antenatal care
FPM: fractional polynomial model
HMU: hospital maternity unit
ICU: intensive care unit
NHS: National Health Service
RR: relative risk
RTD: road travel distance
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 28.03.24; peer-reviewed by WH Dong, Y Liu; revised version received 22.10.24; accepted 29.10.24; published 21.01.25.

Please cite as:
Cantarutti A, Boracchini R, Bellù R, Ronco R, Rea F, Locatelli A, Zanini R, Corrao G
Assessing the Impact of Distance Traveled and Birth Volumes of Hospital Maternity Units on Newborn Outcomes: Population-Based Cohort Study
JMIR Public Health Surveill 2025;11:e58944
URL: https://publichealth.jmir.org/2025/1/e58944
doi:10.2196/58944

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org