

JMIR Public Health and Surveillance

Impact Factor (2022): 8.5
Volume 7 (2021), Issue 3 ISSN 2369-2960 Editor in Chief: Travis Sanchez, PhD, MPH

Contents

Original Papers

- Changes in Workers' Sedentary and Physical Activity Behaviors in Response to the COVID-19 Pandemic and Their Relationships With Fatigue: Longitudinal Online Study ([e26293](#))
Mohammad Koohsari, Tomoki Nakaya, Gavin McCormack, Ai Shibata, Kaori Ishii, Koichiro Oka. 3
- Predicting Age Groups of Reddit Users Based on Posting Behavior and Metadata: Classification Model Development and Validation ([e25807](#))
Robert Chew, Caroline Kery, Laura Baum, Thomas Bukowski, Annice Kim, Mario Navarro. 14
- Effects of a Co-Design–Based Invitation Strategy on Participation in a Preventive Health Check Program: Randomized Controlled Trial ([e25617](#))
Trine Thilsing, Lars Larsen, Anders Sonderlund, Signe Andreassen, Jeanette Christensen, Nanna Svensson, Marie Dahl, Jens Sondergaard. 2 8
- Estimating the Incidence of Conjunctivitis by Comparing the Frequency of Google Search Terms With Clinical Data: Retrospective Study ([e22645](#))
Paola Kammrath Betancor, Linda Tizek, Alexander Zink, Thomas Reinhard, Daniel Böhringer. 42
- Use of #SaludTues Tweetchats for the Dissemination of Culturally Relevant Information on Latino Health Equity: Exploratory Case Study ([e21266](#))
Amelie Ramirez, Rosalie Aguilar, Amanda Merck, Cliff Despres, Pramod Sukumaran, Stacy Cantu-Pawlik, Patricia Chalela. 50
- Context and Barriers to the Prescription of Nonoccupational Postexposure Prophylaxis Among HIV Medical Care Providers: National Internet-Based Observational Study in China ([e24234](#))
Haibo Ding, Zehao Ye, Weiming Tang, Xiaojie Huang, Hui Wang, Sitong Cui, Yongjun Jiang, Wenqing Geng, Junjie Xu, Hong Shang. 65
- Hybrid Approach to Estimation of Underreporting of Tuberculosis Case Notification in High-Burden Settings With Weak Surveillance Infrastructure: Design and Implementation of an Inventory Study ([e22352](#))
Ellen Mitchell, Olusola Adejumo, Hussein Abdur-Razzaq, Chidubem Ogbudebe, Nkem Chukwueme, Samson Olorunju, Mustapha Gidado. 8 1
- The Relationship Between the Global Burden of Influenza From 2017 to 2019 and COVID-19: Descriptive Epidemiological Assessment ([e24696](#))
Stefan Baral, Katherine Rucinski, Jean Twahirwa Rwema, Amrita Rao, Neia Prata Menezes, Daouda Diouf, Adeeba Kamarulzaman, Nancy Phaswana-Mafuya, Sharmistha Mishra. 115

Bridging Learning in Medicine and Citizenship During the COVID-19 Pandemic: A Telehealth-Based Case Study ([e24795](#))

Thiago Cerqueira-Silva, Roberto Carreiro, Victor Nunes, Louran Passos, Bernardo Canedo, Sofia Andrade, Pablo Ramos, Ricardo Khouri, Carolina Santos, Jedson Nascimento, Aurea Paste, Ivan Paiva Filho, Marília Santini-Oliveira, Álvaro Cruz, Manoel Barral-Netto, Viviane Boaventura. . .

1 3 1

Early Improvement of Acute Respiratory Distress Syndrome in Patients With COVID-19 in the Intensive Care Unit: Retrospective Analysis ([e24843](#))

Zhu Zhan, Xin Yang, Hu Du, Chuanlai Zhang, Yuyan Song, Xiaoyun Ran, An Zhang, Mei Yang. 139

The Nature and Extent of Online Marketing by Big Food and Big Alcohol During the COVID-19 Pandemic in Australia: Content Analysis Study ([e25202](#))

Florentine Martino, Ruby Brooks, Jennifer Browne, Nicholas Carah, Christina Zorbas, Kirstan Corben, Emma Saleeba, Jane Martin, Anna Peeters, Kathryn Backholer. 151

The Influence of Social Distancing on COVID-19 Mortality in US Counties: Cross-sectional Study ([e21606](#))

Phoebe Tran, Lam Tran, Liem Tran. 167

Automated Travel History Extraction From Clinical Notes for Informing the Detection of Emergent Infectious Disease Events: Algorithm Development and Validation ([e26719](#))

Kelly Peterson, Julia Lewis, Olga Patterson, Alec Chapman, Daniel Denhalter, Patricia Lye, Vanessa Stevens, Shantini Gamage, Gary Roselle, Katherine Wallace, Makoto Jones. 174

App Use and Usability of a Barcode-Based Digital Platform to Augment COVID-19 Contact Tracing: Postpilot Survey and Paradata Analysis ([e25859](#))

Thomas Scherr, Jenna DeSousa, Carson Moore, Austin Hardcastle, David Wright. 187

Early Detection of Dengue Fever Outbreaks Using a Surveillance App (Mozzify): Cross-sectional Mixed Methods Usability Study ([e19034](#))

Von Herbuela, Tomonori Karita, Thaddeus Carvajal, Howell Ho, John Lorena, Rachele Regalado, Girly Sobrepeña, Kozo Watanabe. 215

A Mobile App and Dashboard for Early Detection of Infectious Disease Outbreaks: Development Study ([e14837](#))

Euijoon Ahn, Na Liu, Tej Parekh, Ronak Patel, Tanya Baldacchino, Tracy Mullavey, Amanda Robinson, Jinman Kim. 231

Combining Wearable Devices and Mobile Surveys to Study Child and Youth Development in Malawi: Implementation Study of a Multimodal Approach ([e23154](#))

Onicio Leal Neto, Simon Haenni, John Phuka, Laura Ozella, Daniela Paolotti, Ciro Cattuto, Daniel Robles, Guilherme Lichand. 247

Diagnostic Accuracy of Detecting Diabetic Retinopathy by Using Digital Fundus Photographs in the Peripheral Health Facilities of Bangladesh: Validation Study ([e23538](#))

Tahmina Begum, Aminur Rahman, Dilruba Nomani, Abdullah Mamun, Alayne Adams, Shafiqul Islam, Zara Khair, Zareen Khair, Iqbal Anwar.

2 6 3

Rapid Surveillance Reports

The Annual American Men's Internet Survey of Behaviors of Men Who Have Sex With Men in the United States: Key Indicators Report 2018 ([e21812](#))

Sarah Wiatrek, Maria Zlotorzynska, Ramona Rai, Patrick Sullivan, Travis Sanchez. 101

Analyzing Cross-country Pandemic Connectedness During COVID-19 Using a Spatial-Temporal Database: Network Analysis ([e27317](#))

Amanda Chu, Jacky Chan, Jenny Tsang, Agnes Tiwari, Mike So. 205

Original Paper

Changes in Workers' Sedentary and Physical Activity Behaviors in Response to the COVID-19 Pandemic and Their Relationships With Fatigue: Longitudinal Online Study

Mohammad Javad Koohsari^{1,2,3}, PhD; Tomoki Nakaya⁴, PhD; Gavin R McCormack^{1,5,6,7}, PhD; Ai Shibata⁸, PhD; Kaori Ishii¹, PhD; Koichiro Oka¹, PhD

¹Faculty of Sport Sciences, Waseda University, Tokorozawa, Japan

²Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia

³Behavioural Epidemiology Laboratory, Baker Heart and Diabetes Institute, Melbourne, Australia

⁴Graduate School of Environmental Studies, Tohoku University, Sendai, Japan

⁵Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

⁶Faculty of Kinesiology, University of Calgary, Calgary, AB, Canada

⁷School of Architecture, Planning and Landscape, University of Calgary, Calgary, AB, Canada

⁸Faculty of Health and Sport Sciences, University of Tsukuba, Tsukuba, Japan

Corresponding Author:

Mohammad Javad Koohsari, PhD

Faculty of Sport Sciences

Waseda University

2-579-15 Mikajima

Tokorozawa, 359-1192

Japan

Phone: 81 4 2947 7189

Email: javadkoohsari@aoni.waseda.jp

Abstract

Background: Sedentary behaviors and physical activity are likely to be affected by the COVID-19 outbreak, and sedentary lifestyles can increase subjective fatigue. The nonpharmaceutical policies imposed as a result of the COVID-19 pandemic may also have adverse effects on fatigue.

Objective: This study has two aims: to examine the changes in sedentary behaviors and physical activity of company workers in response to the COVID-19 pandemic in Japan and to examine relationships between changes in these sedentary behaviors and physical activity and changes in fatigue.

Methods: Data from a nationwide prospective online survey conducted in 2019 and 2020 were used. On February 22, 2019, an email with a link to participate in the study was sent to 45,659 workers, aged 20 to 59 years, who were randomly selected from a database of approximately 1 million individuals. A total of 2466 and 1318 participants, who self-reported their occupation as company workers, answered the baseline and follow-up surveys, respectively. Surveys captured fatigue, workday and daily domain-specific sedentary behaviors and physical activity, and total sedentary behaviors and physical activity. We used multivariable linear regression models to estimate associations of changes in sedentary behaviors and physical activity with changes in fatigue.

Results: Increases in public transportation sitting during workdays, other leisure sitting time during workdays, and other leisure sitting time were associated with an increase in the motivation aspect of fatigue ($b=0.29$, 95% CI 0-0.57, $P=.048$; $b=0.40$, 95% CI 0.18-0.62, $P<.001$; and $b=0.26$, 95% CI 0.07-0.45, $P=.007$, respectively). Increases in work-related sitting time during workdays, total sitting time during workdays, and total work-related sitting time were significantly associated with an increase in the physical activity aspect of fatigue ($b=0.06$, 95% CI 0-0.12, $P=.03$; $b=0.05$, 95% CI 0.01-0.09, $P=.02$; and $b=0.07$, 95% CI 0-0.14, $P=.04$, respectively). The motivation and physical activity aspects of fatigue increased by 0.06 for each 1-hour increase in total sitting time between baseline and follow-up ($b=0.06$, 95% CI 0-0.11, $P=.045$; and $b=0.06$, 95% CI 0.01-0.10, $P=.009$, respectively).

Conclusions: Our findings demonstrated that sedentary and active behaviors among company workers in Japan were negatively affected during the COVID-19 outbreak. Increases in several domain-specific sedentary behaviors also contributed to unfavorable changes in workers' fatigue. Social distancing and teleworking amid a pandemic may contribute to the sedentary lifestyle of

company workers. Public health interventions are needed to mitigate the negative effects of the COVID-19 pandemic or future pandemics on sedentary and physical activity behaviors and fatigue among company workers.

(*JMIR Public Health Surveill* 2021;7(3):e26293) doi:[10.2196/26293](https://doi.org/10.2196/26293)

KEYWORDS

COVID-19; physical inactivity; sitting time; mental health; Japan; prospective design

Introduction

Sedentary lifestyles can increase the risk of many chronic diseases [1] and can increase subjective fatigue in working adults [2,3]. Subjective fatigue refers to “an overwhelming sense of tiredness, lack of energy, or feeling of exhaustion” [4]. Subjective fatigue has several adverse effects on health and well-being, productivity, and safety among workers [5-7]. For example, a population-based study conducted in Germany found that chronic fatigue was negatively associated with quality of life [8].

While the exact causes of fatigue in workplaces have not yet been fully investigated [9], promoting physical activity and interrupting sedentary time may mitigate subjective fatigue at work [10]. For example, a randomized trial study found that breaking prolonged sitting time with walking breaks was associated with lower fatigue levels among a sample of overweight and obese adults [11]. Longer sitting time was also found to be associated with more fatigue in a sample of Swedish adults [12]. Nevertheless, this evidence is mainly based on cross-sectional studies that were conducted among high-risk groups, such as overweight and obese individuals. Further research is warranted to assess whether longitudinal changes in sedentary and active behaviors may influence general workers' fatigue.

The World Health Organization declared COVID-19 a pandemic on March 11, 2020 [13]. Since then, several social distancing policies, such as school and university closures, recommendations for working from home, and bans on large gatherings, have been implemented by governments and local authorities to slow the spread of COVID-19. In Japan, a state of emergency was declared on April 7, 2020, by the government to contain the spread of COVID-19 [14]. Under this state of emergency, prefectural governors were allowed to request that residents restrict their unnecessary trips and stay at home as much as possible; they also asked that public places, such as cinemas and museums, limit their operations. However, in contrast with several other countries, no penalty or legal force under Japanese law existed for those people or companies who disobeyed the lockdown request. Nevertheless, sedentary behaviors and physical activity are likely to be affected by these social distancing policies. A worldwide study of the effects of COVID-19 on physical activity found a rapid decrease in step counts, an indicator of physical activity, across multiple countries [15]. The nonpharmaceutical policies resulting from public health measures in place to reduce person-to-person transmission of COVID-19 may also have adverse effects on people's fatigue. For instance, a study conducted in Poland reported a high level of everyday fatigue (ie, measured through items capturing physical, mental, and social fatigue) during the

COVID-19 home quarantine period compared to before the quarantine period [16]. Since it is currently unknown how long the COVID-19 pandemic will last, workers may need to follow workplace social distancing policies for an unpredictable duration. Even after this pandemic ends, it is likely that workers will continue to be sedentary due to advancements in workplace technology. Given the links between fatigue and an active lifestyle, further research is needed to shed light on how COVID-19 has influenced the sedentary behaviors and physical activities of different population groups, especially company workers who are already exposed to a sedentary lifestyle.

To our knowledge, no study has examined longitudinal changes in sedentary behaviors and physical activity among company workers during the COVID-19 pandemic and their effects on fatigue. Therefore, this study has two aims: (1) examine the changes in sedentary behaviors and physical activity of company workers during the COVID-19 outbreak in Japan and (2) examine the relationships between these changes in sedentary behaviors and physical activity and fatigue.

Methods

Data Source and Participants

Our study included data from a nationwide prospective online survey conducted in 2019 and 2020. The participants were recruited from a group of registered individuals of a Japanese internet research service company (MyVoice Communication, Inc). Approximately 1 million individuals across Japan voluntarily participated in populating this company's database and registered their sociodemographic information. The eligibility criteria for study participation were based on having the same number of participants of each sex (ie, female and male) and each age group (ie, workers in their 20s, 30s, 40s, and 50s) and being a company worker. Given our research budget, we aimed to recruit a total of 3200 workers, aged 20 to 59 years: 1600 participants from each sex category and 800 participants from each age group. On February 22, 2019, an email with a link to participate in the study was sent to 45,659 individuals who were randomly selected from the database by sex and age strata; the company sent emails until the planned sample size was reached. Of these, 2921 individuals were eligible to participate in the follow-up study on July 8, 2020—279 respondents had withdrawn from the company's database—and 1709 completed the follow-up survey (1709/3200, 53.4% of the baseline participants). Only participants who self-reported their occupation as company workers were included in this study: 2466 and 1318 participants at baseline and follow-up, respectively. The follow-up time frame had not been originally intended, and it was in response to the COVID-19 pandemic. Cash rewards valued at ¥140 (US \$1.30) and ¥120 (US \$1.10) were offered as incentives to

participate in the first and second surveys, respectively. All participants signed an online informed consent form. The internet research service company removed participant names, addresses, personal health numbers, contact information, and any other details from the data set that might be used to identify individuals prior to transferring the data to researchers. The Institutional Ethics Committee of Waseda University approved this study (2020-135).

Measures

Fatigue

The Japanese version of the Checklist Individual Strength questionnaire (CIS20-R) was used to assess perceived fatigue during the past 2 weeks [17,18]. The CIS20-R is a multidimensional 20-item questionnaire that measures four aspects of fatigue—subjective fatigue, concentration, motivation, and physical activity—as well as total fatigue [17]. This questionnaire has high internal reliability (Cronbach α = .90 for total fatigue and ranges from .83 to .92 for its subscales) [17], and the questionnaire is valid in the working population [18,19]. Participants scored each item on a 7-point Likert scale ranging from 1 (exactly) to 7 (not at all). The scoring protocol for the CIS20-R has been described in detail elsewhere [19]. In brief, higher scores indicate a higher degree of fatigue, more concentration problems, reduced motivation, and less physical activity. Total fatigue was calculated from the sum of the four aspects of fatigue. We estimated absolute changes in total fatigue (ie, total CIS20-R score) and aspects of fatigue before and during the COVID-19 pandemic by subtracting baseline values from follow-up values.

Sedentary Behaviors and Physical Activity

Domain-specific sedentary behaviors were assessed using a Japanese, 6-item, self-reported questionnaire with a 1-week recall period [20]. Participants were asked to report their daily average sedentary behaviors over the past 7 days separately for workdays—or weekdays for unemployed individuals—and nonworkdays (ie, weekends) across the following six domains: driving or riding by car; using public transport; at work; watching television, videos, and DVDs; using a computer, cell phone, or tablet PC outside of working hours; and during leisure time, excluding watching television, videos, and DVDs. Participants reported the average time (hours and minutes) spent in each of these categories per day. The reliability and validity of this questionnaire have been reported elsewhere [20]. Average daily values of total sedentary time and each domain's sedentary time were calculated with weighting for the number of workdays and nonworkdays. The average workday value of total sedentary time was also calculated by summing all six domains for workdays. Domain-specific physical activity was measured using the Global Physical Activity Questionnaire (GPAQ) [21]. The method of cleaning and scoring GPAQ data has been described in detail elsewhere [22]. Briefly, this questionnaire consists of 16 questions that assess self-reported domain-specific physical activity—during work, transport, and leisure—and sitting time in a typical week. It has acceptable reliability and validity in Japanese adults [23]. Average daily hours of work-related vigorous physical activity, work-related moderate

physical activity, transport-related physical activity, leisure-related vigorous physical activity, leisure-related moderate physical activity, and total physical activity were calculated. We estimated absolute changes in sedentary behaviors and physical activity before and during the COVID-19 pandemic by subtracting baseline values from follow-up values.

Sociodemographic Variables

Participants self-reported their baseline age, sex (ie, female or male), marital status (ie, single or couple), highest education (ie, tertiary, below tertiary, or other), and gross annual individual income (ie, <¥4,000,000 [US \$37,000] or ≥¥4,000,000 [US \$37,000]).

Statistical Analysis

Descriptive statistics, including means and SDs as well as frequencies, were calculated for the baseline sociodemographic variables. Paired *t* tests were used to compare the domain-specific sedentary and physical activity behaviors with total fatigue and aspects of fatigue before and during the COVID-19 outbreak. A separate multivariable linear regression model estimated the associations between changes in sedentary behaviors and physical activities and the changes in each fatigue aspect at the individual level. The regression models were adjusted for the sociodemographic variables (ie, age, sex, marital status, highest education, and gross annual household income). The regression models were also adjusted for baseline fatigue to account for the potential effect of baseline values on change scores. Normality assumptions were checked by the quantile-quantile plots of the residuals. For all point estimates, where *b* is the unstandardized regression coefficient, 95% CIs were estimated. A complete-case analysis was chosen because the proportion of missing data for our variables of interest was low (5%) [24]. We found no significant differences in any measured characteristics between those individuals with missing data and those with complete data. Our analysis excluded those who were unable to engage in physical activity due to health issues. Analyses were conducted using Stata 15.0 (Stata Corp), and the level of significance was set at *P* < .05.

Results

Sample Characteristics

Table 1 shows the baseline sample characteristics (N=2466). The participants had an average age of 39.6 years (SD 10.7), and the majority were male, were single, had a high tertiary education, and had an annual gross household income lower than ¥4,000,000 (US \$37,000). The majority of the participants had sedentary occupations (ie, office-based workers) (1807/2466, 73.3%) and 26.7% (659/2466) had active occupations (eg, manual labor). There were no participants who became unemployed by the time of the follow-up survey. The participants reported the total number of weekly days they worked from home before and during the COVID-19 pandemic. The average number days that the participants worked from home was 0.2 and 1 day per week before and during the COVID-19 pandemic, respectively.

Table 1. Baseline characteristics of study participants in Japan, February 2019.

Characteristic	Value (N=2466)
Age in years, mean (SD)	39.6 (10.7)
Sex, n (%)	
Female	1212 (49.1)
Male	1254 (50.9)
Marital status, n (%)	
Single	1392 (56.4)
Couple	1074 (43.6)
Highest education, n (%)	
Tertiary	1989 (80.7)
Below tertiary	477 (19.3)
Gross annual individual income (¥), n (%)	
<4,000,000 ^a	1381 (56.0)
≥4,000,000 ^a	1085 (44.0)

^a¥4,000,000 is equivalent to US \$37,000.

Changes in Sedentary Behaviors, Physical Activity, and Fatigue

Mean hours per day of sedentary behaviors and physical activity before and during the COVID-19 pandemic are shown in [Tables 2](#) and [3](#). These comparisons were restricted only to participants who answered both surveys. There were several significant differences between workers' domain-specific sedentary behaviors from one period to the next: participants spent more sedentary time in *work-related sitting time during workday*, *TV viewing time during workday*, *PC use sitting time during workday*, *total sitting time during workday*, *work-related sitting time*, *PC use sitting time*, and *total sitting time* during the COVID-19 pandemic (see [Table 2](#)). Participants also reported significantly lower amounts of several domain-specific physical activities during the COVID-19 outbreak: the mean values of vigorous leisure physical activity and total physical activity

were significantly lower during the COVID-19 outbreak (see [Table 3](#)).

At baseline, the Cronbach α for internal consistency was .87, .75, .56, and .28 for the subjective fatigue, concentration, motivation, and physical activity items, respectively. These estimates were .86, .78, .59, and .33 for the follow-up items, respectively. Despite a small difference in mean values between subjective fatigue before and during the COVID-19 outbreak, this difference was nevertheless statistically significant: higher scores refer to a higher degree of fatigue. No significant difference was observed for the means of the three fatigue aspects of concentration, motivation, and physical activity nor for total fatigue (ie, total CIS20-R score) before and during the COVID-19 pandemic (see [Table 4](#)).

Higher scores refer to a higher degree of subjective fatigue, more concentration problems, reduced motivation, less physical activity, and higher total fatigue (ie, total CIS20-R score).

Table 2. Workers' domain-specific sedentary behaviors before (February 2019) and during (July 2020) the COVID-19 outbreak in Japan (n=1086).

Sedentary behavior	Hours per day, mean (SD)		P value ^a	Difference, mean (SD)
	Before	During		
Car sitting time during workday	0.45 (0.86)	0.45 (0.87)	.80	0.01 (0.85)
Public transportation sitting during workday	0.46 (0.72)	0.49 (0.79)	.15	0.03 (0.72)
Work-related sitting time during workday	5.18 (3.26)	5.69 (3.30)	<.001	0.50 (2.93)
TV viewing time during workday	1.56 (1.36)	1.65 (1.36)	.03	0.09 (1.27)
PC use sitting time during workday	1.36 (1.27)	1.47 (1.27)	.01	0.11 (1.40)
Other leisure sitting time during workday	0.52 (0.69)	0.54 (0.73)	.53	0.02 (0.92)
Total sitting time during workday	9.53 (3.96)	10.29 (3.99)	<.001	0.75 (3.95)
Car sitting time	0.50 (0.78)	0.50 (0.76)	.90	0 (0.72)
Public transportation sitting	0.39 (0.58)	0.39 (0.61)	.80	0 (0.57)
Work-related sitting time	3.91 (2.49)	4.22 (2.52)	<.001	0.31 (2.32)
TV viewing time	1.89 (1.54)	1.97 (1.56)	.06	0.08 (1.45)
PC use sitting time	1.61 (1.41)	1.76 (1.45)	<.001	0.14 (1.47)
Other leisure sitting time	0.67 (0.83)	0.69 (0.88)	.48	0.02 (1.09)
Total sitting time	8.96 (3.53)	9.53 (3.59)	<.001	0.56 (3.67)

^aP values were based on paired-sample *t* tests.

Table 3. Workers' domain-specific physical activity before (February 2019) and during (July 2020) the COVID-19 outbreak in Japan (n=1315).

Physical activity	Hours per day, mean (SD)		P value ^a	Difference, mean (SD)
	Before	During		
Work-related vigorous physical activity	0.21 (1.03)	0.19 (0.84)	.40	-0.02 (1.06)
Work-related moderate physical activity	0.43 (1.32)	0.39 (1.17)	.24	-0.05 (1.43)
Transport-related physical activity	0.48 (0.72)	0.44 (0.73)	.10	-0.04 (0.83)
Leisure-related vigorous physical activity	0.25 (0.72)	0.20 (0.53)	.01	-0.05 (0.64)
Leisure-related moderate physical activity	0.37 (0.83)	0.33 (0.72)	.07	-0.04 (0.80)
Total physical activity	1.75 (2.79)	1.55 (2.31)	.01	-0.20 (2.61)

^aP values were based on paired-sample *t* tests.

Table 4. Workers' total fatigue and aspects of fatigue before (February 2019) and during (July 2020) the COVID-19 outbreak in Japan (n=1318).

Fatigue aspect	CIS20-R ^a score, mean (SD)		P value ^b
	Before	During	
Subjective fatigue	31.2 (9.4)	31.9 (9.1)	.002
Concentration	20.1 (4.4)	20.3 (4.4)	.24
Motivation	17.4 (4.2)	17.3 (4.2)	.33
Physical activity	12.3 (3.0)	12.2 (3.0)	.45
Total fatigue	81.0 (16.9)	81.6 (16.6)	.10

^aCIS20-R: Checklist Individual Strength questionnaire.

^bP values were based on paired-sample *t* tests.

Associations of Changes in Sedentary Behaviors and Physical Activity With Fatigue

Table 5 shows the associations of the changes in workers' total fatigue (ie, total CIS20-R score) and aspects of fatigue with those in workers' domain-specific sedentary behaviors. Increases in public transportation sitting during workdays, other leisure sitting times during workdays, and other leisure sitting times were significantly associated with an increase in the motivation aspect of fatigue. Increases in work-related sitting time during

workdays, total sitting time during workdays, and work-related sitting time were significantly associated with an increase in the physical activity aspect of fatigue. The subscale scores for the motivation and physical activity aspects of fatigue increased by 0.03 for each half-hour increase in total sitting time between baseline and follow-up ($b/2=0.03$, 95% CI 0-0.05 and 95% CI 0.01-0.05, respectively). None of the changes in domain-specific physical activity were significantly associated with workers' total fatigue (ie, total CIS20-R score) or fatigue subscale scores (see Table 6).

Table 5. Results of multivariable linear regression models assessing associations between absolute changes in workers' domain-specific sedentary behaviors and changes in workers' total fatigue and aspects of fatigue in Japan (n=1086).

Sedentary behavior	Subjective fatigue, <i>b</i> ^a (95% CI)	Concentration, <i>b</i> (95% CI)	Motivation, <i>b</i> (95% CI)	Physical activity, <i>b</i> (95% CI)	Total fatigue, <i>b</i> (95% CI)
Car sitting time during workday	-0.05 (-0.55 to 0.45)	-0.04 (-0.30 to 0.22)	-0.01 (-0.25 to 0.24)	0.03 (-0.16 to 0.22)	-0.05 (-0.94 to 0.85)
Public transportation sitting during workday	-0.23 (-0.82 to 0.36)	-0.08 (-0.38 to 0.23)	0.29 (0 to 0.57) ^b	0.04 (-0.19 to 0.26)	0.01 (-1.04 to 1.06)
Work-related sitting time during workday	0.02 (-0.13 to 0.16)	0.07 (-0.01 to 0.14)	0.02 (-0.05 to 0.10)	0.06 (0 to 0.12)	0.16 (-0.10 to 0.42)
TV viewing time during workday	0.07 (-0.27 to 0.40)	0.06 (-0.11 to 0.23)	-0.02 (-0.18 to 0.14)	0.04 (-0.09 to 0.17)	0.16 (-0.44 to 0.76)
PC use sitting time during workday	0.05 (-0.25 to 0.35)	-0.13 (-0.28 to 0.03)	-0.02 (-0.16 to 0.13)	0.03 (-0.09 to 0.14)	-0.06 (-0.61 to 0.48)
Other leisure sitting time during workday	0.30 (-0.17 to 0.76)	0.02 (-0.22 to 0.26)	0.40 (0.18 to 0.62)	0.12 (-0.06 to 0.29)	0.76 (-0.08 to 1.59)
Total sitting time during workday	0.03 (-0.08 to 0.14)	0.02 (-0.03 to 0.08)	0.04 (-0.01 to 0.09)	0.05 (0.01 to 0.09)	0.14 (-0.06 to 0.33)
Car sitting time	0.23 (-0.36 to 0.82)	0.07 (-0.24 to 0.37)	0.09 (-0.20 to 0.37)	0.10 (-0.13 to 0.32)	0.50 (-0.55 to 1.56)
Public transportation sitting	-0.27 (-1.01 to 0.47)	-0.13 (-0.51 to 0.26)	0.31 (-0.05 to 0.67)	0.05 (-0.23 to 0.33)	-0.02 (-1.34 to 1.31)
Work-related sitting time	-0.02 (-0.21 to 0.16)	0.07 (-0.02 to 0.17)	0.03 (-0.06 to 0.12)	0.07 (0 to 0.14)	0.15 (-0.18 to 0.48)
TV viewing time	0.11 (-0.19 to 0.40)	0.11 (-0.04 to 0.26)	0.03 (-0.11 to 0.17)	0.07 (-0.04 to 0.18)	0.33 (-0.20 to 0.85)
PC use sitting time	-0.02 (-0.31 to 0.27)	-0.11 (-0.26 to 0.04)	0.03 (-0.11 to 0.17)	0.04 (-0.07 to 0.15)	-0.05 (-0.57 to 0.46)
Other leisure sitting time	0.23 (-0.16 to 0.62)	0.02 (-0.18 to 0.22)	0.26 (0.07 to 0.45)	0.09 (-0.06 to 0.24)	0.52 (-0.18 to 1.21)
Total sitting time	0.03 (-0.09 to 0.14)	0.03 (-0.03 to 0.09)	0.06 (0 to 0.11)	0.06 (0.01 to 0.10)	0.17 (-0.04 to 0.36)

^a*b* is the unstandardized regression coefficient; based on the Checklist Individual Strength questionnaire (CIS20-R) scores. All models were adjusted for age, sex, marital status, highest education, gross annual household income, and baseline fatigue.

^bItalicized values are statistically significant at $P<.05$.

Table 6. Results of multivariable linear regression models assessing associations between absolute changes in workers' domain-specific physical activity and changes in workers' total fatigue and aspects of fatigue in Japan (n=1315).

Physical activity	Subjective fatigue, <i>b</i> ^a (95% CI)	Concentration, <i>b</i> (95% CI)	Motivation, <i>b</i> (95% CI)	Physical activity, <i>b</i> (95% CI)	Total fatigue, <i>b</i> (95% CI)
Work-related vigorous physical activity	−0.23 (−0.59 to 0.14)	−0.02 (−0.21 to 0.17)	−0.08 (−0.26 to 0.10)	0.01 (−0.13 to 0.15)	−0.30 (−0.95 to 0.35)
Work-related moderate physical activity	−0.10 (−0.37 to 0.17)	−0.07 (−0.21 to 0.07)	−0.03 (−0.16 to 0.10)	0.05 (−0.05 to 0.15)	−0.18 (−0.66 to 0.30)
Transport-related physical activity	−0.06 (−0.52 to 0.40)	−0.02 (−0.26 to 0.22)	0.21 (−0.02 to 0.43)	0.11 (−0.06 to 0.28)	0.20 (−0.63 to 1.02)
Leisure-related vigorous physical activity	−0.26 (−0.86 to 0.33)	−0.17 (−0.48 to 0.15)	0.16 (−0.14 to 0.45)	−0.11 (−0.33 to 0.12)	−0.44 (−1.51 to 0.63)
Leisure-related moderate physical activity	−0.30 (−0.78 to 0.18)	−0.16 (−0.41 to 0.09)	0.13 (−0.11 to 0.37)	−0.04 (−0.22 to 0.14)	−0.30 (−1.16 to 0.57)
Total physical activity	−0.12 (−0.27 to 0.03)	−0.05 (−0.13 to 0.02)	0.02 (−0.05 to 0.09)	0.02 (−0.04 to 0.07)	−0.14 (−0.40 to 0.13)

^a*b* is the unstandardized regression coefficient; based on the Checklist Individual Strength questionnaire (CIS20-R) scores. All models were adjusted for age, sex, marital status, highest education, gross annual household income, and baseline fatigue.

Discussion

Principal Findings

This study is one of the first attempts to empirically examine the longitudinal changes in workers' domain-specific sedentary behaviors, physical activity, and fatigue during the COVID-19 outbreak. It also investigated the associations of the changes in workers' sedentary behaviors and physical activity with changes in their fatigue. Compared with before the COVID-19 pandemic, we found that workers spent more time (ie, 5–45 minutes per day) engaged in several domain-specific sedentary behaviors, including work-related sitting, TV viewing time during workdays, PC use sitting time, and total sitting time during the COVID-19 pandemic. They also reported less time engaged in leisure-related vigorous physical activity and total physical activity during the COVID-19 outbreak. Several cross-sectional and longitudinal studies conducted among other population groups echoed similar patterns in how the COVID-19 pandemic unfavorably affected people's active movement behaviors [25–30]. For example, a national study conducted in Canada found that children and youth reported higher sedentary behaviors and lower physical activity during the COVID-19 outbreak; the study found that less than 5% of children and less than 1% of youth met the combined 24-hour movement behavior goals for physical activity, sedentary behavior, and sleep [25]. Another study conducted in China found that adults' average steps per day and average moderate- or vigorous-intensity exercise significantly decreased during the semilockdown period of the COVID-19 pandemic [28]. Moreover, a study conducted among hypertensive older adults found that there was an increase in accelerometer-based sedentary behavior and a decrease in physical activity during the COVID-19 pandemic [29]. Our longitudinal findings are the first to confirm that COVID-19-related social distancing policies negatively affect sedentary and active behaviors among company workers, a population already at risk for being physically inactive [31,32].

Our findings also identified changes in sedentary behaviors and physical activity for each domain during the COVID-19 outbreak. Influenced by the government's self-isolation advice [33], leisure-related vigorous physical activity significantly decreased during the COVID-19 outbreak. The pre-COVID-19 level of leisure-related vigorous physical activity was only 15 minutes per day in our sample, which became 12 minutes per day during the COVID-19 pandemic. Since only a few people engaged in this activity, even in non-COVID-19 times, this pandemic has almost obliterated the vigorous physical activity that took place in the leisure time of the Japanese worker population. We also found that participants reported more work-related sitting time during the COVID-19 outbreak. Company workers usually sit when doing their work, and some of that additional sitting was due to an increase in the use of computers for all work-related meetings during the COVID-19 pandemic. Several companies in Japan have introduced and supported some level of teleworking amid the outbreak of COVID-19. While teleworking may lead to less time traveling for work, it is also likely that teleworking limits the opportunities for employees to accumulate transport-related physical activity and intensifies workers' sedentary activities, especially during a pandemic. Further studies are needed to investigate the influence of teleworking on workers' sedentary behaviors and physical activity. We also found that participants spent more sitting time on watching TV or using the PC. One reason for these increases is that people were more likely to avoid going outside following the Japan government's advice to minimize nonessential trips and reduce train operating hours in some areas. Staying at home limits people's ability to engage in physical activity and socialize with their friends and family and increases screen sitting time. Future studies can identify home-based interventions to improve people's active behaviors during a pandemic.

We found that increases in several domain-specific sedentary behaviors were associated with unfavorable changes in the

motivation and physical activity aspects of fatigue. Several previous studies have shown adverse effects on fatigue from too much sitting [3,11,12]. While yet to be tested in other populations, public health strategies to reduce sedentary behaviors among company workers could have positive effects on worker fatigue. Reducing sedentary behavior has been highlighted as a more feasible strategy during the COVID-19 pandemic than achieving optimal exercise levels [34]. However, teleworking from home, a new norm resulting from COVID-19-related social distancing policies, may impose a new challenge to reducing occupational sitting time in workers [35]. Future research is needed to identify how teleworking may influence sedentary behaviors and how to develop new policies to reduce workers' sitting time in the new era of teleworking from home. Furthermore, we found a significant increase in subjective fatigue among workers during the COVID-19 outbreak. Although the exact reasons for increased subjective fatigue in workers during the COVID-19 pandemic remain to be elucidated, the burden of this pandemic on people's mental health may be involved [36]. Several studies have provided preliminary evidence on the harmful effects of the COVID-19 pandemic on mental health [37-39]. For instance, a national study conducted in the United States found that the prevalence of depression symptoms was more than 3 times higher during the COVID-19 outbreak than before the outbreak [37]. Another study conducted in the Republic of Ireland found that approximately 30% of the participants were screened positive for generalized anxiety disorder or depression during the first week of the COVID-19 lockdown [39]. A sedentary lifestyle, intensified by social distancing policies, may also play a role in the observed increase in worker fatigue since the onset of the COVID-19 pandemic.

Limitations and Strengths

This study has some limitations. Our self-reported measures of sedentary behaviors and physical activity may be subject to memory and recall bias [40]. While this is a national study of participants recruited from throughout Japan, we are unable to confirm the generalizability of the data from our sample to all company workers. Data on employment status, including full-time or part-time employment, were also unavailable. The physical activity aspect of fatigue had low internal reliability at both baseline and follow-up. Since subjective fatigue may also lead to less physical activity and more sedentary behaviors, the direction of the observed associations remains inconclusive. Our study has several unique strengths, including the use of national data, the focus on the less-studied population group of company workers, and the examination of domain-specific sedentary behaviors and physical activity at different intensities.

Conclusions

Following the COVID-19 outbreak, there have been many concerns regarding its effects on people's physical inactivity and chronic diseases [41-43]. Our findings demonstrated that sedentary and active behaviors among company workers in Japan were negatively affected during the COVID-19 outbreak. Increases in several domain-specific sedentary behaviors also contributed to unfavorable changes in workers' fatigue. Social distancing and teleworking amid a pandemic may contribute to the sedentary lifestyle of company workers. Public health interventions are needed to mitigate the negative effects of the COVID-19 pandemic or future pandemics on sedentary and active behaviors and fatigue among company workers. More evidence is needed to identify the magnitude of changes in sedentary and active behaviors in workers during the COVID-19 outbreak and how these changes may influence workers' health and well-being.

Acknowledgments

TN was supported by the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research program (grant 20H00040). GM is supported by a Canadian Institutes of Health Research Foundation Grant (grant FDN-154331). KO is supported by the JSPS Grants-in-Aid for Scientific Research program (grant 20H04113).

Authors' Contributions

MK and KO conceived the idea of the study, analyzed the data, and drafted the paper. TN assisted with the statistical analysis. TN, GM, AS, and KI contributed to the writing of the paper and assisted with the interpretation. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Dempsey P, Matthews C, Dashti S, Doherty A, Bergouignan A, van Roekel EH, et al. Sedentary behavior and chronic disease: Mechanisms and future directions. *J Phys Act Health* 2020 Jan 01;17(1):52-61. [doi: [10.1123/jpah.2019-0377](https://doi.org/10.1123/jpah.2019-0377)] [Medline: [31794961](https://pubmed.ncbi.nlm.nih.gov/31794961/)]
2. de Vries JD, Claessens BJC, van Hooff MLM, Geurts SAE, van den Bossche SNJ, Kompier MAJ. Disentangling longitudinal relations between physical activity, work-related fatigue, and task demands. *Int Arch Occup Environ Health* 2016 Jan;89(1):89-101 [FREE Full text] [doi: [10.1007/s00420-015-1054-x](https://doi.org/10.1007/s00420-015-1054-x)] [Medline: [25952313](https://pubmed.ncbi.nlm.nih.gov/25952313/)]

3. Thorp AA, Kingwell BA, Owen N, Dunstan DW. Breaking up workplace sitting time with intermittent standing bouts improves fatigue and musculoskeletal discomfort in overweight/obese office workers. *Occup Environ Med* 2014 Nov;71(11):765-771. [doi: [10.1136/oemed-2014-102348](https://doi.org/10.1136/oemed-2014-102348)] [Medline: [25168375](https://pubmed.ncbi.nlm.nih.gov/25168375/)]
4. Karlsen K, Larsen JP, Tandberg E, Jørgensen K. Fatigue in patients with Parkinson's disease. *Mov Disord* 1999 Mar;14(2):237-241. [doi: [10.1002/1531-8257\(199903\)14:2<237::aid-mds1006>3.0.co;2-x](https://doi.org/10.1002/1531-8257(199903)14:2<237::aid-mds1006>3.0.co;2-x)] [Medline: [10091615](https://pubmed.ncbi.nlm.nih.gov/10091615/)]
5. Saremi M, Fallah MR. Subjective fatigue and medical errors among nurses in an educational hospital. *Iran Occup Health* 2013;10(4):1-8.
6. Sadeghniiat-Haghighi K, Yazdi Z. Fatigue management in the workplace. *Ind Psychiatry J* 2015;24(1):12-17 [FREE Full text] [doi: [10.4103/0972-6748.160915](https://doi.org/10.4103/0972-6748.160915)] [Medline: [26257477](https://pubmed.ncbi.nlm.nih.gov/26257477/)]
7. Connolly D, Fitzpatrick C, O'Toole L, Doran M, O'Shea F. Impact of fatigue in rheumatic diseases in the work environment: A qualitative study. *Int J Environ Res Public Health* 2015 Oct 28;12(11):13807-13822 [FREE Full text] [doi: [10.3390/ijerph121113807](https://doi.org/10.3390/ijerph121113807)] [Medline: [26516896](https://pubmed.ncbi.nlm.nih.gov/26516896/)]
8. Martin A, Chalder T, Rief W, Braehler E. The relationship between chronic fatigue and somatization syndrome: A general population survey. *J Psychosom Res* 2007 Aug;63(2):147-156. [doi: [10.1016/j.jpsychores.2007.05.007](https://doi.org/10.1016/j.jpsychores.2007.05.007)] [Medline: [17662751](https://pubmed.ncbi.nlm.nih.gov/17662751/)]
9. Williamson A, Friswell R. Fatigue in the workplace: Causes and countermeasures. *Fatigue* 2013 Apr;1(1-2):81-98. [doi: [10.1080/21641846.2012.744581](https://doi.org/10.1080/21641846.2012.744581)]
10. Lerman S, Eskin E, Flower D, George E, Gerson B, Hartenbaum N. Fatigue risk management in the workplace. *J Occup Environ Med* 2012;54(2):231-258. [doi: [10.1097/jom.0b013e318247a3b0](https://doi.org/10.1097/jom.0b013e318247a3b0)]
11. Wennberg P, Boraxbekk C, Wheeler M, Howard B, Dempsey PC, Lambert G, et al. Acute effects of breaking up prolonged sitting on fatigue and cognition: A pilot study. *BMJ Open* 2016 Feb 26;6(2):e009630 [FREE Full text] [doi: [10.1136/bmjopen-2015-009630](https://doi.org/10.1136/bmjopen-2015-009630)] [Medline: [26920441](https://pubmed.ncbi.nlm.nih.gov/26920441/)]
12. Engberg I, Segerstedt J, Waller G, Wennberg P, Eliasson M. Fatigue in the general population—Associations to age, sex, socioeconomic status, physical activity, sitting time and self-rated health: The northern Sweden MONICA study 2014. *BMC Public Health* 2017 Aug 14;17(1):654 [FREE Full text] [doi: [10.1186/s12889-017-4623-y](https://doi.org/10.1186/s12889-017-4623-y)] [Medline: [28806984](https://pubmed.ncbi.nlm.nih.gov/28806984/)]
13. Ghebreyesus T. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. World Health Organization. 2020 Mar 11. URL: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> [accessed 2021-03-07]
14. [COVID-19] Declaration of a state of emergency in response to the novel coronavirus disease. The Prime Minister of Japan and His Cabinet. 2020 Apr 16. URL: https://japan.kantei.go.jp/ongoingtopics/_00020.html [accessed 2021-03-07]
15. Tison GH, Avram R, Kuhar P, Abreau S, Marcus GM, Pletcher MJ, et al. Worldwide effect of COVID-19 on physical activity: A descriptive study. *Ann Intern Med* 2020 Nov 03;173(9):767-770. [doi: [10.7326/m20-2665](https://doi.org/10.7326/m20-2665)]
16. Bartoszek A, Walkowiak D, Bartoszek A, Kardas G. Mental well-being (depression, loneliness, insomnia, daily life fatigue) during COVID-19 related home-confinement-A study from Poland. *Int J Environ Res Public Health* 2020 Oct 12;17(20):7417 [FREE Full text] [doi: [10.3390/ijerph17207417](https://doi.org/10.3390/ijerph17207417)] [Medline: [33053787](https://pubmed.ncbi.nlm.nih.gov/33053787/)]
17. Vercoulen JH, Swanink CM, Fennis JF, Galama JM, van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. *J Psychosom Res* 1994 Jul;38(5):383-392. [doi: [10.1016/0022-3999\(94\)90099-x](https://doi.org/10.1016/0022-3999(94)90099-x)]
18. Aratake Y, Tanaka K, Wada K, Watanabe M, Katoh N, Sakata Y, et al. Development of Japanese version of the Checklist Individual Strength questionnaire in a working population. *J Occup Health* 2007 Nov;49(6):453-460 [FREE Full text] [doi: [10.1539/joh.49.453](https://doi.org/10.1539/joh.49.453)] [Medline: [18075205](https://pubmed.ncbi.nlm.nih.gov/18075205/)]
19. Beurskens A, Bültmann U, Kant I, Vercoulen J, Bleijenberg G, Swaen G. Fatigue among working people: Validity of a questionnaire measure. *Occup Environ Med* 2000 May;57(5):353-357 [FREE Full text] [doi: [10.1136/oem.57.5.353](https://doi.org/10.1136/oem.57.5.353)] [Medline: [10769302](https://pubmed.ncbi.nlm.nih.gov/10769302/)]
20. Ishii K, Shibata A, Kurita S, Yano S, Inoue S, Sugiyama T, et al. Validity and reliability of Japanese-language self-reported measures for assessing adults domain-specific sedentary time. *J Epidemiol* 2018;28(3):149-155. [doi: [10.2188/jea.je20170002](https://doi.org/10.2188/jea.je20170002)]
21. Armstrong T, Bull F. Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *J Public Health* 2006 Mar 2;14(2):66-70. [doi: [10.1007/s10389-006-0024-x](https://doi.org/10.1007/s10389-006-0024-x)]
22. Global Physical Activity Questionnaire (GPAQ) Analysis Guide. Geneva, Switzerland: World Health Organization; 2012. URL: https://www.who.int/ncds/surveillance/steps/resources/GPAQ_Analysis_Guide.pdf [accessed 2021-03-07]
23. Bull F, Maslin T, Armstrong T. Global Physical Activity Questionnaire (GPAQ): Nine country reliability and validity study. *J Phys Act Health* 2009 Nov;6(6):790-804. [doi: [10.1123/jpah.6.6.790](https://doi.org/10.1123/jpah.6.6.790)] [Medline: [20101923](https://pubmed.ncbi.nlm.nih.gov/20101923/)]
24. 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 Dec 06;17(1):162 [FREE Full text] [doi: [10.1186/s12874-017-0442-1](https://doi.org/10.1186/s12874-017-0442-1)] [Medline: [29207961](https://pubmed.ncbi.nlm.nih.gov/29207961/)]
25. Moore SA, Faulkner G, Rhodes RE, Brussoni M, Chulak-Bozzer T, Ferguson LJ, et al. Impact of the COVID-19 virus outbreak on movement and play behaviours of Canadian children and youth: A national survey. *Int J Behav Nutr Phys Act* 2020 Jul 06;17(1):85 [FREE Full text] [doi: [10.1186/s12966-020-00987-8](https://doi.org/10.1186/s12966-020-00987-8)] [Medline: [32631350](https://pubmed.ncbi.nlm.nih.gov/32631350/)]
26. López-Bueno R, Calatayud J, Andersen LL, Balsalobre-Fernández C, Casaña J, Casajús JA, et al. Immediate impact of the COVID-19 confinement on physical activity levels in Spanish adults. *Sustainability* 2020 Jul 16;12(14):5708. [doi: [10.3390/su12145708](https://doi.org/10.3390/su12145708)]

27. Wang Y, Zhang Y, Bennell K, White DK, Wei J, Wu Z, et al. Physical distancing measures and walking activity in middle-aged and older residents in Changsha, China, during the COVID-19 epidemic period: Longitudinal observational study. *J Med Internet Res* 2020 Oct 26;22(10):e21632 [FREE Full text] [doi: [10.2196/21632](https://doi.org/10.2196/21632)] [Medline: [33027035](https://pubmed.ncbi.nlm.nih.gov/33027035/)]
28. He M, Xian Y, Lv X, He J, Ren Y. Changes in body weight, physical activity, and lifestyle during the semi-lockdown period after the outbreak of COVID-19 in China: An online survey. *Disaster Med Public Health Prep* 2020 Jul 14;1-6 [FREE Full text] [doi: [10.1017/dmp.2020.237](https://doi.org/10.1017/dmp.2020.237)] [Medline: [32660669](https://pubmed.ncbi.nlm.nih.gov/32660669/)]
29. Browne RA, Macêdo GAD, Cabral LL, Oliveira GT, Vivas A, Fontes EB, et al. Initial impact of the COVID-19 pandemic on physical activity and sedentary behavior in hypertensive older adults: An accelerometer-based analysis. *Exp Gerontol* 2020 Dec;142:111121 [FREE Full text] [doi: [10.1016/j.exger.2020.111121](https://doi.org/10.1016/j.exger.2020.111121)] [Medline: [33132145](https://pubmed.ncbi.nlm.nih.gov/33132145/)]
30. Hu Z, Lin X, Chiwanda Kaminga A, Xu H. Impact of the COVID-19 epidemic on lifestyle behaviors and their association with subjective well-being among the general population in Mainland China: Cross-sectional study. *J Med Internet Res* 2020 Aug 25;22(8):e21176 [FREE Full text] [doi: [10.2196/21176](https://doi.org/10.2196/21176)] [Medline: [32759103](https://pubmed.ncbi.nlm.nih.gov/32759103/)]
31. Parry S, Straker L. The contribution of office work to sedentary behaviour associated risk. *BMC Public Health* 2013 Apr 04;13:296 [FREE Full text] [doi: [10.1186/1471-2458-13-296](https://doi.org/10.1186/1471-2458-13-296)] [Medline: [23557495](https://pubmed.ncbi.nlm.nih.gov/23557495/)]
32. Waters CN, Ling EP, Chu AHY, Ng SHX, Chia A, Lim YW, et al. Assessing and understanding sedentary behaviour in office-based working adults: A mixed-method approach. *BMC Public Health* 2016 Apr 27;16:360 [FREE Full text] [doi: [10.1186/s12889-016-3023-z](https://doi.org/10.1186/s12889-016-3023-z)] [Medline: [27117178](https://pubmed.ncbi.nlm.nih.gov/27117178/)]
33. Prevention Measures Against Coronavirus Disease 2019 (COVID-19). Tokyo, Japan: Japanese Ministry of Health, Labour and Welfare; 2020 Feb 25. URL: <https://www.mhlw.go.jp/content/10900000/000607599.pdf> [accessed 2021-03-07]
34. Zieff G, Bates L, Kerr Z, Moore J, Hanson E, Battaglini C, et al. Targeting sedentary behavior as a feasible health strategy during COVID-19. *Transl Behav Med* 2020 Oct 29;1-6 [FREE Full text] [doi: [10.1093/tbm/ibaa101](https://doi.org/10.1093/tbm/ibaa101)] [Medline: [33118596](https://pubmed.ncbi.nlm.nih.gov/33118596/)]
35. Chtourou H, Trabelsi K, H'mida C, Boukhris O, Glenn JM, Brach M, et al. Staying physically active during the quarantine and self-isolation period for controlling and mitigating the COVID-19 pandemic: A systematic overview of the literature. *Front Psychol* 2020;11:1708 [FREE Full text] [doi: [10.3389/fpsyg.2020.01708](https://doi.org/10.3389/fpsyg.2020.01708)] [Medline: [33013497](https://pubmed.ncbi.nlm.nih.gov/33013497/)]
36. Salari N, Hosseini-Far A, Jalali R, Vaisi-Raygani A, Rasoulpoor S, Mohammadi M, et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: A systematic review and meta-analysis. *Global Health* 2020 Jul 06;16(1):57 [FREE Full text] [doi: [10.1186/s12992-020-00589-w](https://doi.org/10.1186/s12992-020-00589-w)] [Medline: [32631403](https://pubmed.ncbi.nlm.nih.gov/32631403/)]
37. Ettman CK, Abdalla SM, Cohen GH, Sampson L, Vivier PM, Galea S. Prevalence of depression symptoms in US adults before and during the COVID-19 pandemic. *JAMA Netw Open* 2020 Sep 01;3(9):e2019686 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.19686](https://doi.org/10.1001/jamanetworkopen.2020.19686)] [Medline: [32876685](https://pubmed.ncbi.nlm.nih.gov/32876685/)]
38. De Boni RB, Balanzá-Martínez V, Mota JC, Cardoso TDA, Ballester P, Atienza-Carbonell B, et al. Depression, anxiety, and lifestyle among essential workers: A web survey from Brazil and Spain during the COVID-19 pandemic. *J Med Internet Res* 2020 Oct 30;22(10):e22835 [FREE Full text] [doi: [10.2196/22835](https://doi.org/10.2196/22835)] [Medline: [33038075](https://pubmed.ncbi.nlm.nih.gov/33038075/)]
39. Hyland P, Shevlin M, McBride O, Murphy J, Karatzias T, Bentall RP, et al. Anxiety and depression in the Republic of Ireland during the COVID-19 pandemic. *Acta Psychiatr Scand* 2020 Sep;142(3):249-256. [doi: [10.1111/acps.13219](https://doi.org/10.1111/acps.13219)] [Medline: [32716520](https://pubmed.ncbi.nlm.nih.gov/32716520/)]
40. Ainsworth B, Caspersen C, Matthews C, Mâsse LC, Baranowski T, Zhu W. Recommendations to improve the accuracy of estimates of physical activity derived from self report. *J Phys Act Health* 2012 Jan;9 Suppl 1:S76-S84 [FREE Full text] [doi: [10.1123/jpah.9.s1.s76](https://doi.org/10.1123/jpah.9.s1.s76)] [Medline: [22287451](https://pubmed.ncbi.nlm.nih.gov/22287451/)]
41. Smirmaul B, Arena R. The urgent need to sit less and move more during the COVID-19 pandemic. *J Cardiopulm Rehabil Prev* 2020;40(5):287-289. [doi: [10.1097/hcr.0000000000000538](https://doi.org/10.1097/hcr.0000000000000538)]
42. Peçanha T, Goessler KF, Roschel H, Gualano B. Social isolation during the COVID-19 pandemic can increase physical inactivity and the global burden of cardiovascular disease. *Am J Physiol Heart Circ Physiol* 2020 Jun 01;318(6):H1441-H1446 [FREE Full text] [doi: [10.1152/ajpheart.00268.2020](https://doi.org/10.1152/ajpheart.00268.2020)] [Medline: [32412779](https://pubmed.ncbi.nlm.nih.gov/32412779/)]
43. Pinto AJ, Dunstan DW, Owen N, Bonfá E, Gualano B. Combating physical inactivity during the COVID-19 pandemic. *Nat Rev Rheumatol* 2020 Jul;16(7):347-348 [FREE Full text] [doi: [10.1038/s41584-020-0427-z](https://doi.org/10.1038/s41584-020-0427-z)] [Medline: [32355296](https://pubmed.ncbi.nlm.nih.gov/32355296/)]

Abbreviations

CIS20-R: Checklist Individual Strength questionnaire

GPAQ: Global Physical Activity Questionnaire

Edited by T Sanchez; submitted 05.12.20; peer-reviewed by A Swartz, H Zihao; comments to author 23.01.21; revised version received 02.02.21; accepted 09.03.21; published 26.03.21.

Please cite as:

Koohsari MJ, Nakaya T, McCormack GR, Shibata A, Ishii K, Oka K

Changes in Workers' Sedentary and Physical Activity Behaviors in Response to the COVID-19 Pandemic and Their Relationships With Fatigue: Longitudinal Online Study

JMIR Public Health Surveill 2021;7(3):e26293

URL: <https://publichealth.jmir.org/2021/3/e26293>

doi: [10.2196/26293](https://doi.org/10.2196/26293)

PMID: [33727211](https://pubmed.ncbi.nlm.nih.gov/33727211/)

©Mohammad Javad Koohsari, Tomoki Nakaya, Gavin R McCormack, Ai Shibata, Kaori Ishii, Koichiro Oka. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 26.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Predicting Age Groups of Reddit Users Based on Posting Behavior and Metadata: Classification Model Development and Validation

Robert Chew¹, MSc; Caroline Kery¹, MSc; Laura Baum², MA, MPH; Thomas Bukowski³, MA, MPA; Annice Kim⁴, PhD; Mario Navarro⁵, PhD

¹Center for Data Science, RTI International, Research Triangle Park, NC, United States

²Center for Health Analytics, Media, and Policy, RTI International, Atlanta, GA, United States

³Center for Health Analytics, Media, and Policy, RTI International, Berkeley, CA, United States

⁴Center for Health Analytics, Media, and Policy, RTI International, Research Triangle Park, NC, United States

⁵Office of Health Communications and Education, Center for Tobacco Products, US Food and Drug Administration, Silver Spring, MD, United States

Corresponding Author:

Robert Chew, MSc
Center for Data Science
RTI International
3040 East Cornwallis Road
Research Triangle Park, NC, 27709
United States
Phone: 1 9195415823
Email: rchew@rti.org

Related Article:

This is a corrected version. See correction statement: <https://publichealth.jmir.org/2021/4/e30017>

Abstract

Background: Social media are important for monitoring perceptions of public health issues and for educating target audiences about health; however, limited information about the demographics of social media users makes it challenging to identify conversations among target audiences and limits how well social media can be used for public health surveillance and education outreach efforts. Certain social media platforms provide demographic information on followers of a user account, if given, but they are not always disclosed, and researchers have developed machine learning algorithms to predict social media users' demographic characteristics, mainly for Twitter. To date, there has been limited research on predicting the demographic characteristics of Reddit users.

Objective: We aimed to develop a machine learning algorithm that predicts the age segment of Reddit users, as either adolescents or adults, based on publicly available data.

Methods: This study was conducted between January and September 2020 using publicly available Reddit posts as input data. We manually labeled Reddit users' age by identifying and reviewing public posts in which Reddit users self-reported their age. We then collected sample posts, comments, and metadata for the labeled user accounts and created variables to capture linguistic patterns, posting behavior, and account details that would distinguish the adolescent age group (aged 13 to 20 years) from the adult age group (aged 21 to 54 years). We split the data into training (n=1660) and test sets (n=415) and performed 5-fold cross validation on the training set to select hyperparameters and perform feature selection. We ran multiple classification algorithms and tested the performance of the models (precision, recall, F1 score) in predicting the age segments of the users in the labeled data. To evaluate associations between each feature and the outcome, we calculated means and confidence intervals and compared the two age groups, with 2-sample t tests, for each transformed model feature.

Results: The gradient boosted trees classifier performed the best, with an F1 score of 0.78. The test set precision and recall scores were 0.79 and 0.89, respectively, for the adolescent group (n=254) and 0.78 and 0.63, respectively, for the adult group (n=161). The most important feature in the model was the number of sentences per comment (permutation score: mean 0.100, SD 0.004). Members of the adolescent age group tended to have created accounts more recently, have higher proportions of

submissions and comments in the r/teenagers subreddit, and post more in subreddits with higher subscriber counts than those in the adult group.

Conclusions: We created a Reddit age prediction algorithm with competitive accuracy using publicly available data, suggesting machine learning methods can help public health agencies identify age-related target audiences on Reddit. Our results also suggest that there are characteristics of Reddit users' posting behavior, linguistic patterns, and account features that distinguish adolescents from adults.

(*JMIR Public Health Surveill* 2021;7(3):e25807) doi:[10.2196/25807](https://doi.org/10.2196/25807)

KEYWORDS

Reddit; social media; age; machine learning; classification

Introduction

Public health campaigns are a primary means for government agencies and nongovernmental organizations to raise awareness about important health issues affecting their communities. Running effective public health education campaigns requires organizations to tailor messages and provide outreach that resonates with their target audience. For example, the US Food and Drug Administration (FDA) Center for Tobacco Products has developed separate public education campaigns on the health consequences of tobacco products for at-risk teens [1]; lesbian, gay, bisexual, and transgender young adults [2]; multicultural youth [3]; rural male teens [4]; and adult smokers [5]. With increased media consumption and interpersonal interactions occurring online, social media platforms have become important in both engaging target audiences in public education campaigns and in understanding behaviors and perceptions around emerging public health issues across these target audiences. However, it can be challenging to apply results from social media analyses because there is limited information about who these data represent, and thus it is difficult to glean information from populations of interest. In tobacco prevention and control, being able to segment social media posts by age-based audience segments, would help researchers and public health agencies identify emerging issues and changes in behaviors and attitudes to facilitate public health surveillance and educational outreach to these at-risk populations. This would allow researchers to naturalistically observe of their target audience and how they interact with the discussion of tobacco products. This study presents a model developed to classify users of the popular social media site *Reddit* into those more likely to be older or younger than 21 years of age. These age categories mirror the federal minimum age for the purchase of tobacco products in the United States, which was amended from 18 years to 21 years of age [6].

Reddit is a public network of online communities organized around people's interests. As of 2019, an estimated 11% of US adults use Reddit, with an estimated 22% of US adults aged 18 to 29 years using the platform [7]. Unlike other social media platforms such as Twitter that restrict post length, Reddit was designed for long-form entries, which allows users to provide much greater levels of detail in their posts. In addition, Reddit, similar to traditional forums, allows users to respond to posted material in the form of comment chains. This combination of longer post length and an engaged user base encourages active and nuanced discussions across a myriad of topics relevant to

the public health community. In recent years, researchers have used Reddit data to understand emerging public health issues, including cannabis [8-10], opioid [11,12], and alcohol [13,14] use.

Recent studies have used Reddit to investigate underage electronic nicotine delivery systems use. Brett et al [15] performed a content analysis on Reddit posts to identify perceptions among Redditors about underage use of Juul. Zhan et al [16] coded conversations about youth using Juul on the now-banned UnderageJuul subreddit. For studies focused on underage electronic nicotine delivery systems use, the authors used text indicators in posts or the subreddit community's age as proxies for identifying conversations among youth. However, not everyone posting on subreddits such as UnderageJuul are likely youth, and it is likely that youth are having conversations about vaping in other Reddit communities outside of electronic nicotine delivery systems specific subreddits. For these reasons, a broader approach is needed to identify youth who are discussing electronic nicotine delivery systems across Reddit communities.

One way to identify adolescent conversations about electronic nicotine delivery systems is to examine the characteristics of Reddit users' posting behavior, look for patterns that distinguish adolescents from adults, and develop an algorithm to predict the age of Reddit users. Predicting latent user demographics on social media is a popular area of research, with researchers developing models to infer categories such as gender [17-19], age [20-22], race and ethnicity [23-25], and political orientation [26-28]. Specifically, latent attribute models trained on Reddit data are becoming increasingly popular, with models for inferring geolocation [29], gender [30], and mental health status [31,32], all recently published.

The objective of this study was to develop an age prediction algorithm to classify Reddit users as adolescents or adults, based on publicly available data, to build a tool that would allow for more nuanced analyses of tobacco related discussions occurring on Reddit. This would allow comparisons of adult and adolescent discussions. Our work extends the literature by developing a model with age categories that are meaningful for a broad public health audience.

Methods

Ethics

This study was conducted between January and September 2020. It was determined that the study did not entail research with human participants as defined by 45 CFR 46 by the RTI International review board (STUDY00021174) and the FDA Research Involving Human Subjects Committee (2020-CTP-129).

Identifying Self-Reporting of Age on Reddit

Developing supervised classification models requires ground truth observations labeled with the outcome of interest. To obtain the ages of Reddit users, we used a social media data platform (Brandwatch) to identify publicly available Reddit posts or comments in which users self-reported their age. Brandwatch is a secure cloud computing platform accessible only to registered users, and data that are extracted contain no personal information beyond the publicly posted username. Extracted data were maintained on secure internal servers and no attempts were made to link usernames to individuals. We searched for English language Reddit posts with the phrase “I’m <age>” that were posted between December 1 and December 31, 2019. In the literature, similar approaches have been used to construct large Reddit data sets containing self-reported demographics [33-35]. We were interested in youth (13-17 years), young adult (18-20 years), and adult (21-54 years) age segments. For the youth segment, the search query was “I’m 13” OR “I’m 14” OR “I’m 15,” OR “I’m 16” OR “I’m 17.” A similar strategy was used to identify posts for the young adult (18-20 years) and adult (21-54 years) age categories. The upper age range was chosen to correspond with the target audience’s age range for FDA campaigns, and age range for young and middle age ranges were chosen to reflect the previous (18 years) and current (21 years) minimum age for sale of tobacco products

in the United States. The query pool included both original submissions and comments on posts to maximize the likelihood of identifying users reporting their age.

We reviewed the search results and employed multiple data cleaning steps to prepare the datafile for manual coding. First, we manually reviewed the posts pulled and found that some used “I’m <number>” phrases that did not refer to age such as “I’m 50% done” or “I’m 5ft tall.” To exclude these posts, we dropped all posts where we could not find the phrase “I’m <number>” followed by a space or punctuation. Second, posts from subreddits labeled on Reddit’s list of subreddits as “not suitable for work” or “not suitable for life” were excluded to protect coders from sensitive or harmful content. Finally, if a user posted multiple times during the time period observed, we only included the most recent post.

Manual Labeling of Age from Reddit Posts

Posts for each age range were imported into the SMART annotation app [36] for coding. We trained 3 coders on coding measures and how to use the SMART platform. All coders coded the same test sample of 30 posts or comments across all age groups. There was high interrater reliability across the coders: 82% unadjusted percent agreement and Cohen κ =0.867. Based on the level of agreement for this sample of 30 posts or comments, the full coding was conducted independently.

Coders manually reviewed and labeled self-reported age in posts until they reached a minimum of 625 cases within each of the target age categories. We targeted 625 labeled cases so that we could have at least 500 labeled cases per age category to train and cross validate the algorithm (80%) and 125 labeled cases per age category for a final test set (20%). In Table 1, we report the final number of posts coded for age, as well as the number of posts excluded because they were not relevant or age could not be determined.

Table 1. Number of manually labeled Reddit posts by age category (December 2019).

Category	Coded posts or comments, n
All ages	2156
13-17 years	683
18-20 years	642
21-54 years	831
Excluded ^a	
Cannot determine age	252
Not relevant	154

^aWe excluded posts if age could not be determined (eg, number provided not related to age; multiple ages provided; different language) or if the post was not relevant (eg, user explicitly states they are based outside of the United States; from throwaway accounts).

Originally, we planned on modeling 3 separate age categories (13-17 years, 18-20 years, and 21-54 years), collecting a nearly balanced labeled data set of observations across the 3 ranges of interest; however, after repeated experimentation with the 3-age category model, correctly categorizing members of the middle range (18-20 years) continued to be difficult, and members were often mischaracterized as belonging to either the young or older age range. Given the poor performance for this age range and

the recent increase in the US federal age threshold (from 18 years to 21 years), under which the sale of tobacco products, we collapsed the young (13-17 years) and middle (18-20 years) categories into a single category (13-20 years). The implications of this decision are included in the Discussion.

Collecting User-, Submission-, and Comment-Level Metadata for Manually Labeled Reddit User Accounts

Once the final set of Reddit user accounts were labeled for each age range, we collected publicly available user-, submission- (originating post), and comment-level metadata using Reddit's application programming interface (API). An API is an interface that receives requests for data or services from outside programs and sends back replies. Reddit's API allows for registered apps to have limited access to Reddit metadata.

For each Reddit user, we collected their user-level metadata, which included properties of their account, such as the year it was created and their status on Reddit. No personally identifiable information, such as the user's demographic information or their location, is available through the Reddit API. Some Reddit users had deleted their accounts since we had first identified their age-related post; therefore, we were unable to collect their information. For Reddit users with active accounts, we collected the metadata fields for up to 100 of their most recent submissions and up to 100 of their most recent comments. In cases where the Reddit user did not have 100 submissions or comments, we collected as many as were available. Submissions and comments prior to 2019 were omitted from analyses to prevent the influence of significantly older posts that may have been posted when a user was in a younger age category, and 85 accounts were excluded from the analysis because they were

made private or deleted between the initial collection of self-reported age posts and the API calls for collecting user information.

The Reddit API provides detailed metadata fields for each user profile, submission, and comment. The Reddit API returned 50 metadata fields for each user profile, 12,855 metadata fields for each submission, and 322 metadata fields for each comment for a total of 13,227 metadata fields per account. Although the Reddit API returns many data fields, not all were useful for training the algorithm. First, we excluded metadata fields that had 95% or greater missing data ($n=7888$) or contained no variation ($n=5180$). Second, we excluded metadata fields that were related to page cosmetics such as text color or video width ($n=28$). Third, we excluded several complex data types with high cardinality (ie, very uncommon or unique), such as images, HTML, and URLs ($n=44$).

Variables for Age Prediction

The fields that we examined for predicting age of Reddit users included summary statistics of metadata fields from the API and other derived variables that may help distinguish adolescent from older age groups (eg, subreddit frequencies, emoji use, reading level of text). In total, we moved into the modeling phase with 1523 variables. Table 2 provides a summary of the variable groups created, as well as counts and examples from each group.

Table 2. Variables for modeling age of Reddit users in the training set derived from the comment, post, and user data collected for users whose ages were confirmed by manual labeling.

Variable group	Metadata used	Example	Variables (N=1523), n
Summary statistics	All	Median post score	189
Subreddit frequencies	Posts and comments	Frequency user posted to "Teenagers" subreddit	624
Emoji frequencies	Comments	Frequency of "👍" out of emojis used by user	101
Literary characteristics	Posts and comments	Average Flesch Reading Ease score	28
Patterns in Posting	Posts and comments	Percentage of user's posts that were videos	42
Term usage	Comments	TF-IDF ^a score for the term "school"	539 ^b

^aTF-IDF: term frequency-inverse document frequency.

^bThe number of TF-IDF terms varied across the cross-validation folds based on the comments and submissions vocabulary present in the training portion of each fold. The value presented here is the number of TF-IDF features when calculated on text from the full training set.

Summary Statistics of Metadata Fields From Reddit API

The first set of variables derived to predict age were constructed directly from the numeric and categorical metadata fields. Fields reporting on user-level attributes were used directly, whereas fields for submissions or comments were derived from aggregations of submissions or comments for each user. For each numeric comment or submission field such as "ups" (the number of upvotes for a post), we derived variables for summarizing their mean, standard deviation, median, first quartile, and third quartile. For binary fields, we calculated the

proportion of comments or submissions for which the field's value was 1. This resulted in 189 summary statistics variables.

Variables Derived to Help Distinguish Age Groups

In addition to using fields directly available from the Reddit API, we derived further variables to help distinguish between age groups, such as subreddit frequencies, emoji frequencies, literary characteristics, posting patterns, and term usage in comments. In total, there were 624 derived subreddit frequency variables, 101 emoji frequency variables, 28 literary characteristic variables (14 for comment text, 14 for submission text), 42 posting pattern variables, and 539 term usage variables.

Given the importance of text features in previous age classification modeling for social media [37], more detailed information is provided for these sets of variables. For the first set, we turned to established research on Facebook posts by the World Well-Being Project (WWBP) at the University of Pennsylvania. The WWBP [38] had previously identified terms that were used more often by people in 1 of 4 age ranges (13-18 years, 19-22 years, 23-29 years, >30 years). To incorporate this information, each Reddit comment was split into unique words (unigrams), adjacent words (bigrams), and adjacent-word triplets (trigrams). A variable was created for each of the WWBP age group term lists by calculating how many times a term from the list was present in a user's comment, divided by the total number of n-grams in that user's comments.

The second set of fields used the term frequency-inverse document frequency (TF-IDF) text representation, which gives a higher weight to terms used frequently within a text document that are also used infrequently across documents. To build these variables, we first cleaned the comment text of English stop words, including many prepositions, conjunctions, and pronouns. Although essential for grammatical English, these terms typically contribute less when there is a need to distinguish between differences in language usage across groups of interest. We then lemmatized each word, reducing conjugated words to their stems (eg, the lemma of *running* is *run*) to better recognize repeated usage of terms across conjugations. Finally, we dropped all numbers from the comment text to avoid age self-report terms that are more likely to occur in our sample than the broader Reddit population. From this processed text, we calculated the TF-IDF score for each unigram, bigram, or trigram between 0.25% and 90% of users in their comments. TF-IDF values were derived only from the data being used for training, whether independently for each cross-validation fold on the training set for hyperparameter tuning and model selection, or on the full training set for a final hold-out comparison to the test set. This resulted in approximately 539 variables when applied to the full training set.

Model Training

To assess the validity of our findings, we split the data into training and test sets using an 80/20 split. The samples are stratified by age group to ensure similar proportions of age groups across the training and test sets. We performed 5-fold cross validation on the training set to select appropriate hyperparameters and perform model selection to prevent overfitting on the test data.

Model Evaluation Metrics

To evaluate the model performance, we used area under the receiver operating characteristic curve (AUROC), precision, recall, and F1 score (Multimedia Appendix 1) to assess different aspects of the predictions. All metrics are bounded between 0 and 1, and with all else equal, a higher value indicates a better performing model.

Model Selection

We assessed several classification algorithms with different underlying assumptions to model age. Specifically, we ran variants of the following methods using the sklearn [39] machine

learning framework in Python (Python 3.7): logistic regression with regularization [40], support vector machine [41], *k*-nearest neighbor [42], random forest [43], and gradient boosted trees [44]. Testing several models is motivated by a version of the No Free Lunch Theorem [45], which states that there are no theoretical guarantees that any one machine learning algorithm will work best on all tasks, implicitly promoting an empirical approach to model selection for supervised learning.

Feature Selection

In addition to developing a model with strong predictive performance, we also aimed to create a model that is interpretable and makes errors easier to diagnose. To help develop a tractable model that mitigates the influence of noisy features, we performed all relevant feature selection [46] using the Boruta method [47] in Python. All relevant feature selection aims to find all features connected to classification decision, as opposed to the minimal optimal selection problem that aims to provide a compact set of features that mimics the optimal classifier. For our implementation, we used 100 iterations and $\alpha=.05$.

Feature Importance

To better understand which features contributed to the model predictions, we calculated permutation feature importance scores [43,48] for each feature in the model. The permutation feature importance is the increase in the model's prediction error after permuting the feature, holding all other features constant. Randomly shuffling the values of the feature effectively breaks the association between the permuted feature, other features, and the outcome. A higher score indicates that the model highly relies on that feature to make accurate predictions, since modifying that feature corresponds to a large drop in model performance. One caution when interpreting permutation importance scores is that the reported contribution from individual features may not be accurate if the features are strongly correlated. We constructed a feature correlation matrix to check this condition and did not find high multicollinearity between any of the remaining features (Figure S1 in Multimedia Appendix 2). Permutations for each feature were repeated 5 times, and the mean and standard deviation were reported.

Although permutation feature importance can help us understand which variables contribute most, it does not provide information on how the features vary across age groups. As a final assessment, we calculated the means, confidence intervals, and 2-sample *t* test comparisons for each transformed feature in the final model. The features were transformed prior to calculating means, intervals, and hypothesis testing because visual inspection of the distributions indicated that the normality assumption for the 2-sample *t* test would not be satisfied for a majority of the covariates. A quantile transformation was applied to each feature to make the sample more closely resemble a Gaussian normal distribution. This univariate transformation first estimates the cumulative distribution function of a variable to a uniform distribution and then maps the values to the desired output distribution (ie, normal distribution) using the associated quantile function.

Results

Model Selection

Table 3 summarizes the training set cross-validation model results for each classification algorithm. The reported results use each algorithm's best performing hyperparameter values discovered during hyperparameter tuning. The model with the highest overall F1 score is the gradient boosted tree (mean 0.77,

SD 0.03), followed by random forest (mean 0.75, SD 0.03), and support vector machines (mean 0.74, SD 0.04). The *k*-nearest neighbor model overall had the lowest F1 score (mean 0.59, SD 0.04) of the models considered. Gradient boosted trees were selected as the most appropriate algorithm for this modeling task of those assessed based on their overall higher cross-validation mean and comparable standard deviation for F1 scores across models.

Table 3. Classifier performance for predicting manually labeled age of Reddit users based on the full derived variable set.

Age group	Precision, mean (SD)	Recall, mean (SD)	F1 score, mean (SD)	AUROC ^a , mean (SD)	Support, n
Random forest					
13-20 years	0.75 (0.04)	0.91 (0.01)	0.82 (0.02)	— ^b	202
21-54 years	0.79 (0.03)	0.53 (0.04)	0.63 (0.03)	—	128
Overall	—	—	0.75 (0.03)	0.83 (0.02)	331
k-Nearest neighbors					
13-20 years	0.65 (0.04)	0.98 (0.01)	0.78 (0.03)	—	202
21-54 years	0.84 (0.04)	0.18 (0.04)	0.29 (0.05)	—	128
Overall	—	—	0.59 (0.04)	0.75 (0.03)	331
Support vector machine					
13-20 years	0.75 (0.04)	0.89 (0.03)	0.82 (0.03)	—	202
21-54 years	0.76 (0.06)	0.54 (0.05)	0.63 (0.04)	—	128
Overall	—	—	0.74 (0.04)	0.83 (0.02)	331
Gradient boosted trees					
13-20 years	0.79 (0.04)	0.85 (0.02)	0.82 (0.02)	—	202
21-54 years	0.74 (0.05)	0.64 (0.04)	0.68 (0.03)	—	128
Overall	—	—	0.77 (0.03)	0.84 (0.02)	331

^aAUROC: area under the receiver operating characteristics curve.

^bNot available or not applicable.

Feature Selection

All relevant feature selection reduced the number of features from 1523 to 15 (Table 4), resulting in a reduced model with comparable predictive performance (F1 score 0.78) to a model with all variables included (F1 score 0.79).

As a final diagnostic check, a weighted version of the reduced model was run to assess the impact of imbalanced sample sizes of the outcome categories introduced by collapsing age groups. The version of the models with and without the weighting adjustment performed comparably, suggesting that the class imbalance did not drastically impact model results (Table S1 in Multimedia Appendix 3).

Table 4. Model performance on the test set for predicting age group of Reddit users using gradient boosted trees classifier with all features vs select features.

Age group	Precision	Recall	F1 score	AUROC ^a	Support, n
Model with all features (1523 features)					
13-20 years	0.81	0.87	0.84	— ^b	254
21-54 years	0.77	0.67	0.72	—	161
Overall	—	—	0.79	0.87	415
Reduced model with selected features (15 features)					
13-20 years	0.79	0.89	0.84	—	254
21-54 years	0.78	0.63	0.70	—	161
Overall	—	—	0.78	0.86	415

^aAUROC: area under the receiver operating characteristics curve.

^bNot available or not applicable.

Feature Importance

Figure 1 shows the permutation feature importance scores for each variable in the model. Of the feature categories, the largest number (n=7) were derived from terms used in the comment text (TF-IDF weights for “home,” “look like,” “need,” “school,” and “totally”; frequency of WWBP 23-to-29 word set used, normalized count of WWBP 23-to-29 word set used). Three features describe subreddit usage (proportion of comments and posts in *r/teenagers*, proportion of comments and posts in *r/news*, 75th percentile of subscribers for subreddits posted to) and another 3 characterize the user’s literary style when

commenting (average Coleman Liau Index for comments, number of sentences per comment, proportion of user comments posted in a thread they started). Only 2 features were derived from user account information that does not require collecting additional posts or comments (account creation year, lifetime comment karma).

The most important model feature was the number of sentences per comment (score: mean 0.100; SD 0.004). The least important feature is the proportion of a user’s posts and comments in the *r/news* subreddit (score: mean 0.017; SD 0.002). Although the text features account for the largest number of relevant variables, features across the categories show high feature importance.

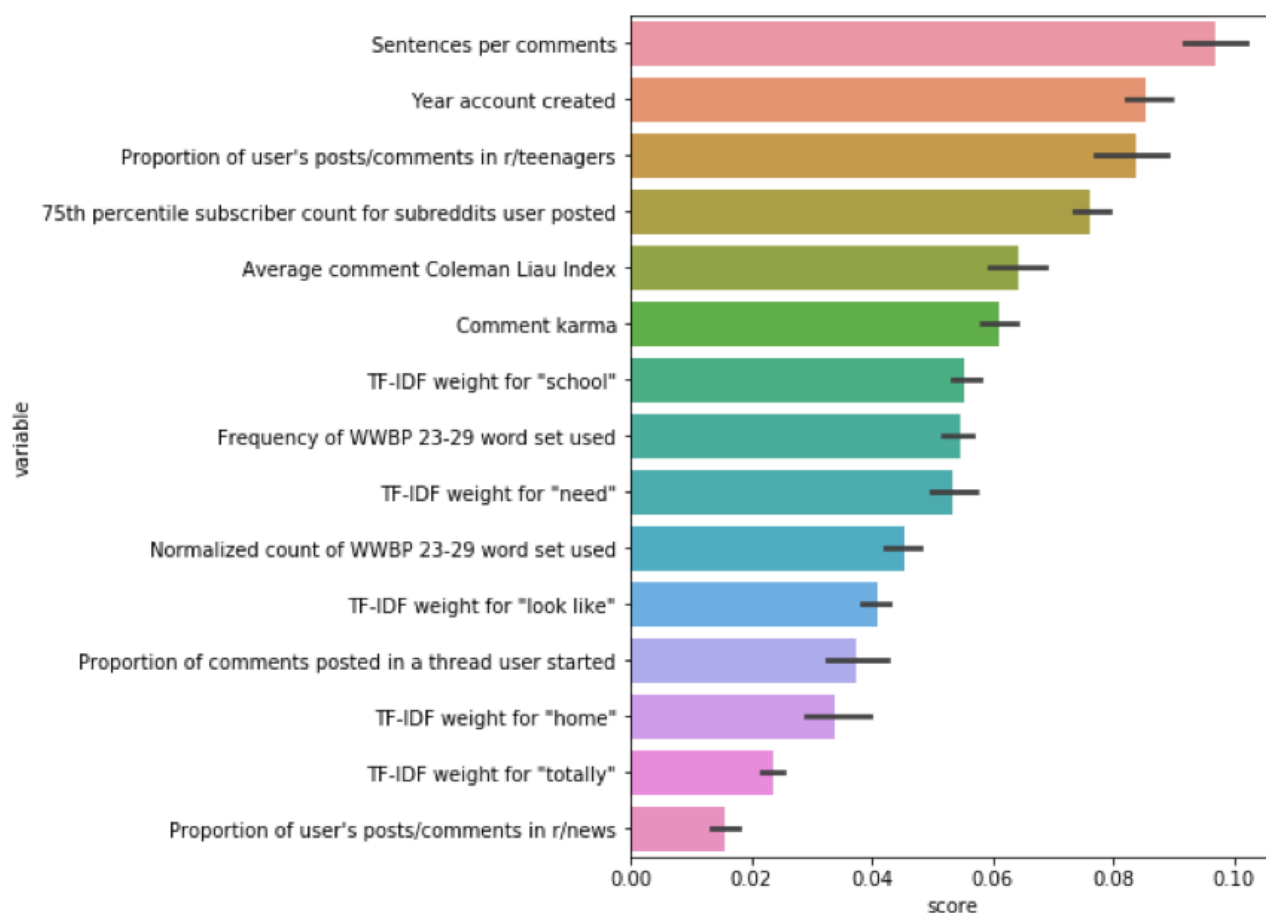
Figure 1. Permutation feature importance scores for top variables in the reduced model for predicting age of Reddit users.

Table 5 summarizes how the most important features varied across age groups. Means for all variables were significantly different, with the exception of the TF-IDF weight for “school,” which exhibited a spike at 0, for which the quantile transformation could not properly account. Table S2 in Multimedia Appendix 4 provides additional nonparametric results for detecting differences in sample distributions (Mann Whitney *U* test and Kolmogorov-Smirnov) that do not compare sample means and are less influenced if the sample transformations fail to meet normality; results from both tests suggest that distributions for the TF-IDF weight for the “school” variable were statistically different (Multimedia Appendix 4) between adolescents and adults in our sample.

For the top variables identified by the feature permutations, when comparing the 13 to 20 years age group, members of the older age group tended to use a larger number of sentences per comment (21-54 years: $\mu=0.37$; 13-20 years: $\mu=-0.43$), higher average comment Coleman Liau Index (21-54 years: $\mu=0.08$; 13-20 years: $\mu=-0.25$), and higher comment karma (21-54 years: $\mu=0.30$; 13-20 years: $\mu=-0.19$). Likewise, when compared with the 21 to 54 years age group, members of the younger 13 to 20 years age group tend to have created an account more recently (21-54 years: $\mu=1.35$; 13-20 years: $\mu=2.96$), have a higher proportion of submissions and comments in the *r/teenagers* subreddit (21-54 years: $\mu=-5.11$; 13-20 years: $\mu=-3.49$), and post more in subreddits with higher subscriber count (21-54 years: $\mu=0.33$; 13-20 years: $\mu=-0.14$).

Table 5. Age group comparisons for the top performing variables (by permutation feature importance).

Variable	Type	Age 13-20 years (n=1014), mean ^a (95% CI)	Age 21-54 years (n=643), mean ^a (95% CI)	t statistic (df)	P value
Sentences per comments	Literary characteristics	-0.43 (-0.52, -0.35)	0.37 (0.27, 0.46)	-12.31 (1489.89)	<.001
Year account created	Summary statistics	2.96 (2.79, 3.14)	1.35 (1.10, 1.59)	10.65 (1272.76)	<.001
Proportion of user's posts or comments in <i>r/teenagers</i>	Subreddit frequencies	-3.49 (-3.67, -3.31)	-5.11 (-5.17, -5.05)	16.69 (1208.10)	<.001
75th percentile subscriber count for subreddits user posted	Summary statistics	-0.14 (-0.23, -0.04)	-0.33 (-0.46, -0.20)	2.41 (1276.15)	.02
Average comment Coleman Liau Index	Literary characteristics	-0.25 (-0.34, -0.16)	0.08 (-0.01, 0.17)	-5.06 (1571.31)	<.001
Comment karma	Summary statistics	-0.19 (-0.25, -0.13)	0.30 (0.22, 0.38)	-9.85 (1309.14)	<.001
TF-IDF ^b weight for "school"	Term usage	-2.46 (-2.65, -2.27)	-2.63 (-2.86, -2.40)	1.14 (1406.18)	.25
Frequency of WWBP ^c 23-29 word set used	Term usage	-1.23 (-1.39, -1.08)	-0.05 (-0.20, 0.10)	-10.97 (1592.33)	<.001
TF-IDF weight for "need"	Term usage	-1.58 (-1.75, -1.41)	-0.31 (-0.48, -0.15)	-10.48 (1577.27)	<.001
Normalized count of WWBP 23-29 word set used	Term usage	-1.29 (-1.44, -1.14)	0.04 (-0.11, 0.19)	-12.35 (1569.17)	<.001
Proportion of comments posted in a thread user started	Summary statistics	-0.74 (-0.90, -0.57)	-1.41 (-1.62, -1.20)	4.95 (1366.37)	<.001
TF-IDF weight for "look like"	Term usage	-3.04 (-3.22, -2.86)	-1.84 (-2.08, -1.61)	-7.85 (1327.70)	<.001
TF-IDF weight for "home"	Term usage	-3.45 (-3.62, -3.28)	-1.89 (-2.14, -1.65)	-10.24 (1252.67)	<.001
TF-IDF weight for "totally"	Term usage	-4.23 (-4.37, -4.09)	-2.95 (-3.19, -2.71)	-8.92 (1092.36)	<.001
Proportion of user's posts or comments in <i>r/news</i>	Subreddit frequencies	-5.03 (-5.10, -4.97)	-4.29 (-4.48, -4.11)	-7.38 (810.43)	<.001

^aQuantile transformed means.^bTF-IDF: term frequency-inverse document frequency.^cWorld Well-Being Project

Discussion

Principal Findings

This study extends the literature on latent user attribute classification on social media by developing a 2-class age classification model for Reddit to better analyze Reddit posts and comments for specific age groups. This classification model can be used to better qualitatively analyze public health issues within the digital space of Reddit. Our results suggest that several different components of a user's Reddit data help distinguish between users 13 to 20 years and those 21 to 54 years. The relevant text features suggest that the adolescent age group generally use language in their comments that align with expected activities and phrasing often associated with younger age cohorts. For example, the adolescent age group were more likely, on average, to use the term "school" and less likely to use the term "home" than the adult age group. This is perhaps most noticeable in the WWBP 23-to-29 term set variables; both the raw frequency and normalized versions of these variables were found to be relevant and more highly associated with the adult age group.

For subreddits, the adolescent age group on average posted more frequently in *r/teenager* and less frequently in *r/news* than the older age group. They also tend to post in subreddits with more subscribers when compared with the older age group. This may be an indication that popular subreddits are popular because they contain many adolescent users, given that the general Reddit user base is skewed young [49]. The adolescent age group tended to have shorter comments than the adult age group. Younger users also tend to have a higher proportion of comments posted in a thread they started. In terms of account characteristics, the adolescent age group tended to have user accounts that were created more recently and also had lower comment karma. This relationship corroborates the intuition that younger users, because of their age, have had less opportunity to accumulate a long Reddit tenure and comment karma than older users.

All feature categories had at least one variable represented in the final model, except for "emoji" and "patterns in posting." Although we expected that Reddit users of different age groups would prefer to use emojis more tightly tied to their current life stage (eg, higher usage of "school" or "backpack" emoji for

younger users), our findings did not support this conclusion. Previous research has found that the number of emojis posted on public Facebook status updates decreases with user's age [50,51]. Additionally, web survey results on emoji use also supports the connection between emoji usage and age, with younger participants using both emoji and emoticons more frequently, having more positive attitudes, and having more motives for their usage than their older counterparts [52]. Likewise, although previous studies have identified certain user posting patterns as being associated with age, none of our variables from this group were found to be relevant. For example, Rosenthal and McKeown [53] found variable sets composed of lexical stylistic features (eg, the number of URL and image links) to be predictive of age groups on LiveJournal. Although most studies predicting age of social media users agree that text features are an important signal, future research could help assess which other types of features tend to be more resilient across social media platforms.

Our findings complement those in literature of age prediction on social media and extends the research findings available for Reddit. De Pril [54] analyzed 2 subreddit communities where users routinely self-report age for relationship advice (*r/relationship_advice*) and to meet other Redditors (*r/r4r*) and examined patterns of users posting behavior to predict 2 age categories (<24 years or >24 years for *r/r4r*; <26 years or >26 years for *r/relationship_advice*) using an ensemble of a multilayer perceptron and a *k*-nearest neighbor classifier, reporting an accuracy of 0.77 and AUROC of 0.86 for *r/r4r* and an accuracy of 0.72 and AUROC of 0.79 for *r/relationship_advice*. Gjurkovic et al [34] modeled age on Reddit using linear regression with TF-IDF terms, reporting a Pearson $r=0.75$ between the labeled and predicted ages. Tiginova et al [35] developed a series of age classification benchmark models for Reddit, reporting a hidden attribute model consisting of a convolutional neural network with attention performing best (AUROC 0.88 across 4 age categories—13-23 years, 24-45 years, 46-65 years, and 66-100 years), suggesting that deep learning models trained on text from user's posts can improve over a bag-of-words approach when predictive performance is more important than interpretation. Although our findings agree with those of previous literature [34,35,54] on Reddit that language features are effective predictors of age, our study suggests that other relevant metadata features, such as subreddit usage, account metadata, and stylistic linguistic features also provide insight into inferring age of Reddit users. Additionally, although difficult to directly compare given differences in the sample composition and age category definitions, our test set (F1 score 0.78) and (AUROC 0.86) perform comparably to previous Reddit age classifications (accuracy 0.72-0.77; AUROC 0.79-0.88 [35,54]).

Our study has several limitations that should be taken into consideration when interpreting the results. First, due to the poor accuracy in predicting the middle age group, we combined the range with that of the lowest age group; therefore, the young group had nearly twice as many users as the older age group (21-54 years). Although the classes were not severely imbalanced, imbalanced class ratios can impact model performance if they provide too great of an incentive to favor

the majority class. We ran a weighted version of the final reduced gradient boosted tree model to understand if a weighting correction would impact performance. The results (Table S1 in Multimedia Appendix 3) confirm that performance is nearly the same with or without weighting for the imbalance. Second, users who self-report age may be systematically different than those who do not report age, which may limit generalizability to all Reddit users. Selection biases stemming from what users choose to post or disclose on social media has been documented on Twitter, with significant differences found between the populations of users who do or do not geo-tag their tweets [55,56] and enable or disable location services [56]. Third, due to the evolving nature of language and platform use on social media, studies of this nature need to be continually updated. Previous research has found that language models used for predicting age and gender on social media tend to degrade over time if not retrained, with larger differences (ie, faster changing language) for younger social media users than older social media users [57]. In a production setting, building a model on training data that is nonrepresentative of new examples, whether due to sample selection biases or changes in the environment over time that change the joint distribution of outcomes and covariates, can severely erode the quality of predictions [58]. Monitoring differences in covariates between the training data and new observations is important for determining when to retrain or modify an existing model. Finally, our work used posts of self-reported ages in English. Many important model features outlined in this work likely will not extend to non-English speaking users (eg, language features), and caution should be used if applying this model to accounts of non-English speaking users.

Areas for future research include updating the algorithm to account for changes in linguistic patterns and social media use, examining how users who self-report age may be systematically different than those who do not report age, and considering how algorithms may be applied outside of their intended use case. For example, it is possible that an age prediction algorithm could be applied to target tobacco product advertising to adolescents. Although we believe this is unlikely, given that social media companies including Reddit have policies prohibiting advertising of tobacco products on their platforms that they monitor and enforce [59,60], researchers should consider tradeoffs when developing algorithms including how algorithms may be applied outside of their intended use case.

Conclusion

Social media platforms are important channels for public health organizations to communicate, and learn about, behaviors and perceptions on emerging health risks such as tobacco and electronic nicotine delivery systems. In particular, being able to identify members of subgroups on social media is essential for effectively promoting behavior change through that medium, as well as for learning about the issues uniquely impacting each target audience. This study presents an age classification model for Reddit users, an important but understudied social media platform in the latent user attribute modeling literature. This model is both interpretable and has competitive predictive performance when compared with previous Reddit age classification results in the literature. Given the persistent

adoption of social media usage especially for teens and young adults, any means of helping to reach this audience is useful for improved digital public health campaigns and formative research insights about adolescent vaping behaviors and perceptions.

Acknowledgments

The authors thank Jennifer Schoden, Ashley Andrews, and Michaela Fogarty for assisting with manual labeling of Reddit posts. We also thank Michael Wenger for his help modifying the SMART platform for our labeling needs.

This work was funded by contract with Center for Tobacco Products, US Food and Drug Administration, US Department of Health and Human Services (No. HHSF223201510002B-Order #75F40119F19020). This publication represents the views of the author(s) and does not represent FDA Center for Tobacco Products position or policy.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Definitions of receiver operating characteristic curve (AUROC), precision, recall, and F1 score.

[[DOCX File, 13 KB](#) - [publichealth_v7i3e25807_app1.docx](#)]

Multimedia Appendix 2

Feature correlation matrix.

[[DOCX File, 149 KB](#) - [publichealth_v7i3e25807_app2.docx](#)]

Multimedia Appendix 3

Class-weighted gradient boosted tree on the reduced variable set.

[[DOCX File, 58 KB](#) - [publichealth_v7i3e25807_app3.docx](#)]

Multimedia Appendix 4

Mann-Whitney U test and Kolmogorov-Smirnov test results comparing each model feature by age group.

[[DOCX File, 59 KB](#) - [publichealth_v7i3e25807_app4.docx](#)]

References

1. Duke JC, Alexander TN, Zhao X, Delahanty JC, Allen JA, MacMonegle AJ, et al. Youth's awareness of and reactions to the real cost national tobacco public education campaign. *PLoS One* 2015;10(12):e0144827 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0144827](#)] [Medline: [26679504](#)]
2. Guillory J, Crankshaw E, Farrelly MC, Alam I, Fiocco L, Curry L, et al. LGBT young adults' awareness of and receptivity to the This Free Life tobacco public education campaign. *Tob Control* 2021 Jan;30(1):63-70. [doi: [10.1136/tobaccocontrol-2019-055239](#)] [Medline: [31941821](#)]
3. Guillory J, Henes A, Farrelly MC, Fiocco L, Alam I, Curry L, et al. Awareness of and receptivity to the fresh empire tobacco public education campaign among hip hop youth. *J Adolesc Health* 2020 Mar;66(3):301-307. [doi: [10.1016/j.jadohealth.2019.09.005](#)] [Medline: [31704108](#)]
4. Walker MW, Evans SA, Wimpy C, Berger AT, Smith AA. Developing smokeless tobacco prevention messaging for at-risk youth: early lessons from "the real cost" smokeless campaign. *Health Equity* 2018;2(1):167-173 [[FREE Full text](#)] [doi: [10.1089/heq.2018.0029](#)] [Medline: [30283864](#)]
5. Every try counts campaign. US Food and Drug Administration. URL: <https://www.fda.gov/tobacco-products/every-try-counts-campaign> [accessed 2020-08-18]
6. Tobacco 21. US Food and Drug Administration. URL: <https://www.fda.gov/tobacco-products/retail-sales-tobacco-products/tobacco-21> [accessed 2020-04-21]
7. Who uses YouTube, WhatsApp and Reddit. Pew Research Center. 2019 Jun 12. URL: <https://www.pewresearch.org/internet/chart/who-uses-youtube-whatsapp-and-reddit/> [accessed 2020-08-18]
8. Sowles SJ, Krauss MJ, Gebremedhn L, Cavazos-Rehg PA. "I feel like I've hit the bottom and have no idea what to do": supportive social networking on Reddit for individuals with a desire to quit cannabis use. *Subst Abus* 2017;38(4):477-482. [doi: [10.1080/08897077.2017.1354956](#)] [Medline: [28704167](#)]
9. Park A, Conway M. Tracking health related discussions on reddit for public health applications. *AMIA Annu Symp Proc* 2017;2017:1362-1371 [[FREE Full text](#)] [Medline: [29854205](#)]

10. Meacham MC, Paul MJ, Ramo DE. Understanding emerging forms of cannabis use through an online cannabis community: an analysis of relative post volume and subjective highness ratings. *Drug Alcohol Depend* 2018 Jul 01;188:364-369 [FREE Full text] [doi: [10.1016/j.drugalcdep.2018.03.041](https://doi.org/10.1016/j.drugalcdep.2018.03.041)] [Medline: [29883950](https://pubmed.ncbi.nlm.nih.gov/29883950/)]
11. Pandrekar S, Chen X, Gopalkrishna G, Srivastava A, Saltz M, Saltz J, et al. Social media based analysis of opioid epidemic using Reddit. *AMIA Annu Symp Proc* 2018;2018:867-876 [FREE Full text] [Medline: [30815129](https://pubmed.ncbi.nlm.nih.gov/30815129/)]
12. Costello KL, Martin JD, Edwards Brinegar A. Online disclosure of illicit information: information behaviors in two drug forums. *J Assoc Inf Sci Technol* 2017 Jun 14;68(10):2439-2448. [doi: [10.1002/asi.23880](https://doi.org/10.1002/asi.23880)]
13. Eliseo-Arras RK. Alcohol and non-suicidal self-injury posts on Reddit. 2018 May 31 Presented at: 41st Annual Scientific Meeting of the Research Society on Alcoholism; June 16-20; San Diego, California, USA p. 19A-270A. [doi: [10.1111/acer.13747](https://doi.org/10.1111/acer.13747)]
14. Tamersoy A, De Choudhury M, Chau DH. Characterizing smoking and drinking abstinence from social media. In: *Proceedings of the ACM Conference on Hypertext and Social Media*. 2015 Presented at: 26th ACM Conference on Hypertext and Social Media; September 1-4; Guzelyurt, Cyprus p. 139-148 URL: <http://europepmc.org/abstract/MED/26640831> [doi: [10.1145/2700171.2791247](https://doi.org/10.1145/2700171.2791247)]
15. Brett EI, Stevens EM, Wagener TL, Leavens ELS, Morgan TL, Cotton WD, et al. A content analysis of JUUL discussions on social media: using Reddit to understand patterns and perceptions of JUUL use. *Drug Alcohol Depend* 2019 Jan 01;194:358-362. [doi: [10.1016/j.drugalcdep.2018.10.014](https://doi.org/10.1016/j.drugalcdep.2018.10.014)] [Medline: [30472576](https://pubmed.ncbi.nlm.nih.gov/30472576/)]
16. Zhan Y, Zhang Z, Okamoto JM, Zeng DD, Leischow SJ. Underage JUUL use patterns: content analysis of Reddit messages. *J Med Internet Res* 2019 Sep 09;21(9):e13038 [FREE Full text] [doi: [10.2196/13038](https://doi.org/10.2196/13038)] [Medline: [31502542](https://pubmed.ncbi.nlm.nih.gov/31502542/)]
17. Culotta A, Ravi NK, Cutler J. Predicting Twitter user demographics using distant supervision from website traffic data. *J Artif Intell Res* 2016 Feb 19;55:389-408. [doi: [10.1613/jair.4935](https://doi.org/10.1613/jair.4935)]
18. Han K, Jo Y, Jeon Y. Photos don't have me, but how do you know me?: analyzing and predicting users on Instagram. 2018 Presented at: 26th Conference on User Modeling, Adaptation and Personalization; July 8-11; Singapore. [doi: [10.1145/3213586.3225232](https://doi.org/10.1145/3213586.3225232)]
19. Sloan L, Morgan J, Housley W, Williams M, Edwards A, Burnap P, et al. Knowing the Tweeters: deriving sociologically relevant demographics from Twitter. *Sociol Res Online* 2013 Aug 31;18(3):74-84. [doi: [10.5153/sro.3001](https://doi.org/10.5153/sro.3001)]
20. Bayot R, Gonçalves T. Age and gender classification of tweets using convolutional neural networks. New York, NY: Springer; 2017 Presented at: International Workshop on Machine Learning, Optimization, and Big Data; Sept 14-17; Volterra, Italy p. 337-348. [doi: [10.1007/978-3-319-72926-8_28](https://doi.org/10.1007/978-3-319-72926-8_28)]
21. Guimaraes RG, Rosa RL, De Gaetano D, Rodriguez DZ, Bressan G. Age groups classification in social network using deep learning. *IEEE Access* 2017;5:10805-10816. [doi: [10.1109/access.2017.2706674](https://doi.org/10.1109/access.2017.2706674)]
22. Morgan-Lopez AA, Kim AE, Chew RF, Ruddle P. Predicting age groups of Twitter users based on language and metadata features. *PLoS One* 2017;12(8):e0183537 [FREE Full text] [doi: [10.1371/journal.pone.0183537](https://doi.org/10.1371/journal.pone.0183537)] [Medline: [28850620](https://pubmed.ncbi.nlm.nih.gov/28850620/)]
23. Chen X, Wang Y, Agichtein E, Wang F. A comparative study of demographic attribute inference in Twitter. In: *Proceedings of ICWSM*. 2015 Presented at: Ninth International AAAI Conference on Web and Social Media; May 26-29; Oxford, England URL: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10541>
24. Mohammady AE, Culotta A. Inferring latent attributes of Twitter users with label regularization. 2015 Presented at: Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; May 31-June 5; Denver, Colorado, USA p. 185-195 URL: <https://www.aclweb.org/anthology/N15-1019> [doi: [10.3115/v1/n15-1019](https://doi.org/10.3115/v1/n15-1019)]
25. Preotiuc-Pietro D, Ungar L. User-level race and ethnicity predictors from Twitter text. 2018 Presented at: Proceedings of the 27th International Conference on Computational Linguistics; August 20-26; Sante Fe, New Mexico, USA p. 1534-1526 URL: <https://www.aclweb.org/anthology/C18-1130/>
26. Al Zamal ZF, Liu W, Ruths D. Homophily and latent attribute inference: inferring latent attributes of Twitter users from neighbors. 2012 Presented at: Sixth International AAAI Conference on Weblogs and Social Media; June 4-8; Dublin, Ireland URL: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4713>
27. Preotiuc-Pietro D, Liu Y, Hopkins D, Ungar L. Beyond binary labels: political ideology prediction of Twitter users. 2017 Presented at: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics; July 30-August 4; Vancouver, Canada p. 729-740 URL: <https://www.aclweb.org/anthology/P17-1068/> [doi: [10.18653/v1/p17-1068](https://doi.org/10.18653/v1/p17-1068)]
28. Vijayaraghavan P, Vosoughi S, Roy D. Twitter demographic classification using deep multimodal multitask learning. In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*. 2017 Presented at: 55th Annual Meeting of the Association for Computational Linguistics; July 30-August 4; Vancouver, Canada p. 478-483 URL: <https://www.aclweb.org/anthology/P17-2076/> [doi: [10.18653/v1/p17-2076](https://doi.org/10.18653/v1/p17-2076)]
29. Harrigian K. Geocoding without geotags: a text-based approach for reddit. In: *Proceedings of the 2018 EMNLP Workshop W-NUT*. 2018 Presented at: 4th Workshop on Noisy User-Generated Text; November 1; Brussels, Belgium p. 17-27 URL: <https://www.aclweb.org/anthology/W18-6103/> [doi: [10.18653/v1/w18-6103](https://doi.org/10.18653/v1/w18-6103)]
30. Wang Z, Jurgens D. It's going to be okay: measuring access to support in online communities. 2018 Presented at: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing; October 31-November 4; Brussels, Belgium p. 33-45 URL: <https://www.aclweb.org/anthology/D18-1004/> [doi: [10.18653/v1/d18-1004](https://doi.org/10.18653/v1/d18-1004)]

31. Cohan A, Desmet B, Yates A, Soldani L, MacAvaney S, Goharian N. SMHD: a large-scale resource for exploring online language usage for multiple mental health conditions. 2018 Presented at: 27th International Conference on Computational Linguistics; August 20-26; Santa Fe, New Mexico, USA p. 1485-1497 URL: <https://www.aclweb.org/anthology/C18-1126/>
32. Wolohan JT, Hiraga M, Mukherjee A, Zeeshan AS, Millard M. Detecting linguistic traces of depression in topic-restricted text: attending to self-stigmatized depression with NLP. 2018 Presented at: First International Workshop on Language Cognition and Computational Models; August 21-25; Santa Fe, New Mexico, USA p. 11-21 URL: <https://www.aclweb.org/anthology/W18-4102/>
33. Welch C, Kummerfeld J, Perez-Rosas V, Mihalcea R. Compositional demographic word embeddings. 2020 Presented at: Conference on Empirical Methods in Natural Language Processing (EMNLP); November 16-20; Online URL: <https://www.aclweb.org/anthology/2020.emnlp-main.334/> [doi: [10.18653/v1/2020.emnlp-main.334](https://doi.org/10.18653/v1/2020.emnlp-main.334)]
34. Gjurokovic M, Karan M, Vukojevic I, Bosnjak M, Snajder J. PANDORA Talks: personality and demographics on Reddit. arXiv. Preprint posted online on April 9, 2020 2021 [FREE Full text]
35. Tiginova A, Mirza P, Yates A, Weikum G. RedDust: a Large Reusable Dataset of Reddit User Traits. 2020 Presented at: Language Resources and Evaluation Conference; May 11-16; Marseille, France p. 6118-6126 URL: <https://www.aclweb.org/anthology/2020.lrec-1.751/>
36. Chew R, Wenger M, Kery C, Nance J, Richards K, Hadley E, et al. SMART: an open source data labeling platform for supervised learning. J Mach Learn Res 2019;20(82):1-5 [FREE Full text]
37. Hinds J, Joinson AN. What demographic attributes do our digital footprints reveal? a systematic review. PLoS One 2018;13(11):e0207112 [FREE Full text] [doi: [10.1371/journal.pone.0207112](https://doi.org/10.1371/journal.pone.0207112)] [Medline: [30485305](https://pubmed.ncbi.nlm.nih.gov/30485305/)]
38. Schwartz HA, Eichstaedt JC, Kern ML, Dziurzynski L, Ramones SM, Agrawal M, et al. Personality, gender, and age in the language of social media: the open-vocabulary approach. PLoS One 2013;8(9):e73791 [FREE Full text] [doi: [10.1371/journal.pone.0073791](https://doi.org/10.1371/journal.pone.0073791)] [Medline: [24086296](https://pubmed.ncbi.nlm.nih.gov/24086296/)]
39. Pedregosa F, Varoquaux G, Gramfort A, Vincent M, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12(85):2825-2830 [FREE Full text]
40. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. New York, NY: Springer; 2001.
41. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995 Sep;20(3):273-297. [doi: [10.1007/BF00994018](https://doi.org/10.1007/BF00994018)]
42. Cover T, Hart P. Nearest neighbor pattern classification. IEEE Trans Inform Theory 1967 Jan;13(1):21-27. [doi: [10.1109/tit.1967.1053964](https://doi.org/10.1109/tit.1967.1053964)]
43. Breiman L. Random forests. Mach Learn 2001;45(1):5-32 [FREE Full text]
44. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Statist 2001 Oct;29(5):1189-1232. [doi: [10.1214/aos/1013203451](https://doi.org/10.1214/aos/1013203451)]
45. Shalev-Shwartz S, Ben-David S. Understanding Machine Learning: From Theory to Algorithms. Cambridge, MA: Cambridge University Press; 2014.
46. Rudnicki W, Wrzesien M, Paja W. All relevant feature selection methods and applications. In: Feature Selection for Data and Pattern Recognition. New York, NY: Springer; 2015:11-28.
47. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. J Stat Soft 2010;36(11):1-13. [doi: [10.18637/jss.v036.i11](https://doi.org/10.18637/jss.v036.i11)]
48. Fisher A, Rudin C, Dominici F. All models are wrong, but many are useful: learning a variable's importance by studying an entire class of prediction models simultaneously. J Mach Learn Res 2019;20(177):1-81 [FREE Full text]
49. Barthel M, Stocking G, Holcomb J, Mitchell A. Reddit news users more likely to be male, young and digital in their news preferences. Pew Research Center. 2016 Feb 25. URL: <https://www.journalism.org/2016/02/25/reddit-news-users-more-likely-to-be-male-young-and-digital-in-their-news-preferences/> [accessed 2020-08-16]
50. Oleszkiewicz A, Karwowski M, Pisanski K, Sorokowski P, Sobrado B, Sorokowska A. Who uses emoticons? data from 86,702 Facebook users. Personality and Individual Differences 2017 Dec;119:289-295. [doi: [10.1016/j.paid.2017.07.034](https://doi.org/10.1016/j.paid.2017.07.034)]
51. Settanni M, Marengo D. Sharing feelings online: studying emotional well-being via automated text analysis of Facebook posts. Front Psychol 2015;6:1045 [FREE Full text] [doi: [10.3389/fpsyg.2015.01045](https://doi.org/10.3389/fpsyg.2015.01045)] [Medline: [26257692](https://pubmed.ncbi.nlm.nih.gov/26257692/)]
52. Prada M, Rodrigues DL, Garrido MV, Lopes D, Cavalheiro B, Gaspar R. Motives, frequency and attitudes toward emoji and emoticon use. Telemat Inform 2018 Oct;35(7):1925-1934. [doi: [10.1016/j.tele.2018.06.005](https://doi.org/10.1016/j.tele.2018.06.005)]
53. Rosenthal S, McKeown K. Age prediction in blogs: a study of style, content, online behavior in pre- and post-social media generations. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. 2011 Presented at: 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies; June 19-24; Portland, Oregon, USA p. A-72 URL: <https://www.aclweb.org/anthology/P11-1077/>
54. De Pril R. User classification based on public Reddit data. Ghent University.: Ghent University; 2019. URL: https://lib.ugent.be/fulltxt/RUG01/002/785/835/RUG01-002785835_2019_0001_AC.pdf [accessed 2020-02-12]
55. Lippincott T, Carrell A. Observational comparison of geo-tagged and randomly-drawn tweets. 2018 Presented at: Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media; June 6; New Orleans, Louisiana, USA p. 50-55 URL: <https://www.aclweb.org/anthology/W18-1107/> [doi: [10.18653/v1/w18-1107](https://doi.org/10.18653/v1/w18-1107)]

56. Sloan L, Morgan J. Who tweets with their location? understanding the relationship between demographic characteristics and the use of geoservices and geotagging on Twitter. PLoS ONE 2015 Nov 6;10(11):e0142209. [doi: [10.1371/journal.pone.0142209](https://doi.org/10.1371/journal.pone.0142209)]
57. Jaidka K, Chhaya N, Ungar L. Diachronic degradation of language models: insights from social media. 2018 Presented at: 56th Annual Meeting of the Association for Computational Linguistics; July 15-20; Melbourne, Australia p. 195-200 URL: <https://www.aclweb.org/anthology/P18-2032/> [doi: [10.18653/v1/p18-2032](https://doi.org/10.18653/v1/p18-2032)]
58. Quionero-Candela J, Sugiyama M, Schwaighofer A, Lawrence N. Dataset Shift in Machine Learning. Cambridge, MA, USA: The MIT Press; 2009.
59. Facebook's policy prohibiting ads for tobacco and related products on Facebook and Instagram. Facebook. URL: https://m.facebook.com/policies/ads/prohibited_content/tobacco [accessed 2021-02-01]
60. Reddit advertising policy: prohibited advertisements. Reddit. URL: <https://advertising.reddithelp.com/en/categories/reddit-advertising-policy/reddit-advertising-policy-prohibited-advertisements> [accessed 2021-02-01]

Abbreviations

API: application programming interface
AUROC: area under the receiver operating characteristic curve
FDA: Food and Drug Administration
TF-IDF: term frequency–inverse document frequency
WWBP: World Well-Being Project

Edited by T Sanchez, G Eysenbach; submitted 17.11.20; peer-reviewed by K Harrigan; comments to author 10.12.20; revised version received 04.02.21; accepted 09.02.21; published 16.03.21.

Please cite as:

Chew R, Kery C, Baum L, Bukowski T, Kim A, Navarro M

Predicting Age Groups of Reddit Users Based on Posting Behavior and Metadata: Classification Model Development and Validation
JMIR Public Health Surveill 2021;7(3):e25807

URL: <https://publichealth.jmir.org/2021/3/e25807>

doi:[10.2196/25807](https://doi.org/10.2196/25807)

PMID:[33724195](https://pubmed.ncbi.nlm.nih.gov/33724195/)

©Robert Chew, Caroline Kery, Laura Baum, Thomas Bukowski, Annice Kim, Mario Navarro. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 16.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Effects of a Co-Design–Based Invitation Strategy on Participation in a Preventive Health Check Program: Randomized Controlled Trial

Trine Thilsing¹, PhD; Lars Bruun Larsen^{1,2}, PhD; Anders Larrabee Sonderlund¹, PhD; Signe Skaarup Andreassen³, MHS; Jeanette Reffstrup Christensen^{1,4}, PhD; Nanna Herning Svensson¹, MPH; Marie Dahl^{5,6}, PhD; Jens Sondergaard¹, PhD

¹Research Unit of General Practice, Department of Public Health, University of Southern Denmark, Odense C, Denmark

²Steno Diabetes Center Zealand, Holbaek, Denmark

³Steno Diabetes Center Odense, Odense, Denmark

⁴Research unit of User Perspectives and Community-based Interventions, Department of Public Health, University of Southern Denmark, Odense C, Denmark

⁵Vascular Research Unit, Department of Surgery, Regional Hospital Central Denmark, Viborg, Denmark

⁶Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

Corresponding Author:

Trine Thilsing, PhD

Research Unit of General Practice

Department of Public Health

University of Southern Denmark

JB Winsløws Vej 9A

Odense C, DK-5000

Denmark

Phone: 45 65507195

Email: tthilsing@health.sdu.dk

Abstract

Background: Preventive primary care programs that aim to reduce morbidity and mortality from lifestyle-related diseases are often affected by low-to-moderate participation rates. Improving participation rates is essential for clinical effectiveness and cost-effectiveness. In 2016-2017, we conducted a pilot study (TOF pilot1) for a preventive primary care intervention (TOF is the Danish abbreviation for “Early Detection and Prevention”). Among 8814 invited patients, 3545 (40.22%) consented to participate, with the highest participation rates among women and patients with higher income, education, and employment.

Objective: The aim of this study was to evaluate the effects of a revised invitation strategy on invitation comprehensibility, the overall participation rate, and participant demography. The new strategy specifically targeted men and patients of low educational attainment.

Methods: This study was embedded in the second TOF pilot study (TOF pilot2, initiated in October 2018) that tested an adjusted intervention. The revised invitation strategy comprised a prenotification postcard and a new invitation that specifically targeted men and patients of low educational attainment. The new invitation was developed in a co-design process that included communication professionals and target-group representatives. The study sample consisted of 4633 patients aged between 29 and 59 years, who resided in one of two municipalities in the Region of Southern Denmark. Eligible patients were randomly assigned to one of four invitation groups. The control group (Group 1) received the original invitation used in TOF pilot1. The intervention groups received the original invitation and the prenotification postcard (Group 2), the new revised invitation and the prenotification postcard (Group 3), or the new invitation but no prenotification postcard (Group 4).

Results: Overall, 2171 (46.86%) patients consented to participate. Compared to the control group, participation rates increased significantly in all three intervention groups (all $P < .001$). Participation across the three intervention groups increased for women and men, as well as for patients with high and low educational attainment and high and low family income. The largest relative increase in participation rates occurred among men, patients with low educational attainment, and patients with low family income. No increase in participation was detected for unemployed patients or patients of non-Danish origin. Most participants found the

original (813/987, 82.37%) and new (965/1133, 85.17%) invitations easy to understand with no significant difference ($P=.08$) in comprehensibility between invitations.

Conclusions: The results suggest that participation in preventive primary care interventions can be greatly increased by implementing a co-design-based invitation strategy that includes prenotification postcards and targeted invitations. Although firm conclusions cannot be made from this study, the observed increased participation rates for men and patients of low socioeconomic status may be relevant in programs that aim to reduce social inequality in health.

Trial Registration: ClinicalTrials.gov NCT03913585; <https://clinicaltrials.gov/ct2/show/NCT03913585>

(*JMIR Public Health Surveill* 2021;7(3):e25617) doi:[10.2196/25617](https://doi.org/10.2196/25617)

KEYWORDS

participation rate; prenotification; invitation letter; co-design; preventive health checks; primary care

Introduction

Background

Preventive primary care health checks that aim to reduce lifestyle-related morbidity and mortality often have only moderate-to-low participation rates (<50%) [1-4]. In addition, the rate of uptake appears to be unevenly spread across the population. Specifically, patients who are most likely to participate in preventive health checks are those who have a higher socioeconomic status, are older, are female, or have a lower than average prevalence of cardiovascular risk factors [3,5,6].

Important factors that facilitate participation include the mode of invitation, patient awareness of the given preventive program, and clarity of the program's purpose [7-9]. For example, a written invitation is often the first point of contact between program providers and the patient, thus representing a vital element in recruitment. Previous studies have assessed the effects of invitation content and mode of distribution on participation rate. A US study targeting Hispanic employees found that participation in a worksite dietary intervention increased when the initial advertisement (a flyer) was supplemented with a personalized letter to individual workers. Tailoring the letter further by including heart disease risk statistics for Hispanics rather than the general population, however, had no additional effect on participation [10]. By contrast, Sallis et al increased participation in a health check intervention from 29.3% to 33.5% by revising the invitation with insights from behavioral science, including simplification, action-focused behavioral instructions, and personal salience, and specifically addressing implementation intentions [11]. Other tactics, including emphasizing support from a funding agency, incorporating endorsements from senior-position health professionals as opposed to junior-position health professionals [12], and tailoring the invitation with information about patient cardiovascular disease risk, had no effect [13]. Finally, in a recent study by Koitsalu et al, higher participation rates in a cancer screening program were associated with the use of prenotification postcards and reminders. The study also assessed invitation length but found no effect on participation [8]. While these studies report on a relatively broad range of specific invitation methods that seem to vary somewhat in effectiveness, the evidence ultimately indicates the potential for boosting intervention participation by focusing on both the overall invitation strategy and individual invitation components.

Specific Basis for Our Study

In a recent study (TOF pilot1), we tested the feasibility and acceptability of a step-wise approach to preventive primary care health checks [14]. TOF is the Danish abbreviation for "Early Detection and Prevention." The TOF intervention centers on a two-step process as follows: (1) the identification and stratification of the at-risk population through a participant risk-assessment questionnaire combined with health information from electronic patient records at general practitioners (GPs); (2) an offer of targeted and cohesive preventive services to the high-risk population.

In the TOF pilot1 study, a random sample of 8814 patients (aged 29-59 years) from 47 GP clinics was invited to take part in the study [2,14,15]. An invitation to participate was sent to each patient's digital mailbox, followed by two reminders 2 weeks apart in the event of nonresponse. The digital mailbox is an online system provided by the Danish government for secure communication between individuals and public authorities/other trusted organizations. Almost all (>92%) Danish citizens aged 15 years or above have a digital mailbox [16].

A total of 3545 (40.22%) invited patients consented to take part in the study. Consistent with other similar studies, women and patients with higher income, education, and employment rates were most likely to participate [2].

In an attempt to increase general participation rates as well as participation specifically among underrepresented patient groups at possible increased risk of lifestyle-related diseases (males and patients with low educational attainment) [17,18], we revised the invitation strategy, taking a pragmatic co-design approach. To this end, we engaged communication professionals as well as target-group representatives [19]. This paper reports on the effects of the revised invitation strategy on invitation comprehensibility, the overall participation rate, and participant demography.

Methods

Context of the Study

This study is nested within the second TOF pilot study (TOF pilot2) that tests the feasibility and acceptability of an adjusted version of the TOF intervention. The study has been registered on ClinicalTrials.gov (NCT03913585).

Adjustments to the Invitation Strategy

Prior to study commencement, the TOF invitation strategy was revised by (1) adding a prenotification postcard sent to prospective participants 2 weeks prior to the digital invitation, and (2) designing a new invitation based on the results of a co-design approach [19]. The prenotification postcard and the original and new TOF invitations (all translated into English) are presented in [Multimedia Appendix 1](#), [Multimedia Appendix 2](#), and [Multimedia Appendix 3](#).

Prenotification Postcard

The purpose of the prenotification postcard was to create awareness about the upcoming invitation and the TOF intervention in general. It urged recipients to keep an eye on their digital mailbox during the next couple of weeks as “something important is on its way.” The prenotification also contained brief information about the full intervention, including the possibility of receiving a preventive health check, and was signed by the patients’ GP, the municipality, and the Region of Southern Denmark.

Invitation

The original and new invitations were designed as one-page PDF files written in Danish. Both included the contact details of the project coordinator and a hyperlink to the study webpage where additional information about the study could be accessed, including a short animated film outlining the individual steps of the intervention. Both invitation versions also included a link to a web-based digital support system through which participants could provide their informed consent to participate. The invitations were addressed to individual patients and were signed by the patients’ GP, the municipality, and the Region of Southern Denmark.

Development of the New Invitation by a Co-Design Approach

The new invitation was developed in a pragmatic co-design process in collaboration with communication professionals, purposely sampled representatives from the target population, and the Men’s Health Society [20], a multidisciplinary organization dedicated to the field of men’s health in all its aspects.

First, three revamped versions of the invitation were developed by communication professionals and on the basis of current knowledge on facilitators and barriers for participation in preventive health checks [9,11]. All versions were designed to include clear and unambiguous information about the intervention, action-focused behavioral instructions (pictogram), and information on target-group selection (scarcity). A pictogram was included to clearly depict the individual steps of the intervention (when, where, and how), with the aim to rehearse the cognitive link between behavior and context [21]. A special effort was made to keep sentences short, concise, and free of jargon [22]. The exact wording of the invitation was inspired by recent Danish preventive programs focusing on men and socially deprived groups [23,24], and by specific recommendations from the Men’s Health Society.

Thereafter, the three new invitations were discussed and tested for content and comprehensibility in a focus group. Focus group members were purposely sampled to include men and people with low educational level. The recruitment process comprised advertisements on Facebook and Instagram, which were targeted at people aged 29 to 59 years with low educational attainment. In addition, attendees at “Meeting Place for Men” in the city of Sønderborg [25] (a social connectedness initiative for men) were contacted directly. The resulting focus group comprised 10 people between 34 and 57 years old (three women and seven men). Eight focus group members had no formal education beyond secondary school, one was a primary school teacher, and one was a printmaker. None of the focus group members were eligible to participate in the TOF pilot2 study as they resided outside the participating municipalities.

A semistructured interview guide was developed by a multidisciplinary research team and communication professionals. The guide included thematic open- and closed-ended questions on facilitators and barriers for receiving information through the digital mailbox, on general perceptions of health, and on the specific content and comprehensibility of the three invitations [26]. Before the meeting, all focus group members were asked to study the three invitations. During the meeting, all participants were encouraged to provide input.

The focus group meeting was filmed and transcribed. Strong action-oriented comments, points, and statements about the content, wording, and design of the three invitations were identified and used in the development of the final invitation.

Study Design

The effect of the revised invitation strategy on overall participation and participant demography was tested in a randomized controlled trial nested within the TOF pilot2 study. Initially, a total of 61 GPs from 22 GP clinics in two municipalities in the Region of Southern Denmark (Haderslev and Middelfart) were invited to take part in the study. Subsequently, the target group was selected from the patient registries of the participating GP clinics and comprised patients born between 1959 and 1988 (aged 29-59 years). The chosen age range resembles age ranges used in previous lifestyle interventions in primary care [27-30] and was determined on the rationale that people in this age group may achieve the greatest health benefits from improvements in lifestyle. Patients were excluded if they lived outside the municipalities of Haderslev or Middelfart, if they did not have a digital mailbox (<5% of the target population) [16], or if their name and address were unlisted. Patients from the municipality of Haderslev who were invited to the first TOF pilot study (TOF pilot1) were also excluded.

Before study commencement, all eligible patients were randomly assigned to one of four invitation groups and thus received one of the following: (1) the original invitation used in the first TOF pilot study (control group, Group 1), (2) a prenotification postcard and the original invitation (Group 2), (3) a prenotification postcard and the new invitation (Group 3), and (4) the new invitation but no prenotification postcard (Group 4).

In order to avoid cross-contamination, patients living together were randomized to the same invitation group. Randomization was done using Stata (refer to the subsection Sample Size Calculation, Randomization, and Statistical Analysis).

On October 8, 2018, the prenotification postcard was sent by standard mail to patients in Groups 2 and 3. Two weeks later on October 22, 2018, invitations were sent to the digital mailboxes of all eligible patients (Groups 1-4). In the event of nonresponse, the invitation was followed up with two digital reminders sent 2 weeks apart. The reminders were identical to the first invitation, except for a brief sentence that informed the participant that this was a reminder. The deadline for providing informed consent to participate was December 3, 2018 (6 weeks after the invitation was first sent out).

Immediately following consent, the participant was redirected from the digital support system to an electronic questionnaire that included the following question on invitation comprehensibility: "The information about the project included in the digital invitation was." The answer options were as follows: "Easy to understand," "Fairly understandable," and "Difficult to understand."

Dependent Variable

Patients were defined as participants or nonparticipants based on whether they had provided informed consent to participate in the study.

Independent Variables

Invitation group (Groups 1-4), invitation type (original/new), and prenotification postcard (yes/no) were defined as described above. Participants' evaluation of invitation comprehensibility was dichotomized as 1 ("easy to understand") or 0 ("fairly understandable/difficult to understand").

Register-Based Data

Information on sex, age, country of origin, educational level, employment status, and family income was retrieved from the national Danish Bureau of Statistics (Statistics Denmark) and linked with individual patients' Danish Personal Identification number (CPR).

Participant age was determined at the time of invitation and categorized in 10-year age brackets. Country of origin was retrieved for the year 2018 and categorized as Danish, Western, or non-Western. Western countries included countries in the European Union and associated countries, as well as the United States, Canada, Australia, and New Zealand. Non-Western countries included the European countries of Albania, Bosnia and Herzegovina, Belarus, Yugoslavia, Kosovo, Macedonia, Moldova, Montenegro, Russia, Serbia, Soviet Union, Turkey, and Ukraine; all countries in Africa, South and Central America, and Asia; and all countries in Oceania (except Australia and New Zealand). Stateless persons were also defined as non-Western. The highest attained educational level was retrieved for October 2018 and categorized as secondary school, high school, vocational education, higher education, or higher education-master's level. Subsequently, the highest educational attainment was dichotomized (low educational attainment [yes/no]: yes = secondary school; no = high school, vocational

education, higher education, or higher education-master's level). Employment status was retrieved for November 2018 and categorized as employed, self-employed, unemployed/on benefits, social welfare recipient, or other. In Denmark, all unemployed workers are eligible to receive social welfare benefits, whereas unemployment benefits are accessible only to citizens who have been unemployed for less than 2 years and who are members of a voluntary unemployment benefit fund. The final group ("others") represents, for example, unemployed persons from a family that relies on one income only. For all analyses, employment status was dichotomized (unemployed [yes/no]: yes = unemployed/on benefits, social welfare recipients, or other; no = employed or self-employed). Family income was retrieved for 2013-2018, defined by the mean annual net income of the household, and was categorized in quartiles. Subsequently, family income was dichotomized (low income [yes/no]: yes = lowest quartile; no = above the lowest quartile).

Sample Size Calculation, Randomization, and Statistical Analysis

Based on results from a previous study that employed a similar approach of prenotification postcard followed by invitation [8], we estimated that the intervention could achieve a 6% increase in the participation rate (ie, 46% compared with 40% in the control group) [2]. The study was therefore designed to detect a 6% difference in the participation rate between control and intervention groups with 80% power and 5% significance. This required a total sample size of 4404 (1101 per group) patients.

Randomization was performed by participant address, assigning random numbers to the cohort using the generate rannum = uniform() command in Stata. Subsequent allocation to create approximately equally sized groups was performed using the egen recruitmentgroup = cut(rannum), group(4) command.

Descriptive statistics have been used to present the study sample. Generalized linear models (binreg) were run to assess the effects of invitation mode and individual invitation elements on participation rates. Adjusted analyses accounted for age and sex.

Chi-square analyses were performed to compare participation rates in each of the three intervention groups to that in the control group overall and in sociodemographic subgroups. Generalized linear models (binreg) were run to calculate risk ratios (RRs) and 95% CIs. RR was chosen over odds ratio (OR) as OR tends to exaggerate the estimate of the relationship between an exposure and an outcome in cases where there is an association [31].

Generalized linear models (binreg) also assessed the association between invitation type (original or new) and level of comprehensibility ("easy to understand" or "fairly understandable/difficult to understand"). The significance level was set at $P < .05$.

All statistical analyses were performed on secure servers at Statistics Denmark using Stata version 16.0 (Statacorp).

Ethics Approval and Consent to Participate

The study was approved by the Research & Innovation Organisation, University of Southern Denmark (18/32742), and

the TOF pilot2 study was registered at ClinicalTrials.gov (NCT03913585) [32]. According to Danish regulations (Act on Research Ethics Review of Health Research Projects [section 14.2]), this study did not need approval from a health research ethics committee as no research on human tissue or other biological material was performed. The study complies with the Helsinki Declaration by requiring informed consent from participants.

Focus group members received compensation for transportation costs but were not remunerated for their participation in the meeting. Participants in the TOF pilot2 study did not receive any remuneration or compensation.

Results

Assignment to Invitation Groups and Sociodemographic Characteristics of the Invitation Groups

A total of 15 GPs from four clinics took part in the study. Of 6347 patients born between 1959 and 1988, 4633 were eligible to participate in the study. Random assignment placed 1151 patients in Group 1 (original invitation, no postcard [control]), 1156 in Group 2 (original invitation plus postcard), 1148 in Group 3 (new invitation plus postcard), and 1178 in Group 4 (new invitation, no postcard).

Table 1 shows the sociodemographic characteristics of all 4633 patients who were invited to take part in the study. Overall, the four invitation groups did not differ in any demographic characteristics.

Table 1. Sociodemographic characteristics and mode of invitation among patients invited to participate in a preventive primary care program for lifestyle-related diseases (TOF pilot2).

Mode of invitation	Group 1: Original invitation (control) (n=1151, 24.84%), n (%)	Group 2: Original invitation + postcard (n=1156, 24.95%), n (%)	Group 3: New invitation + postcard (n=1148, 24.78%), n (%)	Group 4: New invitation (n=1178, 25.43%), n (%)	Total (n=4633, 100%), n (%)
10-year age groups (missing n=0)					
29-39 years	302 (26.24)	315 (27.25)	321 (27.96)	309 (26.23)	1247 (26.92)
40-49 years	441 (38.31)	437 (37.80)	456 (39.72)	442 (37.52)	1776 (38.33)
50-60 years	408 (35.45)	404 (34.95)	371 (32.32)	427 (36.25)	1610 (34.75)
Sex (missing n=0)					
Male	545 (47.35)	533 (46.11)	561 (48.87)	560 (47.54)	2199 (47.54)
Female	606 (52.65)	623 (53.89)	587 (51.13)	618 (52.46)	2434 (52.46)
Country of origin (missing n=21, 0.45%)					
Denmark	1044 (91.26)	1042 (90.37)	1040 (90.99)	1064 (90.78)	4190 (90.85)
Western	32 (2.80)	38 (3.30)	45 (3.94)	42 (3.58)	157 (3.40)
Non-Western	68 (5.94)	73 (6.33)	58 (5.07)	66 (5.63)	265 (5.75)
Highest educational attainment (missing n=155, 3.34%)					
Secondary school	196 (17.61)	189 (16.83)	183 (16.56)	190 (16.71)	758 (16.93)
Highschool, vocational education, higher education, or higher education-master's level	917 (82.39)	934 (83.17)	922 (83.44)	947 (83.29)	3720 (83.07)
Employment status (missing n=15, 0.32%)					
Unemployed/on benefits, social welfare recipients, or other ^a	208 (18.15)	207 (17.94)	213 (18.60)	215 (18.33)	843 (18.25)
Employed or self-employed	938 (81.85)	947 (82.06)	932 (81.40)	958 (81.67)	3775 (81.75)
Family income (missing n=21, 0.45%)					
Lowest quartile	260 (22.73)	276 (23.94)	270 (23.62)	288 (24.57)	1094 (23.72)
Greater than the lowest quartile	884 (77.27)	877 (76.06)	873 (76.38)	884 (75.43)	3518 (76.28)

^a“Other” represents, for example, unemployed persons from a family that relies on one income only.

Participation

A total of 2171 (46.86%) out of 4633 invited patients consented to participate in the study. Participation rates ranged from

39.44% (454/1151) to 50.78% (583/1148) across the four invitation groups. The participation rate for patients who received the original invitation alone (control group) (39.44%)

was comparable to the participation rate obtained in TOF pilot1 (40.22%) [2]. Figure 1 shows the flow of patients from sampling to participation in the TOF pilot2 study.

Compared to participants in Group 1, who received the original invitation alone, participation rates were higher for Group 2 (original invitation plus prenotification postcard), Group 3 (new invitation plus prenotification postcard), and Group 4 (new

invitation alone). These differences in participation rates remained highly significant after adjustments for age and sex (Table 2). The highest rate of participation was achieved in Group 3. Differences in participation rates between the three intervention groups (Groups 2, 3, and 4), however, did not reach statistical significance (Group 2 vs 3: $P=.36$; Group 2 vs 4: $P=.78$; Group 3 vs 4: $P=.23$).

Figure 1. Flow diagram of a preventive primary care program (TOF pilot2) from initial sampling of patients to participation rates in each of four invitation groups. GP: general practitioner.

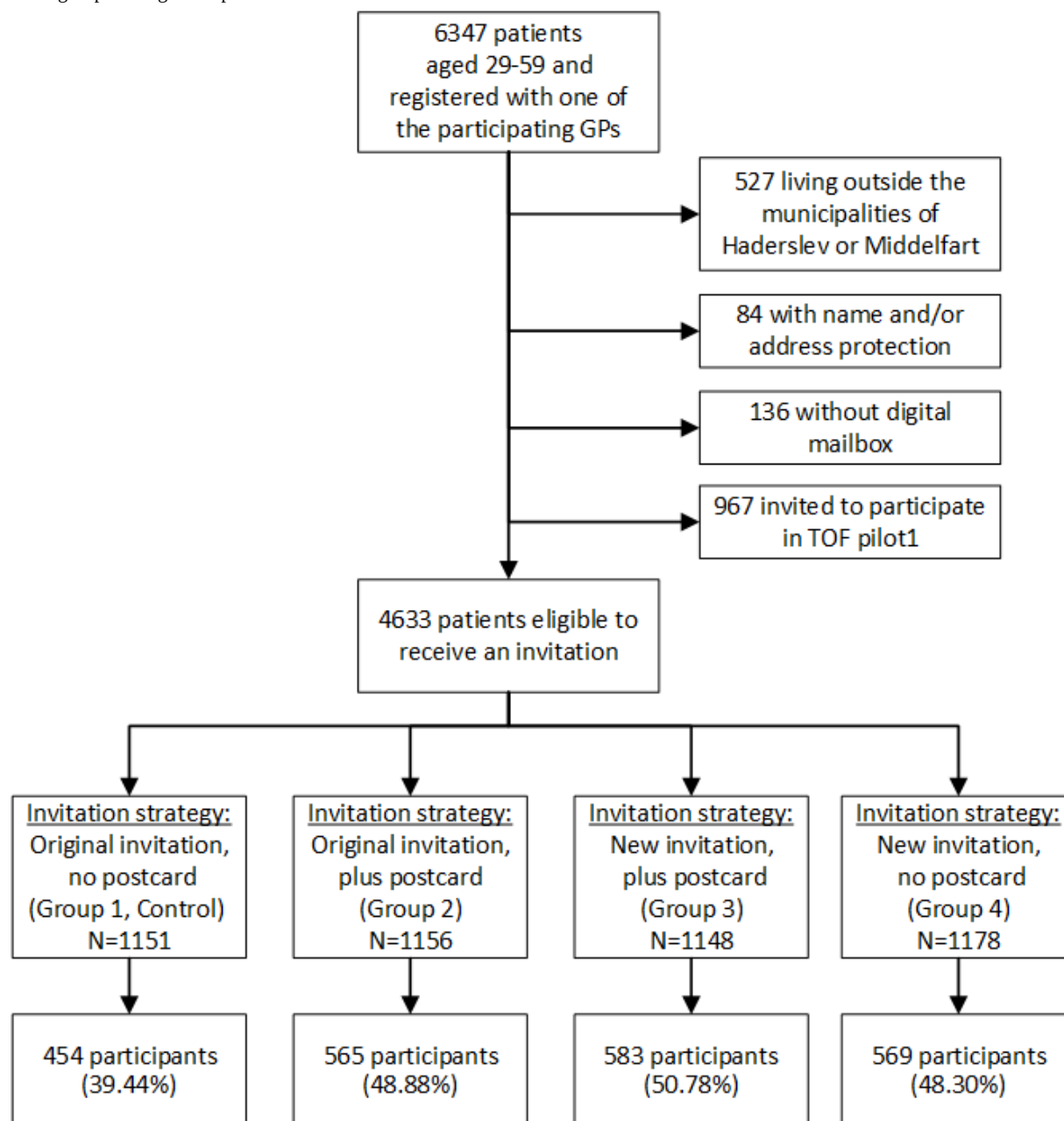


Table 2. Analysis of associations between participation, mode of invitation, and invitation elements among patients invited to take part in a preventive primary care program for lifestyle-related diseases (TOF pilot2).

Variable	Sample size (n)	Model 1 (crude)		Model 2 (adjusted for age and sex)	
		RR ^a (95% CI)	<i>P</i> value	RR (95% CI)	<i>P</i> value
Mode of invitation					
Original invitation (control)	1151	1 (0) ^b	N/A ^c	1 (0) ^b	N/A
Original invitation + postcard	1156	1.24 (1.13-1.36)	<.001	1.23 (1.12-1.34)	<.001
New invitation + postcard	1148	1.29 (1.17-1.41)	<.001	1.29 (1.18-1.41)	<.001
New invitation	1178	1.22 (1.12-1.34)	<.001	1.22 (1.11-1.33)	<.001
Individual invitation elements					
Prenotification postcard					
No	2329	1 (0) ^b	N/A	1 (0) ^b	N/A
Yes	2304	1.13 (1.07-1.21)	<.001	1.13 (1.07-1.20)	<.001
Invitation					
Original	2307	1 (0) ^b	N/A	1 (0) ^b	N/A
New	2326	1.12 (1.05-1.19)	<.001	1.12 (1.06-1.19)	<.001

^aRR: risk ratio.^bReference group.^cN/A: not applicable.

Table 3 shows the difference in participation rates between the intervention groups and the control group by sociodemographic subgroups.

The three intervention groups recorded higher participation rates than the control group for both female and male patients and across all age groups. In addition, patients of any educational level, who were of Danish origin, employed, or with a family

income above the lowest quartile, were more likely to participate if they received one of the three new modes of invitation. Further, with or without the prenotification postcard, the new invitation increased participation rates for patients with a family income in the lowest quartile. By contrast, none of the new modes of invitation affected participation rates for unemployed patients or patients of non-Danish (Western or non-Western) origin.

Table 3. Participation in a preventive primary care program (TOF pilot2) and risk ratios of participation obtained from comparing each of the new modes of invitation (Groups 2-4) to the original one (Group 1, control) overall and by sociodemographic subgroups.

Participation over- all and by sociode- mographic sub- groups	Group 1: Original invi- tation (con- trol)	Group 2: Original invitation + postcard			Group 3: New invitation + postcard			Group 4: New invitation			Total
	Value, n (%)	Value, n (%)	P value ^a	RR ^b (95% CI)	Value, n (%)	P value	RR (95% CI)	Value, n (%)	P value	RR (95% CI)	Value, n (%)
Overall	454 (39.44)	565 (48.88)	<.001	1.24 (1.13-1.36)	583 (50.78)	<.001	1.29 (1.17-1.41)	569 (48.30)	<.001	1.22 (1.12-1.34)	2171 (46.86)
10-year age groups											
29-39 years	75 (24.83)	115 (36.51)	.002	1.47 (1.15-1.88)	134 (41.74)	<.001	1.68 (1.33-2.13)	102 (33.01)	.03	1.33 (1.03-1.71)	426 (34.16)
40-49 years	185 (41.95)	218 (49.89)	.02	1.19 (1.03-1.37)	230 (50.44)	.01	1.20 (1.04-1.39)	217 (49.10)	.03	1.17 (1.01-1.35)	850 (47.86)
50-60 years	194 (47.55)	232 (57.43)	.005	1.21 (1.06-1.38)	219 (59.03)	.001	1.24 (1.09-1.42)	250 (58.55)	.001	1.23 (1.08-1.40)	895 (55.59)
Sex											
Male	179 (32.84)	234 (43.90)	<.001	1.34 (1.15-1.56)	248 (44.21)	<.001	1.35 (1.16-1.57)	242 (43.21)	<.001	1.32 (1.13-1.53)	903 (41.06)
Female	275 (45.38)	331 (53.13)	.007	1.17 (1.04-1.31)	335 (57.07)	<.001	1.26 (1.12-1.41)	327 (52.91)	.008	1.17 (1.04-1.31)	1268 (52.10)
Country of origin											
Denmark	432 (41.38)	530 (50.86)	<.001	1.23 (1.12-1.35)	553 (53.17)	<.001	1.29 (1.17-1.41)	541 (50.85)	<.001	1.23 (1.12-1.35)	2056 (49.07)
Western	8 (25.00)	18 (47.37)	.054	1.89 (0.95-3.77)	16 (35.56)	.32	1.42 (0.69-2.81)	14 (33.33)	.44	1.33 (0.64-2.79)	56 (35.67)
Non-Western	11 (16.18)	16 (21.92)	.39	1.35 (0.68-2.71)	13 (22.41)	.37	1.39 (0.67-2.85)	13 (19.70)	.59	1.22 (0.59-2.52)	53 (20.00)
Highest educational attainment											
Secondary school	49 (25.00)	71 (37.57)	.008	1.50 (1.11-2.04)	80 (43.72)	<.001	1.75 (1.30-2.34)	74 (38.95)	.003	1.56 (1.15-2.10)	274 (36.15)
Highschool, vocational education, higher education, or higher education-master's level	393 (42.86)	484 (51.82)	<.001	1.21 (1.10-1.33)	487 (52.82)	<.001	1.23 (1.12-1.36)	484 (51.11)	<.001	1.19 (1.08-1.31)	1848 (49.68)
Employment status											
Unem- ployed/on benefits, so- cial welfare recipients, or other	76 (36.54)	77 (37.20)	.89	1.02 (0.79-1.31)	78 (36.62)	.99	1.00 (0.78-1.29)	77 (35.81)	.88	0.98 (0.76-1.26)	308 (36.54)

Participation overall and by sociodemographic subgroups	Group 1: Original invitation (control)	Group 2: Original invitation + postcard			Group 3: New invitation + postcard			Group 4: New invitation			Total
	Value, n (%)	Value, n (%)	P value ^a	RR ^b (95% CI)	Value, n (%)	P value	RR (95% CI)	Value, n (%)	P value	RR (95% CI)	Value, n (%)
Employed or self-employed	377 (40.19)	488 (51.53)	<.001	1.28 (1.16-1.42)	504 (54.08)	<.001	1.35 (1.22-1.48)	492 (51.36)	<.001	1.28 (1.16-1.41)	1861 (49.30)
Family income											
Lowest quartile	65 (25.00)	90 (32.61)	.052	1.30 (1.00-1.71)	97 (35.93)	.006	1.44 (1.10-1.87)	100 (34.72)	.01	1.39 (1.07-1.81)	352 (32.18)
Greater than the lowest quartile	386 (43.67)	474 (54.05)	<.001	1.24 (1.12-1.36)	485 (55.56)	<.001	1.27 (1.16-1.40)	468 (52.94)	<.001	1.21 (1.10-1.34)	1813 (51.53)

^aP values for differences in the participation rate between the individual intervention groups (Groups 2-4) and the control group (Group 1).

^bRR: risk ratio.

Invitation Comprehensibility

A total of 2120 (97.65%) participants answered the question on invitation comprehensibility. Of these, 987 had received the original invitation and 1133 had received the new one. The response rates for these two groups were thus 96.86% (987/1019) and 98.35% (1133/1152), respectively.

Among participants who received the new invitation, 965 (85.17%) found it easy to understand. This level of invitation comprehensibility was comparable to that of the original invitation group (n=813, 82.37%, RR 1.03; 95% CI 1.00-1.07; $P=.08$). Including missing data in the “fairly understandable/difficult to understand” group did not alter these results.

Discussion

Focus of the Study

We investigated whether different iterations of invitation strategies might increase the participation rate in a preventive primary care intervention targeting lifestyle-related diseases. We focused on participation in general as well as participation specifically among men and people with low educational attainment.

Effect on Participation

Each of the experimental invitation strategies greatly increased participation rates compared with the control. There were, however, no statistical differences in impact between the three intervention groups (Groups 2-4). Overall, participation rates increased for both women and men, as well as for patients of high and low socioeconomic status. However, the consistently higher RRs for men compared to women, patients with low compared to high educational attainment, and patients with low compared to high family income may indicate a larger relative effect in these groups. Firm conclusions for this effect cannot be made based on the reported results though. Taken together, the impact of an invitation strategy on participation in preventive programs like the one tested here may also have implications

for other settings, such as worksite health promotion programs, which often have low/moderate participation rates [33].

The apparent effect on participation among non-Danish patients was not significant. This may be due to the rather small sample sizes, but cultural characteristics and the fact that invitations and postcards were in Danish most likely also contributed [34,35]. Similarly, the relatively low participation rates for unemployed patients may signify greater participation barriers in this group. For example, unemployment has previously been shown to be associated with poorer mental and physical health, which may impede the likelihood of taking part in research studies [36,37]. We also note, however, that the revised invitation was designed to target men and patients of low educational attainment. As such, the purposely sampled focus group for invitation design did not include unemployed patients or patients of non-Danish origin specifically. Involving patients from these target groups in future design processes might allow for invitation strategy adjustments to accommodate these groups as well.

Specific Invitation Elements

Supplementing both the original and revised invitations with the prenotification postcard increased participation rates. Of particular note, the notoriously hard-to-reach youngest age group (29-39 years old) also responded well to this mode of invitation. These results are consistent with previous research showing higher participation and odds of response to questionnaires when the study invitation is preceded by a prenotification [8,38]. The specific content of the postcard likely contributed further as the teaser sentence “Something important is on its way – remember to keep an eye on your digital mailbox” may have primed participants to be more attentive to their digital mailbox and the upcoming invitation. Indeed, a systematic review on methods to increase responses to postal and electronic questionnaires revealed that using teasers on the envelope can increase participation [38].

The new invitation comprised action-focused behavioral instructions (pictograms), shorter sentences, and jargon - free

language. Formatting invitation content in this way has been demonstrated to facilitate research participation in both the general population as well as hard-to-reach subgroups, including people with intellectual disabilities [22]. In other words, the deliberate focus on creating an easy-to-grasp lay-person invitation may have contributed to engaging a broader audience.

Compared to the original invitation, the heading of the new invitation included a clear incentive (“Get a free health check”) coupled with an action-focused instruction (“Use five minutes on a questionnaire”). Although the effect of such specific wording is difficult to assess, past research would suggest that it may have influenced participation. For instance, Sallis et al increased participation in a National Health Service health check program by using behavioral instructions and concrete statements in the invitation [11].

In this study, most participants found both invitations easy to understand with no significant difference in comprehension between the two. As invitation comprehensibility was only assessed in patients consenting to participate, it cannot be ruled out that nonparticipation may be related to invitation comprehensibility, that is, people might have declined the invitation because it was not sufficiently clear to them. Nonetheless, results from a study on female nonparticipants in a screening program for cardiovascular diseases and diabetes revealed that although the participants believed they understood the screening invitation, they seemed unaware of what the examination entailed [39]. Thus, self-reported understanding of similar written information may be somewhat biased.

While we also assessed other potential factors for nonparticipation besides invitation comprehensibility (eg, motivation and time constraints) [9], these results will be published elsewhere as they are deemed beyond the scope of this article.

Clinical Implications

Preventive programs rely on high uptake to optimize clinical effectiveness and cost-effectiveness [40]. However, in many studies, patient uptake is disproportionately higher for women than men and for patients of high than lower socioeconomic status [2,3,5,6]. To this end, initiatives to increase uptake among men and among men and women of low socioeconomic status should be prioritized to mitigate inequality in health.

Results from this study indicate that targeting men and patients of low educational attainment led to an increase in participation not only in these groups, but also among women and patients of high educational attainment. Despite this, co-design processes, like the one tested here, may still be relevant in efforts focusing on social inequality in health, as the relative increase in participation rates appeared to be higher for men, patients with low educational attainment, and patients with low family income.

Further, in order to reach the most socioeconomically disadvantaged groups, the invitation strategy should probably be combined with other more individual-oriented recruitment approaches [41,42]. Recent evidence suggests that the use of support workers in general practice with the specific aim to increase uptake of preventive health checks may greatly increase

participation among patients from deprived areas and among minority groups [43]. In addition, Cook et al found that invitation by telephone was highly effective for recruiting patients from specific ethnic minority groups [44].

Digital Versus Paper-Based Invitations

In this study, we used a digital mailbox and a web-based digital support system to distribute invitations and obtain informed consent. Although access to both systems required a two-phase log in, results from the TOF pilot studies showed participation rates comparable to those obtained in other studies that relied on paper-based invitations [2,4,11]. This may in part be explained by the fact that the digital mailbox is a trusted and familiar mode of communication between the individual citizen and municipal, regional, and national authorities in Denmark. Another advantage relates to cost. For example, Ebert et al found that web-based invitations were more cost-effective (by a factor of 10) than paper-based ones and that nonrespondents were demographically similar in the two groups, indicating low risk of selection bias [45]. In this study, combining the new invitation with a prenotification postcard did not outperform the new digital invitation alone in terms of participation rates. Therefore, the additional cost associated with distributing the prenotification postcard should be considered in any relevant invitation strategy.

Strengths and Limitations

In this study, sample size calculations were based on the expected overall increase in the participation rate rather than specific participation rates in socioeconomic subgroups. Repeating the study with a larger population might reveal an increase in participation rates for patients of non-Danish origin. In addition, a larger study would reveal whether the observed relative increase in participation rates for men compared to women and for patients with low compared to high educational attainment and family income is replicable.

From the results, it was not possible to assess which specific aspects of the revised invitation and/or prenotification postcard drove the observed effects. Using a more rigorous and theory-based co-design procedure may lead to more insights into these aspects. In addition, it would be relevant to test the potential effect of combining different recruitment strategies in future research.

As nearly all (98%) Danish citizens are registered with a GP [46], the study sample resembled the general Danish population for this particular age group. However, since patient sampling was based on GP clusters, differences in participation rates between these clusters cannot be excluded, and future studies may consider looking into this.

We randomized our sample by household to ensure that patients living together were invited in the same way. This procedure along with sample representability provided the study with high internal as well as external validity. In addition, this study provides important new insights into the potential effects of making a special effort when it comes to invitation strategies for preventive health checks. Further, tailoring the invitations to specific groups by employing co-design procedures may help attenuate inequalities in health.

Conclusion

The results of this study showed that high improvements in participation rates in a preventive health check intervention may be obtained by taking a co-design approach to the invitation

strategy that involves communication professionals and target-group representatives. In particular, the increased participation of men and patients of low socioeconomic status indicates the potential value of such initiatives to mitigate inequalities in health.

Acknowledgments

The authors wish to thank the patients, general practitioners, and municipal health professionals who took part in the study. The study received funding from Trygfonden (grant no 125508), The Committee of Quality and Education in General Practice in the Region of Southern Denmark (grant no 18/527), and The Region of Southern Denmark (grant no 11/13244). The sponsors had no involvement in the review or approval of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Original invitation.

[[PDF File \(Adobe PDF File\), 251 KB - publichealth_v7i3e25617_app1.pdf](#)]

Multimedia Appendix 2

New invitation.

[[PDF File \(Adobe PDF File\), 218 KB - publichealth_v7i3e25617_app2.pdf](#)]

Multimedia Appendix 3

Prenotification postcard.

[[PDF File \(Adobe PDF File\), 160 KB - publichealth_v7i3e25617_app3.pdf](#)]

Multimedia Appendix 4

CONSORT 2010 checklist.

[[DOC File , 219 KB - publichealth_v7i3e25617_app4.doc](#)]

References

1. van den Brekel-Dijkstra K, Rengers AH, Niessen MAJ, de Wit NJ, Kraaijenhagen RA. Personalized prevention approach with use of a web-based cardiovascular risk assessment with tailored lifestyle follow-up in primary care practice--a pilot study. *Eur J Prev Cardiol* 2016 Mar;23(5):544-551. [doi: [10.1177/2047487315591441](#)] [Medline: [26080811](#)]
2. Larsen LB, Sondergaard J, Thomsen JL, Halling A, Sønderlund AL, Christensen JR, et al. Digital Recruitment and Acceptance of a Stepwise Model to Prevent Chronic Disease in the Danish Primary Care Sector: Cross-Sectional Study. *J Med Internet Res* 2019 Jan 21;21(1):e11658 [[FREE Full text](#)] [doi: [10.2196/11658](#)] [Medline: [30664466](#)]
3. Dryden R, Williams B, McCowan C, Themessl-Huber M. What do we know about who does and does not attend general health checks? Findings from a narrative scoping review. *BMC Public Health* 2012 Aug 31;12(1):723 [[FREE Full text](#)] [doi: [10.1186/1471-2458-12-723](#)] [Medline: [22938046](#)]
4. Van der Meer V, Nielen MM, Drenthen AJ, Van Vliet M, Assendelft WJ, Schellevis FG. Cardiometabolic prevention consultation in the Netherlands: screening uptake and detection of cardiometabolic risk factors and diseases--a pilot study. *BMC Fam Pract* 2013 Feb 26;14(1):29 [[FREE Full text](#)] [doi: [10.1186/1471-2296-14-29](#)] [Medline: [23442805](#)]
5. Bender AM, Jørgensen T, Helbech B, Linneberg A, Pisinger C. Socioeconomic position and participation in baseline and follow-up visits: the Inter99 study. *Eur J Prev Cardiol* 2014 Jul;21(7):899-905. [doi: [10.1177/2047487312472076](#)] [Medline: [23233551](#)]
6. Bjerregaard AL, Maindal HT, Bruun NH, Sandbæk A. Patterns of attendance to health checks in a municipality setting: the Danish 'Check Your Health Preventive Program'. *Prev Med Rep* 2017 Mar;5:175-182 [[FREE Full text](#)] [doi: [10.1016/j.pmedr.2016.12.011](#)] [Medline: [28050340](#)]
7. Harte E, MacLure C, Martin A, Saunders CL, Meads C, Walter FM, et al. Reasons why people do not attend NHS Health Checks: a systematic review and qualitative synthesis. *Br J Gen Pract* 2017 Dec 04;68(666):e28-e35. [doi: [10.3399/bjgp17x693929](#)]
8. Koitsalu M, Eklund M, Adolfsson J, Grönberg H, Brandberg Y. Effects of pre-notification, invitation length, questionnaire length and reminder on participation rate: a quasi-randomised controlled trial. *BMC Med Res Methodol* 2018 Jan 05;18(1):3 [[FREE Full text](#)] [doi: [10.1186/s12874-017-0467-5](#)] [Medline: [29304734](#)]

9. de Waard AM, Wändell PE, Holzmann MJ, Korevaar JC, Hollander M, Gornitzki C, SPIMEU Research Group. Barriers and facilitators to participation in a health check for cardiometabolic diseases in primary care: A systematic review. *Eur J Prev Cardiol* 2018 Aug 19;25(12):1326-1340 [FREE Full text] [doi: [10.1177/2047487318780751](https://doi.org/10.1177/2047487318780751)] [Medline: [29916723](https://pubmed.ncbi.nlm.nih.gov/29916723/)]
10. Kiernan M, Phillips K, Fair JM, King AC. Using direct mail to recruit Hispanic adults into a dietary intervention: an experimental study. *Ann Behav Med* 2000;22(1):89-93. [doi: [10.1007/BF02895172](https://doi.org/10.1007/BF02895172)] [Medline: [10892533](https://pubmed.ncbi.nlm.nih.gov/10892533/)]
11. Sallis A, Bunten A, Bonus A, James A, Chadborn T, Berry D. The effectiveness of an enhanced invitation letter on uptake of National Health Service Health Checks in primary care: a pragmatic quasi-randomised controlled trial. *BMC Fam Pract* 2016 Mar 24;17(1):35 [FREE Full text] [doi: [10.1186/s12875-016-0426-y](https://doi.org/10.1186/s12875-016-0426-y)] [Medline: [27009045](https://pubmed.ncbi.nlm.nih.gov/27009045/)]
12. van Wonderen KE, Mohrs J, Ijff M, Bindels PJE, ter Riet G. Two simple strategies (adding a logo or a senior faculty's signature) failed to improve patient participation rates in a cohort study: randomized trial. *J Clin Epidemiol* 2008 Oct;61(10):971-977. [doi: [10.1016/j.jclinepi.2008.05.008](https://doi.org/10.1016/j.jclinepi.2008.05.008)] [Medline: [18762135](https://pubmed.ncbi.nlm.nih.gov/18762135/)]
13. Gidlow CJ, Ellis NJ, Riley V, Chadborn T, Bunten A, Iqbal Z, et al. Randomised controlled trial comparing uptake of NHS Health Check in response to standard letters, risk-personalised letters and telephone invitations. *BMC Public Health* 2019 Feb 21;19(1):224 [FREE Full text] [doi: [10.1186/s12889-019-6540-8](https://doi.org/10.1186/s12889-019-6540-8)] [Medline: [30791884](https://pubmed.ncbi.nlm.nih.gov/30791884/)]
14. Larsen LB, Sonderlund AL, Sondergaard J, Thomsen JL, Halling A, Hvidt NC, et al. Targeted prevention in primary care aimed at lifestyle-related diseases: a study protocol for a non-randomised pilot study. *BMC Fam Pract* 2018 Jul 21;19(1):124 [FREE Full text] [doi: [10.1186/s12875-018-0820-8](https://doi.org/10.1186/s12875-018-0820-8)] [Medline: [30031380](https://pubmed.ncbi.nlm.nih.gov/30031380/)]
15. Larsen LB, Sondergaard J, Thomsen JL, Halling A, Sønderlund AL, Christensen JR, et al. Step-wise approach to prevention of chronic diseases in the Danish primary care sector with the use of a personal digital health profile and targeted follow-up - an assessment of attendance. *BMC Public Health* 2019 Aug 13;19(1):1092 [FREE Full text] [doi: [10.1186/s12889-019-7419-4](https://doi.org/10.1186/s12889-019-7419-4)] [Medline: [31409343](https://pubmed.ncbi.nlm.nih.gov/31409343/)]
16. Statistik om Digital Post. Digitaliseringsstyrelsen. URL: <https://digst.dk/it-loesninger/digital-post/om-loesningen/tal-og-statistik-om-digital-post/> [accessed 2020-11-02]
17. Sommer I, Griebler U, Mahlke P, Thaler K, Bouskill K, Gartlehner G, et al. Socioeconomic inequalities in non-communicable diseases and their risk factors: an overview of systematic reviews. *BMC Public Health* 2015 Sep 18;15(1):914 [FREE Full text] [doi: [10.1186/s12889-015-2227-y](https://doi.org/10.1186/s12889-015-2227-y)] [Medline: [26385563](https://pubmed.ncbi.nlm.nih.gov/26385563/)]
18. Di Tanna GL, Wirtz H, Burrows KL, Globe G. Evaluating risk prediction models for adults with heart failure: A systematic literature review. *PLoS ONE* 2020 Jan 15;15(1):e0224135 [FREE Full text] [doi: [10.1371/journal.pone.0224135](https://doi.org/10.1371/journal.pone.0224135)] [Medline: [31940350](https://pubmed.ncbi.nlm.nih.gov/31940350/)]
19. Steen M, Manshot M, De Koning N. Benefits of Co-design in Service Design Projects. *International Journal of Design* 2011;5(2):53-60 [FREE Full text]
20. Men's Health Society, Denmark. URL: <http://www.sundmand.dk/English.php> [accessed 2020-11-09]
21. Ley P. Communicating with Patients: Improving Communication, Satisfaction, and Compliance. Kent, UK: Croom Helm Ltd; 1988.
22. Chinn D. An empirical examination of the use of Easy Read health information in health consultations involving patients with intellectual disabilities. *J Appl Res Intellect Disabil* 2020 Mar;33(2):232-247. [doi: [10.1111/jar.12657](https://doi.org/10.1111/jar.12657)] [Medline: [31475419](https://pubmed.ncbi.nlm.nih.gov/31475419/)]
23. Thygesen M. Sundhedskampagner og social ulighed i sundhed - Hvordan kortuddannede forstår og bruger sundhedskampagner. University of Southern Denmark. 2013. URL: https://www.kommunikationsforum.dk/log/multimedia/PDF%20og%20andre%20dokumenter/Specialer/Specialer_Sundhedskampagner_og_social_ulighed_i_sundhed.pdf [accessed 2021-02-28]
24. Et forebyggelses- og egenomsorgsprojekt målrettet mænd i et samarbejde mellem Ringkøbing-Skjern Kommune og Region Midtjylland. Sund Mand. 2017. URL: https://www.rm.dk/api/NewESDHBBlock/DownloadFile?agendaPath=%5C%5CRMAPPS0221.onerm.dk%5Ccms01-ext%5CESDH%20Data%5CRM_Internet%5CDagsordener%5CKvalitets- og Efteru%202015%5C27-02-2014%5CAaben_dagsorden&appendixId=63596 [accessed 2020-11-02]
25. Mænds Mødesteder - Sønderborg. 2020. URL: <https://www.mmdanmark.dk/soenderborg/forside> [accessed 2020-11-02]
26. Borch I, Fischer L. Spørg hvorfor: Sådan skaber du meningsfulde kvalitative undersøgelser med kunder og brugere. 1. udgave ed. Nyt Teknisk Forlag. Odense, Denmark: Nyt Teknisk Forlag; 2012.
27. Kastarinen MJ, Puska PM, Korhonen MH, Mustonen JN, Salomaa VV, Sundvall JE, LIHEF Study Group. Non-pharmacological treatment of hypertension in primary health care: a 2-year open randomized controlled trial of lifestyle intervention against hypertension in eastern Finland. *J Hypertens* 2002 Dec;20(12):2505-2512. [doi: [10.1097/00004872-200212000-00031](https://doi.org/10.1097/00004872-200212000-00031)] [Medline: [12473876](https://pubmed.ncbi.nlm.nih.gov/12473876/)]
28. Roderick P, Ruddock V, Hunt P, Miller G. A randomized trial to evaluate the effectiveness of dietary advice by practice nurses in lowering diet-related coronary heart disease risk. *Br J Gen Pract* 1997 Jan;47(414):7-12 [FREE Full text] [Medline: [9115804](https://pubmed.ncbi.nlm.nih.gov/9115804/)]
29. Baron J, Gleason R, Crowe B, Mann J. Preliminary trial of the effect of general practice based nutritional advice. *Br J Gen Pract* 1990 Apr;40(333):137-141 [FREE Full text] [Medline: [2115348](https://pubmed.ncbi.nlm.nih.gov/2115348/)]

30. No authors listed. Effectiveness of health checks conducted by nurses in primary care: final results of the OXCHECK study. Imperial Cancer Research Fund OXCHECK Study Group. *BMJ* 1995 Apr 29;310(6987):1099-1104 [[FREE Full text](#)] [Medline: [7742676](#)]
31. Ranganathan P, Aggarwal R, Pramesh C. Common pitfalls in statistical analysis: Odds versus risk. *Perspect Clin Res* 2015;6(4):222-224 [[FREE Full text](#)] [doi: [10.4103/2229-3485.167092](#)] [Medline: [26623395](#)]
32. An Adjusted Preventive Program Against Lifestyle Related Diseases (TOFpilot2). ClinicalTrials.gov. URL: <https://clinicaltrials.gov/ct2/show/NCT03913585> [accessed 2020-11-02]
33. Robroek SJ, van Lenthe FJ, van Empelen P, Burdorf A. Determinants of participation in worksite health promotion programmes: a systematic review. *Int J Behav Nutr Phys Act* 2009 May 20;6:26 [[FREE Full text](#)] [doi: [10.1186/1479-5868-6-26](#)] [Medline: [19457246](#)]
34. Kwok C, Endrawes G, Lee CF. Cultural Beliefs and Attitudes About Breast Cancer and Screening Practices Among Arabic Women in Australia. *Cancer Nurs* 2016;39(5):367-374. [doi: [10.1097/NCC.0000000000000325](#)] [Medline: [26645110](#)]
35. Ray-Mazumder S. Role of gender, insurance status and culture in attitudes and health behavior in a US Chinese student population. *Ethn Health* 2001;6(3-4):197-209. [doi: [10.1080/13557850120078125](#)] [Medline: [11696931](#)]
36. Kessler RC, House JS, Turner JB. Unemployment and health in a community sample. *J Health Soc Behav* 1987 Mar;28(1):51-59. [Medline: [3571906](#)]
37. Virtanen P, Liukkonen V, Vahtera J, Kivimäki M, Koskenvuo M. Health inequalities in the workforce: the labour market core-periphery structure. *Int J Epidemiol* 2003 Dec;32(6):1015-1021. [doi: [10.1093/ije/dyg319](#)] [Medline: [14681267](#)]
38. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database Syst Rev* 2009 Jul 08(3):MR000008. [doi: [10.1002/14651858.MR000008.pub4](#)] [Medline: [19588449](#)]
39. Dahl M, Lindholt J, Sogaard R, Frost L, Andersen LS, Lorentzen V. An interview-based study of nonattendance at screening for cardiovascular diseases and diabetes in older women: Nonattendees' perspectives. *J Clin Nurs* 2018 Mar;27(5-6):939-948. [doi: [10.1111/jocn.14018](#)] [Medline: [28815826](#)]
40. Economic Modelling for Vascular Checks. Department of Health, Vascular Policy Team. 2008. URL: http://www.em-online.com/download/medical_article/36931_DH_085917.pdf [accessed 2020-11-02]
41. Harkins C, Shaw R, Gillies M, Sloan H, Macintyre K, Scoular A, et al. Overcoming barriers to engaging socio-economically disadvantaged populations in CHD primary prevention: a qualitative study. *BMC Public Health* 2010 Jul 02;10:391 [[FREE Full text](#)] [doi: [10.1186/1471-2458-10-391](#)] [Medline: [20598130](#)]
42. Sinclair A, Alexander HA. Using outreach to involve the hard-to-reach in a health check: what difference does it make? *Public Health* 2012 Feb;126(2):87-95. [doi: [10.1016/j.puhe.2011.11.004](#)] [Medline: [22206983](#)]
43. Kósa K, Katona C, Papp M, Fürjes G, Sándor J, Bíró K, et al. Health mediators as members of multidisciplinary group practice: lessons learned from a primary health care model programme in Hungary. *BMC Fam Pract* 2020 Jan 28;21(1):19 [[FREE Full text](#)] [doi: [10.1186/s12875-020-1092-7](#)] [Medline: [31992209](#)]
44. Cook EJ, Sharp C, Randhawa G, Guppy A, Gangotra R, Cox J. Who uses NHS health checks? Investigating the impact of ethnicity and gender and method of invitation on uptake of NHS health checks. *Int J Equity Health* 2016 Jan 20;15:13 [[FREE Full text](#)] [doi: [10.1186/s12939-016-0303-2](#)] [Medline: [26791963](#)]
45. Ebert JF, Huibers L, Christensen B, Christensen MB. Paper- or Web-Based Questionnaire Invitations as a Method for Data Collection: Cross-Sectional Comparative Study of Differences in Response Rate, Completeness of Data, and Financial Cost. *J Med Internet Res* 2018 Jan 23;20(1):e24 [[FREE Full text](#)] [doi: [10.2196/jmir.8353](#)] [Medline: [29362206](#)]
46. Pedersen KM, Andersen JS, Søndergaard J. General practice and primary health care in Denmark. *J Am Board Fam Med* 2012 Mar;25 Suppl 1:S34-S38 [[FREE Full text](#)] [doi: [10.3122/jabfm.2012.02.110216](#)] [Medline: [22403249](#)]

Abbreviations

GP: general practitioner

OR: odds ratio

RR: risk ratio

TOF: Early Detection and Protection Project (translated Danish abbreviation)

Edited by T Sanchez; submitted 09.11.20; peer-reviewed by N Schoenberg, M Niessen; comments to author 16.12.20; revised version received 22.01.21; accepted 15.02.21; published 10.03.21.

Please cite as:

Thilsing T, Larsen LB, Sonderlund AL, Andreassen SS, Christensen JR, Svensson NH, Dahl M, Sondergaard J

Effects of a Co-Design–Based Invitation Strategy on Participation in a Preventive Health Check Program: Randomized Controlled Trial

JMIR Public Health Surveill 2021;7(3):e25617

URL: <https://publichealth.jmir.org/2021/3/e25617>

doi:[10.2196/25617](https://doi.org/10.2196/25617)

PMID:[33688836](https://pubmed.ncbi.nlm.nih.gov/33688836/)

©Trine Thilsing, Lars Bruun Larsen, Anders Larrabee Sonderlund, Signe Skaarup Andreassen, Jeanette Reffstrup Christensen, Nanna Herning Svensson, Marie Dahl, Jens Sondergaard. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 10.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Estimating the Incidence of Conjunctivitis by Comparing the Frequency of Google Search Terms With Clinical Data: Retrospective Study

Paola Kammrath Betancor¹, Dr med; Linda Tizek²; Alexander Zink², PD, Dr med; Thomas Reinhard¹, Prof Dr; Daniel Böhlinger¹, Prof Dr

¹Eye Center, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany

²Department of Dermatology and Allergy, School of Medicine, Technical University of Munich, Munich, Germany

Corresponding Author:

Paola Kammrath Betancor, Dr med

Eye Center, Medical Center

Faculty of Medicine

University of Freiburg

Killianstraße 5

Freiburg, D-79106

Germany

Phone: 49 76127040010

Email: paola.kammrath.betancor@uniklinik-freiburg.de

Abstract

Background: Infectious conjunctivitis is contagious and may lead to an outbreak. Prevention systems can help to avoid an outbreak.

Objective: We aimed to evaluate if Google search data on conjunctivitis and associated terms can be used to estimate the incidence and if the data can provide an estimation for outbreaks.

Methods: We obtained Google search data over 4 years for the German term for conjunctivitis (“Bindehautentzündung”) and 714 associated terms in 12 selected German cities and Germany as a whole using the Google AdWords Keyword Planner. The search volume from Freiburg was correlated with clinical data from the Freiburg emergency practice (Eye Center University of Freiburg).

Results: The search volume for the German term for conjunctivitis in Germany as a whole and in the 12 German cities showed a highly uniform seasonal pattern. Cross-correlation between the temporal search frequencies in Germany as a whole and the 12 selected cities was high without any lag. Cross-correlation of the search volume in Freiburg with the frequency of conjunctivitis (International Statistical Classification of Diseases and Related Health Problems [ICD] code group “H10.-”) from the centralized ophthalmologic emergency practice in Freiburg revealed a considerable temporal association, with the emergency practice lagging behind the frequency. Additionally, Pearson correlation between the count of patients per month and the count of searches per month in Freiburg was statistically significant ($P=.04$).

Conclusions: We observed a close correlation between the Google search volume for the signs and symptoms of conjunctivitis and the frequency of patients with a congruent diagnosis in the Freiburg region. Regional deviations from the nationwide average search volume may therefore indicate a regional outbreak of infectious conjunctivitis.

(JMIR Public Health Surveill 2021;7(3):e22645) doi:[10.2196/22645](https://doi.org/10.2196/22645)

KEYWORDS

epidemic keratoconjunctivitis; big data; Google search; Freiburg clinical data

Introduction

Conjunctivitis is a very common diagnosis. Viral conjunctivitis is one of the most frequent eye infections worldwide [1]. It may

be extremely contagious and often begins with very uncharacteristic symptoms. They include redness, photophobia, watery eyes, and foreign body sensation. Patients can be infectious even days before having any symptoms and weeks

after symptoms have improved [1]. The differential diagnosis of acute infectious conjunctivitis is difficult, since the signs and symptoms of other inflammatory conjunctival conditions are quite similar [2]. The common differential diagnoses are seasonal allergic conjunctivitis and bacterial conjunctivitis. However, adenoviral keratoconjunctivitis is the most important disease because every infection may lead to keratitis nummularis. Hereby, subepithelial infiltrates can cause a reduction in the visual acuity and increase sensitivity to glare [3].

The correct diagnosis is important because infectious patients must be given sick leave to prevent spread, and the differential diagnoses need to be handled differently. This should be managed by an ophthalmologist because of the huge personal and socioeconomic impacts. The economic significance of epidemic keratoconjunctivitis has been quantified in some studies [4], but is very difficult to assess, as there is no reliable basis for a robust estimation of the affected population [5].

Since 2001, in Germany, adenoviral epidemic keratoconjunctivitis has been a notifiable disease according to the German Protection Against Infection Act (IfSG). If adenovirus is detected, the detecting laboratory (not the ophthalmologist) is obliged to report it. Only if the ophthalmologist has an urgent suspicion of an outbreak in his/her practice or hospital, he/she has the obligation to report it to the public health department, regardless of the pathogen [6]. Nevertheless, it can happen that the ophthalmologist diagnoses the disease on the basis of clinical symptoms alone. Therefore, it is possible that there is high underreporting. In addition, it can be assumed that a large proportion of adenovirus diseases are mild and self-healing and that these patients do not visit an ophthalmologist. From an epidemiological point of view, this is a serious problem because this precludes preventive measures to contain the outbreak [5].

Nowadays, the internet is one of the most used sources of health information. Patients “Google” their symptoms before seeing any doctor and may even end up without seeking professional health care at all [7]. Interestingly, infectious disease activity can be predicted by the frequency of internet searches. Studies related to influenza surveillance already proved that search-term surveillance is a possible supplement for disease-surveillance systems [8,9], whereas other studies showed a correlation of search data and cancer registries [10,11] or identified unmet medical needs in pruritus [12,13]. Internet biosurveillance systems use data from multiple sources. They refer to aggregator news sites like Google, social media sources like Facebook, text-based news sites, and apps [14].

The purpose of this study was to evaluate the potential of the analysis of Google search data to estimate the incidence of epidemic conjunctivitis. Therefore, we examined the relationship between searches for the German term for conjunctivitis (“Bindehautentzündung”) as well as 714 associated terms and actual conjunctivitis occurrence in the emergency practice of Freiburg over 4 years.

Methods

Data Collection

A retrospective longitudinal study using Google AdWords Keyword Planner was carried out to analyze the temporal course of keywords related to the German term for conjunctivitis (“Bindehautentzündung”) across Germany as well as across 12 German cities (Berlin, Hamburg, Munich, Cologne, Frankfurt [Main], Stuttgart, Hannover, Nuremberg, Freiburg, Leipzig, Rostock, and Kassel). Most of the cities were chosen because they are Germany’s largest cities by population (decreasing order: Berlin, Hamburg, Munich, Cologne, Frankfurt, Stuttgart, Leipzig, Hannover, and Nuremberg). Furthermore, Rostock was chosen to examine whether proximity to the coast has an influence on Google search volume. The remaining cities (Freiburg and Kassel) were chosen because a nationwide overview of various regions should be generated, and they are some of the largest cities within these regions. Primarily, the Keyword Planner is used for optimizing marketing campaigns; however, in recent years, it is being used to answer scientific questions. To examine the search volume for a specific topic, words or phrases are entered in the tool, and then, the tool finds the most relevant search terms with their average monthly search volume estimated by Google. In this study, the German term for conjunctivitis (“Bindehautentzündung”) was used to assess the search volume from November 2015 to October 2019. The region and language settings were set so that the Google data were limited solely to users across Germany whose language preference was German. The program identifies all keywords it considers relevant. The tool generally provides all the associated keywords for each city for an easier comparison, but this does not mean that in every city, people use the respective keywords. In our study, all 714 terms were used by people living in Berlin. The numbers of missing terms in the other cities were as follows: Hamburg, 4; Munich, 2; Cologne, 3; Frankfurt (Main), 13; Stuttgart, 9; Leipzig, 24; Hannover, 30; Nuremberg, 27; Freiburg, 90; Rostock, 136; and Kassel, 99. These missing search terms were imputed as 0. We eventually normalized the results per 100,000 inhabitants to account for differences in population size. The numbers of inhabitants in the German cities are provided in [Multimedia Appendix 1](#).

We correlated the search frequencies with the actual patient frequencies on the basis of ICD (International Statistical Classification of Diseases and Related Health Problems) coding for the patients in Freiburg using data from the centralized ophthalmologic emergency practice for the Freiburg region. This anonymized data set, including postal codes, ICD codes, and date of presentation, was kindly provided by the Association of Statutory Health Insurance Physicians of Südbaden (“Kassenärztliche Vereinigung Südbaden”). The ICD codes of all patient visits from November 2015 to October 2019 were searched for the ICD “H10.-” code group (it is primarily for “conjunctivitis”, but also includes “mucopurulent conjunctivitis”, “acute allergic conjunctivitis”, “other acute conjunctivitis”, “not specified acute conjunctivitis”, “chronic conjunctivitis”, “blepharoconjunctivitis”, “other conjunctivitis”, and “not specified conjunctivitis”). In this time period, 28,775

patients visited the emergency practice. Of these, 3882 patients were diagnosed with a code from the ICD “H10.-” code group.

Data Presentation and Statistical Analysis

We performed cross-correlations of the Google search data of Germany in total with the individual German cities. This was done to measure the regional variations of search frequencies for conjunctivitis within Germany. In the second step, we performed a cross-correlation between the search frequencies for Freiburg and the frequencies of the codes in the ICD “H10.-” code group from the centralized ophthalmologic emergency practice for the Freiburg region. This was done to estimate the association between search frequencies and frequencies of actual “H10.-” diagnoses. To achieve this, we had to interpolate the monthly search frequencies to obtain data at day resolution for the sum of all 714 associated terms. Statistical analysis was based on the R platform [15], using the function “qplot” (package ggplot2). For cross-correlation, we used “ccr” (base). Cross-correlation is a method to analyze similarities between two time series. This method provides a degree of similarity and a time lag that would be needed to align the two time series for optimal congruence. For Pearson correlation, we used “cor.test” (base). Pearson correlation assesses the strength of a linear association between two continuous variables. A P value $<.05$ was considered statistically significant [16].

Results

Google search data for the German term for conjunctivitis (“Bindehautentzündung”) from Germany and 12 German cities showed similar peaks of search queries per 100,000 inhabitants (Figure 1A and B). The search volume for the German term for conjunctivitis (“Bindehautentzündung”) in Germany totaled 5963 searches per 100,000 inhabitants, and in Freiburg, the search volume reached 8609 searches per 100,000 inhabitants. Since the general population tended to use a broad spectrum of layman’s terms for the symptoms of conjunctivitis, we repeated the procedure with 714 associated terms related to conjunctivitis that had been identified by Google Keyword Planner (Multimedia Appendix 2). The search volume in Germany as

a whole was lower than in the individual cities (Figure 2). Including all the associated terms, the search volumes were 12,132 searches per 100,000 inhabitants in Germany and 35,926 searches per 100,000 inhabitants in Freiburg. In the 12 cities, the cumulative average search volume of all associated terms per 100,000 inhabitants was 240% higher in comparison to that for Germany as a whole. This hints toward lesser usage of Google in rural areas in comparison with the bigger cities included in our analysis. The curves for the search frequency of all associated terms showed a clear seasonal pattern (Figure 1). Cross-correlation between the temporal search frequencies in Germany as a whole and the 12 selected cities was generally high. The lowest cross-correlation coefficient (r) was 0.6 in Nuremberg and 0.8 for all other cities. This indicates a nationwide seasonal trend.

From 2015 to 2019, the centralized ophthalmologic emergency practice for the Freiburg region (“KV-Notfallpraxis Freiburg”) treated on average 5755 patients per year. Of these, an average of 776 had conjunctivitis every year. The demographic analysis of the patients diagnosed with ICD codes from the “H10.-” ICD group revealed a strong predominance of place of residence in Freiburg and its surroundings (data not shown). On average, 48% of all conjunctivitis cases belonged to Freiburg and its surroundings.

The cumulative Google search frequency of all associated terms in Freiburg revealed a temporal correlation with the density of patients diagnosed with conjunctivitis (ICD “H10.-”) in the centralized ophthalmologic emergency practice for the Freiburg region. The time lag between both time series was 30 days. The peak correlation coefficient was 0.39 (Figure 3). In this analysis, values higher than 0.04 have to be considered statistically significant as values ± 0.04 are the extrema at an alpha level of .05 when one of the functions contains only white noise [17].

Pearson correlation between the count of patients per month from the centralized ophthalmologic emergency practice for the Freiburg region and the corresponding count of searches per month yielded a correlation coefficient of 0.3 ($P=.04$).

Figure 1. Search process for the German term for conjunctivitis (“Bindehautentzündung”) and 714 associated terms. (A) Search history of the term “conjunctivitis” from November 2015 to October 2019 in Germany and in the cities Berlin, Hamburg, Munich, Cologne, Frankfurt (Main), Stuttgart, Leipzig, Hannover, Nuremberg, Freiburg, Rostock, and Kassel per 100,000 inhabitants. (B) Same representation as (A) but for the search process of the term “conjunctivitis” and the 714 associated terms.

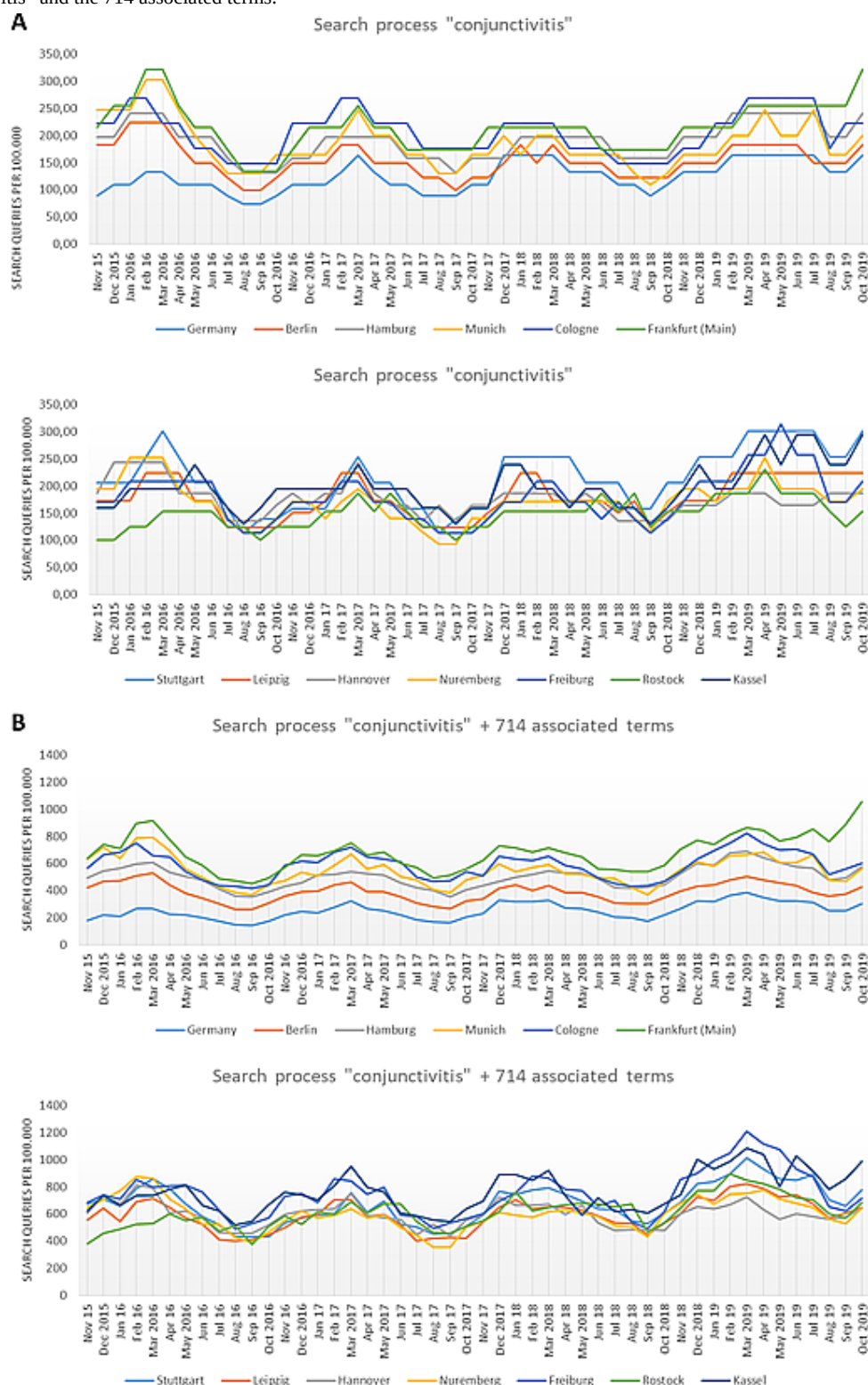


Figure 2. Density plot. Patients diagnosed with conjunctivitis (ICD code group "H10.-") in the centralized ophthalmologic emergency practice for the Freiburg region (light blue) superimposed with normalized cumulative Google search frequencies of all 714 terms for conjunctivitis. All curves show the same recurrent seasonal trend.

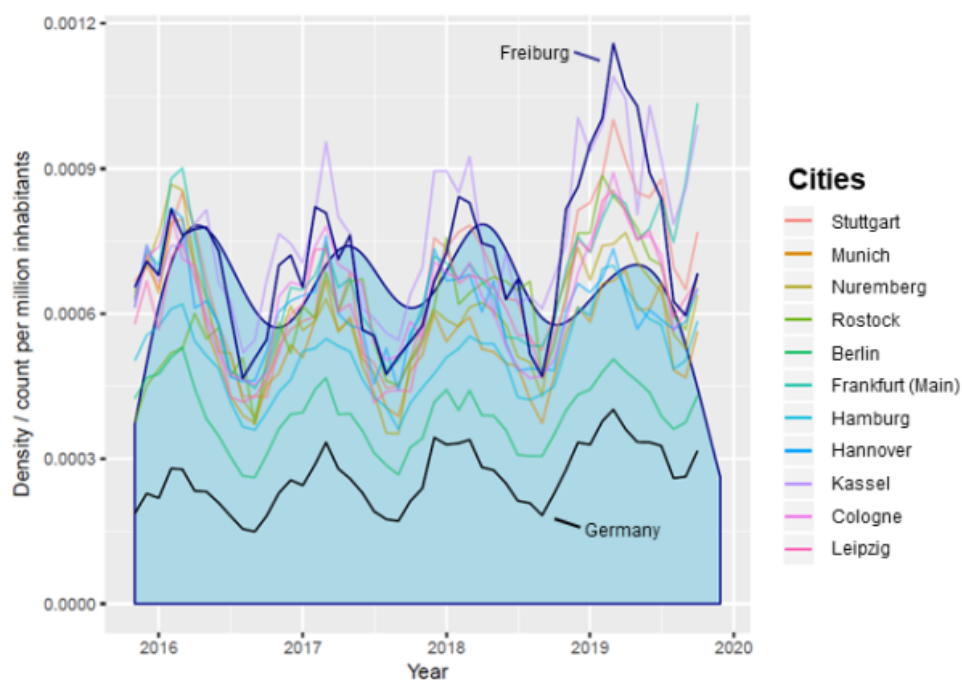
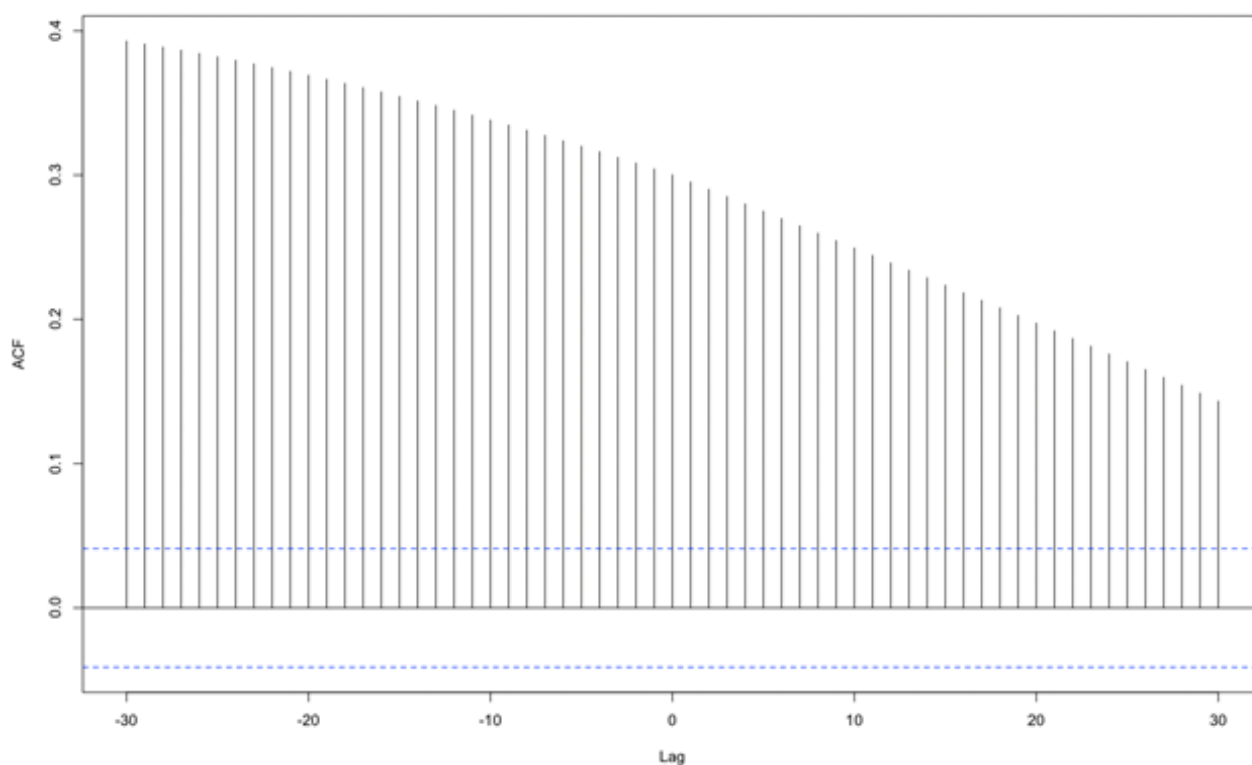


Figure 3. Cross-correlation between Google searches in Freiburg and the density of patients diagnosed with conjunctivitis (ICD code group "H10.-") in the Freiburg region. The visit density in the emergency practice lags behind the Google search. Each vertical bar corresponds to a bivariate correlation of the values at interest between both time series, with the second series shifted for the given time lag (given in days). The blue line depicts the upper limit of the correlation that one would see if one of the functions was only white noise [17]. The search frequencies are available only aggregated per month. Therefore, we had to interpolate these data at day level before applying the cross-correlation function. ACF: autocorrelation function.



Discussion

Principal Findings

The longitudinal cumulative quantity of the Google search volume for all relevant keywords related to the German term for conjunctivitis (“Bindehautentzündung”) was roughly the same in Germany as a whole and in 12 large German cities in relation to the respective inhabitants. Since there was a clear temporal correlation between the cumulative Google search frequency of all 714 associated terms in Freiburg and the density of patients identified with codes in the ICD “H10.-” code group from the centralized ophthalmologic emergency practice for the Freiburg region, and the cross-correlation between the temporal search frequencies in Germany and the 12 selected cities was high, the clinical results from the Freiburg data can most likely be extrapolated to Germany as a whole.

It seems likely that the periodicity of the curves is caused by a seasonal effect since the timing of the highest search rates through the analyzed 4 years was very similar in the 13 curves (Figure 1). Seasonality is well known for infectious conjunctivitis [18]. This seasonality is also observable in the density plot for the patients diagnosed with conjunctivitis (ICD codes in the “H10.-” code group) in the centralized ophthalmologic emergency practice for the Freiburg region (Figure 2). It has to be considered that the ICD code group “H10.-” is used for not only viral or adenoviral conjunctivitis but also “acute allergic conjunctivitis.” Allergic conjunctivitis is frequently observed in spring and summer [19]. It is plausible to suppose that most patients with allergic conjunctivitis do not visit any ophthalmologist or consult the internet because they are familiar with the symptoms that recur each year. Since there is only one climate zone in Germany, the frequency of allergic conjunctivitis should be more or less the same all over the country. The seasonal patterns from Figure 1 may therefore be considered the “normal” background reference for Germany for consulting the internet for the signs and symptoms of all kinds of conjunctivitis. We hypothesize that a small fraction of patients with more severe or more protracted symptoms is more likely to actually see an ophthalmologist. This may explain why cross-correlation between the search frequency in Freiburg and the patient density revealed a substantial time lag between both curves (Figure 3).

Social media are nowadays some of the most common forms of electronic communication and are used to obtain information. However, potential sources of bias need to be considered. First, the restriction to only use data from a single commercial internet search engine may be problematic. However, since Google has an estimated market share of 95% in Germany, this restriction is acceptable at the time of our analysis [20]. Second, people who seek health information on the internet may not be a fully representative subset of the general population. Our data hint in that direction. Interestingly, the normalized search volume in Germany as a whole was much lower than that in the larger cities. This may have two reasons. First, Google usage may be generally higher in urban regions than in the countryside. Rural inhabitants may see an ophthalmologist directly without consulting online media. Second, rural inhabitants may be less

concerned about their ocular health than urban inhabitants. They may ignore the signs and symptoms of conjunctivitis instead of exploring online media for information. Nevertheless, seasonal periodicity was clearly visible in the curve for Germany as a whole, and the correlation with individual cities was generally high without any time lag (Figure 2).

Bias is a bigger issue when active usage of the internet is analyzed, as was noted in a recent related analysis that evaluated social media posts [21]. They were machine classified as related to infectious conjunctivitis. Nevertheless, Deiner et al observed a similar correlation to clinical infectious cases and discussed their method as a possible way to improve the detection of outbreaks [21].

In addition to bias, potential sources of fuzziness of information need to be considered both in the usage of social media posts and analysis of Google. The timing people search for information is potentially not well defined. The search can occur without individuals having any symptoms because someone they know has symptoms, after seeing the doctor, or after having passed the infection. Additionally, the location of the search is extremely variable, since mobile devices give access to the internet nearly everywhere, even when traveling. However, owing to the large amount of data, the concept of detecting regional outbreaks via Google search volume analysis may work despite these sources of error.

There is a strong unmet need to improve the attention of authorities to outbreaks of contagious conjunctivitis to keep them in check. One approach is to improve office tests for adenovirus and tests in patients with severe bilateral conjunctivitis to increase sensitivity and specificity [22]. However, even if all patients with adenovirus conjunctivitis are detected this way, it only leads to diagnosis in the cohort of patients who seek professional health care. In addition, large studies have shown that a substantial proportion of keratoconjunctivitis does not appear to be associated with detectable adenovirus [23]. Moreover, we still have the problem of all those patients who do not visit an ophthalmologist and may be sources of infection for others.

Interestingly, all German cities in our data set showed a more or less synchronous pattern of search frequencies over the observation period. This could mean that this pattern is the “normal” background activity in the absence of an outbreak of epidemic keratoconjunctivitis. It could therefore be possible to automatically detect regional outbreaks from Google search data if regional search volumes behave differently from the national average or other reference regions. However, it is not clear at this point whether such an approach would reach appropriate specificity and sensitivity.

This has been discussed by Wilson et al, along with the advantages and disadvantages of internet-based outbreak surveillance. Looking at the advantages, they pointed out earlier detection, free access to data, and the potential to allow the public to have access to health surveillance information [24]. Another issue raised involves privacy. The Google data generated via the Keyword Planner cannot be attributed to age or gender. For this reason, the data are fully anonymous and cannot be reidentified using secondary data sets. However, it

must be taken into consideration that certain restrictions on freedom may result in the case of a regional outbreak. People should be made aware if their search history is being used to generate surveillance data by competent authorities, even if the data are fully anonymous and reidentification is impossible.

In summary, our analysis demonstrates a clear correlation between the Google search volume for the signs and symptoms of conjunctivitis and the density of patients who present with the signs and symptoms of conjunctivitis. Furthermore, we characterize a seasonal pattern of the Google search volume in

Germany that may serve as a “normal” background reference. With these data, it may be possible to achieve an automated regional outbreak detection system. However, privacy concerns and issues of transparency need to be addressed, and specificity and sensitivity need to be evaluated.

Conclusion

Google search data may be useful in combination with clinical data to assist in estimating the incidence of infectious conjunctivitis in the context of epidemic outbreaks.

Acknowledgments

The article processing charge was funded by the Baden-Wuerttemberg Ministry of Science, Research and Art and the University of Freiburg in the funding program Open Access Publishing.

Authors' Contributions

Conceptualization and design: AZ, DB, and TR; data curation: LT, AZ, DB, and PKB; methodology: DB, LT, AZ, PKB, and TR; software: LT, AZ, and DB; formal analysis: DB, PKB, LT, and AZ; supervision: AZ, DB, and TR; writing—original draft: PKB, DB, LT, AZ, and TR; writing—review & editing: PKB, DB, LT, AZ, and TR.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Population figures of the individual cities.

[DOCX File, 14 KB - [publichealth_v7i3e22645_app1.docx](#)]

Multimedia Appendix 2

Conjunctivitis associated terms. All 714 associated terms were translated into English. In the German language, there are two terms for cataract, “Katarakt” and “grauer Star.” In addition, the term conjunctivitis is described with two words, “Konjunktivitis” and “Bindehautentzündung”.

[DOCX File, 52 KB - [publichealth_v7i3e22645_app2.docx](#)]

References

- Centers for Disease Control and Prevention (CDC). Adenovirus-associated epidemic keratoconjunctivitis outbreaks—four states, 2008–2010. *MMWR Morb Mortal Wkly Rep* 2013 Aug 16;62(32):637–641 [FREE Full text] [Medline: 23945769]
- Yeu E, Hauswirth S. A Review of the Differential Diagnosis of Acute Infectious Conjunctivitis: Implications for Treatment and Management. *Clin Ophthalmol* 2020;14:805–813 [FREE Full text] [doi: 10.2147/OPTH.S236571] [Medline: 32210533]
- Böhringer D, Birnbaum F, Reinhard T. [Cyclosporin A eyedrops for keratitis nummularis after adenovirus keratoconjunctivitis]. *Ophthalmologie* 2008 Jun;105(6):592–594. [doi: 10.1007/s00347-008-1731-1] [Medline: 18365204]
- Pepose JS, Sarda SP, Cheng WY, McCormick N, Cheung HC, Bobbili P, et al. Direct and Indirect Costs of Infectious Conjunctivitis in a Commercially Insured Population in the United States. *Clin Ophthalmol* 2020;14:377–387 [FREE Full text] [doi: 10.2147/OPTH.S233486] [Medline: 32103884]
- Sundmacher R, Hillenkamp J, Reinhard T. [Prospects for therapy and prevention of adenovirus keratoconjunctivitis]. *Ophthalmologie* 2001 Oct;98(10):991–1007; quiz 1008–1009. [doi: 10.1007/s003470170052] [Medline: 11699326]
- Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2018. Robert Koch Institut. URL: https://www.rki.de/DE/Content/Infekt/Jahrbuch/Jahrbuch_2018.pdf?__blob=publicationFile [accessed 2020-02-25]
- Wolf JA, Moreau JF, Patton TJ, Winger DG, Ferris LK. Prevalence and impact of health-related internet and smartphone use among dermatology patients. *Cutis* 2015 Jun;95(6):323–328. [Medline: 26125208]
- Polgreen PM, Chen Y, Pennock DM, Nelson FD. Using internet searches for influenza surveillance. *Clin Infect Dis* 2008 Dec 01;47(11):1443–1448 [FREE Full text] [doi: 10.1086/593098] [Medline: 18954267]
- Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection—harnessing the Web for public health surveillance. *N Engl J Med* 2009 May 21;360(21):2153–5, 2157 [FREE Full text] [doi: 10.1056/NEJMp0900702] [Medline: 19423867]
- Wehner MR, Nead KT, Linos E. Correlation Among Cancer Incidence and Mortality Rates and Internet Searches in the United States. *JAMA Dermatol* 2017 Sep 01;153(9):911–914 [FREE Full text] [doi: 10.1001/jamadermatol.2017.1870] [Medline: 28658470]

11. Tizek L, Schielein MC, R  th M, Szeimies R, Philipp-Dormston WG, Braun SA, et al. Interest in Skin Cancer in Urban Populations: A Retrospective Analysis of Google Search Terms in Nine Large German Cities. *Acta Derm Venereol* 2019 Jul 01;99(9):797-804 [FREE Full text] [doi: [10.2340/00015555-3214](https://doi.org/10.2340/00015555-3214)] [Medline: [31073622](https://pubmed.ncbi.nlm.nih.gov/31073622/)]
12. Tizek L, Schielein M, R  th M, St  nder S, Pereira MP, Eberlein B, et al. Influence of Climate on Google Internet Searches for Pruritus Across 16 German Cities: Retrospective Analysis. *J Med Internet Res* 2019 Jul 12;21(7):e13739 [FREE Full text] [doi: [10.2196/13739](https://doi.org/10.2196/13739)] [Medline: [31301128](https://pubmed.ncbi.nlm.nih.gov/31301128/)]
13. Zink A, Schuster B, R  th M, Pereira MP, Philipp-Dormston WG, Biedermann T, et al. Medical needs and major complaints related to pruritus in Germany: a 4-year retrospective analysis using Google AdWords Keyword Planner. *J Eur Acad Dermatol Venereol* 2019 Jan;33(1):151-156. [doi: [10.1111/jdv.15200](https://doi.org/10.1111/jdv.15200)] [Medline: [30067284](https://pubmed.ncbi.nlm.nih.gov/30067284/)]
14. Hartley DM, Nelson NP, Arthur RR, Barboza P, Collier N, Lightfoot N, et al. An overview of internet biosurveillance. *Clin Microbiol Infect* 2013 Nov;19(11):1006-1013 [FREE Full text] [doi: [10.1111/1469-0691.12273](https://doi.org/10.1111/1469-0691.12273)] [Medline: [23789639](https://pubmed.ncbi.nlm.nih.gov/23789639/)]
15. The R Project for Statistical Computing. R Project. URL: <https://www.r-project.org/> [accessed 2021-02-20]
16. Andrade C. The P Value and Statistical Significance: Misunderstandings, Explanations, Challenges, and Alternatives. *Indian J Psychol Med* 2019;41(3):210-215 [FREE Full text] [doi: [10.4103/IJPSYM.IJPSYM_193_19](https://doi.org/10.4103/IJPSYM.IJPSYM_193_19)] [Medline: [31142921](https://pubmed.ncbi.nlm.nih.gov/31142921/)]
17. Chatfield C. *The Analysis of Time Series: An Introduction*, Sixth Edition. Boca Raton, FL, USA: Chapman & Hall/CRC; 2016.
18. Jing D, Zhao H, Ou R, Zhu H, Hu L, Giri M, et al. Epidemiological Characteristics and Spatiotemporal Analysis of Acute Hemorrhagic Conjunctivitis from 2004 to 2018 in Chongqing, China. *Sci Rep* 2020 Jun 09;10(1):9286 [FREE Full text] [doi: [10.1038/s41598-020-66467-y](https://doi.org/10.1038/s41598-020-66467-y)] [Medline: [32518362](https://pubmed.ncbi.nlm.nih.gov/32518362/)]
19. Azari AA, Barney NP. Conjunctivitis: a systematic review of diagnosis and treatment. *JAMA* 2013 Oct 23;310(16):1721-1729 [FREE Full text] [doi: [10.1001/jama.2013.280318](https://doi.org/10.1001/jama.2013.280318)] [Medline: [24150468](https://pubmed.ncbi.nlm.nih.gov/24150468/)]
20. Marktanteile der Suchmaschinen weltweit nach mobiler und station  rer Nutzung im Januar 2021. Statista. URL: <https://de.statista.com/statistik/daten/studie/222849/umfrage/marktanteile-der-suchmaschinen-weltweit/> [accessed 2020-06-15]
21. Deiner MS, McLeod SD, Chodosh J, Oldenburg CE, Fathy CA, Lietman TM, et al. Clinical Age-Specific Seasonal Conjunctivitis Patterns and Their Online Detection in Twitter, Blog, Forum, and Comment Social Media Posts. *Invest Ophthalmol Vis Sci* 2018 Feb 01;59(2):910-920 [FREE Full text] [doi: [10.1167/iovs.17-22818](https://doi.org/10.1167/iovs.17-22818)] [Medline: [29450538](https://pubmed.ncbi.nlm.nih.gov/29450538/)]
22. Kodama A, Kobayashi F, Yang HY, Fukagawa K, Yazu H, Fujishima H. Combination of Two Rapid Ophthalmic Test Kits for Improved Diagnosis in Cases of Severe Binocular Conjunctivitis. *Diagnostics (Basel)* 2020 Feb 17;10(2) [FREE Full text] [doi: [10.3390/diagnostics10020109](https://doi.org/10.3390/diagnostics10020109)] [Medline: [32079333](https://pubmed.ncbi.nlm.nih.gov/32079333/)]
23. Lee CS, Lee AY, Akileswaran L, Stroman D, Najafi-Tagol K, Kleiboeker S, BAYnovation Study Group. Determinants of Outcomes of Adenoviral Keratoconjunctivitis. *Ophthalmology* 2018 Sep;125(9):1344-1353 [FREE Full text] [doi: [10.1016/j.ophtha.2018.02.016](https://doi.org/10.1016/j.ophtha.2018.02.016)] [Medline: [29602567](https://pubmed.ncbi.nlm.nih.gov/29602567/)]
24. Wilson K, Brownstein JS. Early detection of disease outbreaks using the Internet. *CMAJ* 2009 Apr 14;180(8):829-831 [FREE Full text] [doi: [10.1503/cmaj.090215](https://doi.org/10.1503/cmaj.090215)] [Medline: [19364791](https://pubmed.ncbi.nlm.nih.gov/19364791/)]

Abbreviations

ICD: International Statistical Classification of Diseases and Related Health Problems

Edited by R Kukafka; submitted 21.07.20; peer-reviewed by A Rovetta, M Pereira; comments to author 01.08.20; revised version received 14.08.20; accepted 30.01.21; published 03.03.21.

Please cite as:

Kammrath Betancor P, Tizek L, Zink A, Reinhard T, B  hringer D

Estimating the Incidence of Conjunctivitis by Comparing the Frequency of Google Search Terms With Clinical Data: Retrospective Study

JMIR Public Health Surveill 2021;7(3):e22645

URL: <https://publichealth.jmir.org/2021/3/e22645>

doi:[10.2196/22645](https://doi.org/10.2196/22645)

PMID:[33656450](https://pubmed.ncbi.nlm.nih.gov/33656450/)

  Paola Kammrath Betancor, Linda Tizek, Alexander Zink, Thomas Reinhard, Daniel B  hringer. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 03.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Use of #SaludTues Tweetchats for the Dissemination of Culturally Relevant Information on Latino Health Equity: Exploratory Case Study

Amelie G Ramirez^{1*}, DrPH; Rosalie P Aguilar^{1*}, MS, PhD; Amanda Merck^{1*}, MPH; Cliff Despres¹, BJ; Pramod Sukumaran¹, MSc, MPH, PhD; Stacy Cantu-Pawlik¹, MPH; Patricia Chalela¹, DrPH

Department of Population Health Sciences, Institute for Health Promotion Research, UT Health San Antonio, San Antonio, TX, United States

*these authors contributed equally

Corresponding Author:

Amelie G Ramirez, DrPH

Department of Population Health Sciences

Institute for Health Promotion Research

UT Health San Antonio

7411 John Smith

San Antonio, TX, 78229

United States

Phone: 1 2105626500

Fax: 1 2105626545

Email: ramirezag@uthscsa.edu

Abstract

Background: Latinx people comprise 18% of the US adult population and a large share of youth and continue to experience inequities that perpetuate health disparities. To engage Latinx people in advocacy for health equity based on this population's heavy share of smartphone, social media, and Twitter users, *Salud America!* launched the #SaludTues Tweetchat series. In this paper, we explore the use of #SaludTues to promote advocacy for Latinx health equity.

Objective: This study aims to understand how #SaludTues Tweetchats are used to promote dissemination of culturally relevant information on social determinants of health, to determine whether tweetchats serve to drive web traffic to the *Salud America!* website, and to understand who participates in #SaludTues Tweetchats and what we can learn about the participants. We also aim to share our own experiences and present a step-by-step guide of how tweetchats are planned, developed, promoted, and executed.

Methods: We explored tweetchat data collected between 2014 and 2018 using Symplur and Google Analytics to identify groups of stakeholders and web traffic. Network analysis and mapping tools were also used to derive insights from this series of chats.

Results: We conducted 187 chats with 24,609 reported users, 177,466 tweets, and more than 1.87 billion impressions using the hashtag #SaludTues during this span, demonstrating effective dissemination of and exposure to culturally relevant information. Traffic to the *Salud America!* website was higher on Tuesdays than any other day of the week, suggesting that #SaludTues Tweetchats acted effectively as a website traffic-driving tool. Most participants came from advocacy organizations (165/1000, 16.5%) and other health care-related organizations (162/1000, 16.2%), whereas others were unknown users (147/1000, 14.7%) and individual users outside of the health care sector (117/1000, 11.7%). The majority of participants were located in Texas, California, New York, and Florida, all states with high Latinx populations.

Conclusions: Carefully planned, culturally relevant tweetchats such as #SaludTues can be a powerful tool for public health practitioners and advocates to engage audiences on Twitter around health issues, advocacy, and policy solutions for Latino health equity. Further information is needed to determine the effect that #SaludTues Tweetchats have on self- and collective efficacy for advocacy in the area of Latino health equity.

(*JMIR Public Health Surveill* 2021;7(3):e21266) doi:[10.2196/21266](https://doi.org/10.2196/21266)

KEYWORDS

Latino; social media; social cognitive theory; health communication; health equity; policy; community health; mobile phone

Introduction

Inequities in Latinx Health

In the United States, Latinx people comprise over 18% of the adult population and over a quarter of youth, and they continue to face vast health disparities in obesity, diabetes, heart disease, and certain cancers [1-5]. These disparities are largely the result of systemic inequities in housing, education, employment, health care, and transportation, which disproportionately burden communities of color with unstable housing, poor quality schools, unfair wages, unsafe streets, unhealthy food, numerous other barriers to opportunity, and poorer health outcomes [6,7]. Reducing disparities in health requires advocacy and policy solutions to improve social and environmental factors and create more equitable living conditions [8].

The internet is an important source for distributing information to promote advocacy and policy solutions. More specifically, digital content curation, a systematic, refined process to create tailored online and social health information for an audience, provides an opportunity to empower Latinx social media users to advocate for policy changes [9]. Trends show that 90% of Latinx people aged younger than 50 years are connected online, including high rates of smartphone ownership and engagement with social media [10-13].

Approximately 80% of Latinx people own a smartphone and use it to access the internet [11,14,15]. Latinx people also lead in the use of Twitter (25% vs 24% of non-Latinx Black people and 21% of White people), spending approximately 6 hours a day on social media, which enables social interaction, self-expression, and consumption of news and information [16]. We aimed to reach Latinx social media users as well as health advocates by developing a continuous stream of evidence-based stories, research, resources, and policy change solutions to promote Latinx health equity.

Tweetchats for Health Advocacy and Health Promotion

Tweetchats are social media events that can be used to engage communities and partners in sharing health information [17-20]. Many public and private entities have used Twitter hashtags and tweetchats to promote awareness of a topic, elevate campaigns, share news and updates, drive political participation, and promote conference engagement [17,18,21-29]. For example, a number of local public health departments use Twitter to communicate public health messages as part of specific public health campaigns [17,30]. However, little is known about the effectiveness of using tweetchats to engage audiences, particularly around Latinx health equity and advocacy topics. With this in mind, *Salud America!* launched the #SaludTues Tweetchat series to complement ongoing digital content curation and social media tactics to raise awareness about Latinx health equity issues and solutions.

Salud America! (based at UT Health San Antonio) is a national Latinx-focused organization with an online network of over 400,000 community leaders and advocates, parents, school personnel, health care workers, and researchers with an interest in Latinx health [31]. The team behind *Salud America!* uses digitally curated, culturally relevant, and research-based stories,

videos, tools, and events such as tweetchats to inspire and support healthy changes to policies, systems, and environments [9].

The team's digital content curation is influenced by the Social Cognitive Theory (SCT) by Bandura [32,33]. SCT guides the creation of peer-modeled blog posts and social media messages that inspire community action to start or support social change. At the population level, *Salud America!*'s dissemination approach is largely guided by a combination of SCT and Diffusion of Innovations Theory [32,34]. According to Bandura, adoption of social innovations, such as policy solutions to improve health equity, can be achieved through mass communications coupled with symbolic modeling [32]. This produces a dual path of communications in which both direct exposure to digital content and mediated connections to individuals through social media may influence personal and social change and prompt advocacy [32]. For example, within the context of a Tweetchat, participants may advocate for healthier environments by sharing examples of policy solutions and encouraging others to take local action. Initially, some social media users may be slow to advocate for policy solutions, but as time progresses and more individuals model positive change by sharing policy solutions on social media, the rate of acceptance accelerates [34,35].

In fall 2014, *Salud America!* established a weekly #SaludTues Tweetchat series using its Twitter handle (@SaludAmerica) [9,32,33]. The goals of these chats are to raise awareness about the social and environmental inequities that contribute to disparities in Latinx health and inspire policy solutions and advocacy actions by posing questions and providing positive verbal persuasion, stories, and resources to build self- and collective efficacy. To accomplish this, *Salud America!* disseminates digitally curated content on Latinx health equity from its website [36]. In this case study, we examine data from a series of 2014 to 2018 #SaludTues Tweetchats collected via Symplur, a social media analytics platform for health care [37]. We aim to provide a step-by-step guide on how to organize a successful tweetchat and also describe lessons learned by outlining the process that *Salud America!* uses for organizing its weekly Latino health advocacy-oriented tweetchats. The primary objectives of this paper are as follows. Specifically, we aim to explore the following:

- How are #SaludTues Tweetchats used to promote the dissemination of culturally relevant information on social determinants of health?
- Do tweetchats serve to drive web traffic to the *Salud America!* website?
- Who participates in #SaludTues Tweetchats and what can we learn about the participants?

Methods

The Salud America! Approach to Hosting Tweetchats

The following analysis does not involve human research and was therefore exempt from an institutional review board approval. We begin with a description of our approach to conducting tweetchats (objective 1). *Salud America!* hosts a

60-min #SaludTues Tweetchat on Twitter from 1 to 2 PM EST every Tuesday. This date and time were selected based on Twitter engagement metrics at the time and to avoid competing with other popular tweetchats such as #FoodFri, a popular tweetchat that takes place between 1 and 2 PM EST on Fridays [36]. Although not everyone may be available to participate during this time, we found that midday chats work well for engaging users and cohosts across different US time zones. Tweetchats are accessible to anyone on Twitter by following the #SaludTues hashtag.

Three constructs derived from SCT—the theory behind our digital content curation—are relevant to #SaludTues: (1) peer role modeling, (2) positive verbal persuasion, and (3) perceived self-efficacy [33]. The first two of these constructs were employed to prepare each week's tweetchat materials, with the goal of increasing a person's self-efficacy to engage in message sharing and promoting advocacy for Latinx health equity [9]. For instance, we might prepare an answer tweet with a message about how one of our Salud Heroes (role models) helped promote equity in their community, in response to a question that asks, *how can we advocate for change?* The tweet would include a positive message about the hero, a link to the full-length story, an image of the hero, and perhaps an emoji such as a flexed arm to signal strength and confidence. If the role model is relatable and the message is motivational, the person may experience an increase in self-efficacy for engaging in advocacy actions both online and potentially offline. The Social Ecological Model was also applied to craft messages that help participants understand how multiple dimensions of social and environmental factors influence health and encourage participants to promote social change [38,39]. In addition, the Transtheoretical Model of Behavior Change, also known as the *Stages of Change* model, was used during chats to identify tweetchat participants' levels of readiness for action and to craft messages in response to this.

Planning and Scripting Questions

The *Salud America!* communications team selects weekly tweetchat topics that align with Latinx health equity issues and campaigns. Some topics are specific to *Salud America!*'s content curation scope, which currently focuses on healthy families and schools, healthy neighborhoods and communities, and healthy and cohesive cultures, whereas other topics are based on weekly and monthly observances and partner requests. For example, in February 2017, the tweetchat on *Our Heart Loves Physical*

Activity observed National Heart Month at the request of the US Department of Health and Human Services' Office of Minority Health. A #SaludTues Tweetchat in October 2017 honored National Hispanic Heritage Month and a tweetchat in November 2017 promoted the release of *Salud America!'s Early Childhood Development and Latinx Kids Research Review*. Another tweetchat hosted in October 2018 explored issues and solutions to the lack of Latinx people in clinical trials.

Salud America! serves as a host for each #SaludTues chat [40] and invites organizations and individuals to serve as cohosts. The criteria for selecting cohosts vary but can usually be categorized into one of the following: (1) a subject-matter expert or group, (2) an active advocate or advocacy group, and (3) users that can offer a unique perspective. It is also helpful to find out whether a cohost retweets regularly, if their tweets are culturally sensitive, and whether they typically share engaging tweets that contain images, URL links, and useful media, as these are more likely to be retweeted. Although *Salud America!* seeks to share content from their websites, cohosts are not expected to have a website. Cohosts are asked to: (1) review tweetchat questions and prepare 1-3 responses in advance; (2) promote the event during the week before the Tweetchat; and (3) actively engage during the 60-min Tweetchat by tweeting, retweeting, liking, and replying to other contributors in real time.

As the host, *Salud America!* drafts 6-8 questions and 12-24 responses each week. The first few questions ask participants to consider, define, or share personal experiences with the topic to build relevance and salience. Then, questions implore exploration of evidence, scientific research, and anecdotes of how the topic relates to Latinx health through a social and environmental lens rather than just an individual lens. Finally, questions focus on policy solutions or advocacy actions to address or improve social and environmental factors.

Salud America! prepares responses to the questions with links to digitally curated content from the *Salud America!* website, such as educational resources, stories of peer models, and specific opportunities to take action as well as culturally appropriate photos, graphics, and infographics, often derived from our research reviews. For example, Figure 1 shows a tweet response with an English and Spanish infographic from our research review on safe neighborhoods. *Salud America!* also prepares template retweets and replies to quickly engage with cohosts and participants.

Figure 1. Example of a tweet during #SaludTues Tweetchat.

Similar to a real-life discussion, *Salud America!* works to share, echo, and expand on evidence, redirect focus away from individual behavior change to systemic change, and encourage participants to seize opportunities for advocacy. We offer some examples in the following section.

Promotion and Real-Time Participation

As a host, *Salud America!* drafts promotional materials each week, including a blog post, sample tweets, and graphics. The week before each Tweetchat, *Salud America!* promotes the chat on Twitter at least once a day, escalating to several times a day,

the day before, and the day of the chat. *Salud America!* tags 10 individuals and organizations in each promotional graphic to encourage wider sharing and participation. To identify to-be-tagged individuals and organizations, *Salud America!* searches tweets and profile bios for various keywords relevant to the specific topic. *Salud America!* scrolls through their feed for diverse perspectives on health disparities, social inequities, and policy solutions. For example, a tweet promoting a tweetchat on how to systemically promote the health benefits of swimming tagged 2 individual swim safety advocates, 2 swim safety foundations, 1 drowning prevention coalition, 1 physical activity

coalition, and others. As policy solutions do not occur in a vacuum, it is important to expand promotional efforts beyond Latinx people to include various policy and health equity stakeholders.

Although following the *latest* stream of #SaludTues content on Twitter, the host and cohosts read each tweet and respond accordingly while also tweeting out prescribed responses. On the basis of participants' tweets, the team gauges the level of awareness and readiness for action similar to a real conversation. For example, if a participant appears surprised by the data or unsure about the connection between health disparities and inequities in social and environmental factors, *Salud America!* crafts a response to raise awareness. Similarly, if a participant appears informed and passionate, *Salud America!* crafts a response to build self- and collective efficacy and to inspire action to address social inequities. When crafting replies, *Salud America!* quickly decides whether to encourage more learning by sharing links to research reviews and other informational content on the *Salud America!* website; to inspire action and include links to Salud Hero stories and other examples of policy change on the *Salud America!* website; or to provide positive reinforcement for dissatisfaction with and commitment to addressing social and environmental inequities. The Tweetchat continues for an hour and ends with sending thank you tweets and information on how to engage further with the *Salud America!* network.

Collecting Tweetchat Metrics

All Tweetchats are archived via online software (Wakelet) [41]. These archives capture tweets with the hashtag #SaludTues and are sharable *recaps* that can be accessed by those who miss the chat. For capturing metrics and conducting routine data analysis,

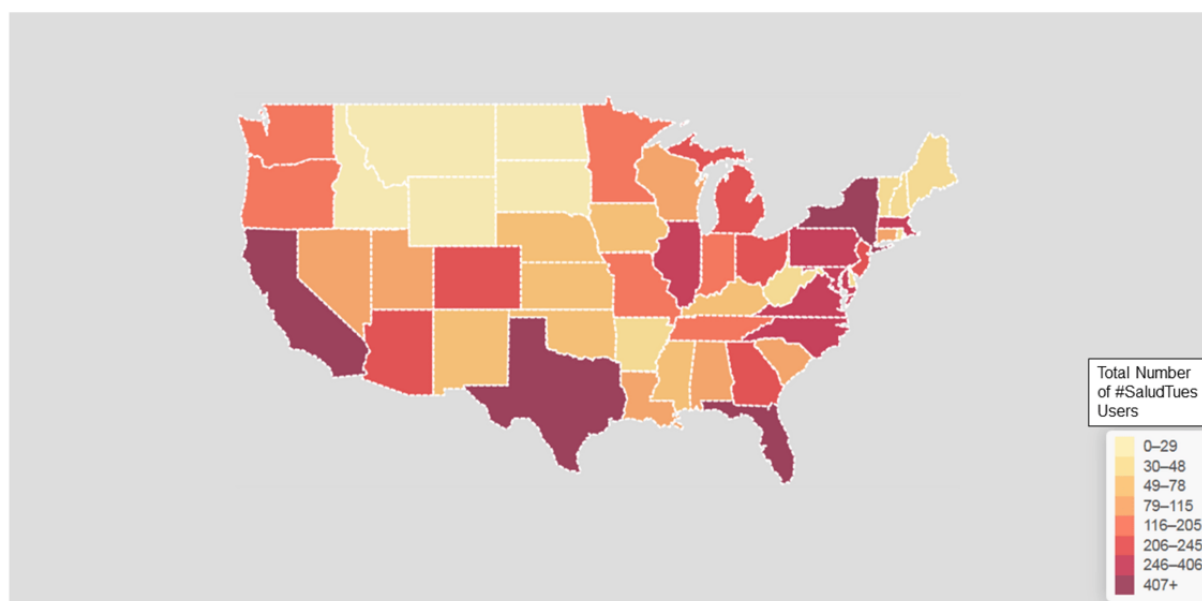
the *Salud America!* communication team used Symplur Signals (Symplur LLC) online software to capture metrics for each chat [37]. Symplur Signals is a paid subscription analytics platform that links directly to the Twitter application program interface [42]. Only tweets with the #SaludTues hashtag are included in the metrics. Metric data includes the number of participants, tweets, and impressions (tweets per participant with hashtag multiplied by a user's number of followers at that time=impressions). After each chat, *Salud America!* compiles Symplur data (total tweets, impressions, users, word bubbles, top tweeters, most retweeted, most mentioned, etc) into a PowerPoint report to share with cohosts.

Data Analysis for 2014-2018 #SaludTues Chats

Descriptive statistics for tweetchats, such as means and totals per month and year, were tracked on a spreadsheet. To determine whether #SaludTues Tweetchats serve as a funnel from Twitter to the *Salud America!* website (objective 2), we used Google Analytics (GA) data to gauge website usage on different days of the week during the first year of our comprehensive website launch (September 1, 2017, to August 31, 2018) [43]. The bulk of the remaining analysis is focused on understanding who participates in #SaludTues Tweetchats as well as where they are tweeting from (objective 3). We also aimed to learn about the terms that were used most during chats.

To understand the geographic exposure of tweetchats, a choropleth map of the United States was accessed through Symplur to determine where #SaludTues users tweet from (Figure 2). This map comprises data collected from users' manually entered locations from their Twitter profiles (disclosing location is voluntary on Twitter).

Figure 2. Choropleth map of the United States showing where Tweetchat users are tweeting from, between November 1, 2014, and August 31, 2018.



To illustrate the variety of #SaludTues trending terms, *Salud America!* used Symplur to create 2 bubble chart visualizations

of 4 tweetchats occurring in April 2018 and all tweetchats occurring over a 4-year period. Symplur Signals contains an

algorithm that identifies the 500 most significant and trending terms in any health care conversation and produces a series of bubbles [44]. The size of bubbles is proportional to the frequency of words appearing in tweetchats (color only aids in differentiation) [44]. As #SaludTues Tweetchats cover a wide variety of topics on Latinx health equity, we chose to examine 1 month of data to gain insight into how topics differ between the two periods.

To understand what type of health care stakeholders participate in tweetchats, *Salud America!* relied on Symplur's health care categories. Symplur software—using publicly available data (Twitter bios and self-identification) and proprietary algorithms combined with human evaluation and quality control—categorizes each Twitter user by their specific role in health care. Users that Symplur cannot identify are classified as *unknown*.

Taking a more traditional look at network analysis, *Salud America!* used Symplur's visualization function to examine 1

month of #SaludTues Tweetchats (from April 2018). Symplur enables the creation of an interactive network analysis of Twitter users most central to conversations by hashtags, the size of their nodes, and how the conversation flows between them and others [45]. Nodes represent Twitter users in a community, and edges indicate direct communications between users. The larger the node, the more frequently mentioned the user [45].

Results

Findings From #SaludTues Tweetchats 2014-2018 and Website Traffic

The results showed that of the 187 chats between 2014 and 2018, the #SaludTues hashtag was used by 24,609 users in 177,466 tweets, garnering more than 1.87 billion impressions (Table 1). Approximately one-third of tweets collected over the 4-year period were promotional tweets leading up to the chat or occurring after the live event.

Table 1. Tweetchat metrics on tweets, users, and impressions collected via Symplur 2014-2018.

Category	Continuous data collected for #SaludTues	Cumulative data from weekly chats for #SaludTues	Promotional tweets for #SaludTues
Total tweets	177,466	125,541	51,925
Retweets	111,491	76,071	35,420
Tweets with links	60,515	39,962	20,553
Tweets with media	48,746	26,779	21,967
Tweets with mentions	141,051	94,902	46,149
Tweets with replies	11,310	10,450	860
Users	24,609	16,952	7657
Impressions	1,870,000,000	1,491,416,579	378,583,421

The table compares 2 methods of collecting data via Symplur. The first column represents continuous data collected from the hashtag between November 1, 2014 and August 31, 2018. The second column represents the sum of data solely for Tuesday chats during the same dates (90-min time frame [11:45 AM-1:15 PM CST]). The difference (column 3) primarily comprises promotional tweets sent before or after tweetchats. During 2017-2018, following the relaunch of *Salud America!*'s website,

24.28% (2020/8318) of *Salud America!* website users visited on Tuesday (the day of #SaludTues chats), which is the highest day of web traffic, any day of the week (Table 2), according to GA. Multimedia Appendix 1 shows a randomly chosen month of 4 #SaludTues Tweetchats in April 2018. During these chats, 292 users participated, generating 2,216 tweets and 26.7 million impressions.

Table 2. Daily referral traffic (*Salud America!* website users) from Twitter (2017-2018); N=8318.

Day of the week ^a	Count, n (%)
Sunday	463 (5.57)
Monday	1293 (15.54)
Tuesday	2020 (24.28)
Wednesday	1458 (17.53)
Thursday	1431 (17.20)
Friday	1166 (14.02)
Saturday	487 (5.86)

^aData collected from Google Analytics from September 4, 2017, to August 31, 2018.

These chats demonstrate a snapshot of the variety of #SaludTues metrics, the breadth of topics and cohosts, and the use of Twitter

as a funnel to *Salud America!*'s website and theory-based health equity content. For the same month, the GA results showed that

#SaludTues: Trending Terms

Some of the most used key terms found in tweetchats were “Latino,” “Latinos,” “health,” “healthy,” and “tweetchat” (Figure 3). Additional key terms individually or in combination offer a glimpse into past and current *Salud America!* topic areas (social support, community health, mental health, safe places to play, access to care, etc) and other points of emphasis for Latinx health (cancer, diabetes, obesity, etc).

Figure 3. A bubble chart visualizing the 100 most frequently used words in tweets with the hashtag #SaludTues between November 1, 2014 and August 31, 2018.



health and safety at home, climate and transportation, and social and emotional learning.

(personal trainers, bloggers, counselors, and dietitians). A few of the accounts marked as unknown did not contain a profile description. The rest of the participants were classified as individual nonhealth-related users (in nonclinical roles), nonhealth organizations, individuals from fields outside of health, government organizations, health care professionals, research and academic institutions, doctors (MDs, DOs, and PhDs who bill directly for services), provider organizations, patient advocates, caregivers, PhD researchers and academics, journalist groups, media organizations, and spam. [Multimedia Appendix 2](#) shows a list of definitions for each segmentation category used by Symplur and seen in [Table 3](#) [46].

Table 3. Top health care stakeholders by tweets for 1000 #SaludTues participants between November 1, 2014, and August 31, 2018 (N=1000).

Health care stakeholder	Count, n (%)
Doctor	33 (3.3)
Health care provider	37 (3.7)
Patient advocate	17 (1.7)
Caregiver	16 (1.6)
Researcher/academic	11 (1.1)
Journalist/media	7 (0.7)
Individual other health	80 (8)
Individual nonhealth	117 (11.7)
Organizational provider	24 (2.4)
Organizational research/academic	34 (3.4)
Organizational government	60 (6)
Organizational advocacy	165 (16.5)
Organizational media	1 (0.1)
Organizational other health care	162 (16.2)
Organizational nonhealth	89 (8.9)
Spam	1 (0.1)
Unknown	147 (14.7)

#SaludTues: Network Analysis

Results of the network analysis revealed that chat host @SaludAmerica is represented as the largest node (circle in the graph), meaning it is one of the most influential profiles within

this network (Figure 5). Other nodes are influential health advocates and groups who have cohosted or participated in one or more tweetchats, such as governmental health agencies, health advocacy groups, physician advocates, and Latinx-focused organizations.

[illegible]

represents a higher degree of communication between the 2 groups. The West Virginia Oral Health Coalition, WalkSafe, Mobility Lab, Health and Human Services Office of the Assistant Secretary for Health Region II, The Praxis Project, The Afterschool Alliance, US Consumer Product Safety Commission, Seguridad del Consumidor, and ChapCare appeared to have had some of the highest degrees of communication with @SaludAmerica during April 2018.

Principal Findings

For nearly 4 years, *Salud America!* successfully built an online platform to engage partners and audiences in tweetchats to raise exposure to Latinx health equity issues, awareness, and action opportunities. These weekly social media events complemented *Salud America!*'s existing digital content curation and provided an additional opportunity to disseminate content to enhance self-efficacy and inspire health advocacy. During tweetchats, we engaged people with specific research, resources, and stories

of health and social issues, relevant to Latinx people interested in or new to health promotion and community action. We also tweeted peer-modeled case studies from across the country, demonstrating the step-by-step process to achieve a specific policy or system change. Throughout these chats, messages were amplified through frequent retweets and the live conversation that took place.

Our analysis also shows that we were successful in engaging a diverse pool of stakeholders from various backgrounds and professions in tweetchats related to Latinx health equity. Engaging diverse users in conversations related to Latinx health equity may be especially important for promoting grassroots advocacy and ultimately the development of public policy aimed at promoting Latinx equity [8,49,50]. Therefore, tweetchats may be of interest in prompting a greater level of grassroots advocacy among Latinx communities; however, further research in this area is needed.

Although the data presented here are specific to #SaludTues Tweetchats, others seeking to establish a strong base of health advocates on Twitter may benefit from adopting similar techniques to those used by our health communications team. To track progress over time and gain additional insights into one's primary audience, it is important to regularly collect metrics on tweetchats. Organizers should start by collecting baseline website metrics before launching a regularly occurring tweetchat. It is also helpful to compare tweetchat metrics from multiple chats over a given period, as described in this paper.

How Are #SaludTues Tweetchats Used for Promoting Dissemination of Culturally Relevant Information on Social Determinants of Health?

Tweetchat questions engaged cohosts and contributors to think critically about health and social issues by using research and examples applicable to contributors' own experiences, such as the aforementioned Salud Hero peer model stories and videos. Tweetchat hosts and cohosts addressed the latest findings and best practices regarding the social determinants of health while encouraging contributors to share their experiences, successes, and actions. We believe our #SaludTues Tweetchats led to a positive exchange of information among chat contributors, showcasing how to more effectively craft health messages and communicate scientific, technical, or social content in concise statements while also serving as reinforcement of positive verbal persuasion to raise awareness and take action beyond the individual level. In addition, the choropleth map shows close alignment with our target audience, the Latinx population, and the areas where most Latinx people in the United States live [51]. The network analysis visualization provides interesting findings that indicate effective broadcasting and exposure of content and messages to multiple partners.

Do #SaludTues Tweetchats Drive Traffic to the Salud America! Website?

Twitter is one of the highest drivers of social media traffic to the *Salud America!* website. Findings from our exploration of tweetchat data and webpage visits over 1 year showed that website traffic from Twitter was at least 7% higher on Tuesdays than on any other day of the week. This suggests that in addition

to exposing thousands of Twitter users to *Salud America!* content, #SaludTues Tweetchats also helped funnel more people to our website. As digital content curation is one of *Salud America!*'s essential programmatic features, website traffic is an important metric.

Who Participates in #SaludTues Tweetchats and What Can We Learn About Participants?

Through our exploration of tweetchat data, we found that we are reaching users from our target audience of Latinx social media users and health advocates. Over 33% of the top 1000 participants were advocacy organizations or other health care organizations. However, approximately 15% of participants were classified as *unknown*. Upon further exploration, we learned that most users classified as *unknown* appeared to be individual or community health advocates (ie, parents, teachers, employers, etc), nonprofits, campaigns, coalitions, and small businesses, all of whom remain very relevant to our work and also play an important role in inspiring policy and system change for health equity. Although we aim to reach as many Latino social media users as possible, our main goal with #SaludTues chats is to amplify information about Latino health equity via Twitter and to offer a Latino health perspective, where there might not otherwise be one.

The #SaludTues network analysis found that *Salud America!* is the central node in the figure, which is expected because the host *Salud America!* tweets out all questions and makes a concerted effort to engage with other participants during the chat. In addition, participants often engage in replying to and retweeting #SaludTues questions, which adds to higher in-degree communication occurring toward @SaludAmerica and an increased degree of centrality of the @SaludAmerica account within the #SaludTues network analysis. Similar results were seen in a previous network analysis conducted by Gomez-Vasquez et al [52], which found that *Salud America!* participates in a large amount of in-degree and out-degree engagement during #SaludTues Tweetchats. Upon closer examination of our network illustration, it appears that many users are not connected to each other. This is most likely due to the diversity of professional sectors from which #SaludTues users come from. Therefore, we believe #SaludTues chats present an important means of connecting users across disciplines and areas of expertise to amplify important health equity themes. Although #SaludTues participants come from different professions, tweets shared during the chat appear to be part of an echo-chamber, with most users holding similar views to others [53]. However, #SaludTues Tweetchats remain an open platform for debate, information seeking, and raising awareness of policy solutions to promote health equity.

Lessons Learned

Several lessons can be taken from the *Salud America!* team's experience with #SaludTues Tweetchats. When seeking cohosts, it is important to look at a potential cohost's Twitter feed, to determine whether they regularly retweet other users or if their feed is primarily their own content. In addition, it is imperative to determine if they regularly reply to others and if their message is consistent with yours or aligns with your goals and audience, as it can help build partnerships. If you want your cohost to

retweet your tweets, you will want someone that already retweets on a regular basis. One might also look at whether the user echoes sentiment, clarifies facts, or provides a unique perspective. Tweetchat hosts may also want an outside perspective to challenge the status quo or may want to avoid confusion and steer clear of conflicting messages. Moreover, it is important to ensure that cohosts are culturally aware and able to discuss health equity issues without victim blaming, stereotypes, or stigmatizing language. It is also best when cohosts are able to share images, graphics, videos, and weblinks as part of their tweets, as these tweets tend to get greater engagement.

Limitations

There are some limitations to this study. At the platform level, Twitter users are not necessarily representative of the general public. At the individual level, participation in tweetchats depends on internet access and knowledge of how to use Twitter, which could be a potential barrier, even though Latinx people are heavy users of social media. In the analysis of #SaludTues Tweetchats, only tweets using the correct hashtag (#SaludTues) were included. Although social impressions are widely cited as an important metric of a tweet's reach, they may not fully assess audience engagement, as it is unknown how many followers are active or how much overlap occurs in followers. Therefore, total impressions may appear higher as a result of some Twitter accounts having some of the same followers. We are also unable to directly correlate greater hashtag impressions with increased funneling to the *Salud America!* website or improved self- and collective efficacy or action to address inequity; however, we do expect those outcomes to have occurred. Follow-up surveys of participants of #SaludTues Tweetchats may help elucidate whether participation has increased their self- and collective efficacy for advocacy. Finally, there are several limitations to using Symplur to collect data over large periods. For instance, Symplur only captured the top 1000 users with its stakeholder segmentation tool. Furthermore, Symplur ranks individuals by

computing individual variables (eg, replies and retweets), which can be manipulated by excessive tweeting or spamming tools and potentially promote artificially high rankings. Despite these limitations, our experience demonstrates that tweetchats can be a powerful tool for the widespread engagement of audiences to raise exposure to, awareness of, and opportunities for promoting advocacy for health equity. Furthermore, although there are some limitations to using Symplur for data analysis, its ability to quickly capture tweets from a particular hashtag and the ease of using this tool make it practical for use in regularly capturing metrics on health-related tweetchats.

Conclusions

From our #SaludTues Tweetchat series between 2014 and 2018, we identified key structural elements that we believe helped with the success of these chats. The preparation and planning of these topics in advance while matching those topics to organizational goals was key. In addition, taking steps to secure influential cohosts and additional participation, scripting engaging tweetchat questions, spending time to promote each chat with specialized messaging and graphics, and effectively tweeting in real time during tweetchats are important elements to a successful chat. Archiving tweetchat data for analysis and having access to analytic tools that provide additional insights are necessary for identifying new communities or important networks of users and key influencers participating in #SaludTues.

It is important to note that substantial effort goes into daily digital content curation for the *Salud America!* website, without which tweetchats as a tool to build self-efficacy for advocacy may not be as powerful. Additional research is warranted to demonstrate the impact of tweetchats on particular audiences, networks, website traffic, locations, sentiment, and topics within Latinx health equity, which could augment the true value of using tweetchats to disseminate important health messages and advocacy actions.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Title, cohosts, users, tweets, impressions, and Twitter acquisitions for four #SaludTues Tweetchats in April 2018.

[[DOCX File, 15 KB](#) - [publichealth_v7i3e21266_app1.docx](#)]

Multimedia Appendix 2

Symplur's stakeholder health care stakeholders based on a user's profile.

[[PDF File \(Adobe PDF File\), 14 KB](#) - [publichealth_v7i3e21266_app2.pdf](#)]

References

1. U.S. Census Bureau QuickFacts: United States. U.S. Census Bureau. URL: <https://www.census.gov/quickfacts/fact/table/US/RHI725218> [accessed 2021-01-20]
2. Llabre M, Schneiderman N, Gallo L, Arguelles W, Daviglus M, Gonzalez F, et al. Childhood trauma and adult risk factors and disease in Hispanics/Latinos in the US. In: Results from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) Sociocultural Ancillary Study. New York: Psychosom Med Lippincott Williams and Wilkins; 2017:172.

3. Shaibi GQ, Greenwood-Ericksen MB, Chapman CR, Konopken Y, Ertl J. Development, implementation, and effects of community-based diabetes prevention program for obese latino youth. *J Prim Care Community Health* 2010 Oct 01;1(3):206-212 [FREE Full text] [doi: [10.1177/2150131910377909](https://doi.org/10.1177/2150131910377909)] [Medline: [21461141](https://pubmed.ncbi.nlm.nih.gov/21461141/)]
4. Siegel R, Naishadham D, Jemal A. Cancer statistics for Hispanics/Latinos, 2012. *CA Cancer J Clin* 2012 Sep 17;62(5):283-298 [FREE Full text] [doi: [10.3322/caac.21153](https://doi.org/10.3322/caac.21153)] [Medline: [22987332](https://pubmed.ncbi.nlm.nih.gov/22987332/)]
5. Flores A. How the U.S. Hispanic population is changing internet. Pew Res Cent. URL: <http://www.pewresearch.org/fact-tank/2017/09/18/how-the-u-s-hispanic-population-is-changing/> [accessed 2021-01-20]
6. Salud America! Coronavirus and Latino Health Equity. Salud America. 2020. URL: <https://salud-america.org/coronavirus-latino-health-equity/> [accessed 2021-01-20]
7. Ramirez AG, Gallion KJ, Despres CE, Adeigbe RT. SaludAmerica!: a national research network to build the field and evidence to prevent Latino childhood obesity. *Am J Prev Med* 2013 Mar;44(3 Suppl 3):178-185 [FREE Full text] [doi: [10.1016/j.amepre.2012.12.005](https://doi.org/10.1016/j.amepre.2012.12.005)] [Medline: [23415181](https://pubmed.ncbi.nlm.nih.gov/23415181/)]
8. Farrer L, Marinetti C, Cavaco Y, Costongs C. Advocacy for health equity: a synthesis review. *Milbank Q* 2015 Jun;93(2):392-437 [FREE Full text] [doi: [10.1111/1468-0009.12112](https://doi.org/10.1111/1468-0009.12112)] [Medline: [26044634](https://pubmed.ncbi.nlm.nih.gov/26044634/)]
9. Ramirez A, Baldwin S, Adeigbe R, Aguilar R, Gallion K, Despres C. SaludToday: Curating Latino health information for a new generation. *Journal of Communication in Healthcare* 2016 Feb 07;9(1):60-69 [FREE Full text] [doi: [10.1080/17538068.2015.1133790](https://doi.org/10.1080/17538068.2015.1133790)]
10. Demographics of social media users and adoption in the United States internet. Pew Research Center. URL: <https://www.pewresearch.org/internet/fact-sheet/social-media/#who-uses-social-media> [accessed 2021-01-20]
11. Hispanics and mobile access to the internet. Pew Research Center. 2016. URL: <https://www.pewresearch.org/hispanic/2016/07/20/3-hispanics-and-mobile-access-to-the-internet/> [accessed 2021-01-20]
12. Huhman M, Berkowitz JM, Wong FL, Prosper E, Gray M, Prince D, et al. The VERB campaign's strategy for reaching African-American, Hispanic, Asian, and American Indian children and parents. *Am J Prev Med* 2008 Jun;34(6 Suppl):194-209. [doi: [10.1016/j.amepre.2008.03.012](https://doi.org/10.1016/j.amepre.2008.03.012)] [Medline: [18471600](https://pubmed.ncbi.nlm.nih.gov/18471600/)]
13. Despres C, Aguilar R, McAlister A, Ramirez AG. Communication for awareness and action on inequitable impacts of Covid-19 on Latinos. *Health Promot Pract* 2020 Nov 07;21(6):859-861. [doi: [10.1177/1524839920950278](https://doi.org/10.1177/1524839920950278)] [Medline: [32762369](https://pubmed.ncbi.nlm.nih.gov/32762369/)]
14. Demographics of mobile device ownership and adoption in the United States. Pew Research Center. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2021-01-20]
15. Digital divide narrows for Latinos internet. Pew Research Center. 2016. URL: <https://www.pewresearch.org/hispanic/2016/07/20/digital-divide-narrows-for-latinos-as-more-spanish-speakers-and-immigrants-go-online/> [accessed 2021-01-20]
16. Digital diversity: a closer look at US Hispanics. Facebook for Business. URL: <https://www.facebook.com/business/news/insights/digital-diversity-a-closer-look-at-us-hispanics> [accessed 2021-01-20]
17. Rabarison KM, Croston MA, Englar NK, Bish CL, Flynn SM, Johnson CC. Measuring audience engagement for public health Twitter chats: insights from #livefitnola. *JMIR Public Health Surveill* 2017 Jun 08;3(2):34 [FREE Full text] [doi: [10.2196/publichealth.7181](https://doi.org/10.2196/publichealth.7181)] [Medline: [28596149](https://pubmed.ncbi.nlm.nih.gov/28596149/)]
18. Scanfeld D, Scanfeld V, Larson EL. Dissemination of health information through social networks: twitter and antibiotics. *Am J Infect Control* 2010 Apr;38(3):182-188 [FREE Full text] [doi: [10.1016/j.ajic.2009.11.004](https://doi.org/10.1016/j.ajic.2009.11.004)] [Medline: [20347636](https://pubmed.ncbi.nlm.nih.gov/20347636/)]
19. Hage AN, Srinivasa RN, Bundy JJ, Gemmete JJ, Johnson EJ, Srinivasa RN, et al. #VascularSurgery. *Ann Vasc Surg* 2018 Nov;53:217-223. [doi: [10.1016/j.avsg.2018.06.001](https://doi.org/10.1016/j.avsg.2018.06.001)] [Medline: [30048687](https://pubmed.ncbi.nlm.nih.gov/30048687/)]
20. Keller B, Labrique A, Jain KM, Pekosz A, Levine O. Mind the gap: social media engagement by public health researchers. *J Med Internet Res* 2014 Jan 14;16(1):8 [FREE Full text] [doi: [10.2196/jmir.2982](https://doi.org/10.2196/jmir.2982)] [Medline: [24425670](https://pubmed.ncbi.nlm.nih.gov/24425670/)]
21. Ince J, Rojas F, Davis C. The social media response to Black Lives Matter: how Twitter users interact with Black Lives Matter through hashtag use. *Ethnic and Racial Studies* 2017 Jul 06;40(11):1814-1830. [doi: [10.1080/01419870.2017.1334931](https://doi.org/10.1080/01419870.2017.1334931)]
22. Aslam AA, Tsou M, Spitzberg BH, An L, Gawron JM, Gupta DK, et al. The reliability of tweets as a supplementary method of seasonal influenza surveillance. *J Med Internet Res* 2014 Nov 14;16(11):250 [FREE Full text] [doi: [10.2196/jmir.3532](https://doi.org/10.2196/jmir.3532)] [Medline: [25406040](https://pubmed.ncbi.nlm.nih.gov/25406040/)]
23. Nagar R, Yuan Q, Freifeld CC, Santillana M, Nojima A, Chunara R, et al. A case study of the New York City 2012-2013 influenza season with daily geocoded Twitter data from temporal and spatiotemporal perspectives. *J Med Internet Res* 2014 Oct 20;16(10):236 [FREE Full text] [doi: [10.2196/jmir.3416](https://doi.org/10.2196/jmir.3416)] [Medline: [25331122](https://pubmed.ncbi.nlm.nih.gov/25331122/)]
24. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: a review and tutorial of applications in medicine and health care. *J Med Internet Res* 2014 Feb 11;16(2):13 [FREE Full text] [doi: [10.2196/jmir.2912](https://doi.org/10.2196/jmir.2912)] [Medline: [24518354](https://pubmed.ncbi.nlm.nih.gov/24518354/)]
25. Bhattacharya S, Srinivasan P, Polgreen P. Engagement with health agencies on Twitter. *PLoS One* 2014;9(11):112235 [FREE Full text] [doi: [10.1371/journal.pone.0112235](https://doi.org/10.1371/journal.pone.0112235)] [Medline: [25379727](https://pubmed.ncbi.nlm.nih.gov/25379727/)]
26. Kavanaugh AL, Fox EA, Sheetz SD, Yang S, Li LT, Shoemaker DJ, et al. Social media use by government: From the routine to the critical. *Government Information Quarterly* 2012 Oct;29(4):480-491. [doi: [10.1016/j.giq.2012.06.002](https://doi.org/10.1016/j.giq.2012.06.002)]
27. Black DR, Dietz JE, Stirratt AA, Coster DC. Do social media have a place in public health emergency response? *J Emerg Manag* 2015;13(3):217-226. [doi: [10.5055/jem.2015.0235](https://doi.org/10.5055/jem.2015.0235)] [Medline: [26150365](https://pubmed.ncbi.nlm.nih.gov/26150365/)]

28. Carroll CL, Bruno K, Ramachandran P. Building community through a #pulmcc twitter chat to advocate for pulmonary, critical care, and sleep. *Chest* 2017 Aug;152(2):402-409. [doi: [10.1016/j.chest.2017.03.003](https://doi.org/10.1016/j.chest.2017.03.003)] [Medline: [28302497](https://pubmed.ncbi.nlm.nih.gov/28302497/)]
29. Wilkinson SE, Basto MY, Perovic G, Lawrentschuk N, Murphy DG. The social media revolution is changing the conference experience: analytics and trends from eight international meetings. *BJU Int* 2015 May;115(5):839-846. [doi: [10.1111/bju.12910](https://doi.org/10.1111/bju.12910)] [Medline: [25130687](https://pubmed.ncbi.nlm.nih.gov/25130687/)]
30. Harris JK, Mueller NL, Snider D, Haire-Joshu D. Local health department use of twitter to disseminate diabetes information, United States. *Prev Chronic Dis* 2013 May 02;10:70 [FREE Full text] [doi: [10.5888/pcd10.120215](https://doi.org/10.5888/pcd10.120215)] [Medline: [23639765](https://pubmed.ncbi.nlm.nih.gov/23639765/)]
31. Success stories. Salud America. 2018. URL: <https://salud-america.org/success-stories/> [accessed 2021-01-20]
32. Bandura A. Social cognitive theory of mass communication. *Media Psychology* 2001 Aug;3(3):265-299. [doi: [10.1207/S1532785XMEP0303_03](https://doi.org/10.1207/S1532785XMEP0303_03)]
33. Bandura A. Health promotion by social cognitive means. *Health Educ Behav* 2004 Apr;31(2):143-164. [doi: [10.1177/1090198104263660](https://doi.org/10.1177/1090198104263660)] [Medline: [15090118](https://pubmed.ncbi.nlm.nih.gov/15090118/)]
34. Rogers EM. Diffusion of Innovations, 4th Edition. New York: Simon and Schuster; 2010.
35. Bandura A. Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes* 1991 Dec;50(2):248-287. [doi: [10.1016/0749-5978\(91\)90022-L](https://doi.org/10.1016/0749-5978(91)90022-L)]
36. #FoodFri. MomsRising. URL: <https://www.momsrising.org/blog/tags/foodfri> [accessed 2021-01-20]
37. Symplur - connecting the dots in healthcare social media. Symplur. URL: <https://www.symplur.com/> [accessed 2021-01-20]
38. Stokols D. Translating social ecological theory into guidelines for community health promotion. *Am J Health Promot* 1996;10(4):282-298. [doi: [10.4278/0890-1171-10.4.282](https://doi.org/10.4278/0890-1171-10.4.282)] [Medline: [10159709](https://pubmed.ncbi.nlm.nih.gov/10159709/)]
39. Green LW, Richard L, Potvin L. Ecological foundations of health promotion. *Am J Health Promot* 1996 Aug 26;10(4):270-281. [doi: [10.4278/0890-1171-10.4.270](https://doi.org/10.4278/0890-1171-10.4.270)] [Medline: [10159708](https://pubmed.ncbi.nlm.nih.gov/10159708/)]
40. Tweetchats!. Salud America. URL: <https://salud-america.org/tweetchats/> [accessed 2021-01-20]
41. SaludAmerica profile and collections. Wakelet. URL: <https://wakelet.com/@SaludAmerica> [accessed 2021-01-20]
42. Symplur. Symplur API. 2018. URL: <https://www.symplur.com/product/symplur-api/> [accessed 2021-01-20]
43. Google Analytics. URL: https://marketingplatform.google.com/about/analytics/?hl=en_US [accessed 2021-01-20]
44. Words. Symplur Help Center. URL: <https://help.symplur.com/en/articles/105339-words> [accessed 2021-01-20]
45. Network analysis. Symplur Help Center. URL: <https://help.symplur.com/en/articles/934991-network-analysis> [accessed 2021-01-20]
46. Healthcare stakeholder segmentation. Symplur Help Center. URL: <https://help.symplur.com/en/articles/103684-healthcare-stakeholder-segmentation> [accessed 2021-01-20]
47. Blondel VD, Guillaume J, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech* 2008 Oct 09;2008(10):P10008. [doi: [10.1088/1742-5468/2008/10/p10008](https://doi.org/10.1088/1742-5468/2008/10/p10008)]
48. Network analysis. Symplur Help Center. URL: <https://help.symplur.com/en/articles/934991-network-analysis> [accessed 2021-01-20]
49. Obar JA, Zube P, Lampe C. Advocacy 2.0: an analysis of how advocacy groups in the United States perceive and use social media as tools for facilitating civic engagement and collective action. *J Info Policy* 2012;2:1. [doi: [10.5325/jinfopoli.2.2012.0001](https://doi.org/10.5325/jinfopoli.2.2012.0001)]
50. Rodríguez-Díaz C. Health is political: advocacy and mobilization for Latinx health. In: Martinez AD, Rhodes SD, editors. *New and Emerging Issues in Latinx Health*. Switzerland: Springer International Publishing; 2020:349-361.
51. Ranking the Latino population in the states. Pew Research Center. URL: <https://www.pewresearch.org/hispanic/2016/09/08/4-ranking-the-latino-population-in-the-states/> [accessed 2021-01-20]
52. Gomez-Vasquez L, Pujols-Bernabel A, Pujols-Bernabel R. Promoting health topics: a case study of #SaludAmerica on Twitter. In: *Proceedings of the 10th International Multi-Conference on Complexity, Informatics and Cybernetics: IMCIC 2019; Volume I: Papers*. 2019. Presented at: The 10th International Multi-Conference on Complexity, Informatics and Cybernetics: IMCIC 2019; March 12-15, 2019; Orlando, FL. p. 128-133 URL: <http://www.iiis.org/CDs2019/CD2019Spring/PapersZ1.htm#/>
53. Williams HT, McMurray JR, Kurz T, Hugo Lambert F. Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change* 2015 May;32:126-138. [doi: [10.1016/j.gloenvcha.2015.03.006](https://doi.org/10.1016/j.gloenvcha.2015.03.006)]

Abbreviations

GA: Google Analytics

SCT: Social Cognitive Theory

Edited by T Sanchez; submitted 10.06.20; peer-reviewed by J Nguyen, P Medina-Ramirez, P Petaipimol; comments to author 25.07.20; revised version received 18.09.20; accepted 31.12.20; published 01.03.21.

Please cite as:

Ramirez AG, Aguilar RP, Merck A, Despres C, Sukumaran P, Cantu-Pawlik S, Chalela P

Use of #SaludTues Tweekchats for the Dissemination of Culturally Relevant Information on Latino Health Equity: Exploratory Case Study

JMIR Public Health Surveill 2021;7(3):e21266

URL: <https://publichealth.jmir.org/2021/3/e21266>

doi: [10.2196/21266](https://doi.org/10.2196/21266)

PMID: [33646131](https://pubmed.ncbi.nlm.nih.gov/33646131/)

©Amelie G Ramirez, Rosalie P Aguilar, Amanda Merck, Cliff Despres, Pramod Sukumaran, Stacy Cantu-Pawlik, Patricia Chalela. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 01.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Context and Barriers to the Prescription of Nonoccupational Postexposure Prophylaxis Among HIV Medical Care Providers: National Internet-Based Observational Study in China

Haibo Ding^{1,2,3,4*}, PhD; Zehao Ye^{1,2,3,4*}, MPH; Weiming Tang⁵, PhD; Xiaojie Huang⁶, PhD; Hui Wang⁷, PhD; Sitong Cui^{1,2,3,4}, MPH; Yongjun Jiang^{1,2,3,4}, PhD; Wenqing Geng^{1,2,3,4}, PhD; Junjie Xu^{1,2,3,4*}, PhD; Hong Shang^{1,2,3,4*}, PhD

¹National Health Commission Key Laboratory of AIDS Immunology (China Medical University), National Clinical Research Center for Laboratory Medicine, The First Affiliated Hospital of China Medical University, Shenyang, Liaoning Province, China

²Key Laboratory of AIDS Immunology, Chinese Academy of Medical Sciences, Shenyang, China

³Key Laboratory of AIDS Immunology of Liaoning Province, Shenyang, China

⁴Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou, China

⁵Project-China, University of North Carolina at Chapel Hill, Guangzhou, China

⁶Center for Infectious Diseases, Beijing You'an Hospital, Capital Medical University, Beijing, China

⁷Shenzhen Third People's Hospital, Shenzhen, China

*these authors contributed equally

Corresponding Author:

Hong Shang, PhD

National Health Commission Key Laboratory of AIDS Immunology (China Medical University)

National Clinical Research Center for Laboratory Medicine

The First Affiliated Hospital of China Medical University

No 155, Nanjing North Street, Heping District

Shenyang, Liaoning Province, 110001

China

Phone: 86 8328 2634

Email: hongshang100@hotmail.com

Abstract

Background: Nonoccupational postexposure prophylaxis (nPEP) is an effective HIV biomedical prevention strategy. The research and use of nPEP are mainly concentrated in the developed world, while little is known about the knowledge, attitudes, and practices of nPEP among HIV medical care providers in developing countries.

Objective: We aimed to assess the nPEP knowledge and prescribing practice among HIV medical care providers in mainland China.

Methods: HIV medical care providers were recruited in China during May and June 2019 through an online survey regarding nPEP-related knowledge, attitudes, and clinical prescription experiences. Multivariable logistic regression was performed to identify factors associated with prescribing nPEP among HIV medical care providers.

Results: A total of 777 eligible participants participated in this study from 133 cities in 31 provinces in China. Of the participants, 60.2% (468/777) were unfamiliar with nPEP and only 53.3% (414/777) of participants ever prescribed nPEP. HIV care providers who worked in a specialized infectious disease hospital (vs general hospital, adjusted odds ratio [aOR] 2.49; 95% CI 1.85-3.37), had practiced for 6-10 years (vs 5 or fewer years, aOR 3.28; 95% CI 2.23-4.80), had practiced for 11 years or more (vs 5 or fewer years, aOR 3.75; 95% CI 2.59-5.45), and had previously prescribed occupational PEP (oPEP, aOR 4.90; 95% CI 3.29-7.29) had a significantly positive association with prescribing nPEP. However, unfamiliarity with nPEP (aOR 0.08; 95% CI 0.05-0.11), believing nPEP may promote HIV high-risk behavior (aOR 0.53; 95% CI 0.36-0.77) or result in HIV drug resistance (aOR 0.53; 95% CI 0.36-0.77) among key populations, and self-reported having no written oPEP guideline in place (aOR 0.53; 95% CI 0.35-0.79) were negatively associated with nPEP prescription behavior.

Conclusions: HIV medical care providers have insufficient nPEP knowledge and an inadequate proportion of prescribing, which may impede the scale-up of nPEP services to curb HIV acquisition. The implementation of tailored nPEP training or retraining to HIV medical care providers would improve this situation.

KEYWORDS

WeChat; nonoccupational postexposure prophylaxis (nPEP); HIV medical care providers; training; key populations; internet; China; social media; barrier; prophylaxis

Introduction

HIV Epidemic in Key Populations

The Joint United Nations Program on HIV/AIDS (UNAIDS) and World Health Organization (WHO) estimate that 38 million people were living with HIV in 2019, with over two-thirds concentrated in low-income developing countries [1,2]. The epidemic of HIV is concentrated in key populations [3], including men who have sex with men (MSM) [4,5]. There has been an increasing number of new HIV infections in China over the past 5 years [6], with approximately 958,000 people reported living with HIV in 2019 [7]. Data based on the HIV Sentinel Surveillance System in China showed that MSM had an HIV infection prevalence rate of 6.9% in 2018 [8].

Effectiveness of Nonoccupational Postexposure Prophylaxis

Nonoccupational postexposure prophylaxis (nPEP) is an effective and cost-effective HIV biomedical prevention strategy [9,10]. There have been no randomized controlled trials for nPEP due to ethical considerations, but a case-control study of occupational postexposure prophylaxis (oPEP) demonstrated an 81% reduction in the odds of HIV transmission [11]. nPEP guidelines have been in use by WHO, European AIDS Clinical Society, United States, and Canada for years to offer guidance on nPEP uptake [12-16], and the research and use on nPEP in the developed world is extensive. However, nPEP services are not widely used in most developing countries with relatively severe HIV epidemics, even though some have released their own guidelines. Additional efforts are needed to target nPEP uptake to end the AIDS epidemic by 2030.

Previous Studies and Existing Gap

HIV medical care providers play an indispensable role in nPEP uptake, especially medication prescription [17]. Previous surveys have reported on HIV care providers prescribing nPEP in developed countries [18-23], most often including factors such as practice specialty, the number of persons living with HIV in treatment, provider familiarity with nPEP, and the nPEP guideline in place [18,19,23]. As these surveys were conducted in developed countries with nPEP guidelines, it is uncertain whether the situation is similar in developing countries without nPEP guidelines. A clear understanding of obstacles encountered by providers in developing countries without nPEP guidelines will be beneficial to the scale-up of nPEP uptake and control of the HIV epidemic.

A positive attitude has emerged recently in China on the use of nPEP for HIV prevention. The Chinese Center for Disease Control and Prevention (China CDC) carried out a pilot program of nPEP among MSM in 7 provinces to promote the uptake of PEP and preexposure prophylaxis (PrEP) between 2018 and 2019 [24]. In addition, China released the Program to Reduce

AIDS (2019-2020) to ensure that the HIV epidemic was controlled at a low level, which encouraged the application of nPEP programs [25]. Considering an increasing body of evidence, China released the nPEP guideline in October 2020 [26]; however, little is known about the knowledge, attitude, and practice of nPEP in HIV medical care providers in China. It is necessary to understand the nPEP perception among HIV medical care providers and barriers associated with prescribing nPEP to provide targeted interventions.

Objectives

We sought to understand nPEP perceptions and practice among HIV medical care providers and factors correlated with nPEP prescription under the current efforts of scale-up of nPEP services.

Methods

Study Design and Participant Enrollment

We conducted a nationwide online survey among HIV medical care providers during May and June 2019. After a presurvey to adjust the questionnaire items, a survey invitation was sent to 937 HIV medical care providers from two WeChat groups, “National clinicians group majors in HIV/AIDS” and “National physician platform for communicating of difficult cases in HIV/AIDS.” These WeChat groups are currently the leading online WeChat-based communication platforms for HIV-related clinicians in China, with the largest number of registered HIV-related clinicians. The investigator released recruitment information via the WeChat groups, including the study aims, procedure, and requirements of the survey. Eligible participants completed an anonymous online survey by scanning the QR (quick response) code link of the online questionnaire. Inclusion criteria included being age 18 years or older, self-reported practicing in HIV-related medical institutions, having treated at least one person living with HIV over the past year, and providing online informed consent to the study content and protocol. Each individual was allowed to access the online survey once. Each internet protocol address is restricted to answer only one questionnaire. A 30-yuan honorarium (approximately US \$4.50) was paid to each participant through WeChat accounts after completion of the 5 to 10 minute questionnaire survey. We used contact information only for releasing rewards and did not disclose it to others.

Data Collection

After providing informed consent, participants completed anonymous online questionnaires on sociodemographic characteristics (age, sex, ethnicity, and educational background), hospital types, technical titles, practice specialty, length of practice, nPEP-related knowledge, attitudes, and clinical prescription experiences ([Multimedia Appendix 1](#)). The 3 questions on nPEP-related knowledge (with possible answers

yes, no, and I don't know) were as follows: Do you think China has issued national clinical guidelines on nPEP? Do you think unprotected anal intercourse (UAI) risk exceeds percutaneous occupational exposure risk? Do you think percutaneous occupational exposure risk exceeds unprotected vaginal intercourse (UVI) exposure risk?

Data on nPEP-related attitudes (with possible answers agree, neutral, and disagree) were also collected as follows: Do you agree that clinicians have enough time to prescribe nPEP? Do you agree that prescribing nPEP in clinical settings is feasible? Do you agree that prescribing nPEP will promote HIV drug resistance? Do you agree that prescribing nPEP will promote high-risk behaviors?

Additionally, we collected nPEP-related experiences, including the experience of encountering key populations seeking nPEP help and nPEP prescribing history. Before submission, participants could review all items of the questionnaire and make sure mandatory items were completed. To evaluate the impact of geographic HIV epidemic level on prescribing nPEP, we categorized regions into high, middle, and low epidemic levels according to the number of HIV/AIDS cases reported in 2017 ([Multimedia Appendix 2](#)). The top one-third of regions were classified as having a high epidemic level, while the bottom one-third were classified as having a low epidemic level. Further, to evaluate the impact of the nPEP pilot program recently conducted by China CDC, we divided the provinces into 2 categories, nPEP and non-nPEP pilot provinces. The study protocol was reviewed and approved by the institutional review board committee of the First Affiliated Hospital of China Medical University ([2019]2015-138-9). We have completed the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) for this study ([Multimedia Appendix 3](#)).

Sample Size Calculation

We calculated the sample size of participants based on the formula of a 2-sided confidence interval for one proportion: $N = Z_{1-\alpha/2}^2 \times P \times (1 - P) / D^2$. For a conservative estimate of sample size, the proportion of nPEP prescription (P) was set to be 0.5.

At a 5% significance level (α) and 5% margin of error (D), the smallest sample size was calculated as 384 observations.

Data Analysis

Category variables were described by frequency and percentage and continuous variables by mean and standard deviation or median and interquartile range (IQR). All the core variables involved in the questionnaire are required. For variables with a missing ratio of less than 5%, we imputed related missing values in the database by mean for continuous variables and mode for categorical variables in the course of data processing. Variables with more than 5% missing ratio would have been deleted, but there were none in this study. For the needs of analysis, we transformed some variables (eg, familiarity of nPEP) into the binary forms yes (extremely familiar, very familiar) or no (generally familiar, not familiar very much, not familiar at all). We used univariable logistic regression to calculate odds ratios (OR) and their 95% confidence intervals for factors associated with prescribing nPEP among HIV medical care providers. Multivariable logistic regression was applied to estimate associations between predictors and nPEP prescribing history after adjustment for age, sex, ethnicity, and educational background. We used SPSS Statistics version 26.0 (IBM Corporation) for analysis. Variables with 2-tailed $P < .05$ were considered statistically significant.

Results

Demographic Characteristics

Of the HIV medical care providers reached, 82.9% (777/937) of eligible participants from 133 cities in the 31 provinces of China participated in this study ([Figure 1A](#)). Participants had a median age of 42 (IQR 36-48) years. A majority of participants were female (417/777, 53.7%), of Han ethnicity (712/777, 91.6%), had undergraduate or above level of education (743/777, 95.6%), had been in practice for more than 5 years (432/777, 55.6%), and had a technical title of attending physician or above (695/777, 89.4%). Approximately half (394/777, 50.7%) of participants worked at specialized hospitals for infectious diseases ([Table 1](#)).

Figure 1. Distribution of participants (A), proportion of nonoccupational postexposure prophylaxis (nPEP) prescriptions (B), HIV epidemic level (C), and nPEP pilot versus nonpilot provinces of 31 total in China (D).

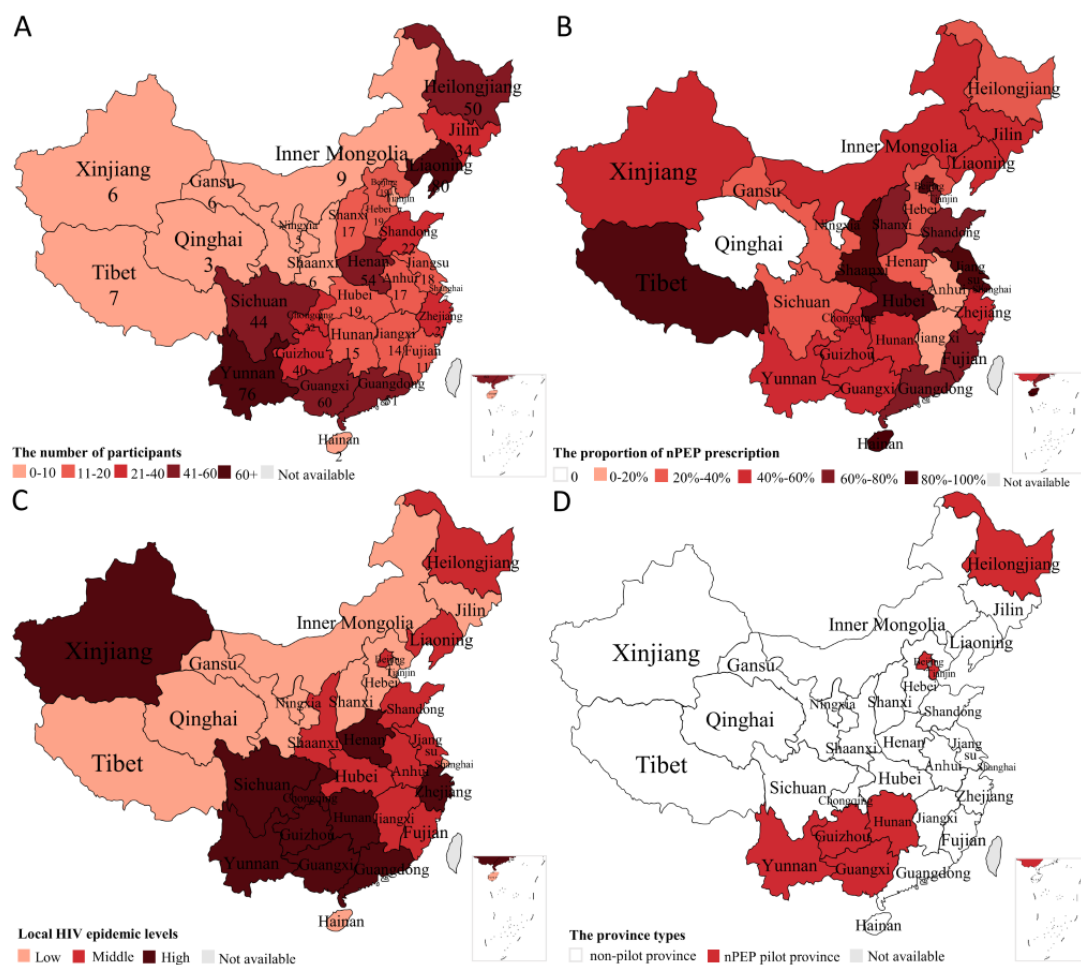


Table 1. Demographics and knowledge of HIV medical care providers in China (n=777).

Variable	Total, n (%)
Age in years	
≤25	3 (0.4)
26-39	302 (38.9)
40-49	331 (42.6)
50-59	131 (16.9)
≥60	10 (1.3)
Ethnicity	
Han	712 (91.6)
Non-Han	65 (8.4)
Sex	
Male	360 (46.3)
Female	417 (53.7)
Educational background	
High school/technical secondary school	5 (0.6)
Junior college	29 (3.7)
Undergraduate or above	743 (95.6)
Administrative regions of China	
North	71 (9.1)
Northeast	164 (21.1)
East	116 (14.9)
South Central	201 (25.9)
Southwest	199 (25.6)
Northwest	26 (3.3)
nPEP^a pilot program province^b	
Yes	267 (34.4)
No	510 (65.6)
Local HIV epidemic level^c	
Low	116 (14.9)
Middle	256 (32.9)
High	405 (52.1)
Hospital type	
Specialized hospital for infectious diseases	394 (50.7)
General hospital	383 (49.3)
Technical title	
General physician	82 (10.6)
Attending physician	274 (35.3)
Associate chief physician	221 (28.4)
Chief physician	200 (25.7)
Clinical practice specialty	
HIV care professional	649 (83.5)
Non-HIV care professional	128 (16.5)

Variable	Total, n (%)
Practice time in years	
≤5	345 (44.4)
6-10	184 (23.7)
≥11	248 (31.9)
Unfamiliar with oPEP^d	
Yes	292 (37.6)
No	485 (62.4)
Unfamiliar with nPEP	
Yes	468 (60.2)
No	309 (39.8)
Do you think China has issued a national clinical guideline on nPEP?	
Yes	551 (70.9)
No	97 (12.5)
I don't know	129 (16.6)
Do you think UAI^e exposure risk between males exceeds percutaneous occupational exposure risk?	
Yes	667 (85.8)
No	61 (7.9)
I don't know	49 (6.3)
Do you think percutaneous occupational exposure risk exceeds UVI^f exposure risk?	
Yes	532 (68.5)
No	199 (25.6)
I don't know	46 (5.9)

^anPEP: nonoccupational postexposure prophylaxis.

^bnPEP pilot programs were conducted by China CDC in provinces Yunnan, Beijing, Tianjin, Heilongjiang, Hunan, Guangxi, and Guizhou.

^cLocal HIV epidemic level was categorized according to the number of HIV/AIDS cases reported in 2017.

^doPEP: occupational postexposure prophylaxis.

^eUAI: unprotected anal intercourse.

^fUVI: unprotected vaginal intercourse.

Knowledge, Experiences, and Attitudes

Overall, only 39.8% (309/777) of participants reported that they were familiar with nPEP, and just 6.8% (53/777) correctly answered all 3 nPEP knowledge-related questions (Table 1). Further, 59.3% (461/777) of participants had provided medical services to fewer than 50 persons living with HIV over the past month, 40.2% (312/777) reported that they had encountered key populations seeking nPEP prescriptions over the past 6 months, and 74.0% (575/777) reported that they had a written oPEP guideline in place (Table 2). Among providers from Northwest China, 69.2% (18/26) were unfamiliar with nPEP and only 23.1% (6/26) had provided HIV care to more than 50 persons living with HIV over the past month.

***A survey of participant opinions on the most suitable population for nPEP prescriptions showed that most participants were inclined to prescribe nPEP to people having a partner living with HIV (543/777, 69.9%), people who had been sexually assaulted (485/777, 62.4%), as well as those with histories of sexually transmitted disease (483/777, 62.2%), unprotected sexual intercourse (466/777, 60.0%), irregular visits to the clinic (454/777, 58.4%), drug injection (452/777, 58.2%), and poor drug adherence (402/777, 51.8%). Moreover, 58.9% (458/777) agreed that they had adequate time to prescribe nPEP, and 27.3% (212/777) and 32.9% (256/777) reported that nPEP may promote HIV drug resistance and HIV high-risk behavior, respectively, among key populations. The problems that primarily concerned participants were high drug cost (452/777, 58.2%) and adverse effects of nPEP (438/777, 56.4%; Table 2).

Table 2. Experiences and attitudes among HIV medical care providers in China (n=777).

Variable	Total, n (%)
Number of persons living with HIV in treatment in the past month	
≤50	461 (59.3)
>50	316 (40.7)
Ever prescribed nPEP ^a	414 (53.3)
Ever prescribed oPEP ^b	197 (25.4)
Self-reported having a written oPEP guideline in place	575 (74.0)
Key populations seeking nPEP help over the past 6 months	
Often/occasionally (more than 1 per month)	312 (40.2)
Never/rarely (less than 1 per month)	465 (59.8)
Having barriers for prescribing nPEP in place	236 (30.4)
nPEP-related attitudes	
Have adequate time to prescribe nPEP	
Agree	458 (58.9)
Neutral	235 (30.2)
Disagree	84 (10.8)
nPEP will promote HIV drug resistance	
Agree	212 (27.3)
Neutral	308 (39.6)
Disagree	257 (33.1)
nPEP will promote HIV risky behavior	
Agree	256 (32.9)
Neutral	302 (38.9)
Disagree	219 (28.2)
Feasible to provide nPEP in place	
Agree	712 (91.6)
Neutral	57 (7.3)
Disagree	8 (1.0)
Worry about being blamed for prescribing nPEP due to no nPEP drug indication	583 (75.0)
Necessary to have expert consensus for nPEP	693 (89.2)
Necessary to establish outpatient for nPEP	620 (79.8)
Concerns about prescribing nPEP	
Increased risk behavior	291 (37.5)
Poor medication adherence	310 (39.9)
HIV drug resistance	238 (30.6)
Side effects	438 (56.4)
High cost	452 (58.2)
No specific guidance for nPEP	159 (20.5)
No nPEP drug indication	126 (16.2)
Resources reduced for HIV-positive patients	121 (15.6)
Other problems	40 (5.1)

^anPEP: nonoccupational postexposure prophylaxis.

^boPEP: occupational postexposure prophylaxis.

Factors Associated With Prescribing nPEP

In total, 53.3% (414/777) of participants had previously prescribed nPEP, among which 38.9% (161/414) reported that they had experienced barriers during the process (Table 2). The proportion prescribing nPEP in each province ranged from 30% to 80% (Figure 1B; Multimedia Appendix 4). The proportions of participants having a history of nPEP prescription were 51.9% (60/414), 56.3% (144/414), and 51.7% (210/414) in provinces with high, middle, and low HIV epidemic levels, respectively (Figure 1C; Table 3), while in the 7 nPEP pilot provinces, 56.2% (150/267) of participants had a history of nPEP prescription (Figure 1D).

Table 3 presents the results of univariable and multivariable logistic regression analyses of factors associated with prescribing nPEP among HIV medical care providers. The forest plot of the results of multivariable logistic regression analysis can be found in Multimedia Appendix 5. After initial adjustment for age, sex, ethnicity, and educational background, we found providers from Northwest regions had a significantly lower proportion of nPEP prescription (vs North China; adjusted odds ratio [aOR] 0.35; 95% CI 0.14-0.89). We further adjusted age, sex, ethnicity, educational background, and administrative regions covariates, and independent factors positively associated

with prescribing nPEP were as follows: practicing in a specialized infectious disease hospital (vs general hospital, aOR 2.49; 95% CI 1.85-3.37), working professionally in HIV care (vs nonprofessional in HIV care, aOR 6.13; 95% CI 3.83-9.81), having a technical title of chief physician (vs general physician, aOR 2.16; 95% CI 1.15-4.05), having 6 to 10 years of practice (vs 5 or fewer years, aOR 3.28; 95% CI 2.23-4.80), having 11 or more years of practice (vs 5 or fewer years, aOR 3.75; 95% CI 2.59-5.45), providing medical services to more than 50 persons living with HIV over the past month (vs 50 or fewer persons living with HIV, aOR 3.89; 95% CI 2.83-5.36), and having previously prescribed oPEP (aOR 4.90, 95% CI 3.29-7.29; each $P < .05$). In contrast, unfamiliar with oPEP (aOR 0.12; 95% CI 0.08-0.16), unfamiliar with nPEP (aOR 0.08; 95% CI 0.05-0.11), unaware that risks of UAI exceed percutaneous occupational exposure risk (aOR 0.63; 95% CI 0.42-0.95); self-reported having no written oPEP guideline in place (aOR 0.53; 95% CI 0.35-0.79), and believing that nPEP may promote HIV high-risk behavior (aOR 0.53; 95% CI 0.36-0.77) or result in HIV drug resistance (aOR 0.53; 95% CI 0.36-0.77) among key populations were negatively associated with nPEP prescription behavior (each $P < .05$). However, practicing in provinces with high HIV epidemic level and nPEP pilot programs were not significantly associated with nPEP prescription behavior.

Table 3. Univariable and multivariable logistic regression analyses of associations of potential predictors and nPEP prescription history among HIV medical care providers in China (n=777).

Variable	nPEP ^a prescription		Crude model	Adjusted model	P value
	Yes, n (%)	No (n, %)	OR ^b (95% CI)	aOR ^c (95% CI)	
Age in years					
<40	153 (50.2)	152 (49.8)	Ref ^d	— ^e	—
40-49	179 (54.1)	152 (45.9)	1.17 (0.86-1.60)	—	—
≥50	82 (58.2)	59 (41.8)	1.38 (0.92-2.07)	—	—
Ethnicity					
Han	374 (52.5)	338 (47.5)	Ref	—	—
Non-Han	40 (61.5)	25 (38.5)	1.45 (0.86-2.43)	—	—
Sex					
Female	233 (55.9)	184 (44.1)	Ref	—	—
Male	181 (50.3)	179 (49.7)	0.80 (0.60-1.06)	—	—
Educational background					
Undergraduate or above	400 (53.8)	343 (46.2)	Ref	—	—
Junior college	13 (44.8)	16 (55.2)	0.61 (0.33-1.47)	—	—
High school/technical secondary school	1 (20.0)	4 (80.0)	0.21 (0.02-1.93)	—	—
Administrative regions of China					
North	45 (63.4)	26 (36.6)	Ref	Ref	—
Northeast	76 (46.3)	88 (53.7)	0.50 (0.28-0.88)	0.51 (0.28-0.91)	.02
East	63 (54.3)	53 (45.7)	0.69 (0.38-1.26)	0.71 (0.38-1.31)	.27
South Central	113 (56.2)	88 (43.8)	0.74 (0.43-1.30)	0.82 (0.46-1.45)	.49
Southwest	107 (53.8)	92 (46.2)	0.67 (0.39-1.17)	0.70 (0.40-1.23)	.22
Northwest	10 (38.5)	16 (61.5)	0.36 (0.14-0.91)	0.35 (0.14-0.89)	.03
nPEP pilot province^f					
No	264 (51.8)	246 (48.2)	Ref	Ref	—
Yes	150 (56.2)	117 (43.8)	1.20 (0.89-1.61)	1.13 (0.82-1.58)	.46
Local HIV epidemic level^g					
Low	60 (51.7)	56 (48.3)	Ref	Ref	—
Middle	144 (56.3)	112 (43.8)	1.20 (0.77-1.86)	1.45 (0.85-2.49)	.17
High	210 (51.9)	195 (48.1)	1.01 (0.67-1.52)	0.50 (0.24-1.03)	.06
Hospital type					
General hospital	158 (41.3)	225 (58.7)	Ref	Ref	—
Specialized hospital for infectious diseases	256 (65.0)	138 (35.0)	2.64 (1.98-3.53)	2.49 (1.85-3.37)	<.001
Technical title					
General physician	38 (46.3)	44 (53.7)	Ref	Ref	—
Attending physician	141 (51.5)	133 (48.5)	1.23 (0.75-2.01)	1.25 (0.75-2.07)	.40
Associate chief physician	105 (47.5)	116 (52.5)	1.05 (0.63-1.74)	0.98 (0.55-1.77)	.96
Chief physician	130 (65.0)	70 (35.0)	2.15 (1.28-3.63)	2.16 (1.15-4.05)	.02
Clinical practice specialty					
Non-HIV care professional	26 (20.3)	102 (79.7)	Ref	Ref	—
HIV care professional	388 (59.8)	261 (40.2)	5.83 (3.69-9.22)	6.13 (3.83-9.81)	<.001

Variable	nPEP ^a prescription		Crude model	Adjusted model	P value
Practice time in years					
≤5	128 (37.1)	217 (62.9)	Ref	Ref	—
6-10	118 (64.1)	66 (35.9)	3.03 (2.09-4.40)	3.28 (2.23-4.80)	<.001
≥11	168 (67.7)	80 (32.3)	3.56 (2.52-5.02)	3.75 (2.59-5.45)	<.001
Unfamiliar with oPEP ^h					
No	345 (71.1)	140 (28.9)	Ref	Ref	—
Yes	69 (23.6)	223 (76.4)	0.13 (0.09-0.18)	0.12 (0.08-0.16)	<.001
Unfamiliar with nPEP					
No	264 (85.4)	45 (14.6)	Ref	Ref	—
Yes	150 (32.1)	318 (67.9)	0.08 (0.06-0.12)	0.08 (0.05-0.11)	<.001
China has issued a national clinical guideline on nPEP					
No	115 (50.9)	111 (49.1)	Ref	Ref	—
Yes/ I don't know	299 (54.3)	252 (45.7)	0.04 (0.01-0.11)	0.03 (0.01-0.09)	<.001
UAI ⁱ risk between males exceeds percutaneous occupational exposure risk					
Yes	366 (54.9)	301 (45.1)	Ref	Ref	—
No/ I don't know	48 (43.6)	62 (56.4)	0.64 (0.42-0.96)	0.63 (0.42-0.95)	.03
Percutaneous occupational exposure risk exceeds UVI ^j exposure risk					
Yes	238 (44.7)	294 (55.3)	Ref	Ref	—
No/I don't know	176 (71.8)	69 (28.2)	3.15 (2.27-4.37)	3.27 (2.33-4.58)	<.001
Number of persons living with HIV in treatment in the past month					
≤50	184 (39.9)	277 (60.1)	Ref	Ref	—
>50	230 (72.8)	86 (27.2)	4.03 (2.95-5.49)	3.89 (2.83-5.36)	<.001
Ever prescribed oPEP					
No	264 (45.5)	316 (54.5)	Ref	Ref	—
Yes	150 (76.1)	47 (23.9)	3.82 (2.65-5.51)	4.90 (3.29-7.29)	<.001
Self-reported having a written oPEP guideline in place					
Yes	350 (60.9)	225 (39.1)	Ref	Ref	—
No	56 (45.9)	66 (54.1)	0.55 (0.37-0.81)	0.53 (0.35-0.79)	.002
Unsure	8 (10.0)	72 (90.0)	0.07 (0.03-0.15)	0.07 (0.03-0.15)	<.001
Key populations seeking nPEP over the past 6 months					
Never/rarely (<1/month)	149 (32.0)	316 (68.0)	Ref	Ref	—
Often/occasionally (>1/month)	265 (84.9)	47 (15.1)	11.96 (8.29-17.25)	13.86 (9.42-20.39)	<.001
nPEP-related attitudes					
Adequate time to prescribe nPEP					
Disagree	44 (52.4)	40 (47.6)	Ref	Ref	—
Neutral	109 (46.4)	126 (53.6)	0.79 (0.48-1.30)	0.83 (0.50-1.38)	.47
Agree	261 (57.0)	197 (43.0)	1.20 (0.76-1.92)	1.24 (0.77-2.00)	.38
Feasible to provide nPEP in place					
Disagree	3 (37.5)	5 (62.5)	Ref	Ref	—
Neutral	11 (19.3)	46 (80.7)	0.40 (0.08-1.93)	0.42 (0.09-2.09)	.29
Agree	400 (56.2)	312 (43.8)	2.14 (0.51-9.01)	2.20 (0.51-9.51)	.29
nPEP will promote HIV drug resistance					

Variable	nPEP ^a prescription		Crude model	Adjusted model	P value
Disagree	161 (62.6)	96 (37.4)	Ref	Ref	—
Neutral	157 (51.0)	151 (49.0)	0.62 (0.44-0.87)	0.65 (0.46-0.92)	.014
Agree	96 (45.3)	116 (54.7)	0.49 (0.34-0.72)	0.53 (0.36-0.77)	.001
nPEP will promote HIV risky behavior					
Disagree	137 (62.6)	82 (37.4)	Ref	Ref	—
Neutral	161 (53.3)	141 (46.7)	0.68 (0.48-0.98)	0.69 (0.48-0.99)	.04
Agree	116 (45.3)	140 (54.7)	0.50 (0.34-0.72)	0.53 (0.36-0.77)	.001
Concern about prescribing nPEP					
No	13 (54.2)	11 (45.8)	Ref	Ref	—
Yes	401 (53.3)	352 (46.7)	0.96 (0.43-2.18)	0.97 (0.43-2.23)	.95
Concern about promoting HIV high-risk behaviors					
No	268 (55.1)	218 (44.9)	Ref	Ref	—
Yes	146 (50.2)	145 (49.8)	0.82 (0.61-1.10)	0.83 (0.61-1.11)	.21
Concern about poor adherence to nPEP					
No	261 (55.9)	206 (44.1)	Ref	Ref	—
Yes	153 (49.4)	157 (50.6)	0.77 (0.58-1.03)	0.81 (0.60-1.09)	.16
Concern about HIV drug resistance					
No	288 (53.4)	251 (46.6)	Ref	Ref	—
Yes	126 (52.9)	112 (47.1)	0.98 (0.72-1.33)	0.97 (0.71-1.32)	.83
Concern about side effects of drugs					
No	181 (53.4)	158 (46.6)	Ref	Ref	—
Yes	233 (53.2)	205 (46.8)	0.99 (0.75-1.32)	0.98 (0.73-1.31)	.88
Concern about cost of nPEP					
No	158 (48.6)	167 (51.4)	Ref	Ref	—
Yes	256 (56.6)	196 (43.4)	1.38 (1.04-1.84)	1.39 (1.03-1.86)	.03
Concern about lack of nPEP clinical guideline					
No	326 (52.8)	292 (47.2)	Ref	Ref	—
Yes	88 (55.3)	71 (44.7)	1.11 (0.78-1.58)	1.19 (0.83-1.71)	.34
Concern about lack of drug indications					
No	338 (51.9)	313 (48.1)	Ref	Ref	—
Yes	76 (60.3)	50 (39.7)	1.41 (0.95-2.08)	1.50 (1.01-2.24)	.045
Concern about reducing treatment resources of HIV-positive patients					
No	377 (57.5)	279 (42.5)	Ref	Ref	—
Yes	37 (30.6)	84 (69.4)	0.33 (0.22-0.50)	0.32 (0.21-0.49)	<.001

^anPEP: nonoccupational postexposure prophylaxis.

^bOR: odds ratio.

^caOR: adjusted odds ratio.

^dRef: reference.

^eN/A: not applicable.

^fnPEP pilot programs were conducted by China CDC in provinces Yunnan, Beijing, Tianjin, Heilongjiang, Hunan, Guangxi, and Guizhou.

^gLocal HIV epidemic level was categorized according to the number of HIV/AIDS cases reported in 2017.

^hoPEP: occupational postexposure prophylaxis.

ⁱUAI: unprotected anal intercourse.

^jUVI: unprotected vaginal intercourse.

Discussion

Principal Findings and Significance

Our study showed that most HIV medical care providers in China were unfamiliar with nPEP, and only a bit more than half of participants had previously prescribed nPEP. We also found that unfamiliarity with nPEP, self-report of having no written PEP-related guideline in place, and less HIV care experience were possibly important barriers to nPEP prescription among HIV medical care providers. This study addresses a gap in the research and shows the negative impact of insufficient knowledge, such as misunderstanding nPEP-related HIV drug resistance and side effects, on the scale-up of nPEP services and subsequent inadequate nPEP prescription by clinicians. It may help public health policymakers learn about HIV medical provider perception of nPEP, thereby providing the opportunity to implement corresponding measures to counter nPEP-related obstacles. Our data also have great significance for further practice after initiating the national nPEP guideline to inform HIV medical care providers in their implementation of nPEP. Additionally, the results of this study are a reference for other countries with similar HIV contexts and insufficient uptake of nPEP services.

Comparison With Prior Work

We found that the proportion (60.2%) of HIV care providers unfamiliar with nPEP was higher than that reported in a previous study from the United States (51.5%) [19]. About 70% of participants incorrectly thought China had already issued a national clinical guideline on nPEP before this survey. This finding may indicate that a high proportion of HIV medical care providers confused oPEP guidelines, released in 2004 [27], with nPEP guidelines or thought the Chinese Guidelines for Diagnosis and Treatment of HIV/AIDS, updated in 2018 [28], were nPEP guidelines. An accordingly high proportion of prescribing nPEP, though, was not found among these providers. This can be attributed to insufficient familiarity with nPEP because of a lack of media advertisements and tailored training. Newly reported HIV cases in China still show an increasing trend [6], however, with strong acceptance of and great demand for nPEP among key populations [29]. This gap could hinder efforts to curb the spread of the HIV epidemic; therefore, intensified publicity through diverse channels and reinforced training or retraining should be offered to improve the knowledge of these providers.

In our study, the proportion of lifetime prescribing of nPEP among HIV care providers (53.3%) was lower than that reported by previous studies from the United States (67.1%) [23], France (58.0%) [30], and Spain (77.3%) [31], which may indicate a huge gap between China and developed countries in the prevention of HIV spread. The gap between the proportion of HIV medical care providers prescribing nPEP and the demand of key populations for nPEP [29] implies that improving the level of nPEP prescription would likely have a remarkable effect on preventing HIV spread. Previous studies found knowledge plays an indispensable role in PrEP prescription behavior [32,33]. Another study found that HIV-related training has a significant correlation with the increased nPEP and PrEP

knowledge and the improved PrEP prescribing practice among HIV care providers [20], which also means a possible effect on nPEP prescription through increasing nPEP knowledge by training. Furthermore, there are many nPEP-related challenges, including risk assessment and management of viral hepatitis, frequent transitions from nPEP to PrEP, and the management of low follow-up rates and poor medication adherence [34], which, if addressed improperly, will bring adverse effects and even harm from nPEP. These challenges will not be resolved in the near-term without targeted nPEP training integrating practical skills exercises into didactic sessions, thereby ultimately delaying the progression of HIV prevention.

Factors Associated With nPEP Prescription

In addition, we identified independent factors positively correlated with nPEP prescription among HIV medical care providers. Compared with providers working in general hospitals, those in specialized infectious disease hospitals had a significantly higher proportion of prescribing nPEP, probably due to more awareness of HIV-related information. HIV-related stigma remains severe in China, however, and key populations are more inclined to visit the general hospital for HIV-related services to protect their privacy and avoid disclosure [35], which may limit access to nPEP services. Thus, for those providers in general hospitals, reinforced targeted training is necessary to improve their perception and enhance willingness to prescribe nPEP. We also found significantly higher proportions of nPEP prescription among HIV care professionals (vs non-HIV care professionals), chief physicians (vs general physicians), providers with more than 5 years of working experience (vs 5 or fewer years), and those having provided HIV care to more than 50 persons living with HIV over the past month (vs 50 or more persons living with HIV). Providers with professional knowledge, high-ranking technical titles, and rich HIV care experience are usually skilled, which can attract more patients. They as well have more opportunities to attend HIV-related international conferences and obtain information on nPEP from other countries. This finding suggests that nPEP-related training should also be focused on young providers to enrich their nPEP knowledge and improve practical skills, which could even be delivered during student medical training. Moreover, the establishment of specific support mechanisms via senior clinicians would help overcome the obstacles to prescribe nPEP faced by young clinicians.

In contrast, we found that unfamiliarity with nPEP, incorrect beliefs that nPEP will promote HIV drug resistance or high-risk behaviors, self-reported lack of written oPEP guideline in working settings, and unfamiliarity with oPEP were all negatively correlated with nPEP prescription among HIV medical care providers. Although the nPEP guideline was released in October 2020 [26], further outreach efforts to clinicians in working settings are needed or existing incorrect perceptions caused by insufficient nPEP knowledge will continue to impede the scale-up of nPEP services. The oPEP guideline was released about 15 years ago, so HIV medical care providers are more familiar with oPEP (62.4%) than nPEP (39.8%).

The associations between the practice of oPEP and nPEP, two methods targeted at different types of HIV exposure, have rarely been explored in previous studies. In our study, the level of prescribing nPEP was higher among HIV medical care providers who had previously prescribed oPEP than that among those who had not. There are many similar features between nPEP and oPEP about assessing HIV exposure risk, principles of treatment, and the types of antiviral drugs. HIV medical care providers who master the oPEP practice may be relatively more familiar with prescribing nPEP. Therefore, training programs combining nPEP with oPEP can create a synergistic effect on both prescription behaviors of HIV care providers. Notably, despite the emphasis of simplifying prescribing practice from the updated WHO guideline for nPEP [13], regardless of HIV exposure types, different types vary in the risk of acquiring HIV and subsequent laboratory test items [36]. Providers confusing the standards of nPEP and oPEP practices may well prescribe nPEP improperly to some individuals at low risk of HIV acquisition [37] or miss some items, such as pregnancy testing and the collection of forensic specimens [36]. It again underlines the necessity to provide targeted nPEP training or retraining based on the nPEP guideline.

Compared with North China (63.4%), we found a surprisingly lower proportion of nPEP prescription in Northwest (38.5%) and Northeast (46.3%) regions where providers have insufficient nPEP familiarity (30.8% and 33.5%, respectively) and less HIV care practice (23.1% and 32.3%, respectively). In the contrast, the HIV epidemic is highly prevalent in Xinjiang Province, located in Northwest China. Hence, more attention should be paid to these regions, especially the Northwest with its limited resources, in future national nPEP training efforts. Given the difficulty of organizing centralized training for HIV medical care providers from various regions, internet-based online training is critical for nPEP implementation. It has clear advantages for transmitting up-to-date knowledge and ideas, particularly for providers in the Northwest regions with insufficient resources for nPEP implementation. Besides traditional didactic sessions, online simulation trainings related to practical skills are also promising methods to offset the gap of resource from regions.

Finally, there was no significant association between high HIV epidemic level and nPEP prescription behavior of HIV medical care providers. This indicates that key populations in those

provinces at high HIV epidemic level may miss the opportunity to obtain nPEP services even after exposure to HIV. Similarly, we did not find the effect of an nPEP pilot program on nPEP prescription behavior of these providers, which may be explained by the relative short duration implementation time and limited number of involved cities. Therefore, it is necessary to further enhance the advertisement of nPEP at a national level to raise the wide attention of HIV medical care providers.

Strengths and Limitations

Our study has many strengths. First, this is a representative cross-sectional study of nPEP perception and prescribing practice among HIV medical care providers in all 31 provinces of China, and the sources of participants from previous studies have been limited. Second, the sample size of this study was larger than those of previous similar studies. Last, as this is the first study of nPEP perception and prescribing practice among HIV medical care providers in China, the results represent a vital reference that could contribute to solving the obstacles to nPEP prescription, popularizing the use of nPEP nationwide, and controlling HIV spread among key populations.

This study also has limitations. First, this study was conducted by two WeChat groups, and our results rely on self-reporting data, which to some extent would cause sampling bias and reporting bias. Second, HIV care providers from Hong Kong, Macao, and Taiwan were not included in the WeChat groups, and the number of samples from western China (ie, Tibet) was insufficient; hence, the results may not well represent the characteristics of HIV medical care providers from these regions. Additionally, given the cross-sectional design, the causal relationships between prescribing nPEP and other factors are uncertain and will require further prospective studies to confirm.

Conclusions

This is the first cross-sectional survey of nPEP-related knowledge, attitudes, and prescribing experience among HIV medical care providers in a country without extensive use of nPEP services. Our results underline the insufficient nPEP knowledge and inadequate proportion of nPEP prescription among these providers. Implementing targeted nPEP training or retraining through the internet, particularly for young providers from general hospitals, should be priorities to eliminate obstacles in popularizing nPEP services and ultimately reducing HIV incidence among national key populations.

Acknowledgments

JX and HD conceived and designed the study. HD collected and cleaned data. JX and ZY analyzed the data and interpreted results. JX and ZY wrote and revised the manuscript. HD and ZY contributed equally as first authors. JX and HS contributed equally as corresponding authors. All authors have read and approved the final manuscript as submitted. The authors thank staff and all clinicians who participated in this study for their efforts. This work was supported by the National Natural Science Foundation of China (81872674), the Mega-Projects of national science research (13th Five-Year Plan [2017ZX10201101-002-007]), and National Science and Technology Major Project (2018ZX10101-001-001-003).

Conflicts of Interest

None declared.

Multimedia Appendix 1

English version of the questionnaire.

[DOCX File, 26 KB - [publichealth_v7i3e24234_app1.docx](#)]

Multimedia Appendix 2

Number of HIV/AIDS cases reported in the 31 provinces of China in 2017.

[XLS File (Microsoft Excel File), 33 KB - [publichealth_v7i3e24234_app2.xls](#)]

Multimedia Appendix 3

Checklist for Reporting Results of Internet E-Surveys.

[DOC File, 64 KB - [publichealth_v7i3e24234_app3.doc](#)]

Multimedia Appendix 4

Proportion of nPEP prescriptions in 31 provinces in China.

[XLS File (Microsoft Excel File), 30 KB - [publichealth_v7i3e24234_app4.xls](#)]

Multimedia Appendix 5

Forest plot of the results of multivariable logistic regression analysis.

[DOC File, 110 KB - [publichealth_v7i3e24234_app5.doc](#)]

References

1. People living with HIV and new HIV infections. The Joint United Nations Programme on HIV/AIDS (UNAIDS). 2019. URL: <http://aidsinfo.unaids.org/> [accessed 2020-08-06]
2. Global Health Observatory (GHO) data. World Health Organization (WHO). 2018. URL: <https://www.who.int/gho/hiv/en/> [accessed 2020-08-06]
3. The Gap Report, 2014. UNAIDS. URL: https://www.unaids.org/sites/default/files/media_asset/UNAIDS_Gap_report_en.pdf [accessed 2020-08-06]
4. Beyrer C, Baral SD, Walker D, Wirtz AL, Johns B, Sifakis F. The expanding epidemics of HIV type 1 among men who have sex with men in low- and middle-income countries: diversity and consistency. *Epidemiol Rev* 2010;32:137-151. [doi: [10.1093/epirev/mxq011](https://doi.org/10.1093/epirev/mxq011)] [Medline: [20573756](#)]
5. Beyrer C, Sullivan P, Sanchez J, Baral SD, Collins C, Wirtz AL, et al. The increase in global HIV epidemics in MSM. *AIDS* 2013 Nov 13;27(17):2665-2678. [doi: [10.1097/01.aids.0000432449.30239.fe](https://doi.org/10.1097/01.aids.0000432449.30239.fe)] [Medline: [23842129](#)]
6. Database for HIV/AIDS. Chinese Center for Disease Control and Prevention. URL: http://www.phsciencedata.cn/Share/ky_sjml.jsp?id=c2ca694e-3995-4c7f-9078-3ed0aaf14556 [accessed 2020-08-06]
7. The progress of HIV prevention. The National Health Commission of the People's Republic of China. 2019. URL: <http://www.nhc.gov.cn/jkj/s3586/201911/c2388ce70bdd404ea6dfcd886591784d.shtml> [accessed 2020-08-06]
8. Key Population Atlas. UNAIDS. 2018. URL: <https://kpatlas.unaids.org/dashboard> [accessed 2020-07-29]
9. Bryant J, Baxter L, Hird S. Non-occupational postexposure prophylaxis for HIV: a systematic review. *Health Technol Assess* 2009 Feb;13(14):1-60 [FREE Full text] [doi: [10.3310/hta13140](https://doi.org/10.3310/hta13140)] [Medline: [19236820](#)]
10. Pinkerton SD, Martin JN, Roland ME, Katz MH, Coates TJ, Kahn JO. Cost-effectiveness of HIV postexposure prophylaxis following sexual or injection drug exposure in 96 metropolitan areas in the United States. *AIDS* 2004 Oct 21;18(15):2065-2073. [doi: [10.1097/00002030-200410210-00011](https://doi.org/10.1097/00002030-200410210-00011)] [Medline: [15577628](#)]
11. Cardo DM, Culver DH, Ciesielski CA, Srivastava PU, Marcus R, Abiteboul D, et al. A case-control study of HIV seroconversion in health care workers after percutaneous exposure. Centers for Disease Control and Prevention Needlestick Surveillance Group. *N Engl J Med* 1997 Nov 20;337(21):1485-1490. [doi: [10.1056/NEJM199711203372101](https://doi.org/10.1056/NEJM199711203372101)] [Medline: [9366579](#)]
12. Smith DK, Grohskopf LA, Black RJ, Auerbach JD, Veronese F, Struble KA, U.S. Department of Health and Human Services. Antiretroviral postexposure prophylaxis after sexual, injection-drug use, or other nonoccupational exposure to HIV in the United States: recommendations from the U.S. Department of Health and Human Services. *MMWR Recomm Rep* 2005 Jan 21;54(RR-2):1-20 [FREE Full text] [Medline: [15660015](#)]
13. Guidelines on post-exposure prophylaxis for HIV and the use of co-trimoxazole prophylaxis for HIV-related infections among adults, adolescents and children: recommendations for a public health approach. World Health Organization. 2014. URL: https://apps.who.int/iris/bitstream/handle/10665/145719/9789241508193_eng.pdf;jsessionid=DF4A1858CDE737192782254FFE89DD9F?sequence=1 [accessed 2020-08-06]
14. EACS Guidelines. The European ADIS Clinical Society (EACS). 2019. URL: https://www.eacsociety.org/files/guidelines-10_0_final_2_2.pdf [accessed 2020-08-06]

15. Centers for Disease Control and Prevention. Announcement: updated guidelines for antiretroviral postexposure prophylaxis after sexual, injection-drug use, or other nonoccupational exposure to HIV—United States, 2016. *MMWR Morb Mortal Wkly Rep* 2016 May 06;65(17):458 [FREE Full text] [doi: [10.15585/mmwr.mm6517a5](https://doi.org/10.15585/mmwr.mm6517a5)] [Medline: [27149423](https://pubmed.ncbi.nlm.nih.gov/27149423/)]
16. Tan DHS, Hull MW, Yoong D, Tremblay C, O'Byrne P, Thomas R, Biomedical HIV Prevention Working Group of the CIHR Canadian HIV Trials Network. Canadian guideline on HIV pre-exposure prophylaxis and nonoccupational postexposure prophylaxis. *Can Med Assoc J* 2017 Nov 27;189(47):E1448-E1458 [FREE Full text] [doi: [10.1503/cmaj.170494](https://doi.org/10.1503/cmaj.170494)] [Medline: [29180384](https://pubmed.ncbi.nlm.nih.gov/29180384/)]
17. Updated recommendations on first-line and second-line antiretroviral regimens and post-exposure prophylaxis and recommendations on early infant diagnosis of HIV: Interim guidance. World Health Organization. 2018. URL: <https://apps.who.int/iris/bitstream/handle/10665/277395/WHO-CDS-HIV-18.51-eng.pdf?ua=1> [accessed 2020-07-31]
18. Rodríguez AE, Castel AD, Parish CL, Willis S, Feaster DJ, Kharfen M, et al. HIV medical providers' perceptions of the use of antiretroviral therapy as nonoccupational postexposure prophylaxis in 2 major metropolitan areas. *J Acquir Immune Defic Syndr* 2013 Nov 01;64 Suppl 1:S68-S79 [FREE Full text] [doi: [10.1097/QAI.0b013e3182a901a2](https://doi.org/10.1097/QAI.0b013e3182a901a2)] [Medline: [24126450](https://pubmed.ncbi.nlm.nih.gov/24126450/)]
19. Henny KD, Duke CC, Buchacz K, Brooks JT, Samandari T, Sutton MY. HIV prescriptions on the frontlines: primary care providers' use of antiretrovirals for prevention in the Southeast United States, 2017. *Prev Med* 2020 Jan;130:105875 [FREE Full text] [doi: [10.1016/j.ypmed.2019.105875](https://doi.org/10.1016/j.ypmed.2019.105875)] [Medline: [31678174](https://pubmed.ncbi.nlm.nih.gov/31678174/)]
20. Henny KD, Duke CC, Geter A, Gaul Z, Frazier C, Peterson J, et al. HIV-related training and correlates of knowledge, HIV screening and prescribing of nPEP and PrEP among primary care providers in southeast United States, 2017. *AIDS Behav* 2019 Nov;23(11):2926-2935 [FREE Full text] [doi: [10.1007/s10461-019-02545-1](https://doi.org/10.1007/s10461-019-02545-1)] [Medline: [31172333](https://pubmed.ncbi.nlm.nih.gov/31172333/)]
21. Landovitz RJ, Combs KB, Currier JS. Availability of HIV postexposure prophylaxis services in Los Angeles County. *Clin Infect Dis* 2009 Jun 01;48(11):1624-1627 [FREE Full text] [doi: [10.1086/598976](https://doi.org/10.1086/598976)] [Medline: [19400685](https://pubmed.ncbi.nlm.nih.gov/19400685/)]
22. Rutland E, Sundaram S, Mani R. The awareness of postexposure prophylaxis for HIV infection following sexual exposure in emergency departments in a regional HIV network. *Int J STD AIDS* 2010 Sep;21(9):636-637. [doi: [10.1258/ijisa.2010.010152](https://doi.org/10.1258/ijisa.2010.010152)] [Medline: [21097736](https://pubmed.ncbi.nlm.nih.gov/21097736/)]
23. John SA, Quinn KG, Pleuhs B, Walsh JL, Petroll AE. HIV post-exposure prophylaxis (PEP) awareness and non-occupational PEP (nPEP) prescribing history among U.S. healthcare providers. *AIDS Behav* 2020 Nov;24(11):3124-3131. [doi: [10.1007/s10461-020-02866-6](https://doi.org/10.1007/s10461-020-02866-6)] [Medline: [32300991](https://pubmed.ncbi.nlm.nih.gov/32300991/)]
24. Chinese Center for Disease Control and Prevention. URL: http://www.chinacdc.cn/zxdt/201912/t20191211_207519.html [accessed 2020-08-06]
25. The program to restrain AIDS (2019-2020). The National Health Commission of the People's Republic of China. 2019. URL: <http://www.nhc.gov.cn/jkj/s3586/201910/40e14fe4c54e41999a147f79558e5eb1.shtml> [accessed 2020-08-06]
26. Chinese Center for Disease Control and Prevention. Guidelines for HIV post-exposure prophylaxis. 2020. URL: <http://ncaids.chinacdc.cn/zxxz/zxxz/202011/W020201116802422550750.pdf> [accessed 2020-12-14]
27. Guidelines for the protection of medical personnel against occupational exposure to HIV. The National Health Commission of the People's Republic of China. 2004. URL: <http://www.nhc.gov.cn/wjw/gfxwj/201304/588fcab93194457cb2cdf3f150b3faac.shtml> [accessed 2020-08-06]
28. AIDS and Hepatitis C Professional Group, Society of Infectious Diseases, Chinese Medical Association, Chinese Center for Disease Control Prevention. [Chinese guidelines for diagnosis and treatment of HIV/AIDS (2018)]. *Zhonghua Nei Ke Za Zhi* 2018 Dec 01;57(12):867-884. [doi: [10.3760/cma.j.issn.0578-1426.2018.12.002](https://doi.org/10.3760/cma.j.issn.0578-1426.2018.12.002)] [Medline: [30486555](https://pubmed.ncbi.nlm.nih.gov/30486555/)]
29. Hou J, Wu Y, Xie L, Meng S, Fu R, Zheng H, et al. Post-exposure prophylaxis: an underutilized biomedical HIV prevention method among gay, bisexual and other men who have sex with men in China. *AIDS Care* 2020 Dec;32(12):1573-1580. [doi: [10.1080/09540121.2020.1742864](https://doi.org/10.1080/09540121.2020.1742864)] [Medline: [32188267](https://pubmed.ncbi.nlm.nih.gov/32188267/)]
30. Laporte A, Jourdan N, Bouvet E, Lamontagne F, Pillonel J, Desenclos J. Post-exposure prophylaxis after non-occupational HIV exposure: impact of recommendations on physicians' experiences and attitudes. *AIDS* 2002 Feb 15;16(3):397-405. [doi: [10.1097/00002030-200202150-00011](https://doi.org/10.1097/00002030-200202150-00011)] [Medline: [11834951](https://pubmed.ncbi.nlm.nih.gov/11834951/)]
31. Almeda J, Allepuz A, Simon B, Blasco JA, Esteve A, Casabona i Barbarà J. [Non-occupational post-exposure HIV prophylaxis. Knowledge and practices among physicians and groups with risk behavior]. *Med Clin (Barc)* 2003 Sep 20;121(9):321-326. [doi: [10.1016/s0025-7753\(03\)73937-2](https://doi.org/10.1016/s0025-7753(03)73937-2)] [Medline: [14499067](https://pubmed.ncbi.nlm.nih.gov/14499067/)]
32. Blackstock OJ, Moore BA, Berkenblit GV, Calabrese SK, Cunningham CO, Fiellin DA, et al. A cross-sectional online survey of HIV pre-exposure prophylaxis adoption among primary care physicians. *J Gen Intern Med* 2017 Jan;32(1):62-70 [FREE Full text] [doi: [10.1007/s11606-016-3903-z](https://doi.org/10.1007/s11606-016-3903-z)] [Medline: [27778215](https://pubmed.ncbi.nlm.nih.gov/27778215/)]
33. Castel AD, Feaster DJ, Tang W, Willis S, Jordan H, Villamizar K, et al. Understanding HIV care provider attitudes regarding intentions to prescribe PrEP. *J Acquir Immune Defic Syndr* 2015 Dec 15;70(5):520-528 [FREE Full text] [doi: [10.1097/QAI.0000000000000780](https://doi.org/10.1097/QAI.0000000000000780)] [Medline: [26247895](https://pubmed.ncbi.nlm.nih.gov/26247895/)]
34. National HIV Curriculum: nonoccupational postexposure prophylaxis. 2020. URL: <https://cdn.hiv.uw.edu/pdf/prevention/nonoccupational-postexposure-prophylaxis/core-concept/all> [accessed 2020-08-06]
35. The worry about privacy disclosure from specialized infectious disease hospitals. Sina. 2015. URL: <http://henan.sina.com.cn/news/hot/2015-06-15/detail-ifxczqar0905361-p3.shtml> [accessed 2020-08-06]

36. Post-exposure prophylaxis to prevent HIV infection: joint WHO/ILO guidelines on post-exposure prophylaxis (PEP) to prevent HIV infection. World Health Organization. 2007. URL: https://apps.who.int/iris/bitstream/handle/10665/43838/9789241596374_eng.pdf?sequence=1 [accessed 2020-08-06]
37. Marzel A, Heinrich H, Schilliger L, Fehr JS, Günthard HF, Kouyos R, et al. Prescription of postexposure prophylaxis for HIV-1 in the emergency room: correct transmission risk assessment remains challenging. *J Acquir Immune Defic Syndr* 2017 Apr 01;74(4):359-366. [doi: [10.1097/QAI.0000000000001265](https://doi.org/10.1097/QAI.0000000000001265)] [Medline: [27906766](https://pubmed.ncbi.nlm.nih.gov/27906766/)]

Abbreviations

aOR: adjusted odds ratio

CHERRIES: Checklist for Reporting Results of Internet E-Surveys

China CDC: Chinese Center for Disease Control and Prevention

IQR: interquartile range

MSM: men who have sex with men

nPEP: nonoccupational postexposure prophylaxis

oPEP: occupational postexposure prophylaxis

OR: odds ratio

PrEP: preexposure prophylaxis

UAI: unprotected anal intercourse

UNAIDS: Joint United Nations Program on HIV/AIDS

UVI: unprotected vaginal intercourse

WHO: World Health Organization

Edited by G Eysenbach; submitted 11.09.20; peer-reviewed by Z Ren, S John; comments to author 18.11.20; revised version received 24.12.20; accepted 21.01.21; published 11.03.21.

Please cite as:

Ding H, Ye Z, Tang W, Huang X, Wang H, Cui S, Jiang Y, Geng W, Xu J, Shang H

Context and Barriers to the Prescription of Nonoccupational Postexposure Prophylaxis Among HIV Medical Care Providers: National Internet-Based Observational Study in China

JMIR Public Health Surveill 2021;7(3):e24234

URL: <https://publichealth.jmir.org/2021/3/e24234>

doi: [10.2196/24234](https://doi.org/10.2196/24234)

PMID: [33704078](https://pubmed.ncbi.nlm.nih.gov/33704078/)

©Haibo Ding, Zehao Ye, Weiming Tang, Xiaojie Huang, Hui Wang, Sitong Cui, Yongjun Jiang, Wenqing Geng, Junjie Xu, Hong Shang. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 11.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Hybrid Approach to Estimation of Underreporting of Tuberculosis Case Notification in High-Burden Settings With Weak Surveillance Infrastructure: Design and Implementation of an Inventory Study

Ellen M H Mitchell^{1*}, MA, PhD; Olusola Adedeji Adejumo^{2*}, MD, MPH; Hussein Abdur-Razzaq³, MD, MPH; Chidubem Ogbudebe⁴, MSc, MPH; Nkem Chukwueme⁵, MD, DPhil; Samson Bamidele Olorunju⁶, MSc; Mustapha Gidado⁷, MSc, MPH, PhD, MD

¹Tropical Infectious Diseases, Department of Public Health, Institute for Tropical Medicine, Antwerp, Belgium

²Mainland Hospital, Yaba Lagos, Nigeria

³Health Research Unit, Lagos State Ministry of Health, Lagos, Nigeria

⁴KNCV TB Foundation Nigeria, Abuja, Nigeria

⁵New York Medical College, New York, NY, United States

⁶Department of Epidemiology and Medical Statistics, Faculty of Public Health, University of Ibadan, Ibadan, Nigeria

⁷KNCV TB Foundation, Koninklijke Centrale Vereniging tot bestrijding der Tuberculose (KNCV), The Hague, Netherlands

*these authors contributed equally

Corresponding Author:

Ellen M H Mitchell, MA, PhD

Tropical Infectious Diseases

Department of Public Health

Institute for Tropical Medicine

155 Nationaalstraat

Antwerp, 2000

Belgium

Phone: 32 0634790509

Email: emitchell@itg.be

Abstract

Background: The greatest risk of infectious disease undernotification occurs in settings with limited capacity to detect it reliably. World Health Organization guidance on the measurement of misreporting is paradoxical, requiring robust, independent systems to assess surveillance rigor. Methods are needed to estimate undernotification in settings with incomplete, flawed, or weak surveillance systems. This study attempted to design a tuberculosis (TB) inventory study that balanced rigor with feasibility for high-need settings.

Objective: This study aims to design a hybrid TB inventory study for contexts without World Health Organization preconditions. We estimated the proportion of TB cases that were not reported to the Ministry of Health in 2015. The study sought to describe TB surveillance coverage and quality at different levels of TB care provision. Finally, we aimed to identify structural-, facility-, and provider-level barriers to notification and reasons for underreporting, nonreporting, and overreporting.

Methods: Retrospective partial digitalization of paper-based surveillance and facility records preceded deterministic and probabilistic record linkage; a hybrid of health facilities and laboratory census with a stratified sampling of HFs with no capacity to notify leveraged a priori knowledge. Distinct extrapolation methods were applied to the sampled health facilities to estimate bacteriologically confirmed versus clinical TB. In-depth interviews and focus groups were used to identify causal factors responsible for undernotification and test the acceptability of remedies.

Results: The hybrid approach proved viable and instructive. High-specificity verification of paper-based records in the field was efficient and had minimal errors. Limiting extrapolation to clinical cases improved precision. Probabilistic record linkage is computationally intensive, and the choice of software influences estimates. Record absence, decay, and overestimation of the private sector TB treatment behavior threaten validity, meriting mitigation. Data management demands were underestimated. Treatment success was modest in all sectors ($R=37.9\%–72.0\%$) and did not align with treatment success reported by the state

(6665/8770, 75.99%). One-fifth of TB providers (36/178, 20%) were doubtful that the low volume of patients with TB treated in their facility merited mastery of the extensive TB notification forms and procedures.

Conclusions: Subnational inventory studies can be rigorous, relevant, and efficient in countries that need them even in the absence of World Health Organization preconditions, if precautions are taken. The use of triangulation techniques, with minimal recourse to sampling and extrapolation, and the privileging of practical information needs of local decision makers yield reasonable misreporting estimates and viable policy recommendations.

(*JMIR Public Health Surveill* 2021;7(3):e22352) doi:[10.2196/22352](https://doi.org/10.2196/22352)

KEYWORDS

tuberculosis; notification; integrated disease surveillance reporting; private sector; inventory study; public health surveillance; epidemiology; infectious disease reporting; infectious disease notification

Introduction

Background

The need to assess the quality and coverage of infectious disease surveillance has increased exponentially in recent years, as emerging infections and antimicrobial resistance crises have highlighted the perils of incomplete, paper-based surveillance systems and lax patient follow-up [1-3]. Tuberculosis (TB) inventory studies are a means to assess the level of misreporting of TB cases in a defined geographical area. Inventory studies help policy makers distinguish between the volume of people with TB who were never reached or treated and those who were treated and simply never reported. Getting these proportions right is critically important to understanding and addressing a country's TB epidemic.

Objectives

The primary aim of this study was to design a methodology that is adequate to determine the magnitude and scope of TB misreporting in Lagos state, a megacity with a highly dynamic private sector. To achieve this, it was necessary to adapt and test an unorthodox approach reflective of this context's unique advantages and challenges. Local stakeholders had a pragmatic need to understand the root causes of undernotification and potential solutions' palatability. Thus, the objectives of this study were to determine the following: (1) number (and proportion) of TB cases treated that were not reported to the Ministry of Health in 2015 among engaged and unengaged facilities; (2) number, proportion, and type of TB facilities that did not comply with the obligation to notify TB cases in 2015 (public and private); (3) characteristics of unnotified versus notified patients with TB; (4) surveillance coverage and quality at different levels of the TB surveillance system; (5) structural-, facility-, and provider-level barriers to notification and reasons for underreporting, nonreporting, and overreporting; and (6) actionable recommendations to improve notification and surveillance quality at the state, local government area (LGA), and facility levels.

As important as inventory studies can be used to shape a country's TB control strategy, the means of obtaining this information is far from simple. In 2012, the World Health Organization (WHO) articulated a series of 7 essential ingredients that must be in place before attempting an inventory study [4]:

1. Case-based data with reliable personal identifiers for record linkage
2. Use of standardized TB case definitions across all care providers
3. Ability to map all care providers outside the existing National Tuberculosis Program (NTP) network
4. Ability to convince all care providers to participate (ie, minimal refusal)
5. Expertise in sampling design, data management, and data analysis
6. At least three independent data sources and sampling of 50% of country areas
7. Statistical research capacity and funding for a research study.

At the time of publication of the WHO inventory study guide in 2012, virtually none of the countries with high TB incidence could fulfill all prerequisites, effectively precluding the measurement of undernotification in the settings where it was most urgent to undertake it. Moreover, developing some of the *essential items* would introduce an ascertainment bias because these elements are known to have a positive association with infectious disease reporting behavior (eg, case-based data systems and constructive relationships with non-NTP care providers) [5]. Universal private sector consent for participation and zero cross-border care seeking were ethically and practically untenable; thus, it was necessary to develop methods to assess the magnitude of biases because of refusal and in- and out-migration.

We chose to combine and adapt methodologies from several WHO-recommended study designs and innovate as needed to design an efficient solution appropriate for high-burden settings with weak surveillance systems. In this paper, we describe the Lagos hybrid method, report how assumptions and techniques are performed, and articulate lessons learned for future replication in high-burden settings where estimates of misreporting are necessary.

Methods

Study Setting

In 2015, Nigeria was considered by many to be at high risk for both under- and overnotification of TB cases because of its sprawling private health sector; yawning treatment coverage gap; ambitious donor-driven case detection targets; and fractured, overlapping infectious disease surveillance systems

[6-10]. Nigeria continues to rank near the top of the list of TB high-burden countries in terms of unmet need for TB treatment, with 91,354 TB cases notified and a TB prevalence rate of 330 per 100,000 people in 2014 [11]. Results from Nigeria's 2012 National TB Prevalence Survey indicated an approximate case detection rate (CDR) of 17%; this would mean that the 100,401 notified TB cases in 2013 represented only 17% of the total estimated 591,000 new TB cases occurring in the same year [7]. According to the 2015 Global TB Report, the WHO estimated that the CDR decreased to 15% [2]. According to projections based on the 2006 census, Lagos's estimated population is 12.5 million, with an average density of 4990 individuals per square kilometer. More than 65% of Lagos's population lives below the poverty line.

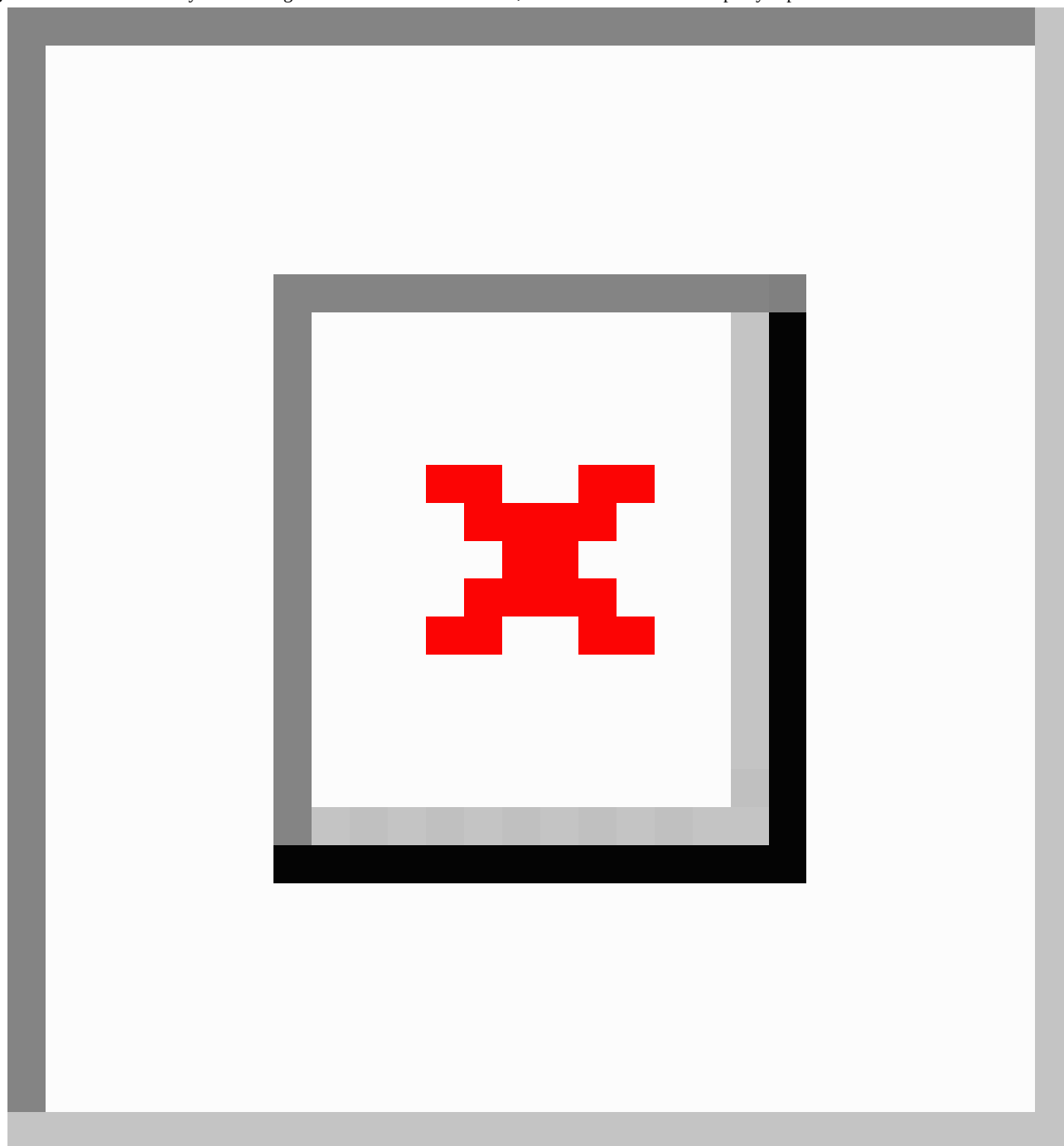
Applying the WHO estimate, the Lagos treatment coverage is 11% because less than 10,000 cases were reported out of the estimated 90,000 cases. Lagos is the largest megacity in Africa in terms of population, divided into 20 administrative districts (referred to as LGA). Border areas are not consistently demarcated, and administrative taxation, transport, and environmental arrangements reflect a functional interdependence between states Ogun and Lagos.

Although Nigeria presents challenges for evaluating disease surveillance systems, it also offers strategic advantages for the study of misreporting such as parallel, independent official

infectious disease reporting systems, multiple sources of TB diagnostic and treatment data, and rigorous recent censuses of the sprawling private sector [12]. The study design was tailored to leverage available data sources and triangulates to improve precision. Specifically, we started from or expanded the WHO inventory study recommendations: (1) retrospective digitalization of paper-based records to create electronic case-based records; (2) use of high-specificity on-site verification procedures to reduce data entry and respondent burden; (3) estimation of the volume of private sector provision of clinical versus bacteriological TB diagnosis using direct and indirect methods; (4) use of buffer zone sensitivity analysis to estimate misreporting in an area *without* well-defined, closed geographical boundaries; (5) estimation of misreporting between administrative levels of the notification system; and (6) study of the underlying rationales and solutions to suboptimal notification behaviors in addition to its quantification.

Figure 1 describes the 2 parallel TB surveillance systems at 2 administrative levels. In parallel to the Lagos TB program, TB outpatient and inpatient treatments are also notifiable through a state-wide monthly disease reporting system known as the Integrated Disease Surveillance and Response (IDSR). The State Tuberculosis and Leprosy Control Program (STBLCP) is the authoritative department within the Lagos State Ministry of Health on TB surveillance matters.

Figure 1. TB surveillance systems in Nigeria in 2015. TB: tuberculosis; TBLS: tuberculosis and leprosy supervisor.



Study Design

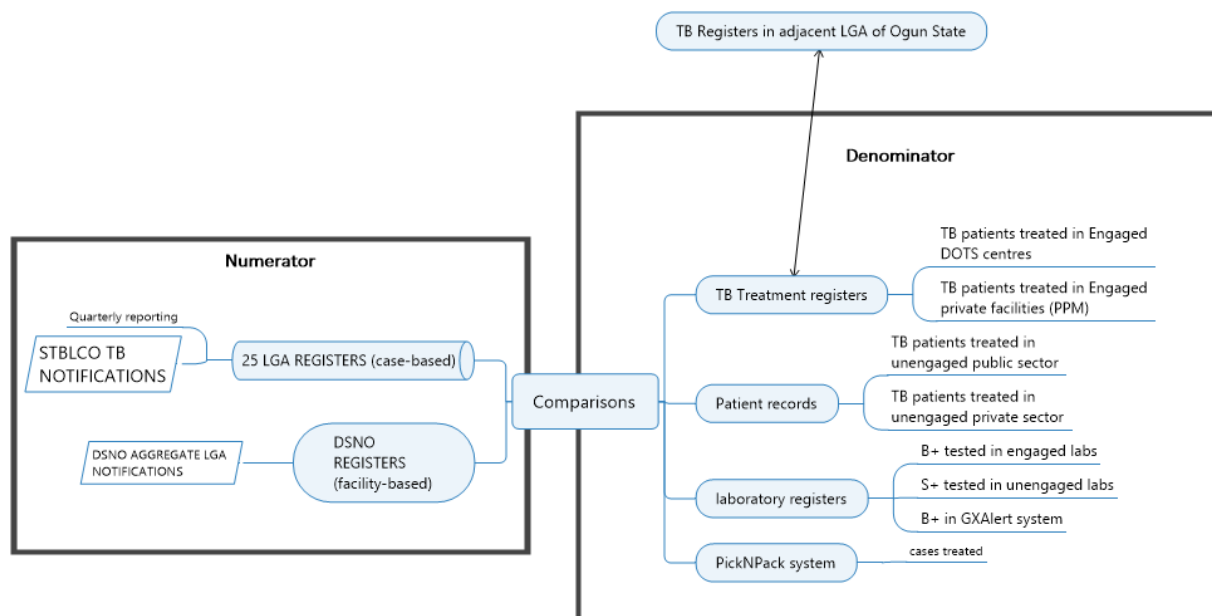
The WHO recommends that inventory study designs in settings without case-based electronic surveillance employ prospective designs in a representative sample of diagnostic sites [4]; however, the addition of data collection systems with standardized case definitions and capacity building required to measure surveillance behavior can introduce bias because these tend to improve surveillance behavior and reporting [13]. Retrospective designs are only recommended in settings with multiple preexisting electronic case-based data sets with standardized case definitions across all providers, an improbable scenario [11]. Therefore, we attempted a hybrid design, quality-assured digitalization of paper-based notifications, and

high-specificity verification of patient source documents to avoid full transcription of source documents in the field. The study triangulated TB case-based data derived from 6 sources instead of 3 (Figure 2).

1. STBLCP aggregate TB case notifications by LGA (State totals)
2. Case-based TB notifications at LGA level (LGA registers)
3. Records of patients with TB treated at engaged public and private facilities in 2015 with Directly Observed Treatment Short-Course (DOTS).
 - from DOTS TB registers
 - from any case notes, lab results, and logbooks of patients diagnosed in other departments

- Records of patients with TB treated in unengaged private and public health facilities (HFs)
4. Smear and/or GeneXpert positive diagnostic test results from engaged laboratories in 2015
 5. Smear and/or GeneXpert positive diagnostic test results from unengaged standalone laboratories.
 6. Patients with TB with Lagos addresses diagnosed or treated in DOTS facilities in 3 contiguous LGA in Ogun in 2015 (abstracted from facility register by data collectors)
 7. Patients with TB with Ogun addresses diagnosed or treated in DOTS facilities in Lagos.

Figure 2. Map of primary data sources for the estimation of the magnitude of underreporting. B+: bacteriologically confirmed TB; DOTS: Directly Observed Treatment Short-Course; DSNO: disease surveillance and notification officer; LGA: local government area; PPM: public-private mix; S+ : smear positive; STBLCO: state tuberculosis and leprosy control officer; TB: tuberculosis.



Sampling

The sampling frames were derived from 5 facility databases: Health Facility Monitoring and Accreditation Agency (HEFAMAA), Millennium Development Goals, Strengthening Health Outcomes through the Private Sector, the State TB program's list of DOTS providers, and the Pick'N'Pack drug management system records [12,14]. We used 2 lists of DOTS centers maintained by the State TB program to create the sampling frame of the engaged DOTS facilities; this included 217 public engaged facilities and 98 private providers involved in the public-private mix (PPM) scheme. For the purposes of the study, we considered a public facility *engaged* if it was listed on the official TB program DOTS facility list or if it was found in the list of facilities in the TB drug commodity management system (Pick'N'Pack) [15]. We aimed to include a census of all engaged facilities; 3 sites reported that they had been engaged in 2016, but these were still classified according to their 2015 engagement status.

Although the PPM program in Lagos is one of the most established and well studied in Nigeria, the actual size of the official engaged PPM program in Lagos is modest, involving 96 private HFs or less than 5% of Lagos's private HF [16-19]. Accurate estimation of the unengaged HFs sample frame was crucial to avoid over- or underestimation and underreporting in the extrapolation phase. [Multimedia Appendix 1](#) [20-23] describes the samples and sampling in detail.

There were 131 public HFs that did not provide TB services and were considered *unengaged*. Among them were 27 sites that had historically offered TB care but had ceased providing TB diagnosis and/or treatment in 2015 because of staffing, managerial, or infrastructure constraints. Some had been classified by the TB program as *semidormant* or *dormant* and had been removed from official lists of DOTS centers. As the study was retrospective, we knew that interventions were planned during the 2015 or 2016 period to resume TB services at dormant sites. Therefore, we sampled 27 nominally *unengaged* public facilities that had historically been DOTS facilities, had laboratory capacity, or had received TB drugs in the recent past. A list maintained by the State TB program of engaged laboratories participating in TB activities in 2015 was used as the sampling frame. Engaged laboratories were trained and equipped to conduct smear microscopy and/or GeneXpert tests. All the engaged laboratories were eligible as they were likely to have a high volume of TB diagnoses, and sampling could have introduced bias.

The HEFAMAA database of 349 registered laboratories was deduplicated and compared with the standalone laboratories in the DOTS database to identify 272 unengaged private laboratories.

Eligibility Criteria of HFs

Facilities were invited to participate if they were functioning as health service providers at the time of the survey (mid-2017)

and if they met the eligibility criteria. The eligibility criteria for the inclusion of HF's in the study are listed in [Textbox 1](#).

Laboratories were eligible for inclusion as engaged if they were standalone diagnostic centers that did not treat and were not

physically housed inside or otherwise connected to an HF; the laboratories inside HF's were searched as part of verification exercises for engaged and unengaged HF's.

Textbox 1. Eligibility and exclusion criteria of the health facilities.

1. The following administrative management structures were included:
 - Public (ie, federal, state, and local government area owned and run)
 - Faith-based
 - Nongovernmental organization (ie, academic and nongovernmental organizations)
 - For-profit
 - Prisons with health facilities
2. Having at least one registered nurse or medical doctor employed
3. The following levels of service delivery were included:
 - Hospitals (general, teaching, specialist, and cottage)
 - Medical centers or primary health clinics
 - Nursing homes
 - Maternity clinics
4. If public, health facilities had to have access to tuberculosis drugs to be considered as treating patients with tuberculosis.

The following health providers or facilities were excluded:

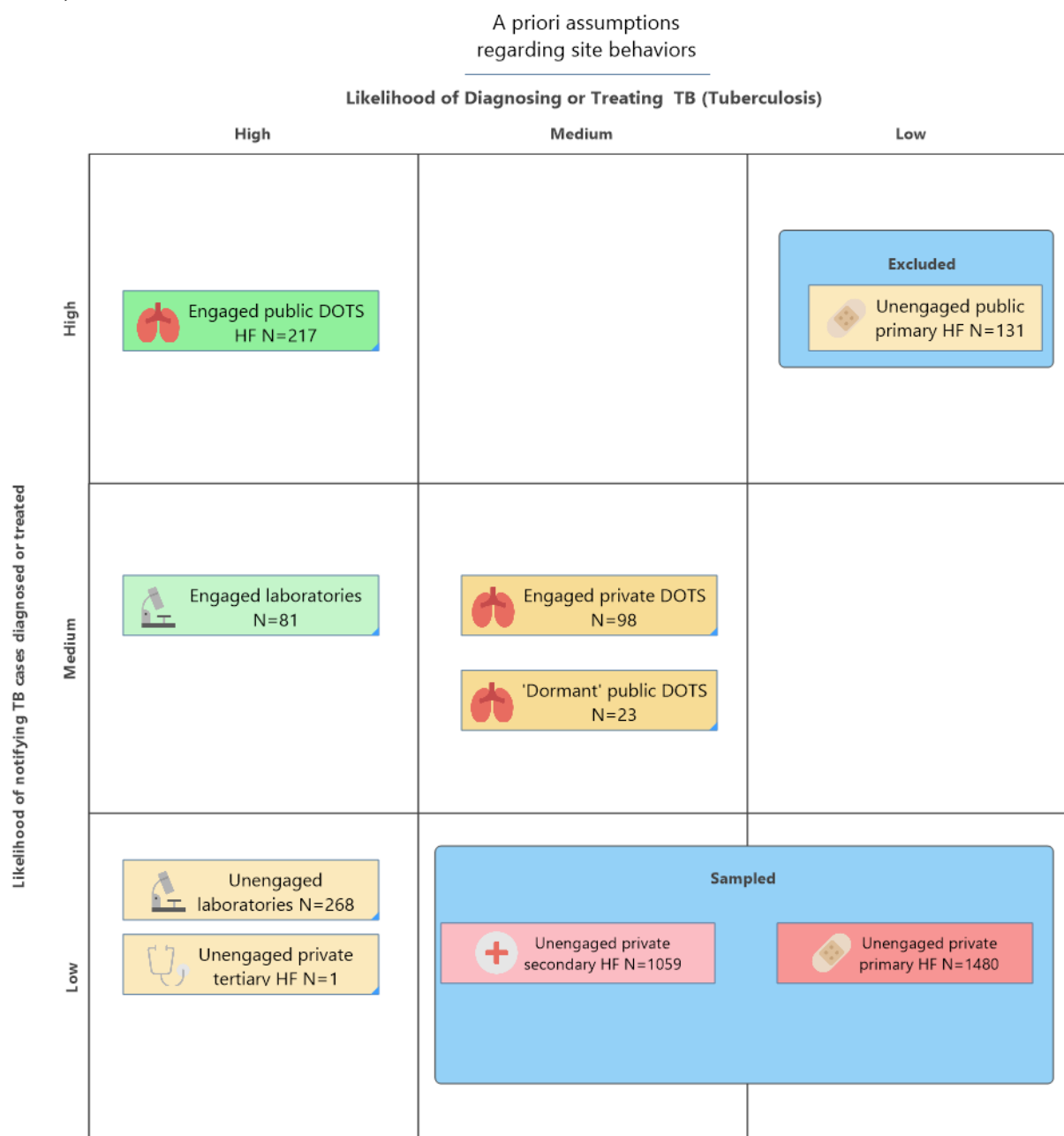
- Community pharmacies
- Patent medical vendors [24]
- Mobile clinics whose location cannot be determined a priori
- Dental and eye clinics
- Dialysis center
- Physiotherapy clinics
- Corporate health facilities

Sampling Rationale

The choice to census or sample a facility stratum was based on 3 factors: the size of the stratum (a proxy for the affordability of a census), probability of offering TB treatment, and the probability of TB case notification capacity ([Figure 3](#)). Previous studies have demonstrated that the volume of TB cases treated in engaged DOTS centers and engaged labs were variable and dynamic over time, TB notification capacity was present, and misreporting was expected to be minimal, so a census of these strata was prudent to minimize error. All facilities engaged as of 2015 were eligible and invited to participate—public DOTS, private DOTS, and laboratories. There was not enough a priori

information about TB diagnostic capacity or volume in unengaged private laboratories to sample them in an unbiased manner, so we chose to census all the unengaged laboratories. Finally, there were 131 unengaged public facilities without staffing, diagnostic capacity, or drugs to treat TB. We excluded these facilities from the study with the exception of 23 public facilities that had offered TB services in the past or were undergoing TB training and engagement during 2015; they were not on any official lists of engaged DOTS centers, and their status was ambiguous. Therefore, we purposefully sampled all sites that had historically been DOTS facilities, had laboratory capacity, or had been targeted for resuming TB diagnosis and/or treatment.

Figure 3. Behavioral assumptions driving the decision to census or sample health facility strata. DOTS: directly observed therapy short-course; HF: health facilities; TB: tuberculosis.



Sampling was necessary for the unengaged private HFs sector because it was not economically feasible to include 2634 HFs. A previous study showed that 32% to 36% of private providers offered TB treatment, and the probability of offering TB care was related to the facility level [12]. Therefore, stratified sampling by level with probability proportional to size appeared viable. We hypothesized that TB case underreporting behavior was common, if not universal, among unengaged private providers. Unengaged private providers had no access to the required NTP recording and reporting (R+R) forms, and the TB program did not accept TB notifications that did not employ national forms and guidelines. We decided to stratify the sample by the HF level, with oversampling of secondary level facilities for 3 reasons. First, we reasoned that larger facilities were more likely to have a more formalized R+R system in place for

Disease Surveillance and Notification Officer and HEFAMAA certification, so we assumed slightly lower levels of undernotification behavior and a bigger sample needed to detect it. Second, more complex facilities have more departments with staff and equipment to diagnose TB but less experience reporting it. Finally, Nigerians' recourse to hospitals for TB care is well documented and obliged a robust estimate in this sector. Among 1142 individuals interviewed for health-seeking behavior through Nigeria's 2012 National TB Prevalence Survey, 45% first sought care from a hospital [20].

Sample Size Calculation for Unengaged Private HFs

The sample size formula recommended by WHO frames TB cases as the sampling unit and requires previous knowledge of the harmonic mean of the cluster size (ie, the average number

of TB cases found in HFs) and the coefficient of between-cluster variation—both parameters unlikely to be known to investigators researching the mostly unregulated private health sectors of high-burden countries. As we sampled only unengaged private HFs, HF was our sampling unit. Applying the standard sample size formula for estimating the population proportion (ie, the proportion of HFs notifying) with specified relative precision, 327 HFs were required (Table 1). Therefore, 380 replacement HFs were drawn from the frame to be substituted in the event of refusal, closures, and other eventualities. We drew the sample proportionally to the distribution of HFs among the LGAs.

Oversampling secondary level facilities was prudent as this stratum had a higher likelihood of TB provision, so even small errors could significantly impact estimates of the overall magnitude of undernotified TB treatment. Although missing TB notifications was unlikely, undercounting treated TB cases in the unengaged private sector was the major validity threat to the study. We also had to power the study such that we would accrue enough unnotified TB cases to describe their characteristics and identify any predictors of non-notification. We aimed to recruit a minimum of 112 unengaged private HF patients with TB treatment documentation.

Table 1. Overview of health facility and laboratory sampling plan.

Type of facility	Numbers of facilities	Sampled quotient, n (%)	Type of sample
Health facility			
Public engaged	217	217 (100.00)	Census (take all)
Private engaged	98	98 (100.00)	Census (take all)
Public unengaged	131	23 (17.55)	Purposive (take all dormant/former DOTS centres)
Private unengaged	2634	327 (12.41)	Stratified sample with selection proportional to size sampled with replacement
Primary	1480	141 (9.53)	N/A ^a
Secondary	1059	186 (17.56)	N/A
Total	3080	665 (21.59)	N/A
Laboratory			
Public engaged	43	43 (100.00)	Census (take all)
Private engaged	33	33 (100.00)	Census (take all)
Public unengaged	4	4 (100.00)	Census (take all)
Private unengaged	268	268 (100.00)	Census (take all)
Total	349	349 (100.00)	N/A

^aN/A: not applicable.

TB Case Definition

For this study's purposes, the definition of a TB case was broad and subdivided into 2 groups: (1) patients with bacteriologically confirmed (B+) TB who had one or more positive smear microscopy test results and/or a positive GeneXpert test result and (2) patients with clinical TB who were diagnosed by a clinician via a chest x-ray, children diagnosed with symptoms compatible with TB and responsive to treatment, and patients of any age treated by a formal medical provider for drug-sensitive TB for a minimum of 1 month, regardless of the site or type of TB or means of diagnosis.

Drug-resistant TB notification was not included in this study because surveillance systems and practices are distinct and treatment provision and treatment are more carefully regulated; thus, misreporting is minimal.

Data Collection

For this study's purposes, the numerator (notified TB) was considered the sum of the cases recorded in the LGA TB notification registers because they contained case-based data.

The LGA registers of notified TB were digitized by trained data entry clerks before fieldwork. These complete case-based notifications were preloaded on the tablets used by data collectors to preclude full-field transcription of patient registers in engaged facilities with a high previsit probability of notification, which was deemed too time-consuming and error-prone; therefore, a hybrid strategy of high-specificity verification was developed and tested.

Data collection instruments were developed, pretested, piloted, and redesigned to reduce respondent burden and focus on variables essential to record linkage. Data dictionaries in the web-based platform were aligned with variables in the national R+R tools. Data on TB cases treated in 2015 were collected from mid-June 2017 to September 2017. Case notification is a human behavior, and the Lagos hybrid method employed HF strata-specific recruitment and data collection approaches in recognition of the different norms, concerns, and motives of each HF stratum.

Data Collection in Engaged HFs

In addition to seeking TB cases in all departments of each facility, data officers verified the existence of each previously

notified TB case (see [Textbox 2](#) for key case variables). For a *definitive match* (verification), the data officers ensured that the tablet and register matched exactly the following 7 variables: first name, last name, age, gender, date treatment started, smear status, and treatment outcome. A strict deterministic 7-variable match algorithm was chosen to maximize the specificity for

linkages. If there was a single discrepancy between a source document and notified case on any of the 7 variables, the data officer deemed it a *nonmatch* and entered the data of patients with TB as a possible *unnotified TB case* to be further scrutinized by probabilistic methods.

Textbox 2. Key case variables.

- First name
- Surname
- Address (ward and state)
- Date of diagnosis
- Age
- Sex
- Smear done
- Chest x-ray done
- Tuberculosis type (smear result and GeneXpert mycobacterium tuberculosis and rifampicin resistance result)
- Tuberculosis location (pulmonary vs extrapulmonary)
- Treatment outcome
- TB and leprosy supervisor assigned identification of patient (if applicable)
- Facility code

Data Collection in Unengaged HFs

The quality and completeness of records in the unengaged sector are known to be highly variable, and a highly adaptive, responsive data collection procedure was needed to gather information. A retrospective approach relies on the existence of patient records in some form. Earlier studies have shown that only 66% of private providers reported having formal patient record systems that summarized the care rendered [21]. Patient information was stored in exercise books (52%), national registers (29%), or forms conceived by the facility staff (15%) [21].

Data collection at unengaged facilities was different from engaged HFs in 3 respects: (1) an incentive (US \$5) was offered to private unengaged HFs for participation; (2) a smaller set of variables was collected to reduce privacy concerns (eg, no HIV data); and (3) given the absence of TB registers, data officers had to search for any available patient or diagnostic primary sources and abstract those TB variables that were possible to abstract in an opportunistic way.

We anticipated high rates of refusal (and/or delayed acceptance) among private unengaged HFs because of name-based data collection requiring disclosure of privileged personal information on clients. We sought to reduce this risk of bias by inviting private providers to informational events where influential stakeholders could dispel rumors and encourage participation. We also sought to give facility managers choices over how, how much, and when data were collected, including the option of partial data provision.

Structured interviews with health care providers were conducted to explore self-reported notification knowledge, attitudes,

behaviors, and support. Health care workers not currently reporting TB cases were invited to indicate how they would respond to the menu of incentives and enablers to notify TB cases. The responses were written on paper instruments by trained data collectors. Information on the structural-, facility-, and provider-level barriers to notification and reasons for underreporting, nonreporting, and overreporting were collected.

A focus group discussion (FGD) of TB and leprosy supervisor was conducted by an experienced facilitator using an FGD guide in a mix of English and Yoruba. A transcript of the audio recording was produced, and framework analysis was used to derive key challenges and recommendations [22].

Ethical Considerations

The study required collecting demographic information for the purposes of record linkage (eg, date, age, sex, first name, and last name). Therefore, strict data security procedures were followed to prevent the inadvertent or deductive disclosure of the identities of TB clients and patients, health workers, and facility staff. Data were encrypted and password protected. No provider names were collected to prevent inadvertent disclosure of the identities of health care providers who are in violation of the obligation to notify TB cases; facilities were only identified using numerical codes. All investigators completed the Nigerian research ethics certification courses. The complete protocol was approved by the Health Research and Ethics Committee of the Lagos State University Teaching Hospital (registration number 04/04/2008).

Data Management

Surveillance data on TB case notifications were single entered into secure relational databases programmed in MySQL

following standard operating procedures (SOPs). Fields contained both hard and soft edits, with automated logic checks to increase data quality. During the study implementation, 14 facilities were visited by 2 independent teams to assess the study procedures' inter-rater reliability for finding cases. Access to the data dashboard was limited to investigators and staff, and web-based access to raw data was limited to those with administrator privileges.

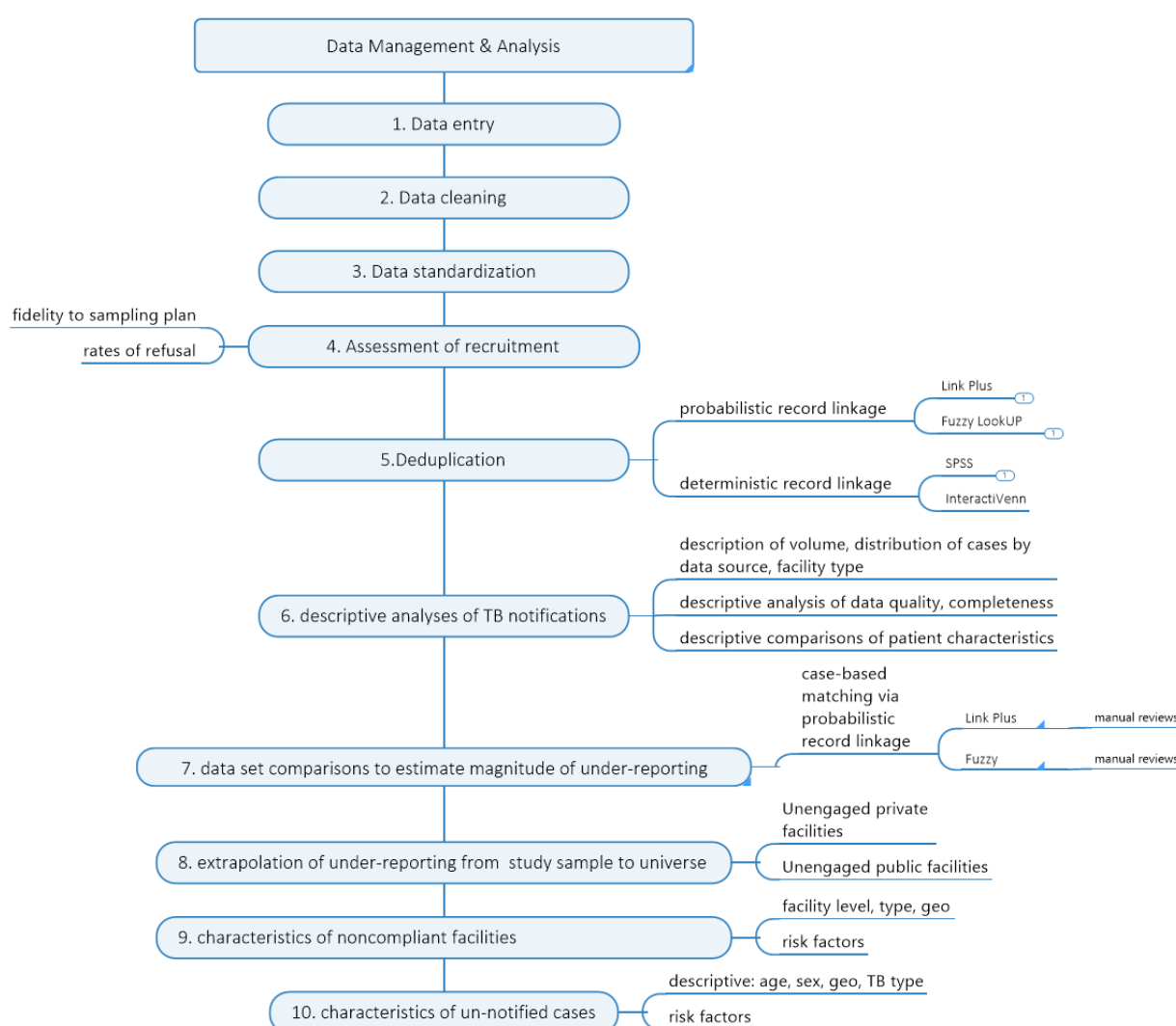
For all patient data collected at any site, the study software automatically generated an alphanumeric case ID code using

first name, last name, age, and sex. Given the ethnic diversity in Lagos, the variance in possible first and last names is higher than in homogeneous settings. This greatly enhances the efficiency of record linkage via name-based matching.

Data Analysis

As the magnitude estimates are highly sensitive to the bias in sampling, participation, and data quality, transparency in data management is crucial. [Figure 4](#) summarizes the order of the tasks.

Figure 4. Overview of the sequence of data management and analysis steps. TB: tuberculosis.



An inventory of all information collected from any source was compiled. Missing files, consents, or duplicate facility visits were investigated by the data manager. Data were cleaned and validated by the data manager and the principal investigators in January 2018. An audit trail was maintained to indicate changes to the data sets during cleaning.

In this study, the identification and removal of duplicates through record linkage was a precursor for comparison between sources. Probabilistic record linkages were conducted to deduplicate the data sets ([Multimedia Appendix 2 \[11,18\]](#)).

Study Objective 1: Number and Proportion of TB Cases Treated

The calculation of the sum of all unique TB cases diagnosed in Lagos in 2015 (the denominator) required extensive record linkage to avoid double counting of patients with TB appearing in multiple data sets. The engaged facility register data were compared with the LGA register data using probabilistic record linkage to identify TB cases that had not been notified.

Estimation of Total Unnotified Cases From Facility Samples

The proportion of unnotified TB cases in each sampled stratum was extrapolated to construct the total number of unreported cases. The record linkage of B+ TB cases diagnosed in laboratories was compared with B+ patients in different databases to estimate the number of B+ patients with TB who were treated in the unengaged sectors or were untreated.

As Nigeria does not have a system of unique personal identifiers for TB, probabilistic record linkage for smear and GeneXpert positive TB cases was essential to compare laboratory and facility records to establish the number of B+ TB cases treated in the unengaged sector (or either untreated). The comparison of laboratory and register cases was conducted with Excel Fuzzy LookUp and CDC LinkPlus because of the limited number of potential record linkage variables and the ability of Fuzzy LookUp to compare information across variables instead of only between variables. LinkPlus failed to identify matches when the order of names was reversed, whereas Fuzzy LookUp could consider similarities in a group of variables (eg, first and last names jointly; [Multimedia Appendix 3](#)).

Extrapolation of Total Patients With TB Volume From the Unengaged Private Sample

To calculate the size of the population of patients with TB treated in the unengaged sector, 2 approaches were used. To estimate the total number of B+ patients with TB treated outside

the engaged HF, a direct measure of total B+ TB diagnoses was feasible because of the inclusion of all laboratories in the state. The number of B+ patients with TB treated in the unengaged private sector was presumed to be the total number of B+ patient test results remaining after removing samples linked to treated patients with TB in TB treatment registers (accounting for cross-border diagnoses). This was a reasonable approach because the study was effectively a census of all known laboratories in Lagos, and therefore, it captured the universe of samples testing positive for TB.

To calculate the number of clinically diagnosed patients with TB treated by the unengaged private sector, we took the number of unique clinically diagnosed patients with TB resident in Lagos found by each stratum (level) of the facility and produced weighted strata-specific averages that were then applied to the universe and summed to obtain the total. Adding the total B+ cases and the clinically diagnosed cases together yielded the estimate of total TB treated by the unengaged private sector ([Figure 5](#)).

As pretreatment loss to follow-up could not be measured in the HFs that were not sampled, we assumed the continental average pretreatment loss to follow-up applied to Lagos residents testing positive for TB. A systematic review found pretreatment loss to follow-up averages of 18% (95% CI 13-22%) in Africa [[23](#)].

In addition to assessing undernotification, this inventory study included a comparison of the volume and characteristics of TB cases in 3 layers of the TB surveillance system. For comparisons of 3 or fewer data sets, Euler and Venn diagrams were generated to maintain proportionality. EulerAPE v3 software was used for area-accurate proportional Euler diagrams [[25](#)]. For comparisons of 4 or more data sets, spherical or symmetrical Venn diagrams were necessary. Interactivenn and Jvenn were used for spherical or Edwards-Venn diagrams ([Figure 6](#)) [[26,27](#)].

Figure 5. Process used to estimate the total number of tuberculosis cases treated in the unengaged private sector. B+: bacteriologically confirmed tuberculosis; HF: health facility; TB: tuberculosis.

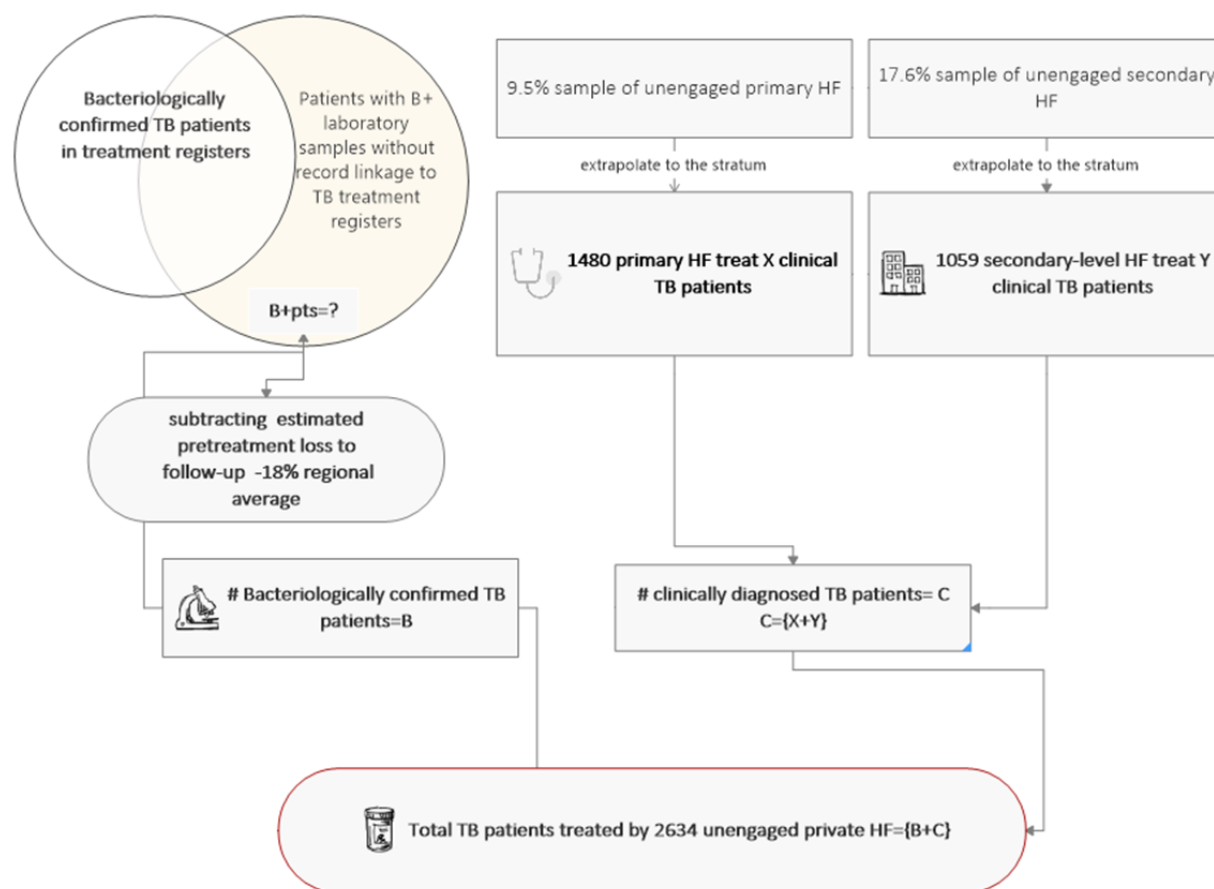
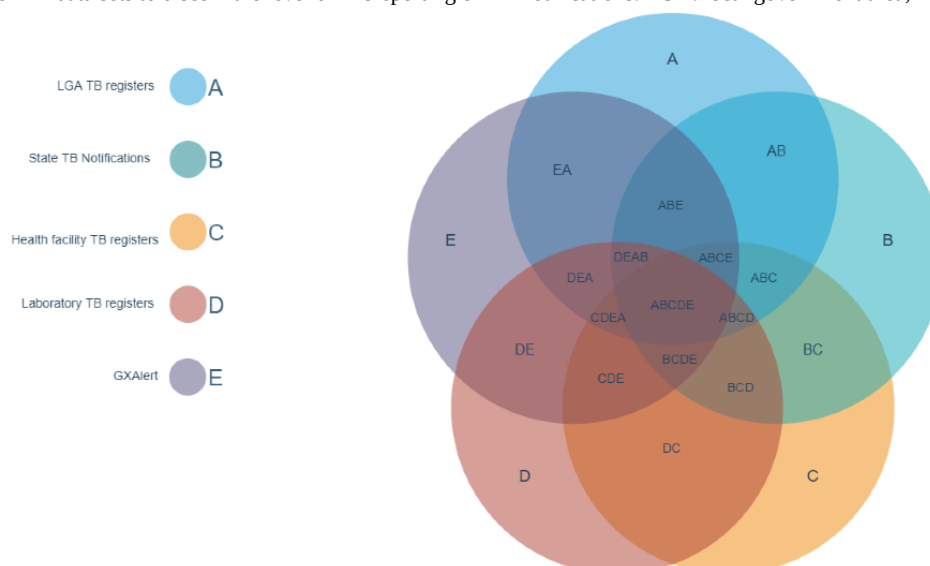


Figure 6. Comparison of TB data sets to discern the level of misreporting of TB notifications. LGA: local government area; TB: tuberculosis.



Adjusting for Cross-Border Diagnosis and Treatment

The WHO inventory study methods may assume no health care provision across geographic borders, an unlikely situation. Due to the elongated shape and riverine coastal areas of Lagos, it is often more convenient to cross the border into Ogun than to traverse Lagos for health care. To account for cross-border care

seeking in estimates of undernotification, it was necessary to measure the magnitude of the cross-border diagnosis and treatment phenomena between Lagos and Ogun. We conducted probabilistic record linkage among TB treatment registers in 3 contiguous LGA in Ogun (bordering Lagos) and smear-positive lab samples in Lagos, and a buffer zone analysis was conducted

to estimate the magnitude of misreporting because of cross-border care seeking.

Study Objective 2: Facility Compliance With Obligations to Notify

As the LGA register and facility register data were case-based, these data permitted facility-level analyses following aggregation of the data. Syntax was written for analyses using the SPSS (IBM version 25). The number, proportion, and type of TB facilities that did not comply with the obligation to notify in 2015 (public and private) were calculated to help pinpoint areas for intervention.

We characterized the engaged facilities as *active*, *semidormant*, or *dormant*, based on whether they were providing TB treatment in 2015. HFs that served an average of less than 1 patient per month were considered *semidormant*. Active and semidormant HFs were included in an analysis of the completeness of notifications. They were subsequently *compliant* or *noncompliant* with reporting requirements. Categorization of *noncompliant* HFs includes accurate reporting (ie, zero difference between facility and LGA register), underreporting (ie, reporting fewer cases than patients recorded in the treatment register), and overreporting (ie, reporting cases that could not be verified in the facility register). Multiple imputations were conducted to estimate the magnitude of misreporting because of refusals among eligible HFs.

Study Objective 3: Comparison of Characteristics of Notified and Unnotified Patients With TB

Beyond the estimation of the magnitude of underreporting, it is necessary to understand the differences between TB cases reported and TB cases unreported. Differences may include variation in the quality of diagnosis, treatment, patient population served, geographic area, and types of facilities that opt to notify TB and those who do not. Descriptive statistics of notified TB cases (eg, LGA distribution, types of TB age, gender, treatment outcomes, and source of notifications [facility type]) were generated to compare service delivery by sector.

TB case characteristics in each data set were described and compared between and among data sets to discern if there were any differences in the type of people or type of TB treated in the various sectors. The magnitude and variance of the cases reported were examined to identify the central tendencies and outliers. Percentages and means were calculated for numerical variables; the chi-square test or Fisher exact test was used to compare categorical variables, whereas two-tailed *t* tests were used to compare continuous variables. For all statistical tests, $P < .05$ was considered statistically significant.

To discern the implications of underreporting, we conducted descriptive analyses to detect differences among notification systems, facilities, and cases. Data were analyzed in a stepwise fashion according to descending aggregation levels: LGA, facility, and person. The movement of B+ patients within and between HF types was mapped to explore retention, referral, and treatment choices.

Study Objective 4: TB Surveillance Coverage and Quality

We compared participation levels in the 2 state-run disease surveillance systems in terms of TB to understand their geographical scope and structural differences and make inferences about the quality and reliability of the estimates generated.

Study Objectives 5 and 6: Barriers to Notification, Reasons for Misreporting, and Potential Solutions

To derive actionable recommendations to improve notification and surveillance quality at the state, LGA, and facility levels, we analyzed each survey separately. We then triangulated the results of all the data sets (eg, health care worker survey, registers, and FGDs) to assess discrepancies between self-reports and facility-level reports. Syntax was written for analyses using the SPSS IBM version 25. Health care worker interview data were appended to the corresponding facility visit report. The magnitude and variance of the responses were examined to identify central tendencies, and outliers were considered for further exploration. Percentages, means, medians, and interquartile ranges were generated to describe survey samples. The chi-square test or Fisher exact test was used to compare categorical variables; *t* tests were used to compare mean values between groups with parametric continuous variables (eg, trust scores and stigma scores). $P < .05$ was considered statistically significant for all statistical tests, and 95% confidence intervals were generated for all point estimates. Odds ratios (ORs) were generated to explore associations. Scales were assessed for reliability with Cronbach alpha, and items with poor performance were deleted.

The focus group transcripts with the local government TB and leprosy supervisor on challenges with the surveillance system were analyzed using the framework method and summative content analysis by a multidisciplinary team with knowledge of Nigerian and Lagosian culture and languages [22].

Results

Pilot

The first pilot suggested that many study procedures and instruments were overly complex and insufficiently clear for data collectors and respondents. A thorough redesign was necessary. The verification of cases in DOTS facilities yielded 100% verification, suggesting low sensitivity and specificity to discrepant information. Instruments, procedures, training, SOPs, supervision, and community entries were revised internet field challenges, obliging a paper-based back-up system and multiple internet connectivity options. Instruments and data collection forms were shortened to improve data quality, reduce respondent burden, and enhance respondent assent. An unacceptable rate of participation (0%) among unengaged private facilities in the pilot obliged the addition of a small monetary incentive (US \$5).

In the second pilot study, the revised SOPs, shortened instruments, and retrained data officers were tested. Although fewer challenges were identified, the web app was updated to correct misclassifications, misspellings, and validation rules.

Questionnaires had a lower respondent burden, and the time per facility was reduced to acceptable levels.

Study Results

As the survey was implemented, participation was robust, and survey implementation proceeded per protocol. In the original population proportional to size sample, a total of 605 HFs and 349 laboratories were included. Recruitment followed the per protocol sampling plan in terms of facility type and facility level, but geographical distribution varied in laboratories because of the high frequency of substitution owing to ineligibility. Ultimately, via sampling with replacement, 608 HFs (100%) and 328 (93.9%) laboratories participated. Public unengaged facilities that were purposefully selected because of ambiguity regarding their engagement status were less likely to participate (6/23, 26% refusal); 1 in 10 unengaged private HFs approached was out of business (45/381, 11.8%) or was unwilling to participate (36/381, 9.4%). An unengaged private HF that was unwilling to participate was replaced with an HF similar in level and LGA. Refusal and proportion granting partial access to patient records (15/381, 3.9%) among unengaged HFs were lower than anticipated, suggesting that selection bias was minimized. The use of a monetary incentive in this stratum may have mitigated the risk of bias of refusal.

Among the 608 facilities enrolled in the study, 564 (92.7%) granted full access to all facility records and 44 (7.3%) granted partial access to facility records. Health care worker survey refusal was moderate (13%).

In *dormant* facilities, there were often no eligible respondents; 3 of the 5 tertiary level public institutions refused to participate in the health care worker survey, which reduced the representativeness of this specific stratum, which is prone to misreporting.

As it was not feasible to include all 2561 unengaged HFs in Lagos, our misreporting estimates are extrapolated from a sample. The sample estimates (7%- 10%) case notification in the unengaged private HFs were below the 3% notification observed. Moreover, the estimates of the frequency of TB treatment provision (32%-26%) were triple the frequency (11.4%) observed in the survey. Valid generalization from a sample to a population relies on a solid understanding of the heterogeneity of the sites and strict fidelity to sampling with replacement per protocol. Knowledge of all types of heterogeneity and variance among unengaged sites was incomplete because of the dynamism of this sector. HFs frequently changed ownership, location, name, services, and level, and although a valid sampling frame was assembled in 2016, changes occurred in the 9 months between the generation of the frame and when data were collected.

Number and Proportion of TB Cases Treated

The calculation of the total volume of TB treated in the unengaged private sector relies on 2 specific assumptions: (1) that B+ TB samples that could not be linked were treated by unengaged private providers and (2) that pretreatment loss to follow-up occurs at the average rate. The use of these parameters allowed us to overcome potential bias concerning record decay at the facilities, but it also introduced a degree of uncertainty into the estimate. Partial access to patient records in 3% (3/98) of DOTs facilities may have also affected our estimate of TB case undernotification. However, as this affected only 3 sites, we decided to adjust for it.

Table 2 summarizes the performance of the Lagos hybrid methodology as compared with the WHO inventory methods.

Table 2. Findings and lessons derived from the implementation of the Lagos hybrid inventory methodology.

WHO ^a inventory methods	Lagos hybrid inventory methodology	Findings and lessons
Electronic case-based national TB ^b surveillance system	Retrospective digitalization of paper-based surveillance to create electronic case-based surveillance system	Quality-assured digitalization requires robust database design and data management. Relational databases are required where each element has a unique ID, including each facility.
Electronic case-based database with records for patients with TB	Digitalization of facility and laboratory records with no notification behavior	The inability to distinguish between sites that diagnosed no patients with TB and those who kept no records of patients with TB treated is a weakness of retrospective designs. The decay of records was an issue. Some patients' paper records were physically damaged in 5 out of 701 sites through poor warehousing or force majeure, impacting the quality of both verification and case finding. As 4 of these sites were engaged, this may have led to the underverification of notified cases, leading to an inflated estimate of overreporting. This can be mitigated by triangulation and minimizing the period between reference year (2015) and data collection (2017).
Electronic case-based database with records for patients with TB	High-specificity on-site verification in engaged HFs ^c with the likelihood of notification behavior	On-site human verification of notified cases using a deterministic 7 variable matching algorithm in engaged DOTS ^d centers with a high likelihood of notification proved a viable alternative to 100% field-based data entry. It was acceptable to providers because TB registers did not need to be removed from the premises, on-site data entry was minimized, and the total respondent burden was reduced.
Census of all HF (retrospective) or random sampling of all HF (prospective)	Hybrid mix of census and stratified probability proportional to size sampling methods among HF strata	The census of laboratories paired with the sampling of private HFs was robust because it allowed for triangulation of self-report, as well as extrapolation. However, the frequency of TB service provision (32%-36%) in the unengaged private HFs was significantly overestimated, so the point estimate of misreporting of clinical TB in the unengaged private sector has large uncertainty bounds. This large sampling error would have compromised the estimate of misreporting of B+ ^e TB also, were it not for the ability to rely upon the TB diagnoses from a census of all laboratories.
Standardized TB case definitions	Broad case definition for TB (all other forms); standardized case definition for B+ TB	The use of a broader definition of clinical TB permitted a more complete accounting of TB treatment coverage that includes overdiagnosis and overtreatment. Documentation of the frequency of diagnosis of TB without bacteriological testing in the unengaged sector is important information for public health stakeholders. The clinical diagnosis was a very small proportion of TB treatment found (11%).
Presence of unique identifiers for record linkage for deterministic record linkage	Use of multi-variable probabilistic record linkage algorithms with sensitivity analysis, combined with an independent review	Probabilistic record linkage with WHO-recommended software underestimated notification due to low sensitivity for name reversal. Use of Excel add-in (Fuzzy Lookup) allowed for matching across and between variables but is not syntax driven and provides no audit trail. The ability to account for name order reversal is important to avoid bias in misreporting estimates.
National in scope or sampled "self-contained" geographical areas	Subnational in scope, but using buffer zone sensitivity analysis to permit estimation and adjustment for cross-border health care seeking	Buffer zone sensitivity analysis is straightforward to conduct and permits robust subnational and urban inventory studies.
No recommendation to study misreport between levels of the TB surveillance system	Comparisons of aggregated data to identify misreporting between administrative levels of the notification system	The addition of within-surveillance system misreporting enhanced the value of the study for local stakeholders.
No recommendation to include the study of the underlying rationales and solutions to suboptimal notification behaviors in addition to its quantification	In-depth interviews with health care providers and focus group discussions with surveillance offers were undertaken	Focus group discussions and in-depth interviews provided context for quantitative findings of misreporting, critical to engage stakeholders in identifying roots of and solutions to the notification problem.

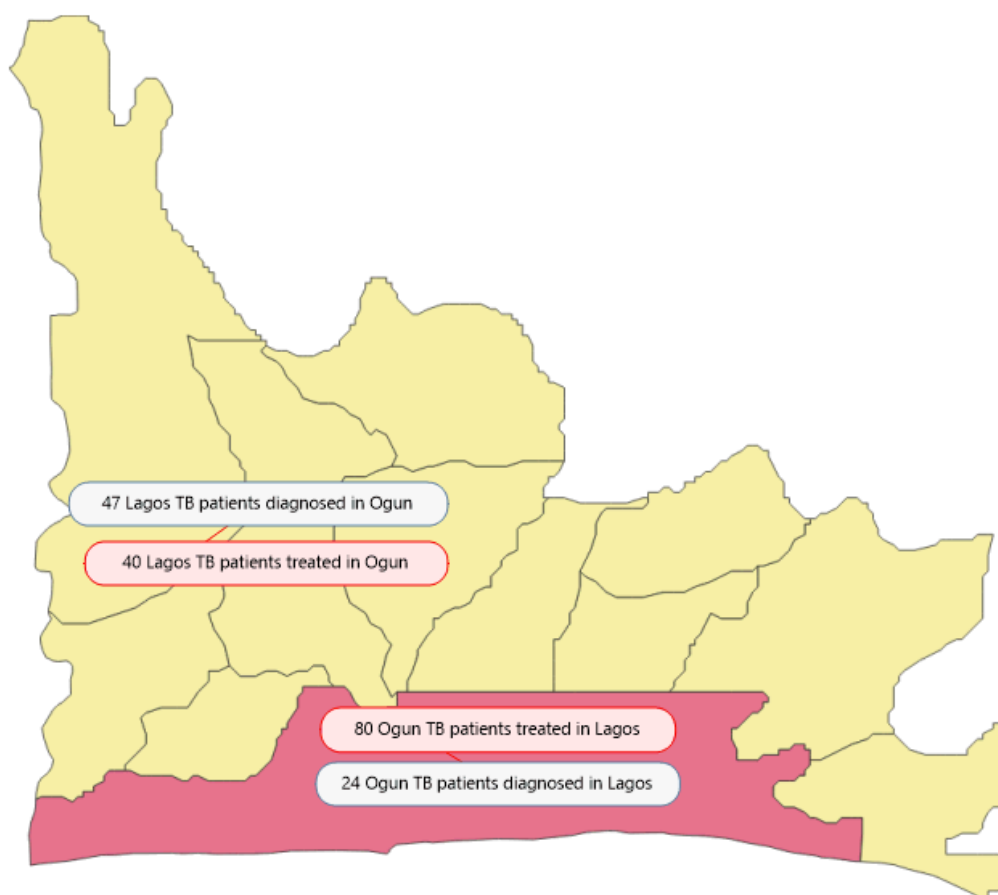
^aWHO: World Health Organization.^bTB: tuberculosis.^cHF: health facility.^dDOTS: Directly Observed Treatment Short-Course.^eB+: bacteriologically confirmed tuberculosis

Adjusting for Cross-Border Diagnosis and Treatment: Buffer Zone Sensitivity Analysis

In the Ogun laboratory data, 47 Lagos residents were diagnosed, and in the sampled HFs, 80 were treated (see Figure 7). Residents with Ogun addresses in the Lagos TB treatment

registers were also identified. The results suggested a low level of cross-border care seeking, with slightly more Lagos residents being diagnosed in Ogun and slightly more Ogun residents seeking treatment in Lagos. The absolute numbers of cross-border TB treatments were deemed so small, so as to have minimal influence on the analysis.

Figure 7. Schematic of cross-border TB diagnosis and treatment. TB: tuberculosis.



Comparison of Characteristics of Patients With TB Notified and Unnotified

Patients with TB who were never notified to the program had similar age and sex distribution to those who were notified. However, they differed in other aspects. Individuals with unknown treatment status were more likely to have had their TB diagnosed via the GeneXpert test than with smear microscopy (19% vs 12%; OR 0.60, 95% CI 0.52-0.71). This suggests that persons living with HIV may be less likely to start treatment for TB and/or that pretreatment loss to follow-up was higher in sites equipped with GeneXpert mycobacterium tuberculosis and rifampicin resistance machines. Treatment success was modest in all sectors ($R=37.9\%-72.0\%$) and did not align with treatment success reported by the state (6665/8770, 75.99%).

TB Surveillance Coverage and Quality

We obtained aggregated LGA level monthly TB totals of inpatients and outpatients from the IDSR office. We sought IDSR facility total monthly TB notification data, but we were

unable to obtain them from LGA officers during the course of the study or during their supervision meeting, which limited the granularity of the analyses. As the inpatient and outpatient department values contain duplicates and therefore cannot be directly compared with LGA, notification comparisons were of limited value. As the IDSR and STBLCP data were aggregated, it was not possible to reconcile them fully. Descriptive analyses are included in Multimedia Appendix 4.

Barriers to Notification, Reasons for Misreporting, and Potential Solutions

Notification was rarely a top priority for TB care providers. When asked what might be done to improve the TB notification system, some instead advocated for regular drug supplies; GeneXpert machines; laboratory personnel; health education materials for patients; more clinical and laboratory staff; and greater efforts at community mobilization, early detection, and financial support for patients with TB. Among those who identified reporting as uncomfortable, the reasons for discomfort with disease notification included practical, logical, strategic, and economic concerns. The most common reason given by

both TB providers and nonproviders was the lack of access to notification forms and registers. A quarter of TB providers (26%) were doubtful that the volume of patients with TB they treated in their facility merited mastery of the extensive TB notification forms and procedures.

Discussion

The WHO strongly recommends that high-burden countries undertake measurement of underreporting while simultaneously setting the methodological bar so high, so as to preclude most TB programs in high-burden settings from attempting one [4]. This study attempted to design an inventory study that struck a balance among competing demands for rigor, ethics, efficiency, and feasibility for high-burden settings with a lack of standardized case definitions; electronic case-based surveillance infrastructure; or closed catchment areas. Combining elements of WHO-endorsed study designs and improvising techniques to address validity threats allowed this study to capture the level of misreporting in a high-need setting. The use of the hybrid method led to robust participation.

Solid participation of HFs may be a consequence of minimizing respondent burden, iterative piloting, and incentives for unengaged private HFs. Incentives for health care workers who work in high refusal strata (eg, tertiary care public HF) should also be considered.

Although the methodology proved robust and responsive to stakeholder needs, multiple methodological trade-offs should be borne in mind by researchers undertaking the hybrid method. First, the inability to distinguish between HF that diagnosed no patients with TB and those who did not have records of patients with TB treated was a shortcoming of the chosen retrospective design. This flaw was offset by the census of all laboratories and could be used as the basis for B+ TB caseloads. However, for clinically diagnosed cases, zero reports could not be verified. Relatedly, the retrospective design was susceptible to ascertainment bias. It is crucial to collect data as immediately after the reference period as possible, as the physical decay of paper records is a threat. The existence of written records for consultation varied widely and is likely correlated with notification behavior and nonrandom bias that lab data triangulation overcame.

Although sampling assumptions of underreporting (90%-93%) proved an underestimate (97%), the frequency of TB treatment in the unengaged private HFs was overestimated, so the sample size was adequate. The inclusion criteria for private unengaged HFs did not include providers such as patient medicine vendors, corporations offering occupational health services, or mobile clinics. These providers were not thought to diagnose or treat TB in 2015. However, community screening plays a very

important role in TB diagnosis in many urban areas, including Lagos, in 2020, and exclusion of these entities should be considered.

One of the strengths of this methodology over the WHO-recommended approaches is that it frames TB case notification as an institutional behavior influenced by sector-specific motivations, policies, politics, infrastructure, and challenges. Using HF strata as sampling units leverages social scientific research and stakeholders' knowledge of notification attitudes and practices. Our methodological use of prior knowledge of notification behaviors reflects a growing call to use Bayesian approaches in inventory studies [28]. In contrast, the WHO inventory study methods draw upon wildlife ecology, assuming random movement of cases and making no attempt to gather data on motives, interests, or abilities of persons working in HFs. One WHO method (capture-recapture) even assumes that TB cases have an equal probability of notification (homogeneity capture), a stance that belies the different stakes providers have to engage in public health surveillance. Framing TB notification as a necessary but labor-intensive, poorly incentivized choice allows researchers to shift the measurement focus from mere counting cases to understanding the knowledge, behavior, and needs of gatekeepers. Our results confirm that many TB providers considered notification an additional burden with unclear utility.

Although our study methods are likely to be applicable to other urban areas, comprehensive, iterative piloting is needed for replication. The probabilistic record linkage possible in Lagos is unlikely to be appropriate for more ethnically homogeneous areas with lower levels of diversity in first and last names.

We show that in settings with weak surveillance systems, hybrid approaches may even be preferable to designs that assume notification is random and oblige significant intervention at HFs because the Lagos hybrid method does not distort *natural* R+R behaviors, is acceptable, and is feasible in settings where inventory studies are most needed.

Conclusions

This study adapted the WHO undernotification estimation methodology and procedures to take into account known surveillance system strengths, weaknesses, seasonality, and (crucially) local stakeholders' needs for actionable information on the causes of TB misreporting. Participatory and locally tailored methods met policy makers' needs for causal information and feasible recommendations. Subnational inventory studies in high-burden settings with weak paper-based surveillance, dynamic private sectors, and open borders are both feasible and informative. The WHO preconditions should not discourage high-burden countries from measuring misreporting as part of the process of solving notification challenges.

Acknowledgments

The authors would like to thank the following persons for constructive criticism of the aims, methodology, process and/or analysis: Eniola Ashimiu Erinoshio, Director of Disease Control, Lagos State Ministry of Health; Mrs Ojomo, Lagos State Laboratory Quality Control Officer; Omotayo Bello, Director of Health HEFAMAA; Dr Abdulsalam, Lagos State epidemiologist; Susan Mitchell, Abt Associates-SHOPS Project; Charalampos Sismanidis, WHO; Nico Nagelkerke, Koninklijke Centrale Vereniging

tot bestrijding der Tuberculose (KNCV) emeritus; Hassan Abiodun, African Research on Family Health; Susan van den Hof, Rijksinstituut voor Volksgezondheid en Milieu (Dutch: National Institute for Health and Environment); Suzanne Verver, Nederlandse Organisatie voor Wetenschappelijk Onderzoek; Sandra Kik, independent consultant; Nico Kalisvaart, KNCV; Sara Massaut, Wemos; D'Arcy Richardson, independent consultant; Jens Levy, KNCV; Susan Bergson, independent consultant; Adebukola Agnes Adegbola; KNCV; Kingsley Umahoin, KNCV; Christina Mergenthaler, Royal Tropical Institute; Victor Adepoju; KNCV; Debby Nongo, United States Agency for International Development (USAID); Ellen-Jane Burgrust, KNCV; Johan Verhoef, KNCV Emeritus; and Mukadi YaDiul, USAID. This study was made possible by the generous support of the American people through the USAID. The Global Health Bureau, Office of Health, Infectious Disease and Nutrition, USAID financially supports this report through Challenge TB under the terms of agreement No AID-OAA-A-14-00029. The funder played no role in the design, analysis, or reporting of this study.

Authors' Contributions

EMHM and OAA co-designed the study, coanalyzed data, and cowrote the manuscript. AHQ and SBO designed the information systems interface and relational data systems and conducted quality assurance. HAR and OI provided input to the methodology and study design. CO and NC provided input to the study protocol and refined SOPs following piloting. MG informed the theoretical framework. All authors have approved this manuscript for submission and publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sampling frame deduplication methods and outcomes.

[[DOCX File, 143 KB](#) - [publichealth_v7i3e22352_app1.docx](#)]

Multimedia Appendix 2

Sampling plan.

[[DOCX File, 266 KB](#) - [publichealth_v7i3e22352_app2.docx](#)]

Multimedia Appendix 3

Standard operating procedures.

[[PDF File \(Adobe PDF File\), 392 KB](#) - [publichealth_v7i3e22352_app3.pdf](#)]

Multimedia Appendix 4

Descriptive analysis of TB surveillance in Lagos, 2015.

[[DOCX File, 23 KB](#) - [publichealth_v7i3e22352_app4.docx](#)]

References

1. Ansumana R, Keitell S, Roberts GM, Ntoumi F, Petersen E, Ippolito G, et al. Impact of infectious disease epidemics on tuberculosis diagnostic, management, and prevention services: experiences and lessons from the 2014–2015 Ebola virus disease outbreak in West Africa. *International Journal of Infectious Diseases* 2017 Mar;56:101-104. [doi: [10.1016/j.ijid.2016.10.010](#)]
2. Marston BJ, Dokubo EK, van Steelandt A, Martel L, Williams D, Hersey S, et al. Ebola Response Impact on Public Health Programs, West Africa, 2014–2017. *Emerg. Infect. Dis* 2017 Dec;23(13). [doi: [10.3201/eid2313.170727](#)]
3. Bell BP, Damon IK, Jernigan DB, Kenyon TA, Nichol ST, O'Connor JP, et al. Overview, Control Strategies, and Lessons Learned in the CDC Response to the 2014–2016 Ebola Epidemic. *MMWR Suppl* 2016 Jul 08;65(3):4-11. [doi: [10.15585/mmwr.su6503a2](#)] [Medline: [27389903](#)]
4. World Health Organization. *Assessing Tuberculosis Under-Reporting Through Inventory Studies*. Switzerland: World Health Organization; 2012.
5. Anyaehie U, Nwakoby B, Chikwendu C, Dim C, Uguru N, Oluka C, et al. Constraints, challenges and prospects of public-private partnership in health-care delivery in a developing economy. *Ann Med Health Sci Res* 2014 Jan;4(1):61-66 [FREE Full text] [doi: [10.4103/2141-9248.126615](#)] [Medline: [24669333](#)]
6. Omoleke SA, Ajibola O, Ajiboye JO, Raji RO. Quagmire of epidemic disease outbreaks reporting in Nigeria. *BMJ Glob Health* 2018 Feb 15;3(1):e000659. [doi: [10.1136/bmjgh-2017-000659](#)]
7. Abubakar A, Idris S, Nguku P, Sabitu K, Sambo M. Assessment of integrated disease surveillance and response strategy implementation in selected Local Government Areas of Kaduna state. *Ann Nigerian Med* 2013;7(1):14. [doi: [10.4103/0331-3131.119981](#)]

8. Nnebue C, Onwasigwe C, Onyeonoro U, Adogu PU. Awareness and knowledge of disease surveillance and notification by health-care workers and availability of facility records in Anambra state, Nigeria. *Niger Med J* 2012;53(4):220. [doi: [10.4103/0300-1652.107557](https://doi.org/10.4103/0300-1652.107557)]
9. Iwu A, Diwe K, Duru C, Uwakwe K. Assessment of disease reporting among health care workers in a South Eastern State, Nigeria. *Int J Community Med Public Health* 2016;2766-2774. [doi: [10.18203/2394-6040.ijcmph20163359](https://doi.org/10.18203/2394-6040.ijcmph20163359)]
10. Aruna O, Nsofor I, Oyediran K. Tuberculosis in Nigeria Rapid Assessment of Infectious Disease Surveillance and Reporting. Chapel Hill, USA. www.measureevaluation.org. 2018. URL: <https://www.measureevaluation.org/resources/publications/tr-18-292>
11. World Health Organization. Global Tuberculosis Report 2015. Global Tuberculosis Report 2015. Geneva, Switzerland: World Health Organization; 2015 Mar 24. URL: http://www.who.int/tb/publications/global_report/en/ [accessed 2015-03-24]
12. Nigeria Private Health Sector Assessment. SHOPS Project. 2012. URL: https://www.shopsplusproject.org/sites/default/files/resources/Nigeria%20Private%20Health%20Sector%20Assessment%2009_10_2012.pdf [accessed 2021-02-08]
13. Hotchkiss DR, Eisele TP, Djibuti M, Silvestre EA, Rukhadze N. Health system barriers to strengthening vaccine-preventable disease surveillance and response in the context of decentralization: evidence from Georgia. *BMC Public Health* 2006 Jul 05;6(1):175 [FREE Full text] [doi: [10.1186/1471-2458-6-175](https://doi.org/10.1186/1471-2458-6-175)] [Medline: [16822316](https://pubmed.ncbi.nlm.nih.gov/16822316/)]
14. National Agency for the Control of AIDS. Federal Republic of Nigeria: GLOBAL AIDS RESPONSE Country Progress Report, Nigeria GARPR 2014. Abuja, Nigeria: Federal Republic of Nigeria; 2014 Jan 01. URL: http://www.unaids.org/sites/default/files/country/documents/NGA_narrative_report_2014.pdf
15. Gulma K, Abubakar M, Mohammed B, Musa B, Nata'ala S, Yusuf M. Stock status assessment of health commodities used in the management of HIV/AIDS and tuberculosis in kano state, Nigeria. *J Pharma Res* 2015;14(2):33-41 [FREE Full text] [doi: [10.18579/jpcrk/2015/14/2/79075](https://doi.org/10.18579/jpcrk/2015/14/2/79075)]
16. Daniel OJ, Adedeji Adejumo O, Abdur-Razzaq HA, Ngozi Adejumo E, Salako AA. Public-private mix for TB and TB-HIV care in Lagos, Nigeria. *Int J Tuberc Lung Dis* 2013 Sep 01;17(9):1195-1198. [doi: [10.5588/ijtld.12.0759](https://doi.org/10.5588/ijtld.12.0759)] [Medline: [23928168](https://pubmed.ncbi.nlm.nih.gov/23928168/)]
17. Adejumo O, Daniel O, Abdur-Razzaq H, Shogbamimu Y, Femi-Adebayo T, Adepoju V, et al. Trend of tuberculosis case notification and treatment outcome in Lagos State, Nigeria: a 5-year retrospective study. *Trans R Soc Trop Med Hyg* 2017 Jul 01;111(7):300-307. [doi: [10.1093/trstmh/trx060](https://doi.org/10.1093/trstmh/trx060)] [Medline: [29165702](https://pubmed.ncbi.nlm.nih.gov/29165702/)]
18. Adejumo O, Daniel O, Adejumo E, Oluwale E, Olumuyiwa O. Treatment Outcomes of Tuberculosis Patients Managed at the Public and Private Dots Facilities in Lagos Nigeria. *IJTDH* 2015 Jan 10;10(2):1-9. [doi: [10.9734/ijtldh/2015/19933](https://doi.org/10.9734/ijtldh/2015/19933)]
19. Adejumo OA, Daniel OJ, Gidado M, Otesanya AF, Adejumo EN, O. Jaiyesimi E, et al. Are Tuberculosis Patients Managed According to the National Guidelines in Lagos State Nigeria? *IJCM* 2016;07(01):16-24. [doi: [10.4236/ijcm.2016.71003](https://doi.org/10.4236/ijcm.2016.71003)]
20. Department of Public Health. Guidelines for Community Tuberculosis Care in Nigeria National Tuberculosis and Leprosy. Abuja, Nigeria: Federal Republic of Nigeria; 2012.
21. Johnson D, Woodman B, Baruwa S, Toriola M, Chatterji M, Kinnan C. A Census of Private Health Facilities in Six States of Nigeria. Bethesda, Maryland USA: Strengthening Health Outcomes through the Private Sector Project, Abt Associates; 2014.
22. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol* 2013 Sep 18;13:117 [FREE Full text] [doi: [10.1186/1471-2288-13-117](https://doi.org/10.1186/1471-2288-13-117)] [Medline: [24047204](https://pubmed.ncbi.nlm.nih.gov/24047204/)]
23. MacPherson P, Houben RM, Glynn JR, Corbett EL, Kranzer K. Pre-treatment loss to follow-up in tuberculosis patients in low- and lower-middle-income countries and high-burden countries: a systematic review and meta-analysis. *Bull. World Health Organ* 2013 Nov 22;92(2):126-138. [doi: [10.2471/blt.13.124800](https://doi.org/10.2471/blt.13.124800)]
24. Beyeler N, Liu J, Sieverding M. A systematic review of the role of proprietary and patent medicine vendors in healthcare provision in Nigeria. *PLoS One* 2015 Jan 28;10(1):e0117165 [FREE Full text] [doi: [10.1371/journal.pone.0117165](https://doi.org/10.1371/journal.pone.0117165)] [Medline: [25629900](https://pubmed.ncbi.nlm.nih.gov/25629900/)]
25. Micallef L, Rodgers P. eulerAPE: drawing area-proportional 3-Venn diagrams using ellipses. *PLoS One* 2014 Jul 17;9(7):e101717 [FREE Full text] [doi: [10.1371/journal.pone.0101717](https://doi.org/10.1371/journal.pone.0101717)] [Medline: [25032825](https://pubmed.ncbi.nlm.nih.gov/25032825/)]
26. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* 2015 May 22;16(1):169 [FREE Full text] [doi: [10.1186/s12859-015-0611-3](https://doi.org/10.1186/s12859-015-0611-3)] [Medline: [25994840](https://pubmed.ncbi.nlm.nih.gov/25994840/)]
27. Bardou P, Mariette J, Escudié F, Djemiel C, Klopp C. jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics* 2014 Aug 29;15(1):293 [FREE Full text] [doi: [10.1186/1471-2105-15-293](https://doi.org/10.1186/1471-2105-15-293)] [Medline: [25176396](https://pubmed.ncbi.nlm.nih.gov/25176396/)]
28. Straetmans M, Bakker M, Alba S, Mergenthaler C, Rood E, Andersen P, et al. Completeness of tuberculosis (TB) notification: inventory studies and capture-recapture analyses, six European Union countries, 2014 to 2016. *Euro Surveill* 2020 Mar;25(12):10 [FREE Full text] [doi: [10.2807/1560-7917.ES.2020.25.12.1900568](https://doi.org/10.2807/1560-7917.ES.2020.25.12.1900568)] [Medline: [32234122](https://pubmed.ncbi.nlm.nih.gov/32234122/)]

Abbreviations

B+: bacteriologically confirmed tuberculosis

CDR: case detection rate
DOTS: Directly Observed Treatment Short-Course
FGD: focus group discussions
HEFAMAA: Health Facility Monitoring and Accreditation Agency
HF: health facility
IDSr: Integrated Disease Surveillance and Response
KNCV: Koninklijke Centrale Vereniging tot bestrijding der Tuberculose
LGA: local government area
NTP: national tuberculosis program
OR: odds ratio
PPM: public-private mix
R+R: recording and reporting
SOP: standard operating procedure
STBLCP: State Tuberculosis and Leprosy Control Program
TB: tuberculosis
USAID: United States Agency for International Development
WHO: World Health Organization

Edited by M Focsa; submitted 09.07.20; peer-reviewed by T Turk, A Kibballi Madhukeshwar; comments to author 20.08.20; revised version received 03.12.20; accepted 08.12.20; published 15.03.21.

Please cite as:

Mitchell EMH, Adejumo OA, Abdur-Razzaq H, Ogbudebe C, Chukwueme N, Olorunju SB, Gidado M
Hybrid Approach to Estimation of Underreporting of Tuberculosis Case Notification in High-Burden Settings With Weak Surveillance
Infrastructure: Design and Implementation of an Inventory Study
JMIR Public Health Surveill 2021;7(3):e22352
URL: <https://publichealth.jmir.org/2021/3/e22352>
doi:[10.2196/22352](https://doi.org/10.2196/22352)
PMID:[33720030](https://pubmed.ncbi.nlm.nih.gov/33720030/)

©Ellen M H Mitchell, Olusola Adedeji Adejumo, Hussein Abdur-Razzaq, Chidubem Ogbudebe, Nkem Chukwueme, Samson Bamidele Olorunju, Mustapha Gidado. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 15.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Rapid Surveillance Report

The Annual American Men's Internet Survey of Behaviors of Men Who Have Sex With Men in the United States: Key Indicators Report 2018

Sarah Wiatrek¹, MPH; Maria Zlotorzynska¹, MPH, PhD; Ramona Rai¹, MPH; Patrick Sullivan¹, DVM, PhD; Travis Sanchez¹, MPH, DVM

Rollins School of Public Health, Emory University, Atlanta, GA, United States

Corresponding Author:

Sarah Wiatrek, MPH
Rollins School of Public Health
Emory University
1518 Clifton Rd
Atlanta, GA
United States
Phone: 1 470 603 5783
Email: swiatre@emory.edu

Abstract

The American Men's Internet Survey (AMIS) is an annual web-based behavioral survey conducted in the United States of men who have sex with men (MSM). This rapid surveillance report describes the sixth cycle of data collection (September-December 2018; AMIS 2018). The key indicators were the same as those previously reported for past AMIS cycles. The AMIS methodology has not substantively changed since AMIS 2017. MSM were recruited from a variety of websites using banner advertisements and email blasts. In addition, participants from AMIS 2017 who agreed to be recontacted for future research were emailed a link to AMIS 2018. Men were eligible to participate if they were aged ≥ 15 years, resided in the United States, provided a valid US ZIP code, and reported ever having sex with a man or identified as gay or bisexual. The analysis was limited to those who reported having oral or anal sex with a male partner in the past 12 months. We examined demographic and recruitment characteristics using multivariable regression modeling ($P < .05$) stratified by the participants' self-reported HIV status. The AMIS 2018 round of data collection resulted in 10,129 completed surveys from MSM representing every US state, Puerto Rico, and Guam. Most participants were non-Hispanic White, aged between 15 and 24 years, living in urban areas in the southern United States, and recruited from general social networking websites. The plurality (4230/10,129, 41.76%) of participants was in the youngest age group, 15-24 years, followed by the 40 years and older age group (3088/10,129, 30.49%). The self-reported HIV prevalence was 6.08% (616/10,129). Compared with HIV-negative or unknown status participants, HIV-positive participants were more likely to have had anal sex without a condom with a male partner in the past 12 months (adjusted odds ratio [aOR] 2.02, 95% CI 1.63-2.50) and more likely to have had anal sex without a condom with a serodiscordant or an unknown status partner (aOR 3.90, 95% CI 3.27-4.66). The reported use of marijuana in the past 12 months was higher among HIV-positive participants than among HIV-negative or unknown status participants (aOR 1.39, 95% CI 1.15-1.68). The reported use of methamphetamines and other illicit substances in the past 12 months was higher among HIV-positive participants than among HIV-negative or unknown status participants (aOR 3.42, 95% CI 2.41-4.87 and aOR 1.90, 95% CI 1.56-2.32, respectively). Most HIV-negative or unknown status participants (6838/9513, 71.88%) reported ever taking an HIV test previously, and 52.51% (4995/9513) of the participants reported undergoing HIV testing in the past 12 months. HIV-positive participants were more likely to report testing and diagnosis of sexually transmitted infections than HIV-negative or unknown status participants (aOR 3.50, 95% CI 2.89-4.24 and aOR 2.61, 95% CI 2.10-3.25, respectively).

(JMIR Public Health Surveill 2021;7(3):e21812) doi:[10.2196/21812](https://doi.org/10.2196/21812)

KEYWORDS

HIV; internet; men who have sex with men; sexually transmitted infections; surveillance; survey

Introduction

The American Men's Internet Survey (AMIS) is an annual web-based behavioral survey conducted in the United States of men who have sex with men (MSM). AMIS was developed to produce timely data from large-scale monitoring of behavior trends among MSM recruited on the web. It was designed to complement the Centers for Disease Control and Prevention's National HIV Behavioral Surveillance (NHBS) system, which collects data on MSM in major US cities every 3 years through venue-based recruitment [1]. An increasing number of MSM are meeting sexual partners through the internet and may have different patterns of sexual risk and HIV testing behaviors compared with MSM recruited through physical venues. AMIS is able to generate annual snapshots of behaviors in a large sample of internet-using MSM with broad geographic diversity as a supplement to venue-based studies, such as the NHBS system. We were also able to collect, update, and share state-level data with public health authorities to inform issues of local relevance by using AMIS.

The methods and past AMIS cycle data (AMIS 2013, AMIS 2014, AMIS 2015, AMIS 2016, and AMIS 2017) have been previously published [2-6].

This supplemental report has updated the existing information with the data collected in AMIS 2018. The methods in AMIS 2018 have not changed from the previously published methods, unless otherwise noted. An in-depth analysis, discussion, and limitations of multiyear trends for indicators reported herein have been published and include data for the first 4 cycles of AMIS (AMIS 2013 to AMIS 2016) [7].

Methods

Recruitment and Enrollment

Similar to the previous year's recruitment process, AMIS participants were recruited through convenience sampling from a variety of websites using banner advertisements or email blasts to members of the website (hereafter referred to generically as *ads*). For AMIS 2018, data were collected from September 2018 to December 2018. The survey was not incentivized. Men who clicked on the ads were taken directly to the survey website hosted on a secure server administered by SurveyGizmo. Recruitment was also done by emailing participants from the previous cycle of AMIS (AMIS 2017) who consented to be recontacted for future studies. To be eligible for the survey, participants had to be aged ≥ 15 years, be cisgender male, reside in the United States, and report that they either had oral or anal sex with a male partner at least once in the past or identify as gay or bisexual (hereafter referred to as MSM). Persons who were aged < 15 years or refused to provide their age were not asked any other screening questions. MSM who met the eligibility criteria and consented to participate in the study started the web-based survey immediately. The full questionnaire for AMIS 2018 is presented in [Multimedia Appendix 1](#).

Several data cleaning steps were performed on the raw data set of eligible responses to obtain the final analysis data set, in the

same manner as in previous AMIS cycles [2-6]. Briefly, these steps were as follows: deduplication; limiting to surveys deemed successful, that is, observations with no missing values for the first question of at least two consecutive sections; limiting to participants who reported having oral or anal sex with a male partner in the past 12 months; and ZIP code validation. These steps are further described in detail.

First, to deduplicate survey responses, demographic data for near-complete ($> 70\%$) survey responses with nonunique internet protocol addresses were compared, and responses that showed a 100% match for age, race, ethnicity, ever having sex with a woman, and email address were considered to be duplicate responses. Only the observation with the highest survey completion was retained. The data set was, then, limited to those surveys that were deemed successful. Finally, the data set was restricted to include participants who reported having oral or anal sex in the past 12 months and who provided a valid US ZIP code. ZIP codes were validated in the same manner as done in AMIS 2017 [6]. Valid US ZIP codes were those that could be matched to the ZIP code of county crosswalk files created by the US Department of Housing and Urban Development [8]. Any ZIP codes that could not be matched to this list were then hand validated by checking against the ZIP code locator tool on the US Postal Service website [9]. ZIP codes that could not be found were classified as invalid.

Human Subjects Protections

The study was conducted in compliance with federal regulations governing the protection of human subjects and was reviewed and approved by our institution's human subjects research review board. No incentive was provided to the participants. Data sets for analyses were stored on secure data servers with access only granted to study staff. The study data are protected under a federal certificate of confidentiality that prevents legal action to force data release.

Measures and Analyses

For the AMIS 2018 analyses, participants were categorized as either AMIS 2017 participants who took the survey again or new participants from the website or app based on the target audience and purpose: gay social networking ($n=2$), gay general interest ($n=1$), general social networking ($n=4$), and geospatial social networking ($n=2$). Recruitment outcomes and demographic characteristics for the AMIS 2018 participants are presented in the first two tables, and thereafter, they are recategorized to how they were originally recruited in AMIS 2017. We did not provide the names of the websites or apps to preserve operator and client privacy, particularly when a category has only one operator. Participants whose data were eligible, unduplicated, and successful and who provided consent; reported having male-male sex in the past 12 months; and provided a valid US ZIP code were included in analyses of participant characteristics and behavior.

To facilitate comparisons, the key indicators and analytic approach used in AMIS were designed to mirror those used by the NHBS system [10]. Population density was defined in the same manner as defined in AMIS 2017 and was based on the National Center for Health Statistics Rural-Urban classification

scheme for counties [11]. The self-reported HIV status was categorized as HIV-positive, HIV-negative, or unknown, consistent with surveillance reports produced by the NHBS system [10]. In total, 3 substance use behaviors in the past 12 months were assessed: use of nonprescribed marijuana, use of methamphetamines, and use of any illicit drug other than marijuana or methamphetamines. All other indicators assessed remained unchanged from AMIS 2017 [6].

The analysis methods for AMIS 2018 did not substantively differ from those previously published but are repeated in this report for clarity. Overall, chi-square tests were used to identify whether participant characteristics differed significantly among recruitment sources. Multivariable logistic regression modeling was used to determine significant differences in behaviors based on the self-reported HIV status while controlling for race or ethnicity, age group, NHBS city residency, and type of recruitment website. The metropolitan statistical areas included in the NHBS system in 2018 were as follows: Atlanta, Georgia; Boston, Massachusetts; Chicago, Illinois; Dallas, Texas; Denver, Colorado; Detroit, Michigan; Houston, Texas; Los Angeles, California; Memphis, Tennessee; Miami, Florida; Nassau-Suffolk, New York; New Orleans, Louisiana; New York City, New York; Newark, New Jersey; Philadelphia, Philadelphia; Portland, Oregon; San Diego, California; San Francisco, California; San Juan, Puerto Rico; Seattle, Washington; Virginia Beach-Norfolk, Virginia; and Washington, District of Columbia. HIV testing behaviors were only examined among those who did not report that they were living with HIV, and these data were presented in participant characteristics. The

multivariable logistic regression results were presented as adjusted odds ratios (aORs) with 95% CI and Wald chi-square *P* values to denote an independently significant difference in the behavior for each subgroup compared with a reference group. Statistical significance was set at $P=.05$.

Results

Recruitment Outcomes

AMIS 2018 was conducted from September 2018 to December 2018 and resulted in 91,142 persons screened for eligibility (Table 1). Of the 3713 participants who completed the AMIS 2017 survey and were emailed links to the AMIS 2018 survey, 39.94% (1483/3713) completed the screening. Almost half (42,011/91,142, 46.09%) of the participants who completed the screening process were eligible to participate. The most common reason for ineligibility was not ever having male-male sex or not identifying as gay or bisexual. Almost all (40,847/42,011, 97.23%) of the participants who were eligible consented to participate in the survey. A total of 6595 (6595/40,847, 16.15%) surveys were likely from duplicate participants. Among unduplicated surveys, 35.75% (12,246/34,252) were considered successful. Most successful surveys were from men who reported having sex with another male in the past 12 months (10,232 /12,246, 83.55%). Almost all of these surveys (10,129/10,232, 98.99%) provided a valid US ZIP code. Overall, the completion rate was 11.1%, with an analytical sample consisting of 10,129 surveys from 91,142 screened participants. The median survey completion time, including eligibility screening, was 20.7 minutes (IQR=16.5-27.4).

Table 1. Recruitment outcomes for the American Men's Internet Survey, United States, 2018.

Recruitment outcomes	Total	Gay social networking (n=2) ^a	General gay interest (n=1) ^a	General social networking (n=4) ^a	Geospatial social networking (n=2) ^a	AMIS ^b 2017 participants
Screened ^c , N	91,142	1288	197	80,768	7406	1483
Ineligible^d, n (%)	49,131 (53.91)	279 (21.66)	140 (71.07)	46,689 (57.81)	1857 (25.07)	166 (11.19)
Not >15 years of age ^e	22,659 (46.12)	115 (41.22)	13 (9.29)	21,507 (46.06)	951 (51.21)	73 (43.98)
Not male ^e	38,832 (79.04)	226 (81)	48 (34.29)	36,830 (78.88)	1590 (85.62)	138 (83.13)
Not MSM ^{e,f}	48,398 (98.51)	257 (92.11)	53 (37.86)	46,233 (99.02)	1697 (91.38)	158 (95.18)
Nonresident ^e	36,714 (74.73)	220 (78.85)	131 (93.57)	34,637 (74.19)	1597 (86)	129 (77.71)
Eligible ^c , n (%)	42,011 (46.09)	1009 (78.34)	57 (28.93)	34,079 (42.19)	5549 (74.93)	1317 (88.81)
Consented ^g , n (%)	40,847 (97.23)	964 (95.54)	55 (96.49)	33,087 (97.09)	5429 (97.84)	1312 (99.62)
Unduplicated ^h , n (%)	34,252 (83.85)	859 (89.11)	51 (92.73)	27,527 (83.20)	4639 (85.45)	1176 (89.63)
Success ⁱ , n (%)	12,246 (35.75)	489 (56.93)	39 (76.47)	8150 (29.61)	2537 (54.69)	1031 (87.67)
MSM in the past 12 months ^j , n (%)	10,232 (83.55)	434 (88.75)	32 (82.05)	6424 (78.82)	2395 (94.40)	947 (91.85)
Valid ZIP code ^k , n (%)	10,129 (98.99)	430 (99.08)	32 (100)	6351 (98.86)	2375 (99.16)	941 (99.37)

^aRefers to the number of websites or apps in this category.

^bAMIS: American Men's Internet Survey.

^cProportion of participants who started the screening questionnaire.

^dProportion of total participants screened. Participants who did not complete the screening questionnaire were considered ineligible.

^eProportion of ineligible participants, including those who did not respond to the question.

^fMSM: men who have sex with men or identify as gay or bisexual.

^gProportion of eligible participants.

^hProportion of participants who consented. Deduplication removes participants who were marked as duplicates using the internet protocol address and demographic data matching.

ⁱProportion of unduplicated participants. Success removes participants who do not pass the test for completeness.

^jProportion of successes.

^kProportion of men who had sex with men in the past 12 months. Valid US ZIP codes were those that could be matched to the ZIP code for county crosswalk files created by the US Department of Housing and Urban Development. Any ZIP codes that could not be matched to this list were then hand validated by checking against the ZIP code locator tool on the US Postal Service website. ZIP codes that could not be found were classified as invalid.

Participant Characteristics

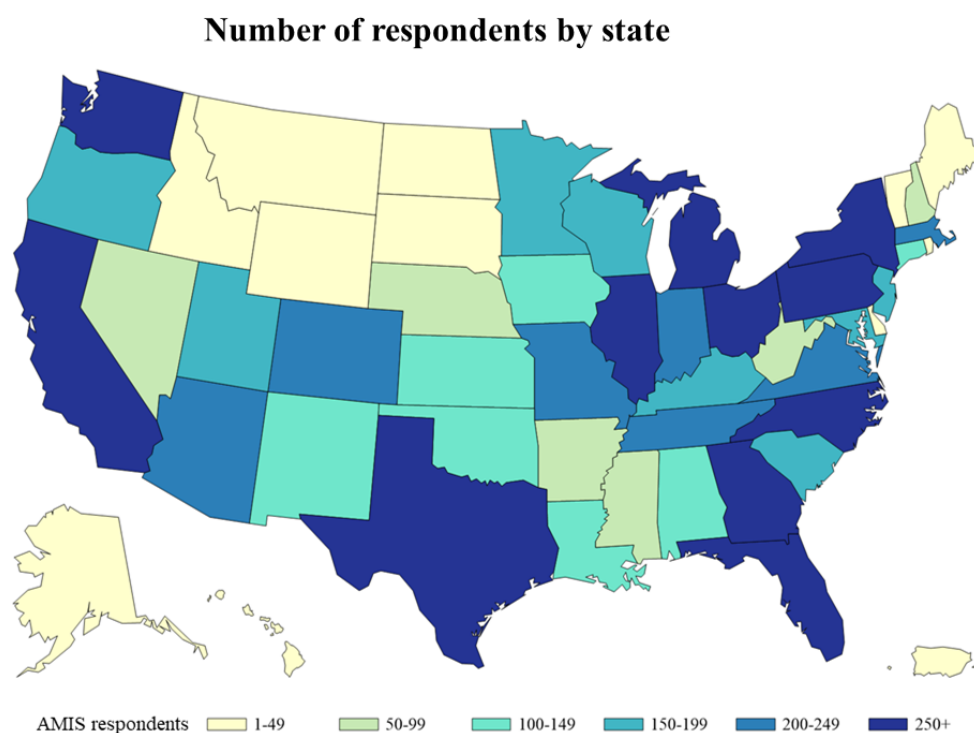
In total, 69.22% (7011/10,129) of the participants included in this report were non-Hispanic White and 41.76% (4230/10,129) were aged 15-24 years; the most common region of residence was the south, followed by the west (Table 2). Participants were recruited from all US states, and there were at least 100 participants each from 29 states and the District of Columbia

(Figure 1). About one-third (3338/10,129, 32.95%) of participants resided in an NHBS city, and about the same proportion (3680/10,129, 36.33%) lived in an urban county. Overall, 6.08% (616/10,129) of participants were living with HIV, 66.39% (6725/10,129) were HIV negative, and 27.52% (2788/10,129) had an unknown HIV status. All participant characteristics differed significantly based on the recruitment source, except NHBS city resident (Table 2).

Table 2. Characteristics of men who have sex with men in the American Men's Internet Survey by recruitment type, United States, 2018.

Participant characteristics	Total	Gay social net- working (n=2) ^a	General gay in- terest (n=1) ^a	General social networking (n=3) ^a	Geospatial so- cial networking (n=2) ^a	AMIS ^b 2016 participants	P value ^c
Race or ethnicity, n (%)							<.001
Black, non-Hispanic	553 (5.46)	36 (8.4)	<5 (—) ^d	337 (5.3)	127 (5.3)	52 (5.5)	
Hispanic	1630 (16.09)	32 (7.4)	<5 (—)	1220 (19.21)	269 (11.33)	106 (11.26)	
White, non-Hispanic	7011 (69.22)	333 (77.4)	24 (75)	4128 (65)	1801 (75.83)	725 (77.05)	
Other or multiple races	749 (7.39)	16 (3.7)	<5 (—)	546 (8.6)	133 (5.6)	51 (5.4)	
Age (years), n (%)							<.001
15-24	4230 (41.76)	18 (4.19)	5 (15.63)	3919 (61.71)	153 (6.44)	135 (14.35)	
25-29	1308 (12.91)	19 (4.42)	<5 (—)	968 (15.24)	197 (8.29)	123 (13.07)	
30-39	1503 (14.84)	63 (14.65)	5 (15.63)	774 (12.19)	484 (20.38)	177 (18.81)	
≥40	3088 (30.49)	330 (76.74)	21 (65.63)	690 (10.86)	1541 (64.88)	506 (53.77)	
Region, n (%)							.02
Northeast	1632 (16.11)	89 (20.70)	7 (21.88)	1007 (15.86)	351 (14.78)	178 (18.92)	
Midwest	2198 (21.70)	93 (21.63)	<5 (—)	1397 (22)	509 (21.43)	195 (20.72)	
South	3865 (38.16)	163 (37.91)	11 (34.38)	2441 (38.43)	906 (38.15)	344 (36.56)	
West	2426 (23.95)	85 (19.77)	10 (31.25)	1502 (23.65)	606 (25.52)	223 (23.70)	
US-dependent areas	8 (0.08)	<5 (—)	<5 (—)	<5 (—)	<5 (—)	<5 (—)	
NHBS^e city resident, n (%)							.12
Yes	3338 (32.95)	133 (30.93)	15 (46.88)	2022 (31.84)	782 (32.93)	386 (41.02)	
No	6791 (67.05)	297 (69.07)	17 (53.13)	4329 (68.16)	1593 (67.07)	555 (58.98)	
Population density^f, n (%)							<.001
Urban	3680 (36.33)	129 (30)	17 (53.13)	2187 (34.44)	906 (38.15)	441 (46.87)	
Suburban	2110 (20.83)	112 (26.05)	5 (15.63)	1351 (21.27)	477 (20.08)	165 (17.53)	
Small or medium metropolitan	3317 (32.75)	138 (32.09)	6 (18.75)	2181 (34.34)	715 (30.11)	277 (29.44)	
Rural	1013 (10)	51 (11.86)	<5 (—)	627 (9.87)	274 (11.54)	57 (6.06)	
Self-reported HIV status, n (%)							<.001
Positive	616 (6.08)	42 (9.8)	<5 (—)	205 (3.2)	255 (10.74)	110 (11.7)	
Negative	6725 (66.39)	300 (69.77)	25 (78.13)	3758 (59.17)	1868 (78.65)	774 (82.25)	
Unknown	2788 (27.52)	88 (20.47)	<5 (—)	2388 (37.6)	252 (10.61)	57 (6.06)	
Total, n (%)	10,129 (100)	430 (4.26)	32 (0.32)	6351 (62.70)	2375 (23.45)	941 (9.29)	N/A ^g

^aRefers to the number of websites or apps in this category.^bAMIS: American Men's Internet Survey.^cA chi-square test for the difference in characteristics between recruitment types.^dPercentage is not reported due to an insufficient n.^eNHBS: National HIV Behavioral Surveillance.^fThe National Center for Health Statistics urban or rural category could not be assigned to 10 participants living in US territories.^gN/A: not applicable.

Figure 1. The number of men who have sex with men who participated in the American Men's Internet Survey (AMIS) by state, 2018.

Sexual Behaviors

Approximately two-thirds (6926/10,129, 68.37%) of participants reported having anal sex without a condom with another male in the past 12 months, and about one-fifth (2390/10,129, 23.59%) of participants reported doing so with a partner of a discordant or an unknown HIV status (Table 3). Compared with HIV-negative or unknown status participants, those who were living with HIV were significantly more likely to report anal intercourse without a condom (aOR 2.02, 95% CI 1.63-2.50),

including with male partners who were of a discordant or an unknown status (aOR 3.90, 95% CI 3.27-4.66). Stratified by the serostatus group, anal intercourse without a condom differed significantly for only HIV-negative or unknown status participants by race or ethnicity, age group, and recruitment website. Anal intercourse without a condom with partners of a discordant or an unknown HIV status differed significantly by race or ethnicity, residence in an NHBS city, and recruitment type for HIV-negative or unknown status participants only.

Table 3. Sexual behaviors with male partners of men who have sex with men in the American Men's Internet Survey, United States, 2018.

Participant characteristics by HIV status	Participants (N)	Sexual behaviors with male partners in the past 12 months			
		Anal intercourse without a condom		Anal intercourse without a condom with a partner of a discordant or an unknown HIV status	
		n (%)	P value ^a	n (%)	P value ^a
HIV positive	616	499 (81)	<.001 ^b	338 (54.87)	<.001 ^b
Race or ethnicity					
Black, non-Hispanic	97	69 (71.13)	.31	39 (40.20)	.09
Hispanic	73	58 (79.45)	.92	41 (56.16)	.35
White, non-Hispanic	406	341 (83.99)	Ref ^c	240 (59.11)	Ref ^c
Other or multiple races	29	21 (72.41)	.40	13 (44.83)	.42
Age (years)					
15-24	43	37 (86.05)	.76	24 (55.81)	.86
25-29	40	37 (92.50)	.12	24 (60)	.29
30-39	104	86 (82.69)	.36	60 (57.69)	>.99
≥40	429	339 (79.02)	Ref ^c	230 (53.61)	Ref ^c
NHBS^d city resident					
Yes	255	203 (79.61)	.89	137 (53.73)	.89
No	361	296 (81.99)	Ref ^c	201 (55.68)	Ref ^c
Recruitment type					
Gay social networking	59	45 (76.27)	>.99	33 (55.93)	.83
General gay interest	7	5 (71.43)	.52	4 (57.14)	.95
General social networking	260	205 (78.85)	Ref ^c	128 (49.23)	Ref ^c
Geospatial social networking	290	244 (84.14)	.10	173 (59.66)	.43
HIV negative or unknown status	9513	6427 (67.56)	Ref ^{b,c}	2052 (21.57)	Ref ^{b,c}
Race or ethnicity					
Black, non-Hispanic	456	294 (64.47)	.32	130 (28.51)	.008
Hispanic	1557	1049 (67.37)	.05	367 (23.57)	.79
White, non-Hispanic	6605	4526 (68.52)	Ref ^c	1357 (20.55)	Ref ^c
Other or multiple races	720	449 (62.36)	.03	166 (23.05)	.78
Age (years)					
15-24	4187	2626 (62.72)	<.001	799 (19.08)	.08
25-29	1268	965 (76.10)	<.001	292 (23.03)	.08
30-39	1399	1072 (76.63)	<.001	308 (22.02)	.37
≥40	2659	1764 (66.34)	Ref ^c	653 (24.56)	Ref ^c
NHBS city resident					
Yes	3083	2103 (68.21)	.41	732 (23.74)	.002
No	6430	4324 (67.25)	Ref ^c	1320 (20.53)	Ref ^c
Recruitment type					
Gay social networking	503	280 (55.66)	<.001	108 (21.47)	.73
General gay interest	42	25 (59.52)	.57	6 (14.29)	.23

Participant characteristics by HIV status	Participants (N)	Sexual behaviors with male partners in the past 12 months			
		Anal intercourse without a condom		Anal intercourse without a condom with a partner of a discordant or an unknown HIV status	
		n (%)	P value ^a	n (%)	P value ^a
General social networking	6587	4385 (66.57)	Ref ^c	1266 (19.22)	Ref ^c
Geospatial social networking	2366	1725 (72.91)	<.001	668 (28.23)	<.001

^aWald chi-square *P* values from the multivariate logistic regression model comparing behavior (yes vs no) between groups with specific characteristics and a reference group.

^bWald chi-square *P* values from the multivariate logistic regression model comparing behavior (yes vs no) among HIV-positive participants and HIV-negative or unknown serostatus participants. Model controlled for race or ethnicity, age, NHBS system city residency, and recruitment type.

^cRef: The reference group being compared to within the multivariate logistic regression models.

^dNHBS: National HIV Behavioral Surveillance.

Substance Use Behaviors

In total, 31.93% (3235/10,129) of participants reported using marijuana, 2.31% (234/10,129) reported using methamphetamines, and 20.43% (2069/10,129) reported using other illicit substances in the past 12 months (Table 4). Compared with participants with HIV negative or unknown status, HIV-positive participants were significantly more likely to report the use of marijuana (aOR 1.39, 95% CI 1.15-1.68), methamphetamines (aOR 3.42, 95% CI 2.41-4.87), and other

illicit substances (aOR 1.90, 95% CI 1.56-2.32) in the past 12 months. Among HIV-positive participants, the use of marijuana varied significantly by age, and the use of other illicit substances varied significantly by race or ethnicity. In addition, the use of marijuana, methamphetamines, and other illicit substances differed significantly by age and race or ethnicity among HIV-negative or unknown status participants. In this group, the use of marijuana and other illicit substances differed significantly by residence in an NHBS city, and the use of marijuana differed significantly by the recruitment website.

Table 4. Substance use behaviors of men who have sex with men in the American Men's Internet Survey, United States, 2018.

Participant characteristics by HIV status	Participants (N)	Used marijuana		Used methamphetamines		Used other substances	
		n (%)	P value ^a	n (%)	P value ^a	n (%)	P value ^a
HIV positive	616	189 (30.68)	.001 ^b	50 (8.12)	<.001 ^b	169 (27.44)	<.001 ^b
Race or ethnicity							
Black, non-Hispanic	97	18 (18.56)	.12	<5 (—) ^c	.21	8 (8.24)	<.001
Hispanic	73	22 (30.14)	.91	5 (6.85)	.85	21 (28.77)	.15
White, non-Hispanic	406	137 (33.74)	Ref ^d	37 (9.11)	Ref ^d	130 (32.02)	Ref ^d
Other or multiple races	29	8 (27.58)	.81	<5 (—)	.95	7 (24.14)	.77
Age (years)							
15-24	43	22 (51.16)	.002	<5 (—)	.81	12 (27.91)	.82
25-29	40	14 (35)	.88	<5 (—)	.24	13 (32.50)	.62
30-39	104	42 (40.38)	.96	11 (10.58)	.19	36 (34.62)	.20
≥40	429	111 (25.87)	Ref ^d	34 (7.93)	Ref ^d	108 (25.17)	Ref ^d
NHBS^e city resident							
Yes	255	81 (31.76)	.22	24 (9.41)	.19	70 (27.45)	.46
No	361	108 (29.92)	Ref ^d	26 (7.20)	Ref ^d	99 (27.42)	Ref ^d
Recruitment type							
Gay social networking	59	15 (25.42)	.63	<5 (—)	.97	16 (27.12)	.84
General gay interest	7	<5 (—)	.42	<5 (—)	.96	<5 (—)	.37
General social networking	260	73 (28.08)	Ref ^d	22 (8.97)	Ref ^d	61 (23.46)	Ref ^d
Geospatial social networking	290	100 (34.48)	.06	26 (8.9)	.96	89 (30.69)	.93
HIV negative or unknown status	9513	3046 (32.02)	Ref ^{b,d}	184 (1.93)	Ref ^{b,d}	1900 (19.97)	Ref ^{b,d}
Race or ethnicity							
Black, non-Hispanic	456	122 (26.75)	.002	6 (1.32)	.18	60 (13.16)	<.001
Hispanic	1557	544 (34.94)	.86	42 (2.70)	.002	337 (21.64)	.10
White, non-Hispanic	6605	2071 (31.36)	Ref ^d	117 (1.77)	Ref ^d	1301 (19.70)	Ref ^d
Other or multiple races	720	269 (37.36)	.02	13 (1.81)	.81	173 (24.03)	<.001
Age (years)							
15-24	4187	1611 (38.48)	<.001	47 (1.12)	.009	861 (20.56)	.18
25-29	1268	427 (33.68)	.01	22 (1.74)	.70	310 (24.45)	<.001
30-39	1399	468 (33.45)	.03	41 (2.93)	.01	328 (23.45)	.005
≥40	2659	540 (20.31)	Ref ^d	74 (2.78)	Ref ^d	401 (15.08)	Ref ^d
NHBS city resident							
Yes	3083	1100 (35.68)	<.001	57 (1.85)	.35	731 (23.71)	<.001
No	6430	1946 (30.26)	Ref ^d	127 (1.98)	Ref ^d	1169 (18.18)	Ref ^d
Recruitment type							
Gay social networking	503	99 (19.68)	.002	21 (4.17)	.10	80 (15.90)	.33
General gay interest	42	19 (45.23)	.003	<5 (—)	.93	11 (26.19)	.18
General social networking	6587	2294 (34.83)	Ref ^d	84 (1.28)	Ref ^d	1343 (20.39)	Ref ^d
Geospatial social networking	2366	632 (26.71)	.15	78 (3.30)	.65	464 (19.61)	.75

^aWald chi-square *P* values from the multivariate logistic regression model comparing behavior (yes vs no) between groups with specific characteristics and a reference group.

^bWald chi-square *P* values from the multivariate logistic regression model comparing behavior (yes vs no) among HIV-positive participants and HIV-negative or unknown serostatus participants. Model controlled for race or ethnicity, age, National HIV Behavioral Surveillance system city residency, and recruitment type.

^cPercentage is not reported due to an insufficient *n*.

^dRef: The reference group being compared to within the multivariate logistic regression models.

^eNHBS: National HIV Behavioral Surveillance.

HIV Testing Behaviors

HIV testing behaviors were examined among participants who were not HIV positive (Table 5). Most participants (6836/9513, 71.85%) were previously tested for HIV infection, and 52.51%

(4995/9513) of the participants were tested in the past 12 months. HIV testing behavior, both ever tested and tested in the past 12 months, differed significantly by age, residence in an NHBS city, and type of recruitment website.

Table 5. HIV testing behaviors of HIV-negative or unknown status men who have sex with men in the American Men's Internet Survey, United States, 2018.

Participant characteristics	Participants (N)	HIV testing behaviors			
		Ever tested for HIV		Tested for HIV in the past 12 months	
		<i>n</i> (%)	<i>P</i> value ^a	<i>n</i> (%)	<i>P</i> value ^a
Race or ethnicity					
Black, non-Hispanic	456	339 (74.34)	.46	256 (56.14)	.50
Hispanic	1557	1028 (66.02)	.87	799 (51.32)	.66
White, non-Hispanic	6605	4871 (73.75)	Ref ^b	3480 (52.69)	Ref ^b
Other or multiple races	720	489 (67.92)	.93	377 (52.36)	.73
Age (years)					
15-24	4187	2115 (55.01)	<.001	1692 (40.41)	<.001
25-29	1268	1057 (83.36)	.11	752 (59.31)	.001
30-39	1399	1268 (90.64)	<.001	901 (64.40)	<.001
≥40	2659	2398 (90.18)	Ref ^b	1650 (62.05)	Ref ^b
NHBS^c city resident					
Yes	3083	2339 (75.87)	<.001	1797 (58.29)	<.001
No	6430	4499 (69.97)	Ref ^b	3198 (49.74)	Ref ^b
Recruitment type					
Gay social networking	503	414 (82.31)	<.001	261 (51.88)	.03
General gay interest	42	38 (90.47)	.22	19 (45.24)	.07
General social networking	6587	4243 (64.41)	Ref ^b	3063 (46.50)	Ref ^b
Geospatial social networking	2366	2128 (89.94)	.08	1640 (69.32)	<.001
Total	9513	6838 (71.88)		4995 (52.51)	

^aWald chi-square *P* values from the multivariate logistic regression model comparing behavior (yes vs no) between groups with specific characteristics and a reference (Ref) group.

^bRef: The reference group being compared to within the multivariate logistic regression models.

^cNHBS: National HIV Behavioral Surveillance.

Sexually Transmitted Infection Testing and Diagnosis

In total, 42.59% (4314/10,129) of participants reported sexually transmitted infection (STI) testing in the past 12 months, and 10.08% (1022/10,129) of participants reported a diagnosis of STI in the past 12 months. Compared with HIV-negative or unknown status participants, HIV-positive participants were

significantly more likely to report STI testing (aOR 3.50, 95% CI 2.89-4.24) and STI diagnosis (aOR 2.61, 95% CI 2.10-3.25) in the past 12 months (Table 6). The most common STI diagnosis among HIV-positive participants was syphilis (78/616, 12.7%), followed by gonorrhea (64/616, 10.4%) and chlamydia (62/616, 10.1%). Gonorrhea was the most common STI diagnosis among HIV-negative or unknown status participants

(542/9513, 5.69%), followed by chlamydia (520/9513, 5.47%) and syphilis (282/9513, 2.96%). STI testing and diagnosed with any STI significantly differed by age, residence in an NHBS

city, and recruitment website among HIV-negative or unknown status participants. Among HIV-positive participants, STI testing only significantly differed by residence in an NHBS city.

Table 6. Sexually transmitted infection testing and diagnosis of men who have sex with men in the American Men's Internet Survey, United States, 2018.

Participant characteristics by HIV status	Participants (N)	STI ^a history in the past 12 months			
		Tested for any STI		Diagnosed with any STI	
		n (%)	P value ^b	n (%)	P value ^b
HIV positive	616	448 (72.7)	<.001 ^c	135 (21.9)	<.001 ^c
Race or ethnicity					
Black, non-Hispanic	97	73 (75)	.46	24 (25)	.68
Hispanic	73	56 (77)	.89	23 (32)	.19
White, non-Hispanic	406	290 (71.4)	Ref ^d	80 (19.7)	Ref ^d
Other or multiple races	29	21 (72)	.53	6 (21)	.47
Age (years)					
15-24	43	37 (86)	.09	13 (30)	.28
25-29	40	30 (75)	.72	11 (28)	.92
30-39	104	88 (84.6)	.28	33 (31.7)	.19
≥40	429	293 (68.3)	Ref ^d	78 (18.2)	Ref ^d
NHBS^e city resident					
Yes	255	202 (79.2)	.004	62 (24.3)	.32
No	361	246 (68.1)	Ref ^d	73 (20.2)	Ref ^d
Recruitment type					
Gay social networking	59	41 (70)	.93	10 (17)	.54
General gay interest	7	5 (71)	.99	<5 (—)	.48
General social networking	260	184 (70.8)	Ref ^d	56 (21.5)	Ref ^d
Geospatial social networking	290	218 (75.2)	.28	67 (23.1)	.84
HIV negative or unknown status	9513	3866 (40.64)	Ref ^{c,d}	887 (9.32)	Ref ^{c,d}
Race or ethnicity					
Black, non-Hispanic	456	192 (42.1)	.90	49 (10.8)	.77
Hispanic	1557	642 (41.23)	.41	169 (10.85)	.25
White, non-Hispanic	6605	2669 (40.41)	Ref ^d	587 (8.89)	Ref ^d
Other or multiple races	720	293 (40.7)	>.99	68 (9.4)	.61
Age (years)					
15-24	4187	1419 (33.89)	<.001	330 (7.88)	.09
25-29	1268	631 (49.76)	<.001	161 (12.70)	<.001
30-39	1399	676 (48.32)	<.001	153 (10.94)	.44
≥40	2659	1140 (42.87)	Ref ^d	243 (9.14)	Ref ^d
NHBS^e city resident					
Yes	3083	1425 (46.22)	<.001	389 (12.62)	<.001
No	6430	2441 (37.96)	Ref ^d	498 (7.74)	Ref ^d
Recruitment type					
Gay social networking	503	183 (36.4)	.72	41 (8.2)	.96
General gay interest	42	12 (29)	.06	<5 (—) ^f	.62
General social networking	6587	2476 (37.59)	Ref ^d	545 (8.27)	Ref ^d

Participant characteristics by HIV status	Participants (N)	STI ^a history in the past 12 months			
		Tested for any STI		Diagnosed with any STI	
		n (%)	P value ^b	n (%)	P value ^b
Geospatial social networking	2366	1188 (50.21)	<.001	296 (12.51)	.02

^aSTI: sexually transmitted infection (includes chlamydia, gonorrhea, and syphilis).

^bWald chi-square *P* values from the multivariate logistic regression model comparing behavior (yes vs no) between groups with specific characteristics and a reference (Ref) group.

^cWald chi-square *P* values from the multivariate logistic regression model comparing behavior (yes vs no) among HIV-positive participants and HIV-negative or unknown serostatus participants. Model controlled for race or ethnicity, age, NHBS system city residency, and recruitment type.

^dRef: The reference group being compared to within the multivariate logistic regression models.

^eNHBS: National HIV Behavioral Surveillance.

^fPercentage is not reported due to an insufficient n.

Discussion

The sixth round of data collection for AMIS was successfully implemented and resulted in over 10,000 surveys from a diverse sample of internet-using MSM residing in all US states. A majority of eligible and enrolled participants were recruited from general social networking, were White, non-Hispanic, aged between 15 and 24 years, and reported being HIV negative. There were notable differences in key behavioral indicators by self-reported HIV status. Compared with HIV-negative or unknown status participants, HIV-positive participants were

more likely to have had anal sex without a condom with a male partner in the past year and more likely to have had anal sex without a condom with a serodiscordant or an unknown status partner. The reported use of marijuana, methamphetamines, and other illicit substances in the past year was higher among HIV-positive participants than among HIV-negative or unknown status participants. HIV-positive participants were also more likely to report testing and diagnosis of STIs than HIV-negative or unknown status participants. When stratified by HIV status, some significant differences in these behavioral indicators by demographics and recruitment websites were also observed.

Acknowledgments

The study was funded by a grant from the National Institutes of Health (P30AI050409)—the Emory Center for AIDS Research.

Conflicts of Interest

TS and PS are members of the Editorial Board of JMIR Public Health and Surveillance. However, they had no involvement in the editorial decision for this manuscript. The manuscript was reviewed and handled by an independent editor.

Multimedia Appendix 1

American Men's Internet Survey 2018 questionnaire.

[[DOCX File, 406 KB](#) - [publichealth_v7i3e21812_app1.docx](#)]

References

1. Finlayson T, Le B, Smith A, Bowles K, Cribbin M, Miles I, et al. HIV risk, prevention, and testing behaviors among men who have sex with men--National HIV Behavioral Surveillance System, 21 U.S. cities, United States, 2008. *MMWR Surveillance Summaries* 2011;60(14):1-34 [[FREE Full text](#)] [Medline: [22031280](#)]
2. Sanchez TH, Sineath RC, Kahle EM, Tregear SJ, Sullivan PS. The annual American men's internet survey of behaviors of men who have sex with men in the United States: protocol and key indicators report 2013. *JMIR Public Health Surveill* 2015;1(1):e3 [[FREE Full text](#)] [doi: [10.2196/publichealth.4314](#)] [Medline: [27227126](#)]
3. Sanchez T, Zlotorzynska M, Sineath C, Kahle E, Sullivan P. The annual American men's internet survey of behaviors of men who have sex with men in the United States: 2014 key indicators report. *JMIR Public Health Surveill* 2016 May 25;2(1):e23 [[FREE Full text](#)] [doi: [10.2196/publichealth.5476](#)] [Medline: [27244770](#)]
4. Zlotorzynska M, Sullivan P, Sanchez T. The annual American men's internet survey of behaviors of men who have sex with men in the United States: 2015 key indicators report. *JMIR Public Health Surveill* 2017 Mar 25;3(1):e13 [[FREE Full text](#)] [doi: [10.2196/publichealth.7119](#)] [Medline: [28356240](#)]
5. Zlotorzynska M, Sullivan P, Sanchez T. The annual American men's internet survey of behaviors of men who have sex with men in the United States: 2016 key indicators report. *JMIR Public Health Surveill* 2019 Feb 20;5(1):e11313 [[FREE Full text](#)] [doi: [10.2196/11313](#)] [Medline: [30785405](#)]

6. Zlotorzynska M, Cantu C, 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: 2017 key indicators report. *JMIR Public Health Surveill* 2020 Apr 13;6(2):e16847 [FREE Full text] [doi: [10.2196/16847](https://doi.org/10.2196/16847)] [Medline: [32281937](https://pubmed.ncbi.nlm.nih.gov/32281937/)]
7. Sanchez TH, Zlotorzynska M, Sineath RC, Kahle E, Tregear S, Sullivan PS. National trends in sexual behavior, substance use and HIV testing among United States men who have sex with men recruited online, 2013 through 2017. *AIDS Behav* 2018 Aug;22(8):2413-2425. [doi: [10.1007/s10461-018-2168-4](https://doi.org/10.1007/s10461-018-2168-4)] [Medline: [29948340](https://pubmed.ncbi.nlm.nih.gov/29948340/)]
8. HUD USPS ZIP code crosswalk files. URL: https://www.huduser.gov/portal/datasets/usps_crosswalk.html [accessed 2017-02-15]
9. USPS ZIP code lookup. URL: https://tools.usps.com/go/ZipLookupAction_input [accessed 2017-02-15]
10. Centers for Disease Control Prevention. Hiv infection risk, prevention, and testing behaviors among men who have sex with men. National HIV Behavioral Surveillance, 20 U.S. Cities. 2014. URL: <https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-special-report-number-15.pdf> [accessed 2021-02-16]
11. Ingram DD, Franco SJ. 2013 NCHS Urban-rural classification scheme for counties. *Vital Health Stat* 2014 Apr(166):1-73 [FREE Full text] [Medline: [24776070](https://pubmed.ncbi.nlm.nih.gov/24776070/)]

Abbreviations

AMIS: American Men's Internet Survey

aOR: adjusted odds ratio

MSM: men who have sex with men

NHBS: National HIV Behavioral Surveillance System

STI: sexually transmitted infection

Edited by G Eysenbach; submitted 26.06.20; peer-reviewed by S Badawy, D Rao; comments to author 08.08.20; revised version received 11.10.20; accepted 17.01.21; published 04.03.21.

Please cite as:

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;7(3):e21812

URL: <https://publichealth.jmir.org/2021/3/e21812>

doi:[10.2196/21812](https://doi.org/10.2196/21812)

PMID:[33496669](https://pubmed.ncbi.nlm.nih.gov/33496669/)

©Sarah Wiatrek, Maria Zlotorzynska, Ramona Rai, Patrick Sullivan, Travis Sanchez. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 04.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

The Relationship Between the Global Burden of Influenza From 2017 to 2019 and COVID-19: Descriptive Epidemiological Assessment

Stefan David Baral^{1*}, MPH, MD; Katherine Blair Rucinski^{1*}, MPH, PhD; Jean Olivier Twahirwa Rwema^{1*}, MPH, MD; Amrita Rao^{1*}, ScM; Neia Prata Menezes^{1*}, MPH; Daouda Diouf^{2*}, MS; Adeeba Kamarulzaman^{3*}, MBBS, FRACP; Nancy Phaswana-Mafuya^{4*}, PhD; Sharmistha Mishra^{5,6*}, MD, PhD

¹Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

²ENDA Santé, Dakar, Senegal

³Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

⁴Department of Environmental Health, Faculty of Health Sciences, University of Johannesburg, Johannesburg, South Africa

⁵Department of Medicine, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

⁶Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

*all authors contributed equally

Corresponding Author:

Stefan David Baral, MPH, MD

Department of Epidemiology

Bloomberg School of Public Health

Johns Hopkins University

615 N. Wolfe Street

Baltimore, MD, 21205

United States

Phone: 1 410 502 8975

Email: sbaral@jhu.edu

Abstract

Background: SARS-CoV-2 and influenza are lipid-enveloped viruses with differential morbidity and mortality but shared modes of transmission.

Objective: With a descriptive epidemiological framing, we assessed whether recent historical patterns of regional influenza burden are reflected in the observed heterogeneity in COVID-19 cases across regions of the world.

Methods: Weekly surveillance data reported by the World Health Organization from January 2017 to December 2019 for influenza and from January 1, 2020 through October 31, 2020, for COVID-19 were used to assess seasonal and temporal trends for influenza and COVID-19 cases across the seven World Bank regions.

Results: In regions with more pronounced influenza seasonality, COVID-19 epidemics have largely followed trends similar to those seen for influenza from 2017 to 2019. COVID-19 epidemics in countries across Europe, Central Asia, and North America have been marked by a first peak during the spring, followed by significant reductions in COVID-19 cases in the summer months and a second wave in the fall. In Latin America and the Caribbean, COVID-19 epidemics in several countries peaked in the summer, corresponding to months with the highest influenza activity in the region. Countries from regions with less pronounced influenza activity, including South Asia and sub-Saharan Africa, showed more heterogeneity in COVID-19 epidemics seen to date. However, similarities in COVID-19 and influenza trends were evident within select countries irrespective of region.

Conclusions: Ecological consistency in COVID-19 trends seen to date with influenza trends suggests the potential for shared individual, structural, and environmental determinants of transmission. Using a descriptive epidemiological framework to assess shared regional trends for rapidly emerging respiratory pathogens with better studied respiratory infections may provide further insights into the differential impacts of nonpharmacologic interventions and intersections with environmental conditions. Ultimately, forecasting trends and informing interventions for novel respiratory pathogens like COVID-19 should leverage epidemiologic patterns in the relative burden of past respiratory pathogens as prior information.

KEYWORDS

SARS-CoV-2; COVID-19; influenza; descriptive epidemiology; epidemiology; assessment; relationship; flu; virus; burden; global health; public health; transmission; pattern

Introduction

The COVID-19 pandemic has affected each region of the world differently. As of December 2020, Europe and the Americas have registered the highest numbers of COVID-19 cases and deaths, while Southeast Asia, sub-Saharan Africa, and the Western Pacific have experienced comparatively milder epidemics [1]. Heterogeneity in COVID-19 burden has also been observed within cities, across cities, and across regions within a country [2,3]. This heterogeneity likely stems from differences in underlying population structure with respect to age distribution, population and housing density, access to health care, burden of comorbidities, socioeconomic and structural barriers to engagement in health, environmental factors, and the breadth and intensity of public health interventions. Furthermore, there remains active investigation into the heterogeneity of the effects of both COVID-19 and the associated response on potential syndemics, including infectious diseases such as HIV and tuberculosis, and noncommunicable conditions including mental health, reproductive health, and malnutrition [4].

Similar to SARS-CoV-2, the virus that causes COVID-19, influenza is a lipid-enveloped respiratory virus but with a more established evidence base for both effective nonpharmacologic intervention strategies and transmission dynamics [5-7]. Importantly, influenza is associated with a shorter incubation period and with lower mortality than the currently estimated SARS-CoV-2 infection case fatality rate [8-10]. The relative risks of indoor versus outdoor transmission and risks for contact, droplet-based, and aerosolization-mediated transmission are better established for influenza [11-13]. Influenza seasonality and spatial distribution are not fully understood but likely reflect individual, population, and environmental determinants [14]. Moreover, the annual burden and associated mortality secondary to pandemic and nonpandemic seasonal influenza strains vary significantly. The 2009 H1N1 pandemic was estimated to cause over 60 million infections in the United States and resulted in higher mortality among youth but fewer deaths overall, as older people were more likely to be immune [15]. In 2017, there were an estimated 45 million cases of mostly H3N2 and Influenza B in the United States associated with higher mortality, particularly among older adults given limited pre-existing immunity and limited vaccine effectiveness [16]. Environmental factors have been shown to explain, in part, the differences in influenza seasonality between temperate regions where cases peak during winter months and tropical regions, which may have multiple peaks or intermittent influenza activity [17,18]. However, even within specific regions, some countries show influenza transmission patterns that appear distinct from regional trends [19,20].

There have been early reports suggesting shared individual and structural determinants of infection between influenza and

SARS-CoV-2. In the United States, there is evidence of higher COVID-19 incidence, morbidity, and mortality among African American communities secondary to underlying inequities, including structural racism [21]. These racial disparities were also observed during the 2009 H1N1 influenza pandemic with similar drivers including higher risk of exposure due to inability to social distance, higher burden of comorbidities, and limited access to health care among African American and other racialized communities [22-24]. Similar early trajectories of COVID-19 and 2009 H1N1 have been noted in South Africa, with higher burdens initially among higher-income communities reporting more travel, with rapid epidemiologic transition to lower-income communities [25]. Finally, increases in COVID-19 in late May and June in some areas of the Southern hemisphere including countries across South America, sub-Saharan Africa, East Asia, and the Pacific were consistent with the timing of the traditional influenza season [26]. Specifically, in South Africa, epidemic growth is consistent with the influenza season and with epicenters of COVID-19 appearing to be similar to areas with increased influenza transmission [27,28].

When this novel coronavirus emerged in early 2020, regional and national forecasting exercises drew on emerging data on the trajectory of detected COVID-19 cases in the epicenters of China and Italy to translate their reproductive rates into best- and worst-case scenarios for populations worldwide. However, despite early comparison of the average mortality of influenza and SARS-CoV-2, there has not been a systematic evaluation of whether historical distributions in influenza burden are relevant in understanding the burden and distribution of SARS-CoV-2 infection. Descriptive epidemiological assessment within and across both influenza and COVID-19 may provide early insight into overlapping individual, structural, and environmental determinants of infection that have implications for public health responses. Moreover, documenting consistencies in the burden of disease across both pathogens may facilitate the generation of testable hypotheses, facilitating future analytic comparisons as well as future pandemic preparedness. In this study, we investigated global influenza transmission patterns from 2017 onward to describe shared trends with the epidemiology of COVID-19.

Methods

Data Abstraction and Assessment of Eligibility

Influenza

We used publicly available data to characterize temporal patterns of influenza across seven regions between 2017 and 2019. Organization of countries into regions was based on World Bank classifications (ie, East Asia and the Pacific, Europe and Central Asia, Latin America and the Caribbean, Middle East and North Africa, North America, South Asia, and sub-Saharan

Africa). Within each region, we assessed country-specific weekly reported cases of influenza, including both Influenza A and B, from January 1, 2017, to December 31, 2019. Data were abstracted from FluNet, a global web-based tool developed by the World Health Organization for influenza virologic surveillance, which captures weekly laboratory-confirmed influenza cases. FluNet aggregates data from the National Influenza Centres of the Global Influenza Surveillance and Response System and other national influenza reference laboratories. Since more temperate zones in the Northern and Southern Hemispheres have generally opposing seasonal influenza trends, we report influenza case counts by calendar year from January 1 to December 31, which also facilitates comparisons with COVID-19. Finally, influenza case counts for the calendar year of 2020 were excluded given that public health measures focused on COVID-19 prevention and mitigation likely decreased seasonal influenza cases substantially for this time period [29]. The suspension of FluView, the core influenza reporting system in the United States, further precluded the exclusion of 2020 influenza case counts.

We assessed patterns of missing data for each country and by year to determine eligibility for inclusion in analyses. For each year, countries where $\geq 90\%$ of available weekly case reports were complete (assessed as the total number of nonmissing records over the total number of all included weeks for that year) were included in the analyses. For countries where data were $< 90\%$ complete, we made the following determination:

- Countries where data missingness appeared to follow a seasonal pattern (eg, data were missing only for weeks outside of the known influenza season, likely reflecting country-specific seasonal influenza reporting guidelines) were retained in analyses.
- Countries where weekly case counts were $< 90\%$ complete and did not appear to follow a seasonal pattern were denoted as missing.

COVID-19

We abstracted the daily number of COVID-19 cases reported between January 1 and October 31, 2020, for all countries and corresponding World Bank regions using data from the World Health Organization Coronavirus Situation Report. As we are in the first year of COVID-19, it was not possible to assess missingness of data similar to the methods used for influenza.

Regional Influenza and COVID-19 Burden

Influenza

We assessed the overall burden of influenza for all included countries across and within each region. To do so, we summed the absolute number of weekly influenza cases for each region, country, and year (January 1 to December 31 of each year) given the temporal variability in influenza seasons worldwide. Countries were ranked in descending order within each region from the highest to lowest number of absolute cases overall and for 2017, 2018, and 2019.

COVID-19

We assessed, and similarly ranked in descending order, cumulative cases of COVID-19 as reported through October 31, 2020, within each region.

Comparison of Regional Burden of Influenza and COVID-19

We used regional plots for both influenza and COVID-19 cases to visualize cross-pathogen comparisons. For each region, we plotted the mean number of weekly influenza cases for the five highest ranked countries across 2017, 2018, and 2019. Influenza cases were averaged to ensure observed trends were robust to one-off spikes or changes to reporting systems in select years. Daily COVID-19 cases for the five highest ranked countries within each region were plotted using a 7-day moving average. We similarly plotted influenza and COVID-19 cases using per-capita proportions for these countries to facilitate standardized comparisons across countries and regions. Plots were visually inspected to confirm temporal patterns, and we described similarities or differences across and within all regions.

Within-Country Comparison of Influenza and COVID-19

We used plots to visually compare within-country trends for both influenza and COVID-19 over time. We assessed the top 20 countries with the highest burden of COVID-19 for each region, further restricting plots to those countries where influenza data were eligible for inclusion. As with regional comparisons, we plotted the mean number of weekly influenza cases across 2017, 2018, and 2019, and daily COVID-19 cases were plotted using a 7-day moving average.

Results

Influenza cases were assessed for 171 countries. For each year, more than half (ranging from $n=95$, 55.6% to $n=97$, 56.7%) of all countries reported $\geq 90\%$ complete data ([Multimedia Appendix 1](#)). The proportion of countries where data were $< 90\%$ complete but followed a pattern of seasonal missingness were also included in influenza analyses and ranged from 22.8% ($n=39$) in 2017 to 24.0% ($n=41$) in 2019. Differences in data completion and patterns of missingness were evident across all regions. The overall influenza burden between 2017 and 2019 varied by region and by countries within all seven identified regions ([Table 1](#)). Regions with the highest absolute burden of influenza included North America (903,554 cases), Europe and Central Asia (599,820 cases), and East Asia and the Pacific (389,657 cases). Countries with the highest proportion of regional influenza cases were as follows: East Asia and the Pacific (China: 302,126/389,657, 77.5%); Europe and Central Asia (the United Kingdom: 106,087/599,020, 17.7%), Latin America and the Caribbean (Mexico: 18,976/93,172, 20.2%), Middle East and North Africa (Qatar: 27,363/67,610, 40.5%), North America (the United States: 749,472/903,554, 82.9%), South Asia (India: 14,895/31,992, 46.6%), and sub-Saharan Africa (South Africa: 3534/18,930, 18.7%). Across regions, countries with the highest overall burden of influenza from 2017 to 2019 included the United States (749,472 cases), China

(302,126 cases), Canada (154,082 cases), the United Kingdom (106,087 cases), and Norway (71,727 cases).

Regions with the highest absolute burden of COVID-19 included Latin America and the Caribbean (11,045,364 cases), Europe and Central Asia (9,197,241 cases), South Asia (9,107,653 cases), and North America (9,081,471 cases; [Table 2](#)). Within each region, countries with the highest proportion of COVID-19 regional cases were as follows: East Asia and the Pacific (Indonesia: 406,945/1,191,632, 34.2%); Europe and Central Asia (Russia: 1,618,116/9,197,241, 17.6%), Latin America and the Caribbean (Brazil: 5,494,376/11,045,364, 49.7%), Middle East and North Africa (Iran: 604,952/2,662,465, 22.7%), North America (the United States: 8,852,730/9,081,471, 97.5%), South Asia (India: 8,137,119/9,107,653, 89.3%), and sub-Saharan Africa (South Africa: 723,682/1,139,329, 63.5%; [Table 2](#)). Specific countries with the highest overall burden of COVID-19 cases as of October 31, 2020, included the United States (8,852,730 cases), India (8,137,119 cases), Brazil (5,494,376 cases), Russia (1,618,116 cases), and France (1,299,278 cases; [Table 2](#)).

Seasonal variation in influenza cases was observed within all regions. Seasonal trends were most pronounced across countries for North America, Europe and Central Asia, the Middle East and North Africa, East Asia and the Pacific, and Latin America and the Caribbean ([Multimedia Appendix 2](#)). Seasonal trends were also evident but largely inconsistent across countries within South Asia and sub-Saharan Africa. These dynamics were

mainly present when trends were assessed by per capita proportions instead of case counts ([Multimedia Appendix 3](#)).

COVID-19 epidemics generally followed trends similar to those observed for influenza from 2017 to 2019 in the regions with the most pronounced influenza seasonal trends ([Multimedia Appendix 3](#)). This was more apparent in Europe and Central Asia with an early wave of COVID-19 epidemics in several countries from March to May 2020 followed by significant drops in transmission in the summer and a second wave in the fall. With the exception of the United States that has seen sustained transmission over the year, North America (Canada and Bermuda) have had COVID-19 epidemics largely similar to those observed in European countries. In Latin America and the Caribbean, COVID-19 epidemics in Brazil, Argentina, Peru, and Colombia peaked during the summer and started decreasing in fall corresponding to influenza trends in the region. Countries in sub-Saharan Africa, South Asia, and the Middle East and North Africa have shown more heterogeneity in their COVID-19 epidemics.

Similarities between COVID-19 and influenza trends were more marked within individual countries irrespective of region ([Multimedia Appendix 4](#)). Peaks of COVID-19 epidemics in exemplary countries, including Australia, Brazil, India, South Asia, and South Africa, coincided with months of peak influenza activity in those countries from 2017 to 2019 ([Multimedia Appendix 5](#)).

Table 1. Influenza cases for top 15 countries by World Bank region, overall and by year from January 1, 2017, to December 31 2019.^a

Regions	Overall influenza cases, n (%)	2019 influenza cases, n (%)	2018 influenza cases, n (%)	2017 influenza cases, n (%)
Sub-Saharan Africa	18,930 (100.0)	7521 (100.0)	5785 (100.0)	5624 (100.0)
South Africa	3534 (18.7)	1164 (15.6)	1160 (20.1)	1210 (21.5)
Ghana	2383 (12.6)	1172 (15.5)	742 (12.8)	469 (8.3)
Senegal	1783 (9.4)	483 (6.4)	480 (8.3)	820 (14.6)
Madagascar	1117 (5.9)	451 (6.0)	298 (5.2)	368 (6.5)
Cameroon	1098 (5.8)	332 (4.4)	319 (5.5)	447 (7.9)
Ivory Coast	1057 (5.6)	458 (6.1)	262 (4.5)	337 (6.0)
Togo	988 (5.2)	501 (6.7)	254 (4.4)	233 (4.1)
Zambia	971 (5.1)	426 (5.7)	391 (6.8)	154 (2.7)
Tanzania	876 (4.6)	242 (3.2)	330 (5.7)	304 (5.4)
Mauritius	767 (4.1)	570 (7.6)	102 (1.8)	95 (1.7)
Uganda	621 (3.3)	179 (2.4)	— ^b	404 (7.2)
Mali	5441 (2.9)	235 (3.1)	144 (2.5)	165 (2.9)
Ethiopia	534 (2.8)	214 (2.8)	203 (3.5)	117 (2.1)
Kenya	461 (2.4)	220 (2.9)	216 (3.7)	—
Burkina Faso	398 (2.1)	—	236 (4.1)	—
Niger	—	144 (1.9)	—	117 (2.1)
Central African Republic	—	—	177 (3.1)	109 (1.9)
East Asia and Pacific	389,657 (100.0)	158,173 (100.0)	101,937 (100.0)	129,547 (100.0)
China	302,126 (77.5)	122,757 (77.6)	80,297 (78.8)	99,072 (76.5)
Japan	29,449 (7.6)	10,029 (6.3)	9076 (8.9)	10,344 (8.0)
Australia	28,775 (7.4)	14,002 (8.9)	4264 (4.2)	10,509 (8.1)
Republic of Korea	4958 (1.3)	1702 (1.1)	1952 (1.9)	1304 (1.0)
Thailand	3640 (0.9)	1568 (1.0)	784 (0.8)	1288 (1.0)
Singapore	3258 (0.8)	1154 (0.7)	1106 (1.1)	998 (0.8)
New Zealand	2377 (0.6)	957 (0.6)	475 (0.5)	945 (0.7)
Malaysia	2270 (0.6)	1115 (0.7)	493 (0.5)	662 (0.5)
Indonesia	2251 (0.6)	496 (0.3)	710 (0.7)	1045 (0.8)
Lao People's Democratic Republic	2039 (0.5)	590 (0.4)	661 (0.6)	788 (0.6)
Mongolia	1746 (0.4)	909 (0.6)	304 (0.3)	533 (0.4)
Vietnam	1493 (0.4)	355 (0.2)	345 (0.3)	793 (0.6)
Myanmar	1340 (0.3)	806 (0.5)	—	431 (0.3)
New Caledonia	1036 (0.3)	541 (0.3)	291 (0.3)	—
Democratic People's Republic of Korea	851 (0.2)	—	474 (0.5)	—
Philippines	—	346 (0.2)	—	268 (0.2)
Cambodia	—	—	234 (0.2)	226 (0.2)
South Asia	31,992 (100.0)	17,313 (100.0)	5401 (100.0)	9278 (100.0)
India	14,895 (46.6)	10,428 (60.2)	1880 (34.8)	2587 (27.9)
Nepal	5678 (17.7)	2569 (14.8)	957 (17.7)	2152 (23.2)
Sri Lanka	3794 (11.9)	486 (2.8)	947 (17.5)	2361 (25.4)

Regions	Overall influenza cases, n (%)	2019 influenza cases, n (%)	2018 influenza cases, n (%)	2017 influenza cases, n (%)
Bangladesh	3647 (11.4)	1783 (10.3)	673 (12.5)	1191 (12.8)
Bhutan	1487 (4.6)	679 (3.9)	364 (6.7)	444 (4.8)
Pakistan	1002 (3.1)	1002 (5.8)	—	—
Maldives	916 (2.9)	88 (0.5)	393 (7.3)	435 (4.7)
Afghanistan	573 (1.8)	278 (1.6)	187 (3.5)	108 (1.2)
Middle East and North Africa	67,610 (100.0)	31,348 (100.0)	21,224 (100.0)	15,038 (100.0)
Qatar	27,363 (40.5)	8365 (26.7)	11,450 (53.9)	7548 (50.2)
Iran	9445 (14.0)	7387 (23.6)	1430 (6.7)	628 (4.2)
Kuwait	6601 (9.8)	6601 (21.1)	—	—
Oman	5788 (8.6)	1376 (4.4)	1388 (6.5)	3024 (20.1)
Egypt	4986 (7.4)	2339 (7.5)	2072 (9.8)	575 (3.8)
Israel	4543 (6.7)	1796 (5.7)	1602 (7.5)	1145 (7.6)
Saudi Arabia	4302 (6.4)	2423 (7.7)	1179 (5.6)	700 (4.7)
Malta	1385 (2.0)	—	880 (4.1)	505 (3.4)
Jordan	891 (1.3)	366 (1.2)	273 (1.3)	252 (1.7)
Tunisia	860 (1.3)	203 (0.6)	315 (1.5)	342 (2.3)
Bahrain	488 (0.7)	197 (0.6)	202 (1.0)	89 (0.6)
Lebanon	481 (0.7)	262 (0.8)	123 (0.6)	96 (0.6)
Algeria	477 (0.7)	33 (0.1)	310 (1.5)	134 (0.9)
Latin America and Caribbean	93,712 (100.0)	33,308 (100.0)	31,606 (100.0)	28,798 (100.0)
Mexico	18,976 (20.2)	6963 (20.9)	5667 (17.9)	6346 (22.0)
Argentina	17,658 (18.8)	6477 (19.4)	4615 (14.6)	6566 (22.8)
Brazil	16,477 (17.6)	3459 (10.4)	7012 (22.2)	6006 (20.9)
Chile	15,783 (16.8)	7371 (22.1)	4192 (13.3)	4220 (14.7)
Colombia	3549 (3.8)	1316 (4.0)	1425 (4.5)	808 (2.8)
Nicaragua	3501 (3.7)	1419 (4.3)	1320 (4.2)	762 (2.6)
Paraguay	3378 (3.6)	1498 (4.5)	932 (2.9)	948 (3.3)
Bolivia	3371 (3.6)	430 (1.3)	2068 (6.5)	873 (3.0)
Peru	1651 (1.8)	349 (1.0)	1061 (3.4)	241 (0.8)
Ecuador	1393 (1.5)	615 (1.8)	778 (2.5)	—
Costa Rica	1266 (1.4)	505 (1.5)	—	591 (2.1)
Guatemala	863 (0.9)	—	322 (1.0)	228 (0.8)
Cuba	846 (0.9)	463 (1.4)	238 (0.8)	145 (0.5)
Panama	846 (0.9)	382 (1.1)	—	262 (0.9)
Haiti	809 (0.9)	358 (1.1)	340 (1.1)	—
Jamaica	—	479 (1.4)	—	—
Honduras	—	—	322 (1.0)	—
Dominican Republic	—	—	203 (0.6)	—
Uruguay	—	—	—	197 (0.7)
El Salvador	—	—	—	288 (1.0)
North America	903,554 (100.0)	310,724 (100.0)	334,093 (100.0)	258,737 (100.0)
United States	749,472 (82.9)	267,528 (86.1)	267,611 (80.1)	214,333 (82.8)

Regions	Overall influenza cases, n (%)	2019 influenza cases, n (%)	2018 influenza cases, n (%)	2017 influenza cases, n (%)
Canada	154,082 (17.1)	43,196 (13.9)	66,482 (19.9)	44,404 (17.2)
Europe and Central Asia	599,020 (100.0)	224,554 (100.0)	235,606 (100.0)	138,870 (100.0)
United Kingdom	106,087 (17.7)	42,448 (18.9)	44,120 (18.7)	19,519 (14.1)
Norway	71,727 (12.0)	22,282 (9.9)	32,966 (14.0)	16,479 (11.9)
France	62,672 (10.5)	25,405 (11.3)	21,610 (9.2)	15,657 (11.3)
Spain	49,971 (8.3)	17,232 (7.7)	19,973 (8.5)	12,766 (9.2)
Russian Federation	49,289 (8.2)	19,340 (8.6)	15,314 (6.5)	14,635 (10.5)
Sweden	46,472 (7.8)	13,823 (6.2)	21,407 (9.1)	11,242 (8.1)
Switzerland	35,856 (6.0)	14,096 (6.3)	11,918 (5.1)	9842 (7.1)
Denmark	33,599 (5.6)	11,835 (5.3)	17,120 (7.3)	4644 (3.3)
Italy	13,539 (2.3)	6361 (2.8)	4382 (1.9)	2796 (2.0)
Portugal	11,904 (2.0)	6634 (3.0)	3984 (1.7)	—
Netherlands	11,430 (1.9)	5166 (2.3)	2836 (1.2)	3428 (2.5)
Austria	11,311 (1.9)	4097 (1.8)	5143 (2.2)	2071 (1.5)
Ireland	10,964 (1.8)	4201 (1.9)	4984 (2.1)	1779 (1.3)
Slovenia	10,502 (1.8)	3576 (1.6)	3020 (1.3)	3906 (2.8)
Croatia	9216 (1.5)	4823 (2.1)	—	1796 (1.3)
Latvia	—	—	2785 (1.2)	3121 (2.2)

^aExcludes countries with missing influenza data.

^bCountry was not in top 15 for this year.

Table 2. COVID-19 cases for top 15 countries by World Bank region, from January 1 to October 31, 2020.

Region	COVID-19 cases, n (%)
Sub-Saharan Africa	1,139,329 (100.0)
South Africa	723,682 (63.5)
Ethiopia	95,789 (8.4)
Nigeria	62,691 (5.5)
Kenya	53,797 (4.7)
Ghana	48,055 (4.2)
Cameroon	21,793 (1.9)
Cote d'Ivoire	20,628 (1.8)
Madagascar	16,968 (1.5)
Zambia	16,415 (1.4)
Senegal	15,593 (1.4)
Sudan	13,804 (1.2)
Namibia	12,907 (1.1)
Mozambique	12,777 (1.1)
Uganda	12,410 (1.1)
Guinea	12,020 (1.1)
East Asia and Pacific	1,191,632 (100.0)
Indonesia	406,945 (34.2)
Philippines	378,933 (31.8)
Japan	100,392 (8.4)
China	91,893 (7.7)
Singapore	58,003 (4.9)
Myanmar	51,496 (4.3)
Malaysia	30,889 (2.6)
Australia	27,582 (2.3)
Republic of Korea	26,511 (2.2)
French Polynesia	7262 (0.6)
Guam	4579 (0.4)
Thailand	3780 (0.3)
New Zealand	1601 (0.1)
Vietnam	1177 (0.1)
Papua New Guinea	589 (0.0)
South Asia	9,107,653 (100.0)
India	8,137,119 (89.3)
Bangladesh	406,364 (4.5)
Pakistan	332,186 (3.6)
Nepal	168,235 (1.8)
Afghanistan	41,334 (0.5)
Maldives	11,643 (0.1)
Sri Lanka	10,424 (0.1)
Bhutan	348 (0.0)
Middle East and North Africa	2,662,465 (100.0)

Region	COVID-19 cases, n (%)
Iran	604,952 (22.7)
Iraq	470,633 (17.7)
Saudi Arabia	346,880 (13.0)
Morocco	215,294 (8.1)
Qatar	132,343 (5.0)
United Arab Emirates	131,508 (4.9)
Kuwait	125,337 (4.7)
Oman	114,434 (4.3)
Egypt	107,385 (4.0)
Bahrain	81,466 (3.1)
Lebanon	79,529 (3.0)
Jordan	69,306 (2.6)
West Bank and Gaza	64,741 (2.4)
Libya	60,628 (2.3)
Tunisia	58,029 (2.2)
Latin America and the Caribbean	11,045,364 (100.0)
Brazil	5,494,376 (49.7)
Argentina	1,143,800 (10.4)
Colombia	1,053,122 (9.5)
Mexico	912,811 (8.3)
Peru	897,594 (8.1)
Chile	508,571 (4.6)
Ecuador	167,147 (1.5)
Bolivia	141,484 (1.3)
Panama	132,045 (1.2)
Dominican Republic	126,332 (1.1)
Costa Rica	107,570 (1.0)
Guatemala	107,339 (1.0)
Honduras	96,150 (0.9)
Venezuela	91,280 (0.8)
Puerto Rico	65,743 (0.2)
North America	9,081,471 (100.0)
United States	8,852,730 (97.5)
Canada	228,542 (2.5)
Bermuda	199 (0.0)
Europe and Central Asia	9,197,241 (100.0)
Russian Federation	1,618,116 (17.6)
France	1,299,278 (14.1)
Spain	1,212,190 (13.2)
United Kingdom	989,749 (10.8)
Italy	647,674 (7.0)
Germany	518,753 (5.6)
Belgium	441,807 (4.8)

Region	COVID-19 cases, n (%)
Ukraine	387,481 (4.2)
Turkey	373,154 (4.1)
Netherlands	340,974 (3.7)
Poland	340,834 (3.7)
Czech Republic	323,673 (3.5)
Israel	314,244 (3.4)
Romania	235,586 (2.6)
Switzerland	153,728 (1.7)

Discussion

In this descriptive epidemiological study, we found ecological consistency between the seasonal influenza trends from 2017 to 2019 and the COVID-19 trends seen to date. Specifically, this similarity was highest in regions with more pronounced influenza seasonality, including North America, Europe and Central Asia, and Latin America and the Caribbean, where peaks of COVID-19 cases coincided with periods of high influenza activity. In East Asia and the Pacific and sub-Saharan Africa, there was more marked heterogeneity in COVID-19 pandemics in countries of these regions. However, within-country comparisons of 2017-2019 influenza epidemics and COVID-19 cases were evident irrespective of region. The ecological consistency in the seasonal trends of COVID-19 and the recent burden of influenza is suggestive of likely shared individual, structural, and environmental determinants of infection, which has implications for public health.

Outside of public health interventions, there is generally less consensus on the determinants that potentiate or inhibit COVID-19 transmission. However, given decades of study, there is more information available on the individual, population, and environmental determinants of influenza. At the individual level, underlying immunity is related to previous exposure to similar strains and vaccine uptake and effectiveness. Effectiveness of vaccines depends on circulating influenza strains (ie, pandemic and nonpandemic or seasonal) secondary to antigenic drift and shift of influenza viruses. Population-level characteristics include vaccine coverage and shield immunity, indoor air quality, population density, social norms, and health care infrastructure [30-32]. Environmental factors affecting influenza include temperature, air pollution, and humidity [30,33]. Given how rapidly COVID-19 has emerged, conclusions of different studies exploring the role of environmental conditions have differed significantly. One study found no associations between latitude and temperature with COVID-19 epidemic growth using a 2-week exposure period in March 2020 across 144 geopolitical areas [34]. However, a “living” review leveraging data from December 2019 onward from nearly 4000 locations worldwide showed significant differences in COVID-19 burden associated with temperature, humidity, pollution, and other environmental factors [35]. Collectively, these studies suggest the potential for at least some seasonal variation secondary to changes in environmental conditions in the transmission of SARS-CoV-2 similar to other

respiratory pathogens, including influenza and other seasonal coronaviruses [36-39].

Influenza seasonality is more pronounced in more temperate areas of the Northern and Southern Hemispheres, consistent with opposing fall and winter seasons. A recent assessment of the seasonality of infectious pathogens included individual-level determinants such as growing immunity throughout an influenza season secondary to vaccination or exposure that results in achieving effective levels of shield immunity, limiting transmission [40]. There has been some study of existing immunity to COVID-19 mediated via cross-reactivity with other coronaviruses [40,41]. However, decreased populationwide immunity could result in sustained transmission even during times with less transmission of other respiratory pathogens. Influenza seasonality may also be secondary to population-level changes, including potentially higher crowding in colder weather, increased use of respite care and housing, people returning to work and school, and increased use of transit services that may collectively facilitate transmission. Environmental conditions may further affect individual susceptibility to infection as well as contact or respiratory droplet transmission patterns. Although there is likely a relationship between testing infrastructure and influenza reporting, the consistency in the COVID-19 and influenza trends among countries within each region was notable. COVID-19 emerged in East Asia and the Pacific though there was relatively limited COVID-19 spread in several settings in March and April 2020, potentially secondary to widely implemented nonpharmacologic interventions [29]. Notably, and consistent with influenza season, there was a resurgence of COVID-19 in Australia starting in July and ending in September 2020 [42-44].

A hallmark of COVID-19 has been the role of social determinants of health with socioeconomic status, living conditions, access to health care, and underlying comorbidities being consistent predictors of COVID-19 incidence and mortality worldwide [45,46]. Moreover, structural factors including racism have been documented in several settings including the United States, the United Kingdom, Sweden, and Canada but are likely relevant in more settings not reporting on the racial identity of those affected [21,47,48]. Notably, similar disparities were observed in the 2009 H1N1 pandemic in the United States, though there has been limited documentation of socioeconomic disparities of influenza in most settings [22,49]. Although disparities are well documented for other infectious diseases such as HIV, these disparities evolve across cycles of

transmission spanning years. Given how rapidly disparities emerged with COVID-19, interventions to address disparities may be observable over shorter time horizons and provide insight into effective interventions to address pervasive social health inequities. Secondary to intersections of differential access to health care and medical mistrust, the development of a generally effective vaccine has not mitigated disparities [50–52]. Leveraging frameworks focused on optimizing implementation of influenza vaccination strategies including increased equity will be critical in informing SARS-CoV-2 vaccines when available [53,54]. Moreover, there has been substantial research into studying correlates of protection for influenza vaccine development as well as the role of cellular immunity in complementing the humoral response, which are both actively being studied for COVID-19 [55–58]. Finally, COVID-19 is rapidly increasing our understanding of postviral morbidities including chronic fatigue syndrome and chronic myofascial pain syndromes as well as potential cardiovascular complications [59]. Although postviral complications have been described for other viruses, the known exposure to COVID-19 as an etiological agent may provide further insight into the mechanisms and potential interventions [60–65].

This study has several limitations. The risk of ecological fallacy is relevant where individuals in generally lower burden regions or countries may have high risks of acquisition and transmission secondary to microepidemics, including in those in congregate living settings, refugee and migrant work camps, long-term care facilities, homeless shelters, and prisons. The missingness in data aggregated by FluNet is another limitation. We assumed that countries where influenza data was missing in a seasonal pattern was a reflection of few case counts or reduced reporting outside of the typical influenza season, and indeed, seasonal “missingness” was most apparent for countries in the European region, where the influenza season has been well documented. However, by excluding countries with >10% data missingness that did not follow a seasonal influenza pattern, several countries in the Pacific and in sub-Saharan Africa (eg, Sudan) were not included in the comparison of rankings despite high COVID-19 and influenza burden. There have been examples of COVID-19 resurgence outside of traditional influenza season, such as in South Africa, potentially due to the emergence of new variants of COVID-19, highlighting the complex intersection of transmission drivers in specific subnational areas [66]. Finally, there have been dramatic decreases in the burden of influenza in the 2020 season likely due to an intersection of nonpharmacological interventions’ effects combined with decreases in international travel. It is unclear how this will affect vulnerability to future seasons of influenza [67].

In addition, absolute influenza and COVID-19 case counts were reported rather than positivity rates; assuming that biases are present in interpreting positivity rates (eg, laboratory infrastructure implementing broad testing strategies), these biases are likely reflected in the total number of cases across both influenza and COVID-19. Heterogeneity in testing capacity,

test kit availability, and guidelines for prioritizing testing eligibility for COVID-19 may additionally challenge direct comparisons between COVID-19 and influenza case counts across countries. Testing capacity for COVID-19 has been highly dynamic, with continued limited capacity to meet screening and testing demands still evident in some countries [68]. In the case of influenza, countrywide variations in diagnostic testing strategies may belie differences in case detection. For instance, rapid influenza diagnostic tests are commonly used in clinical settings in many countries yet have documented lower sensitivity relative to other diagnostic methods [69].

Relatedly, contact tracing strategies for COVID-19 may vary across and within countries. As the pandemic has progressed, strategies have adapted to testing individuals with exposure to confirmed cases, irrespective of their symptomology [70]. This approach is not typically used for influenza. Finally, evidence suggests that nonpharmaceutical interventions for COVID-19 may concurrently limit influenza spread, due to their shared modes of transmission [29,71]. Although COVID-19 transmission has likely persisted due to unknown baseline population immunity, limited influenza detection may reflect decreased propensity for testing outside of the influenza season in some countries, as previously highlighted. Nonetheless, these data further underscore the potential utility to leverage our understanding and mitigation strategies for other respiratory pathogens to understand and combat COVID-19 and future pandemics due to respiratory pathogens.

The results presented here suggest shared regional differential distribution in the global burden of COVID-19 with that of influenza. This ecological overlap suggests that the underlying dynamics of influenza and COVID-19 transmission are similar, as they relate to spatial, seasonal, and population structure, and the disparities that lead to disproportionate vulnerabilities in various population groups. In considering the host–pathogen–environment framework commonly used for infectious diseases, emerging infections represent novel pathogens, but the populations and environment under which they emerge and spread are the same. Likewise, the similarities observed here suggest the potential utility of informing COVID-19 prevention and mitigation interventions, with the vast data available characterizing critical issues including the relative risks of indoor versus outdoor transmission; risks of aerosolization for influenza; and even optimal delivery of vaccines, chemoprophylaxis, and treatment strategies. Ultimately, regional and national public health systems preparing and responding to the current and future waves of COVID-19 may benefit from evaluating the relative burden of past respiratory pathogens and the impact of nonpharmacologic, pharmacologic, and prophylactic intervention strategies as prior information when forecasting and implementing interventions, and when evaluating the transmission impact of interventions applied to date in the context of potential seasonal effects on transmission.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Categories of missingness for 171 countries reporting influenza cases between January 2017 and December 2019 extracted from FluNet, overall and by region.

[DOCX File, 22 KB - [publichealth_v7i3e24696_app1.docx](#)]

Multimedia Appendix 2

Weekly trends in new COVID-19 cases (left panel, solid lines) and influenza cases (right panel, dashed lines) presented by World Bank region. Daily COVID-19 cases were extracted from the World Health Organization and are presented as a 7-day moving average between January 1 and October 31, 2020. Weekly influenza cases were extracted from FluNet and averaged by week across all reporting years (2017-2019).

[PDF File (Adobe PDF File), 2469 KB - [publichealth_v7i3e24696_app2.pdf](#)]

Multimedia Appendix 3

Weekly trends in new COVID-19 cases per capita (left panel, solid lines) and influenza cases per capita (right panel, dashed lines) presented by World Bank region. Daily COVID-19 cases were extracted from the World Health Organization and are presented as a 7-day moving average, per 100,000 persons, between January 1 and October 31, 2020. Weekly influenza cases were extracted from FluNet and averaged by week, per 100,000 persons, across all reporting years (2017-2019).

[PDF File (Adobe PDF File), 2429 KB - [publichealth_v7i3e24696_app3.pdf](#)]

Multimedia Appendix 4

Weekly trends in new COVID-19 cases (solid lines) and influenza cases (dashed lines) presented by country. Daily COVID-19 cases were extracted from the World Health Organization and are presented as a 7-day moving average between January 1 and October 31, 2020. Weekly influenza cases were extracted from FluNet and averaged by week across all reporting years (2017-2019).

[ZIP File (Zip Archive), 6293 KB - [publichealth_v7i3e24696_app4.zip](#)]

Multimedia Appendix 5

Weekly trends in new COVID-19 cases (solid lines) and influenza cases (dashed lines) presented for exemplary countries. Daily COVID-19 cases were extracted from the World Health Organization and are presented as a 7-day moving average between January 1 and October 31, 2020. Weekly influenza cases were extracted from FluNet and averaged by week across all reporting years (2017-2019).

[PDF File (Adobe PDF File), 1074 KB - [publichealth_v7i3e24696_app5.pdf](#)]

References

1. Weekly epidemiological update - 29 December 2020. World Health Organization. 2020. URL: <https://www.who.int/publications/m/item/weekly-epidemiological-update---29-december-2020> [accessed 2021-01-01]
2. Devlin H. Regional differences in Covid-19 transmission rate emerge in England. The Guardian. 2020 May 15. URL: <https://www.theguardian.com/world/2020/may/15/regional-differences-in-covid-19-transmission-rate-emerge-in-england> [accessed 2021-01-01]
3. CDC COVID-19 Response Team. Geographic differences in COVID-19 cases, deaths, and incidence - United States, February 12-April 7, 2020. MMWR Morb Mortal Wkly Rep 2020 Apr 17;69(15):465-471. [doi: [10.15585/mmwr.mm6915e4](https://doi.org/10.15585/mmwr.mm6915e4)] [Medline: [32298250](https://pubmed.ncbi.nlm.nih.gov/32298250/)]
4. Baral S, Rao A, Twahirwa Rwema JO, Lyons C, Cevik M, Kågesten AE, et al. Competing health risks associated with the COVID-19 pandemic and response: a scoping review. medRxiv. Preprint posted online January 8, 2021. [doi: [10.1101/2021.01.07.21249419](https://doi.org/10.1101/2021.01.07.21249419)] [Medline: [33442703](https://pubmed.ncbi.nlm.nih.gov/33442703/)]
5. Ahmed F, Zviedrite N, Uzicanin A. Effectiveness of workplace social distancing measures in reducing influenza transmission: a systematic review. BMC Public Health 2018 Apr 18;18(1):518 [FREE Full text] [doi: [10.1186/s12889-018-5446-1](https://doi.org/10.1186/s12889-018-5446-1)] [Medline: [29669545](https://pubmed.ncbi.nlm.nih.gov/29669545/)]
6. Long Y, Hu T, Liu L, Chen R, Guo Q, Yang L, et al. Effectiveness of N95 respirators versus surgical masks against influenza: a systematic review and meta-analysis. J Evid Based Med 2020 May;13(2):93-101 [FREE Full text] [doi: [10.1111/jebm.12381](https://doi.org/10.1111/jebm.12381)] [Medline: [32167245](https://pubmed.ncbi.nlm.nih.gov/32167245/)]
7. Feng S, Shen C, Xia N, Song W, Fan M, Cowling BJ. Rational use of face masks in the COVID-19 pandemic. Lancet Respir Med 2020 May;8(5):434-436 [FREE Full text] [doi: [10.1016/S2213-2600\(20\)30134-X](https://doi.org/10.1016/S2213-2600(20)30134-X)] [Medline: [32203710](https://pubmed.ncbi.nlm.nih.gov/32203710/)]

8. Lessler J, Reich NG, Brookmeyer R, Perl TM, Nelson KE, Cummings DA. Incubation periods of acute respiratory viral infections: a systematic review. *Lancet Infect Dis* 2009 May;9(5):291-300 [FREE Full text] [doi: [10.1016/S1473-3099\(09\)70069-6](https://doi.org/10.1016/S1473-3099(09)70069-6)] [Medline: [19393959](https://pubmed.ncbi.nlm.nih.gov/19393959/)]
9. Baud D, Qi X, Nielsen-Saines K, Musso D, Pomar L, Favre G. Real estimates of mortality following COVID-19 infection. *Lancet Infect Dis* 2020 Jul;20(7):773 [FREE Full text] [doi: [10.1016/S1473-3099\(20\)30195-X](https://doi.org/10.1016/S1473-3099(20)30195-X)] [Medline: [32171390](https://pubmed.ncbi.nlm.nih.gov/32171390/)]
10. Faust JS, Del Rio C. Assessment of deaths from COVID-19 and from seasonal influenza. *JAMA Intern Med* 2020 Aug 01;180(8):1045-1046. [doi: [10.1001/jamainternmed.2020.2306](https://doi.org/10.1001/jamainternmed.2020.2306)] [Medline: [32407441](https://pubmed.ncbi.nlm.nih.gov/32407441/)]
11. Yang W, Elankumaran S, Marr LC. Concentrations and size distributions of airborne influenza A viruses measured indoors at a health centre, a day-care centre and on aeroplanes. *J R Soc Interface* 2011 Aug 07;8(61):1176-1184 [FREE Full text] [doi: [10.1098/rsif.2010.0686](https://doi.org/10.1098/rsif.2010.0686)] [Medline: [21300628](https://pubmed.ncbi.nlm.nih.gov/21300628/)]
12. Marr LC, Tang JW, Van Mullekom J, Lakdawala SS. Mechanistic insights into the effect of humidity on airborne influenza virus survival, transmission and incidence. *J R Soc Interface* 2019 Jan 31;16(150):20180298 [FREE Full text] [doi: [10.1098/rsif.2018.0298](https://doi.org/10.1098/rsif.2018.0298)] [Medline: [30958176](https://pubmed.ncbi.nlm.nih.gov/30958176/)]
13. Yang W, Marr LC. Dynamics of airborne influenza A viruses indoors and dependence on humidity. *PLoS One* 2011;6(6):e21481 [FREE Full text] [doi: [10.1371/journal.pone.0021481](https://doi.org/10.1371/journal.pone.0021481)] [Medline: [21731764](https://pubmed.ncbi.nlm.nih.gov/21731764/)]
14. Moriyama M, Hugentobler WJ, Iwasaki A. Seasonality of respiratory viral infections. *Annu Rev Virol* 2020 Sep 29;7(1):83-101. [doi: [10.1146/annurev-virology-012420-022445](https://doi.org/10.1146/annurev-virology-012420-022445)] [Medline: [32196426](https://pubmed.ncbi.nlm.nih.gov/32196426/)]
15. Hancock K, Veguilla V, Lu X, Zhong W, Butler EN, Sun H, et al. Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus. *N Engl J Med* 2009 Nov 12;361(20):1945-1952. [doi: [10.1056/NEJMoa0906453](https://doi.org/10.1056/NEJMoa0906453)] [Medline: [19745214](https://pubmed.ncbi.nlm.nih.gov/19745214/)]
16. Allen JD, Ross TM. H3N2 influenza viruses in humans: viral mechanisms, evolution, and evaluation. *Hum Vaccin Immunother* 2018;14(8):1840-1847 [FREE Full text] [doi: [10.1080/21645515.2018.1462639](https://doi.org/10.1080/21645515.2018.1462639)] [Medline: [29641358](https://pubmed.ncbi.nlm.nih.gov/29641358/)]
17. Tamerius JD, Shaman J, Alonso WJ, Bloom-Feshbach K, Uejio CK, et al. Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. *PLoS Pathog* 2013 Mar;9(3):e1003194 [FREE Full text] [doi: [10.1371/journal.ppat.1003194](https://doi.org/10.1371/journal.ppat.1003194)] [Medline: [23505366](https://pubmed.ncbi.nlm.nih.gov/23505366/)]
18. Hirve S, Newman LP, Paget J, Azziz-Baumgartner E, Fitzner J, Bhat N, et al. Influenza seasonality in the tropics and subtropics - when to vaccinate? *PLoS One* 2016;11(4):e0153003 [FREE Full text] [doi: [10.1371/journal.pone.0153003](https://doi.org/10.1371/journal.pone.0153003)] [Medline: [27119988](https://pubmed.ncbi.nlm.nih.gov/27119988/)]
19. Saha S, Chadha M, Shu Y, Group of Asian Researchers on Influenza (GARI). Divergent seasonal patterns of influenza types A and B across latitude gradient in Tropical Asia. *Influenza Other Respir Viruses* 2016 May;10(3):176-184. [doi: [10.1111/irv.12372](https://doi.org/10.1111/irv.12372)] [Medline: [26781162](https://pubmed.ncbi.nlm.nih.gov/26781162/)]
20. Xu Z, Li Z, Hu W. Global dynamic spatiotemporal pattern of seasonal influenza since 2009 influenza pandemic. *Infect Dis Poverty* 2020 Jan 03;9(1):2 [FREE Full text] [doi: [10.1186/s40249-019-0618-5](https://doi.org/10.1186/s40249-019-0618-5)] [Medline: [31900215](https://pubmed.ncbi.nlm.nih.gov/31900215/)]
21. Millett GA, Jones AT, Benkeser D, Baral S, Mercer L, Beyrer C, et al. Assessing differential impacts of COVID-19 on black communities. *Ann Epidemiol* 2020 Jul;47:37-44 [FREE Full text] [doi: [10.1016/j.annepidem.2020.05.003](https://doi.org/10.1016/j.annepidem.2020.05.003)] [Medline: [32419766](https://pubmed.ncbi.nlm.nih.gov/32419766/)]
22. Quinn SC, Kumar S, Freimuth VS, Musa D, Casteneda-Angarita N, Kidwell K. Racial disparities in exposure, susceptibility, and access to health care in the US H1N1 influenza pandemic. *Am J Public Health* 2011 Feb;101(2):285-293 [FREE Full text] [doi: [10.2105/AJPH.2009.188029](https://doi.org/10.2105/AJPH.2009.188029)] [Medline: [21164098](https://pubmed.ncbi.nlm.nih.gov/21164098/)]
23. Bibbins-Domingo K. This time must be different: disparities during the COVID-19 pandemic. *Ann Intern Med* 2020 Aug 04;173(3):233-234 [FREE Full text] [doi: [10.7326/M20-2247](https://doi.org/10.7326/M20-2247)] [Medline: [32343767](https://pubmed.ncbi.nlm.nih.gov/32343767/)]
24. Kirby T. Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. *Lancet Respir Med* 2020 Jun;8(6):547-548 [FREE Full text] [doi: [10.1016/S2213-2600\(20\)30228-9](https://doi.org/10.1016/S2213-2600(20)30228-9)] [Medline: [32401711](https://pubmed.ncbi.nlm.nih.gov/32401711/)]
25. Phaswana-Mafuya N, Shisana O, Gray G, Zungu N, Bekker L, Kuonza L, et al. The utility of 2009 H1N1 pandemic data in understanding the transmission potential and estimating the burden of COVID-19 in South Africa to guide mitigation strategies. *S Afr Med J* 2020 May 25;110(7):576-577. [doi: [10.7196/SAMJ.2020.v110i7.14935](https://doi.org/10.7196/SAMJ.2020.v110i7.14935)] [Medline: [32880325](https://pubmed.ncbi.nlm.nih.gov/32880325/)]
26. Collins D. Peru's coronavirus response was 'right on time' - so why isn't it working? *The Guardian*. 2020 May 20. URL: <https://www.theguardian.com/global-development/2020/may/20/peru-coronavirus-lockdown-new-cases> [accessed 2021-01-01]
27. Weekly respiratory pathogens surveillance report - week 23. National Institute for Communicable Diseases. 2020. URL: <https://www.nicd.ac.za/wp-content/uploads/2020/06/Weekly-Pathogens-surveillance-report-week-23.pdf> [accessed 2021-01-01]
28. National COVID-19 daily report. National Institute for Communicable Diseases. 2020 Jun 15. URL: <https://www.nicd.ac.za/wp-content/uploads/2020/06/COVID19-Daily-Report-National-Public-15June2020.pdf> [accessed 2021-01-01]
29. Cowling BJ, Ali ST, Ng TWY, Tsang TK, Li JCM, Fong MW, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. *Lancet Public Health* 2020 May;5(5):e279-e288 [FREE Full text] [doi: [10.1016/S2468-2667\(20\)30090-6](https://doi.org/10.1016/S2468-2667(20)30090-6)] [Medline: [32311320](https://pubmed.ncbi.nlm.nih.gov/32311320/)]
30. Lofgren E, Fefferman NH, Naumov YN, Gorski J, Naumova EN. Influenza seasonality: underlying causes and modeling theories. *J Virol* 2007 Jun;81(11):5429-5436 [FREE Full text] [doi: [10.1128/JVI.01680-06](https://doi.org/10.1128/JVI.01680-06)] [Medline: [17182688](https://pubmed.ncbi.nlm.nih.gov/17182688/)]

31. Luongo JC, Fennelly KP, Keen JA, Zhai ZJ, Jones BW, Miller SL. Role of mechanical ventilation in the airborne transmission of infectious agents in buildings. *Indoor Air* 2016 Oct;26(5):666-678 [FREE Full text] [doi: [10.1111/ina.12267](https://doi.org/10.1111/ina.12267)] [Medline: [26562748](https://pubmed.ncbi.nlm.nih.gov/26562748/)]
32. Fuhrmann C. The effects of weather and climate on the seasonality of influenza: what we know and what we need to know. *Geography Compass* 2010 Jul;4(7):718-730. [doi: [10.1111/j.1749-8198.2010.00343.x](https://doi.org/10.1111/j.1749-8198.2010.00343.x)]
33. Lowen AC, Steel J. Roles of humidity and temperature in shaping influenza seasonality. *J Virol* 2014 Jul;88(14):7692-7695 [FREE Full text] [doi: [10.1128/JVI.03544-13](https://doi.org/10.1128/JVI.03544-13)] [Medline: [24789791](https://pubmed.ncbi.nlm.nih.gov/24789791/)]
34. Jüni P, Rothenbühler M, Bobos P, Thorpe KE, da Costa BR, Fisman DN, et al. Impact of climate and public health interventions on the COVID-19 pandemic: a prospective cohort study. *CMAJ* 2020 May 25;192(21):E566-E573 [FREE Full text] [doi: [10.1503/cmaj.200920](https://doi.org/10.1503/cmaj.200920)] [Medline: [32385067](https://pubmed.ncbi.nlm.nih.gov/32385067/)]
35. Xu R, Rahmandad H, Gupta M, DiGennaro C, Ghaffarzadegan N, Amini H, et al. The modest impact of weather and air pollution on COVID-19 transmission. *medRxiv*. Preprint posted online May 24, 2020.
36. Sajadi MM, Habibzadeh P, Vintzileos A, Shokouhi S, Miralles-Wilhelm F, Amoroso A. Temperature, humidity, and latitude analysis to estimate potential spread and seasonality of coronavirus disease 2019 (COVID-19). *JAMA Netw Open* 2020 Jun 01;3(6):e2011834 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.11834](https://doi.org/10.1001/jamanetworkopen.2020.11834)] [Medline: [32525550](https://pubmed.ncbi.nlm.nih.gov/32525550/)]
37. Carleton T, Meng KC. Causal empirical estimates suggest COVID-19 transmission rates are highly seasonal. *medRxiv*. Preprint posted online March 30, 2020. [doi: [10.1101/2020.03.26.20044420](https://doi.org/10.1101/2020.03.26.20044420)]
38. Ficitola GF, Rubolini D. Climate affects global patterns of Covid-19 early outbreak dynamics. *medRxiv*. Preprint posted online April 20, 2020.
39. Neher RA, Dyrda R, Druelle V, Hodcroft EB, Albert J. Potential impact of seasonal forcing on a SARS-CoV-2 pandemic. *Swiss Med Wkly* 2020 Mar 09;150:w20224. [doi: [10.4414/sm.w.2020.20224](https://doi.org/10.4414/sm.w.2020.20224)] [Medline: [32176808](https://pubmed.ncbi.nlm.nih.gov/32176808/)]
40. Cohen J. Why do dozens of diseases wax and wane with the seasons—and will COVID-19? *Science*. 2020 Mar 13. URL: <https://www.sciencemag.org/news/2020/03/why-do-dozens-diseases-wax-and-wane-seasons-and-will-covid-19> [accessed 2021-01-01]
41. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* 2020 Jun 25;181(7):1489-1501.e15 [FREE Full text] [doi: [10.1016/j.cell.2020.05.015](https://doi.org/10.1016/j.cell.2020.05.015)] [Medline: [32473127](https://pubmed.ncbi.nlm.nih.gov/32473127/)]
42. Rourke A, Davidson H. Hong Kong imposes restrictions as Covid-19 cases rise across Asia. *The Guardian*. 2020 Jul 27. URL: <https://www.theguardian.com/world/2020/jul/27/global-report-china-reports-biggest-covid-19-surge-in-months-as-australia-sets-new-infections-record> [accessed 2021-01-01]
43. Packham C. Australia suffers 'disturbing' COVID-19 surge, state curbs wedding parties. *Reuters*. 2020 Jul 16. URL: <https://www.reuters.com/article/us-health-coronavirus-australia-idUSKCN24I08U> [accessed 2021-01-01]
44. Packham C. Australia tightens restrictions as COVID-19 cases top 10,000. *Reuters*. 2020. URL: <https://www.reuters.com/article/us-health-coronavirus-australia-idUSKCN24F0AD> [accessed 2021-01-01]
45. Mishra S, Kwong JC, Chan AK, Baral SD. Understanding heterogeneity to inform the public health response to COVID-19 in Canada. *CMAJ* 2020 Jun 22;192(25):E684-E685. [doi: [10.1503/cmaj.201112](https://doi.org/10.1503/cmaj.201112)] [Medline: [32493741](https://pubmed.ncbi.nlm.nih.gov/32493741/)]
46. Okonkwo NE, Aguwa UT, Jang M, Barré IA, Page KR, Sullivan PS, et al. COVID-19 and the US response: accelerating health inequities. *BMJ Evid Based Med* 2020 Jun 03 [FREE Full text] [doi: [10.1136/bmjebm-2020-111426](https://doi.org/10.1136/bmjebm-2020-111426)] [Medline: [32493833](https://pubmed.ncbi.nlm.nih.gov/32493833/)]
47. Yaya S, Yeboah H, Charles CH, Otu A, Labonte R. Ethnic and racial disparities in COVID-19-related deaths: counting the trees, hiding the forest. *BMJ Glob Health* 2020 Jun;5(6) [FREE Full text] [doi: [10.1136/bmjgh-2020-002913](https://doi.org/10.1136/bmjgh-2020-002913)] [Medline: [32513864](https://pubmed.ncbi.nlm.nih.gov/32513864/)]
48. Hansson E, Albin M, Rasmussen M, Jakobsson K. Stora skillnader i överdödlighet våren 2020 utifrån födelseland. *Läkartidningen* 2020;117.
49. Weinman AL, Sullivan SG, Vijaykrishna D, Markey P, Levy A, Miller A, et al. Epidemiological trends in notified influenza cases in Australia's Northern Territory, 2007-2016. *Influenza Other Respir Viruses* 2020 Sep;14(5):541-550. [doi: [10.1111/irv.12757](https://doi.org/10.1111/irv.12757)] [Medline: [32445270](https://pubmed.ncbi.nlm.nih.gov/32445270/)]
50. Hall LL, Xu L, Mahmud SM, Puckrein GA, Thommes EW, Chit A. A map of racial and ethnic disparities in influenza vaccine uptake in the Medicare Fee-for-Service Program. *Adv Ther* 2020 May;37(5):2224-2235 [FREE Full text] [doi: [10.1007/s12325-020-01324-y](https://doi.org/10.1007/s12325-020-01324-y)] [Medline: [32274750](https://pubmed.ncbi.nlm.nih.gov/32274750/)]
51. Loiacono MM, Mahmud SM, Chit A, van Aalst R, Kwong JC, Mitsakakis N, et al. Patient and practice level factors associated with seasonal influenza vaccine uptake among at-risk adults in England, 2011 to 2016: an age-stratified retrospective cohort study. *Vaccine X* 2020 Apr 09;4:100054 [FREE Full text] [doi: [10.1016/j.jvacx.2020.100054](https://doi.org/10.1016/j.jvacx.2020.100054)] [Medline: [32072152](https://pubmed.ncbi.nlm.nih.gov/32072152/)]
52. Akpalu Y, Sullivan SJ, Regan AK. Association between health insurance coverage and uptake of seasonal influenza vaccine in Brazos County, Texas. *Vaccine* 2020 Feb 24;38(9):2132-2135. [doi: [10.1016/j.vaccine.2020.01.029](https://doi.org/10.1016/j.vaccine.2020.01.029)] [Medline: [32008878](https://pubmed.ncbi.nlm.nih.gov/32008878/)]
53. Wang L, Chen J, Marathe A. A framework for discovering health disparities among cohorts in an influenza epidemic. *World Wide Web* 2019 Nov;22(6):2997-3020 [FREE Full text] [doi: [10.1007/s11280-018-0608-8](https://doi.org/10.1007/s11280-018-0608-8)] [Medline: [31777450](https://pubmed.ncbi.nlm.nih.gov/31777450/)]

54. Webster F, Gidding H, Matthews V, Taylor R, Menzies R. What isn't measured isn't done - eight years with no progress in Aboriginal and Torres Strait Islander adult influenza and pneumococcal vaccination. *Aust N Z J Public Health* 2019 Dec;43(6):558-562. [doi: [10.1111/1753-6405.12944](https://doi.org/10.1111/1753-6405.12944)] [Medline: [31617660](https://pubmed.ncbi.nlm.nih.gov/31617660/)]
55. Ward BJ, Pillet S, Charland N, Trepanier S, Couillard J, Landry N. The establishment of surrogates and correlates of protection: useful tools for the licensure of effective influenza vaccines? *Hum Vaccin Immunother* 2018 Mar 04;14(3):647-656 [FREE Full text] [doi: [10.1080/21645515.2017.1413518](https://doi.org/10.1080/21645515.2017.1413518)] [Medline: [29252098](https://pubmed.ncbi.nlm.nih.gov/29252098/)]
56. Sekine T, Perez-Potti A, Rivera-Ballesteros O, Strålin K, Gorin JB, Olsson A, Karolinska COVID-19 Study Group. Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. *bioRxiv*. Preprint posted online June 29, 2020. [doi: [10.1101/2020.06.29.174888](https://doi.org/10.1101/2020.06.29.174888)]
57. Trombetta CM, Montomoli E. Influenza immunology evaluation and correlates of protection: a focus on vaccines. *Expert Rev Vaccines* 2016 Aug;15(8):967-976. [doi: [10.1586/14760584.2016.1164046](https://doi.org/10.1586/14760584.2016.1164046)] [Medline: [26954563](https://pubmed.ncbi.nlm.nih.gov/26954563/)]
58. Gallais F, Velay A, Wendling MJ, Nazon C, Partisani M, Sibilia J, et al. Intrafamilial exposure to SARS-CoV-2 induces cellular immune response without seroconversion. *medRxiv*. Preprint posted online June 22, 2020. [doi: [10.1101/2020.06.21.20132449](https://doi.org/10.1101/2020.06.21.20132449)]
59. Long B, Brady WJ, Koyfman A, Gottlieb M. Cardiovascular complications in COVID-19. *Am J Emerg Med* 2020 Jul;38(7):1504-1507 [FREE Full text] [doi: [10.1016/j.ajem.2020.04.048](https://doi.org/10.1016/j.ajem.2020.04.048)] [Medline: [32317203](https://pubmed.ncbi.nlm.nih.gov/32317203/)]
60. Maggio MC, Cimaz R, Alaimo A, Comparato C, Di Lisi D, Corsello G. Kawasaki disease triggered by parvovirus infection: an atypical case report of two siblings. *J Med Case Rep* 2019 Apr 24;13(1):104 [FREE Full text] [doi: [10.1186/s13256-019-2028-5](https://doi.org/10.1186/s13256-019-2028-5)] [Medline: [31014402](https://pubmed.ncbi.nlm.nih.gov/31014402/)]
61. Puttaraksa K, Pirttinen H, Karvonen K, Nykky J, Naides S, Gilbert L. Parvovirus B19V nonstructural protein NS1 induces double-stranded deoxyribonucleic acid autoantibodies and end-organ damage in nonautoimmune mice. *J Infect Dis* 2019 Apr 16;219(9):1418-1429 [FREE Full text] [doi: [10.1093/infdis/jiy614](https://doi.org/10.1093/infdis/jiy614)] [Medline: [30346568](https://pubmed.ncbi.nlm.nih.gov/30346568/)]
62. Iglesias-Gamarra A, Restrepo JF, Matteson EL. Small-vessel vasculitis. *Curr Rheumatol Rep* 2007 Aug;9(4):304-311. [doi: [10.1007/s11926-007-0049-3](https://doi.org/10.1007/s11926-007-0049-3)] [Medline: [17688840](https://pubmed.ncbi.nlm.nih.gov/17688840/)]
63. Crudele GDL, Amadasi A, Marasciuolo L, Rancati A, Gentile G, Zoja R. A case report of lethal post-viral lymphocytic myocarditis with exclusive location in the right ventricle. *Leg Med (Tokyo)* 2016 Mar;19:1-4. [doi: [10.1016/j.legalmed.2016.01.005](https://doi.org/10.1016/j.legalmed.2016.01.005)] [Medline: [26980245](https://pubmed.ncbi.nlm.nih.gov/26980245/)]
64. Smatti MK, Cyprian FS, Nasrallah GK, Al Thani AA, Almishal RO, Yassine HM. Viruses and autoimmunity: a review on the potential interaction and molecular mechanisms. *Viruses* 2019 Aug 19;11(8) [FREE Full text] [doi: [10.3390/v11080762](https://doi.org/10.3390/v11080762)] [Medline: [31430946](https://pubmed.ncbi.nlm.nih.gov/31430946/)]
65. Klok FA, Boon GJAM, Barco S, Endres M, Geelhoed JJM, Knauss S, et al. The Post-COVID-19 Functional Status scale: a tool to measure functional status over time after COVID-19. *Eur Respir J* 2020 Jul;56(1). [doi: [10.1183/13993003.01494-2020](https://doi.org/10.1183/13993003.01494-2020)] [Medline: [32398306](https://pubmed.ncbi.nlm.nih.gov/32398306/)]
66. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, et al. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv*. Preprint posted online December 22, 2020. [doi: [10.1101/2020.12.21.20248640](https://doi.org/10.1101/2020.12.21.20248640)]
67. Olsen SJ, Azziz-Baumgartner E, Budd AP, Brammer L, Sullivan S, Pineda RF, et al. Decreased influenza activity during the COVID-19 pandemic - United States, Australia, Chile, and South Africa, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Sep 18;69(37):1305-1309. [doi: [10.15585/mmwr.mm6937a6](https://doi.org/10.15585/mmwr.mm6937a6)] [Medline: [32941415](https://pubmed.ncbi.nlm.nih.gov/32941415/)]
68. Omori R, Mizumoto K, Chowell G. Changes in testing rates could mask the novel coronavirus disease (COVID-19) growth rate. *Int J Infect Dis* 2020 May;94:116-118 [FREE Full text] [doi: [10.1016/j.ijid.2020.04.021](https://doi.org/10.1016/j.ijid.2020.04.021)] [Medline: [32320809](https://pubmed.ncbi.nlm.nih.gov/32320809/)]
69. Chartrand C, Pai M. How accurate are rapid influenza diagnostic tests? *Expert Rev Anti Infect Ther* 2012 Jun;10(6):615-617. [doi: [10.1586/eri.12.49](https://doi.org/10.1586/eri.12.49)] [Medline: [22734950](https://pubmed.ncbi.nlm.nih.gov/22734950/)]
70. Centers for Disease Control and Prevention. 2020. URL: <https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/overview.html> [accessed 2021-01-01]
71. Wiemken TL, Shacham E. Identifying potential undocumented COVID-19 using publicly reported influenza-like-illness and laboratory-confirmed influenza disease in the United States: an approach to syndromic surveillance? *Am J Infect Control* 2020 Aug;48(8):961-963 [FREE Full text] [doi: [10.1016/j.ajic.2020.05.007](https://doi.org/10.1016/j.ajic.2020.05.007)] [Medline: [32437754](https://pubmed.ncbi.nlm.nih.gov/32437754/)]

Edited by G Eysenbach; submitted 01.10.20; peer-reviewed by A Safarnejad, P Nguyen; comments to author 20.10.20; revised version received 18.01.21; accepted 21.01.21; published 02.03.21.

Please cite as:

Baral SD, Rucinski KB, Twahirwa Rwema JO, Rao A, Prata Menezes N, Diouf D, Kamarulzaman A, Phaswana-Mafuya N, Mishra S
The Relationship Between the Global Burden of Influenza From 2017 to 2019 and COVID-19: Descriptive Epidemiological Assessment
JMIR Public Health Surveill 2021;7(3):e24696

URL: <https://publichealth.jmir.org/2021/3/e24696>

doi: [10.2196/24696](https://doi.org/10.2196/24696)

PMID: [33522974](https://pubmed.ncbi.nlm.nih.gov/33522974/)

©Stefan David Baral, Katherine Blair Rucinski, Jean Olivier Twahirwa Rwema, Amrita Rao, Neia Prata Menezes, Daouda Diouf, Adeeba Kamarulzaman, Nancy Phaswana-Mafuya, Sharmistha Mishra. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 02.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Bridging Learning in Medicine and Citizenship During the COVID-19 Pandemic: A Telehealth-Based Case Study

Thiago Cerqueira-Silva^{1,2*}, MS; Roberto Carreiro^{3*}, MSc; Victor Nunes¹, MD; Louran Passos⁴, MD; Bernardo F Canedo⁵, MD, PhD; Sofia Andrade⁵, MSc, MD; Pablo Ivan P Ramos³, MSc, PhD; Ricardo Khouri^{1,2}, MSc, PhD; Carolina Barbosa Souza Santos^{1,6}, MSc; Jedson Dos Santos Nascimento⁴, MD, PhD; Aurea Angélica Paste¹, MD; Ivan De Mattos Paiva Filho^{7,8}, MSc, MD; Marília Santini-Oliveira⁹, MD, PhD; Álvaro Cruz^{1,6}, MD, PhD; Manoel Barral-Netto^{1,2*}, MD, PhD; Viviane Boaventura^{1,2*}, MD, PhD

¹Universidade Federal da Bahia, Salvador, Brazil

²Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil

³Centro de Integração de Dados e Conhecimentos para Saúde-CIDACS, Fundação Oswaldo Cruz, Salvador, Brazil

⁴Hospital Santa Izabel, Salvador, Brazil

⁵Universidade do Estado da Bahia, Salvador, Brazil

⁶Fundação ProAR, Salvador, Brazil

⁷Universidade Salvador - UNIFACS, Salvador, Brazil

⁸Secretaria Municipal de Saúde de Salvador, Salvador, Brazil

⁹Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz, Salvador, Brazil

*these authors contributed equally

Corresponding Author:

Viviane Boaventura, MD, PhD

Universidade Federal da Bahia

Av. Adhemar de Barros, s/nº - Ondina

Salvador, 40170-110

Brazil

Phone: 55 (71) 3283 6017

Email: viviane.boaventura@fiocruz.br

Abstract

Background: COVID-19 presented great challenges for not only those in the field of health care but also those undergoing medical training. The burden on health care services worldwide has limited the educational opportunities available for medical students due to social distancing requirements.

Objective: In this paper, we describe a strategy that combines telehealth and medical training to mitigate the adverse effects of the COVID-19 pandemic.

Methods: A toll-free telescreening service, Telecoronavirus, began operations in March 2020. This service was operated remotely by supervised medical students and was offered across all 417 municipalities (14.8 million inhabitants) in the Brazilian state of Bahia. Students recorded clinical and sociodemographic data by using a web-based application that was simultaneously accessed by medical volunteers for supervision purposes, as well as by state health authorities who conducted epidemiological surveillance and health management efforts. In parallel, students received up-to-date scientific information about COVID-19 via short educational videos prepared by professors. A continuously updated triage algorithm was conceived to provide consistent service.

Results: The program operated for approximately 4 months, engaging 1396 medical students and 133 physicians. In total, 111,965 individuals residing in 343 municipalities used this service. Almost 70,000 individuals were advised to stay at home, and they received guidance to avoid disease transmission, potentially contributing to localized reductions in the spread of COVID-19. Additionally, the program promoted citizenship education for medical students, who were engaged in a real-life opportunity to fight the pandemic within their own communities. The objectives of the education, organization, and assistance domains of the Telecoronavirus program were successfully achieved according to the results of a web-based post-project survey that assessed physicians' and students' perceptions.

Conclusions: In a prolonged pandemic scenario, a combination of remote tools and medical supervision via telehealth services may constitute a useful strategy for maintaining social distancing measures while preserving some practical aspects of medical education. A low-cost tool such as the Telecoronavirus program could be especially valuable in resource-limited health care scenarios, in addition to offering support for epidemiological surveillance actions.

(*JMIR Public Health Surveill* 2021;7(3):e24795) doi:[10.2196/24795](https://doi.org/10.2196/24795)

KEYWORDS

medical education; surveillance; COVID-19; education; telehealth; training; impact; medical student; triage; epidemiology; monitoring

Introduction

SARS-CoV-2 emerged as a novel coronavirus in December 2019, linked to a cluster of respiratory diseases [1]. The pathogen was identified to be the causative agent of COVID-19, a disease that rapidly acquired the status of a pandemic with over 85.3 million cases reported globally as of January 2021, resulting in 1.85 million total deaths [1,2]. COVID-19 is characterized by fever, cough, and dyspnea, which can progress to severe pneumonia, requiring mechanical ventilation [3]. At the onset of the pandemic, in a scenario where no effective or safe antiviral treatments nor vaccines for SARS-CoV-2 infection were available [4,5], nonpharmacological interventions were central to COVID-19 mitigation strategies in most countries. The nonpharmacological interventions enforced included stay-at-home orders, strict lockdowns, contact tracing, and social distancing recommendations—all of which are evidence-based strategies designed to effectively reduce the rate of spread of SARS-CoV-2 [6]. However, one caveat of implementing social distancing is its potential to delay health assistance for patients with acute infection who fear seeking health care services [7,8]. In addition, policies designed to avoid crowding, limitations on the use of public transportation, and shortages in the availability of personal protective equipment restricted routine medical school activities during times of isolation [9].

Toward the end of March 2020, several European countries were already facing the collapse of health systems due to the scarcity of human and material resources amidst increased demand [10,11]. Additional emergency measures included recruiting senior medical students to perform frontline work to reduce the burden on formally trained health care workers [12]. The implementation of telemedicine solutions and automated risk screening solutions via web apps were also proposed as solutions to mitigate the pandemic [13].

To respond to the challenges presented by the COVID-19 pandemic, we proposed a telescreening solution, termed *Telecoronavirus*, to reduce the burden on the local public health system in Bahia and to limit the spread of disease. Medical students were invited to voluntarily participate in a task force that served the population of the Brazilian state of Bahia, consisting of 14.8 million individuals across 417 municipalities, with a territorial extension comparable to that of France. The Telecoronavirus program was operational as of March 24, 2020, that is, 18 days after the first COVID-19 case was confirmed in the state of Bahia. At that time, a total of 63 COVID-19 cases had been confirmed in 12 (2.9%) cities in the state of Bahia [14]. This paper details the strategy employed by the

Telecoronavirus program, which can be adapted for use in future waves of COVID-19, as well as other epidemics.

Methods

To provide screening assistance to the population during the pandemic and offer high-quality information regarding coronavirus symptoms, we engaged supervised medical students through a combination of telehealth services and medical training. The volunteer program, termed Telecoronavirus, involved students in their final three years of undergraduate activities and physicians, faculty members, and researchers. The program was also designed to (1) reduce the burden on the hospital and health services infrastructure by preventing patients with mild symptoms from making unnecessary visits and (2) limit the spread of the virus by reducing the circulation of patients with symptoms that did not warrant immediate medical attention.

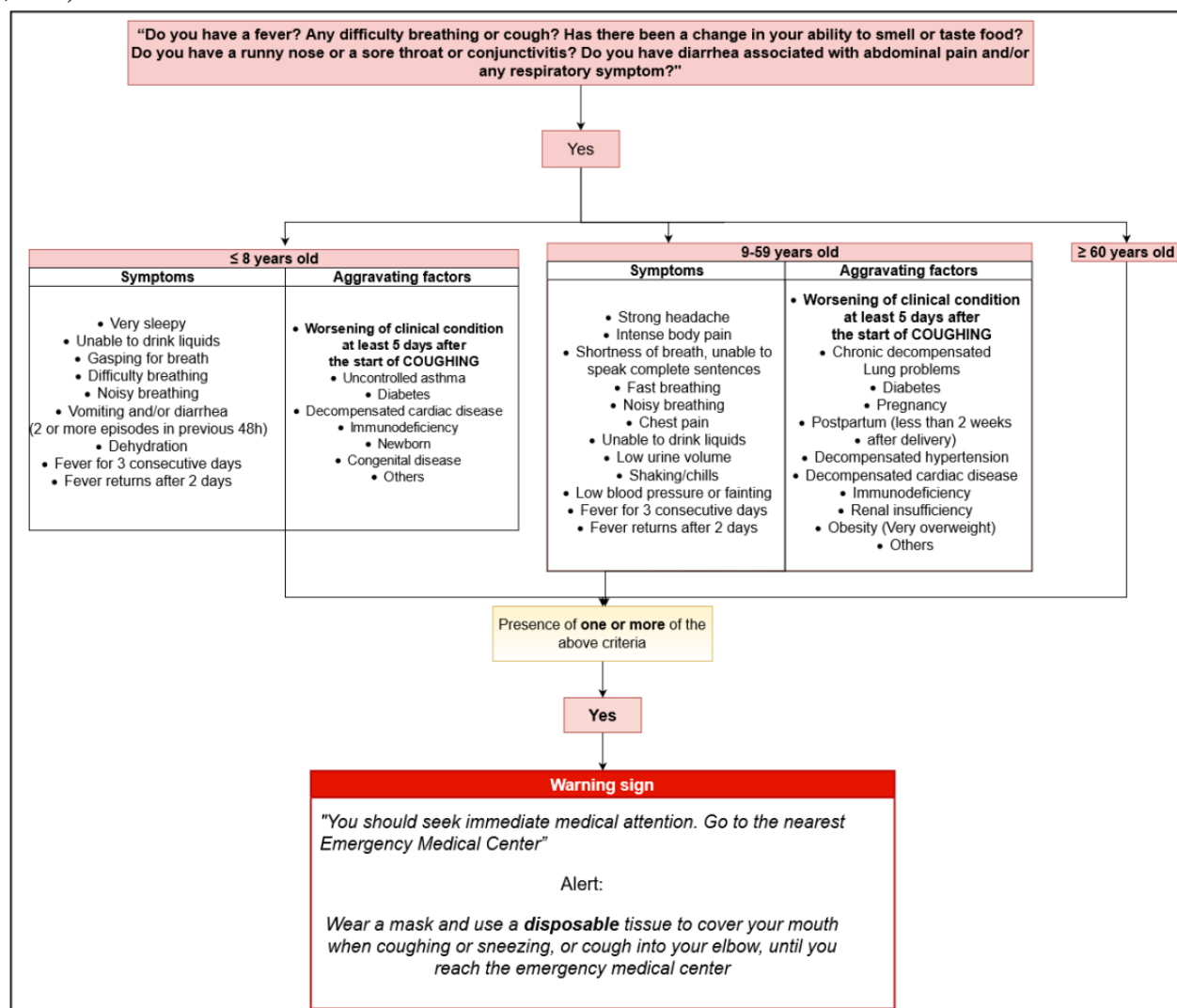
We created an organizational structure capable of interfacing academic institutions, government and regulatory bodies, and civil society. The recruitment of volunteers (students and physicians) was performed remotely using an online educational platform. Students and physicians had their identities respectively confirmed through official lists provided by medical schools and the regional medical board of the state of Bahia (Conselho Regional de Medicina do Estado da Bahia). Recruited students were required to complete video-based classes with instructions on the program's operational objectives and the screening algorithm employed, followed by a test to validate their understanding.

The Telecoronavirus screening algorithm was elaborated by a committee formed of members with expertise in pneumology, infectious diseases, emergency care, and telehealth, in addition to a representative from the regional medical board. A flowchart ([Multimedia Appendix 1](#)) was designed for volunteers to follow, instructing them to collect clinical data and provide general guidance on hygiene and social distancing by using colloquial language to aid users' comprehension and to standardize the information provided by the service.

The flowchart was constantly updated to reflect the most recent emerging scientific evidence on the novel coronavirus, undergoing 20 revisions during the project. Each version of the flowchart generated was identified with a date and link (plus a quick response or QR code), making it easy for all volunteers to access the most up-to-date version. Updated links were sent daily via a messaging application prior to commencing each shift at the service. [Figure 1](#) provides a summary of the final

version (number 20) of the flowchart, detailing clinical conditions that could lead to potentially poor outcomes.

Figure 1. Summary of the Telecoronavirus flowchart used by volunteers, detailing clinical conditions that could lead to potentially poor outcomes (Bahia, Brazil).



In parallel, professors prepared short educational videos synthesizing up-to-date scientific evidence to improve students' knowledge regarding COVID-19. In total, 20 videos were produced, exploring a range of diverse subjects such as epidemiology, clinical manifestations, special conditions (comorbidities and pregnancy), differential diagnoses, interpretation of diagnostic testing, airway management, psychological impacts, and socioethical aspects related to COVID-19.

A statewide toll-free number was set up, and operations were started on March 24, 2020. Call center employees redirected calls from potential patients with COVID-19 (ie, users) to the mobile phones of medical students who were remotely supervised by 133 physicians (residents or physician volunteers whose specialties were not directly applicable to frontline assistance). Two physicians supervised each group of 20 volunteers.

A custom web-based application was designed to record the users' demographic and clinical data, including address, sex, age, symptoms, comorbidities, and other clinical conditions

related to poor outcomes, as shown in Figure 1. Users with one or more of the clinical conditions described in the flowchart were advised to seek medical attention. The application incorporated tools to simultaneously allow data recording by the students, thereby allowing supervisors to examine the data input. The generated electronic database complied with ethical guidelines, and state health authorities were able to simultaneously access data for epidemiological surveillance purposes and aid the management of health resources. The total monthly cost for the Telecoronavirus program was about US \$10,000, including the expenses for a call center service, publicity, and technical support.

At the end of each evaluation, users were advised to either stay at home or seek emergency medical assistance. All users received educational guidance on how to reduce the spread of the virus.

In order to obtain volunteers' opinions about their experience with the Telecoronavirus program, a web-based post-project survey (Multimedia Appendix 2) was administered to assess physicians' and students' perceptions about the three domains

of the Telecoronavirus program: education, organization, and assistance. The education domain included learning about COVID-19, telemedicine skills, telehealth skills (survey questions 6, 11-14). The organization domain comprised remote supervision of the telehealth service and flowchart quality (survey questions 4-8). The assistance domain addressed the perception that patients followed recommendations, the social impact of the Telecoronavirus service, social commitment and citizenship, and sense of accomplishment (survey questions 9, 10, and 14).

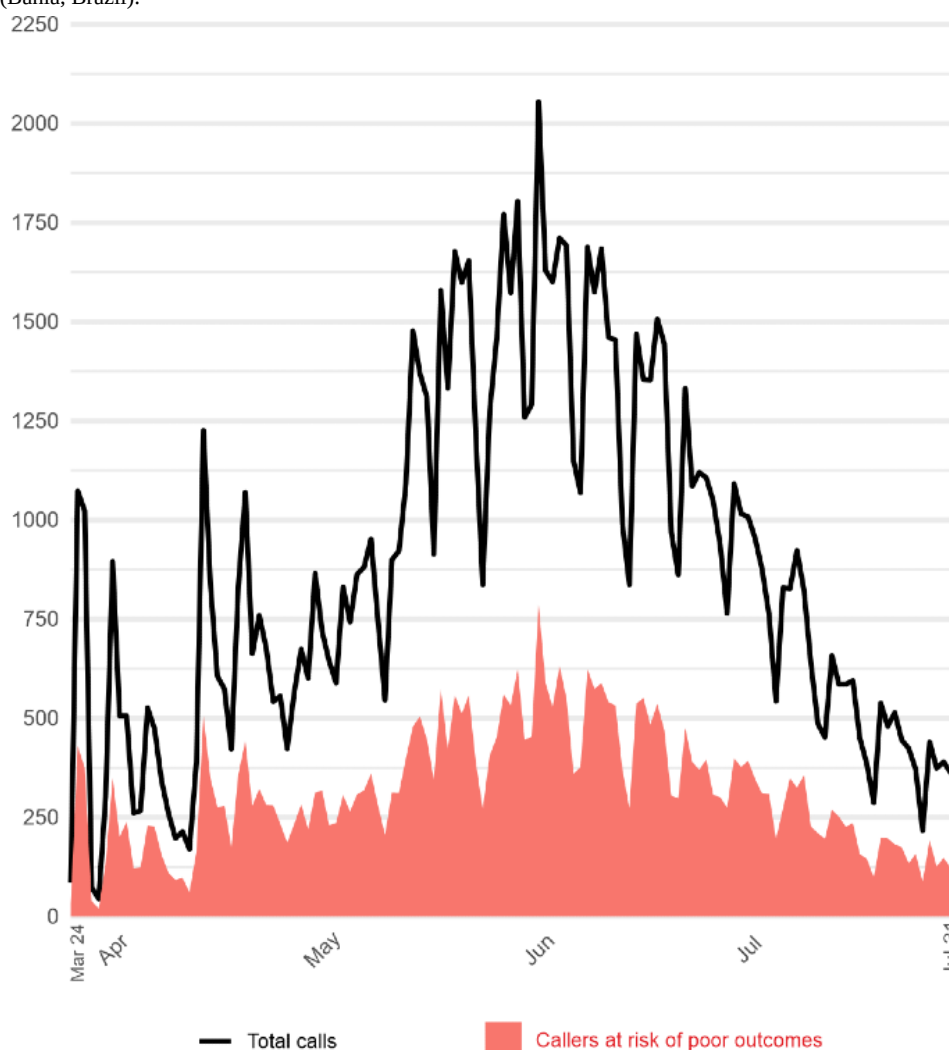
The survey aimed to evaluate the quality of service offered by the program, as well as its impact on technical education and citizenship learning. All volunteers were invited to submit their responses via the web-based survey. Survey responses are presented as mean and SD values, derived from a continuous rating scale ranging from 1 to 10, or as absolute values (expressed as n/N) with corresponding frequency (expressed as percentages) for categorical questions. Considering that this work entails an evaluation of public health service and that only anonymous data were collected via the surveys, the present case study was exempted from an ethics review in accordance with national research ethics guidelines (Resolution 466/12, Brazilian Health Council).

Results

The Telecoronavirus program operated between March 24 and July 31, 2020. Users could contact the center via a statewide toll-free number from anywhere within the state of Bahia. In partnership with the state government of Bahia, the service was publicized across the state using billboards and SMS text messages, in addition to multiple interviews broadcast on various television channels.

By July 31, that is, day 130 of commencing the operation, a total of 111,965 users who resided in 343 of the state's municipalities (ie, more than 80%) had contacted the service. With respect to the daily call volume, the highest demand occurred between May 14 and June 10, with a peak of 2055 calls received on June 1 (Figure 2). The majority of users were women (67,626/111,965, 60.4%) aged between 20 and 59 years (90,243/111,965, 80.6%). On average, the consultation or screening time lasted 8 minutes per call. A total of 41,123 (36.7%) callers were identified as being at risk of poor outcomes; they received recommendations to seek health care assistance at a local emergency medical unit.

Figure 2. Total number of calls received daily by Telecoronavirus, including the numbers of callers deemed at risk of poor outcomes between March 24 and July 31, 2020 (Bahia, Brazil).



The Telecoronavirus program recruited 1396 medical students as volunteers from 12 public and private medical schools across the state of Bahia. Among these, 1119 (80.2%) students were in their final two years of medical studies, and 921 (66%) were female. The supervisor group comprised 133 volunteer physicians, including residents and medical doctors.

The majority of volunteers reported that their experience bridged together the practical experiences of telescreening and provided them with learning resources, such as short videos and scientific texts, while also helping them acquire knowledge about COVID-19 by summarizing emerging scientific evidence (Table 1).

The participating physicians and students reported that the service greatly contributed to the development of telemedicine- or telehealth-related skills, as indicated by scores of 9.0 and 8.8, respectively, on a scale from 0 to 10. Furthermore, 99.6% (802/805) of volunteers indicated they would be interested in performing volunteer work in the health area again (Table 1).

Volunteers reported that the Telecoronavirus program was well-organized and that it operated smoothly. They found the screening algorithms to be clear and easy to use, and they positively evaluated the remote supervision method employed. The volunteers' perception was that patients followed the medical advice offered and adhered to the general guidance provided regarding hygiene and social distancing (Table 1).

Table 1. Evaluations from physician and student volunteers regarding the development of telemedicine- or telehealth-related skills while using the Telecoronavirus program (Bahia, Brazil).

Question topics	Evaluation scores (range 0-10)	
	Students (n=805)	Physicians (n=55)
Remote supervision of telehealth service, mean (SD)	8.63 (1.40)	— ^a
Flowchart quality, mean (SD)	9.34 (1.07)	9.02 (1.05)
Perception that patients followed recommendations, mean (SD)	8.63 (1.00)	—
Social impact of Telecoronavirus service, mean (SD)	9.16 (1.01)	9.25 (1.55)
Learning about COVID-19, mean (SD)	8.84 (1.34)	—
Telemedicine skills, mean (SD)	9.09 (1.07)	8.80 (1.64)
Telehealth skills, mean (SD)	9.33 (0.95)	—
Social commitment and citizenship, positive answer, n (%)	743 (92)	46 (84)
Sense of accomplishment, positive answer, n (%)	773 (96)	52 (95)

^aNot applicable, as these questions were addressed only to medical students.

Discussion

Principal Findings

Our experience revealed that the remote supervision of medical students by physicians, supported by clear guidelines, enabled us to offer a high-quality telescreening service in the face of scarce medical staff and materials in a pandemic scenario requiring social distancing. The Telecoronavirus program proved to be a simple, low-cost (approximate cost US \$0.36 per user), and far-reaching strategy (serving 82% of the state's municipalities), which can be readily applied in the context of other situations where health services are in high demand. Our experience shows that establishing partnerships between governmental and academic actors can provide valuable and innovative solutions for handling public health emergencies.

COVID-19 is predicted to remain a public health concern in the near future and will continue to impose challenges to both health care practices and medical training. A combination of remote tools with medical supervision can be applied to provide practical experience while still maintaining social distancing measures. The experimental method adopted effectively allowed medical students to contribute to mitigating the impact of the pandemic in conditions that required workforce reductions.

Considering that almost 70,000 users were advised to stay at home and received guidance on how to mitigate disease transmission, the Telecoronavirus program may have also contributed to a reduction in the spread of the virus by preventing the unnecessary circulation of mild cases. Modeling-based studies have provided evidence that undocumented infections (including mild cases) can be the source of almost 80% of reported COVID-19 cases [15]. Our approach was capable of leveraging technical aspects of medical knowledge, the potential of telemedicine services, and assistance provided by well-trained volunteers to provide free-of-charge health guidance to the state's population. This citizenship-promoting activity brought medical students closer to the real-life activities and challenges faced by the public health system in times of distress and vulnerability.

During the SARS epidemic of 2002, several provinces in China and Taiwan maintained hotline consultation services that provided medical counseling, including instructions on infection prevention and tracking of suspected cases.[16]. The main disadvantage of these services was the reallocation of health professionals to telescreening activities, reducing the number of regular medical personnel. As the present initiative recruited medical students who were supervised by physicians who were not working at the frontline, it was not necessary to reallocate any frontline medical workers.

Similar initiatives involving medical student engagement in an attempt to mitigate the COVID-19 pandemic have been previously described [17,18]. Medical students from the Ochsner Clinical School (New Orleans, LA, USA) participated in a hotline service, considerably reducing caller wait times and providing support for a symptom tracking program [17]. A similar project was developed in the state of Nevada (USA), wherein medical students participated in a telescreening service to assist underserved rural populations [18]. In both cases, students gathered in a physical environment on their university campus to carry out telescreening activities. In contrast, in the Telecoronavirus program, patient calls were redirected to medical students' mobile phones. Supervision was performed remotely, thereby eliminating the need for transportation and public gatherings, which further reduced the potential risk of SARS-CoV-2 infection among students.

In addition to technical aspects, the personal and social benefits of performing volunteer activities and contributing to the mitigation of the burden imposed by the pandemic in their respective communities may have raised social consciousness among the participating students and physicians.

As a strategy to advise patients suspected of COVID-19, telescreening conducted by medical students offers several advantages compared to automated online screening services: (1) the execution of a screening protocol by a qualified group minimizes interpretation bias and can provide rapid responses in unexpected situations, (2) users are not required to possess a minimum level of formal education level or be digitally literate in order to access the service, and (3) no internet access is required [19]. These points are particularly important in low- and middle-income countries, such as Brazil, where despite recent improvements, a digital divide remains throughout the country's vast territory, as well as across specific age groups, particularly among the elderly [20]. Additionally, the fact that health care team members provide a more humanized service increases adherence to social isolation guidelines and recommendations to seek medical care.

Limitations

Our work presented some limitations due to the challenging conditions imposed by the real-time mobilization during the

pandemic. The screening algorithm specifically targeted respiratory diseases. However, in addition to the COVID-19 pandemic, the state of Bahia faced two other concomitant outbreaks of acute viral infections: dengue and chikungunya. The absence of specific flowchart directions for suspected arbovirus infections resulted in the systematic guidance of patients to seek medical care.

It was unfortunately not possible to obtain formalized feedback from users, who could have provided valuable information to improve the quality of the service. However, some users did call back to notify volunteers that their symptoms had improved, and some even made a point to thank the volunteers for their service, according to reports by volunteers via the messaging application to their workgroups.

Although psychological assistance was made available to all project volunteers, our flowchart did not include protocols to address users' mental distress. To mitigate this deficiency, we provided practical guides and video lessons designed to help volunteers navigate situations in which users were experiencing psychological duress, as applicable within the context of telehealth services.

Several patients reported an inability to self-isolate at home due to social and economic conditions. In the final month of the program's operation, this issue was partially mitigated through referrals of patients to government shelters made available in the state's capital.

Conclusions

Remote tools for the early detection of respiratory infection outbreaks, provided at no cost to users, will be of paramount relevance to the strengthening and expansion of classical epidemiological surveillance activities.

Additionally, establishing connections between the provision of health care services and medical student education and experiential learning may also apply to patients with other illnesses, both acute and chronic, reducing the need for face-to-face consultations and the risk of infection transmission among patients, healthcare team members, and administrative staff.

Acknowledgments

The authors wish to thank Camila Vasconcelos, Felipe Torres, and Daniele Canedo for technical advice; Andris K Walter for English language revisions, critical review of the manuscript, and translating the *Telecoronavirus* flowchart. The authors are also grateful to the Government of the State of Bahia, the Bahia Medical Association, and FESF-tech for technical support, as well as the support provided by the educational institutions Faculdade de Medicina da Bahia, Universidade Federal da Bahia, Escola Bahiana de Medicina e Saúde Pública, Universidade do Estado da Bahia, Universidade Estadual de Feira de Santana, Universidade Estadual de Santa Cruz, Universidade Estadual do Sudoeste da Bahia, Universidade Federal do Recôncavo da Bahia, Universidade Federal do Sul da Bahia, UNIFACS, Uni-FTC, Uni-FG, and Unime. This study was funded by Fundação Oswaldo Cruz (Fiocruz) and Governo do Estado da Bahia-Brazil.

Conflicts of Interest

AC has received research grants and honoraria for advisory or educational services from the following companies: Sanofi, GSK, AstraZeneca, Chiesi, Mylan, Boehringer-Ingelheim, and Eurofarma. The remaining authors declare no conflicts.

Multimedia Appendix 1

Telecoronavirus flowchart (complete version) used by volunteers for screening possible COVID-19 cases (Bahia, Brazil).

[PDF File (Adobe PDF File), 154 KB - [publichealth_v7i3e24795_app1.pdf](#)]

Multimedia Appendix 2

Online post-project survey (Telecoronavirus program, Bahia, Brazil).

[DOCX File, 19 KB - [publichealth_v7i3e24795_app2.docx](#)]

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020 Feb 20;382(8):727-733. [doi: [10.1056/nejmoa2001017](#)]
2. COVID-19 United States Cases by County. Johns Hopkins University & Medicine - Coronavirus Resource Center. URL: <https://coronavirus.jhu.edu/us-map> [accessed 2021-02-26]
3. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* 2020 Feb;395(10223):497-506. [doi: [10.1016/s0140-6736\(20\)30183-5](#)]
4. Cao B, Wang Y, Wen D, Liu W, Wang J, Fan G, et al. A trial of lopinavir–ritonavir in adults hospitalized with severe Covid-19. *N Engl J Med* 2020 May 07;382(19):1787-1799. [doi: [10.1056/nejmoa2001282](#)]
5. Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, et al. Remdesivir for the treatment of Covid-19 — final report. *N Engl J Med* 2020 Nov 05;383(19):1813-1826. [doi: [10.1056/nejmoa2007764](#)]
6. Fong MW, Gao H, Wong JY, Xiao J, Shiu EYC, Ryu S, et al. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings-social distancing measures. *Emerg Infect Dis* 2020 May;26(5):976-984 [FREE Full text] [doi: [10.3201/eid2605.190995](#)] [Medline: [32027585](#)]
7. Marijon E, Karam N, Jost D, Perrot D, Frattini B, Derkenne C, et al. Out-of-hospital cardiac arrest during the COVID-19 pandemic in Paris, France: a population-based, observational study. *The Lancet Public Health* 2020 Aug;5(8):e437-e443. [doi: [10.1016/s2468-2667\(20\)30117-1](#)]
8. Lynn R, Avis J, Lenton S, Amin-Chowdhury Z, Ladhani S. Delayed access to care and late presentations in children during the COVID-19 pandemic: a snapshot survey of 4075 paediatricians in the UK and Ireland. *Arch Dis Child* 2021 Feb;106(2):e8 [FREE Full text] [doi: [10.1136/archdischild-2020-319848](#)] [Medline: [32586927](#)]
9. Rose S. Medical Student Education in the Time of COVID-19. *JAMA* 2020 Jun 02;323(21):2131-2132. [doi: [10.1001/jama.2020.5227](#)] [Medline: [32232420](#)]
10. Armocida B, Formenti B, Ussai S, Palestra F, Missoni E. The Italian health system and the COVID-19 challenge. *The Lancet Public Health* 2020 May;5(5):e253. [doi: [10.1016/s2468-2667\(20\)30074-8](#)]
11. Legido-Quigley H, Mateos-García JT, Campos VR, Gea-Sánchez M, Muntaner C, McKee M. The resilience of the Spanish health system against the COVID-19 pandemic. *The Lancet Public Health* 2020 May;5(5):e251-e252. [doi: [10.1016/s2468-2667\(20\)30060-8](#)]
12. Amante A, Balmer C. Italy rushes new doctors into service as coronavirus deaths rise above 2,500. 2020 Mar 17. URL: <https://www.reuters.com/article/us-health-coronavirus-italy/italy-rushes-to-promote-new-doctors-to-relieve-coronavirus-crisis-idUSKBN214245> [accessed 2021-02-26]
13. Herriman M, Meer E, Rosin R, Lee V, Washington V, Volpp KG. Asked and answered: building a chatbot to address Covid-19-related concerns. *NEJM Catal Innov Care Deliv* 2020 Jun 18 [FREE Full text] [doi: [10.1056/CAT.20.0230](#)]
14. Daily Bulletins - Covid-19 (Webpage in Portuguese). Governo Do Estado. URL: <http://www.saude.ba.gov.br/temasdesaude/coronavirus/boletins-diarios-covid-19/> [accessed 2021-02-26]
15. Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* 2020 May 01;368(6490):489-493 [FREE Full text] [doi: [10.1126/science.abb3221](#)] [Medline: [32179701](#)]
16. Gao Y, Liu R, Zhou Q, Wang X, Huang L, Shi Q, COVID-19 Evidence and Recommendations Working Group. Application of telemedicine during the coronavirus disease epidemics: a rapid review and meta-analysis. *Ann Transl Med* 2020 May;8(10):626 [FREE Full text] [doi: [10.21037/atm-20-3315](#)] [Medline: [32566563](#)]
17. Santos J, Chang D, Robbins K, Cam EL, Garbuzov A, Miyakawa-Liu M, et al. Answering the call: medical students reinforce health system frontlines through Ochsner COVID-19 hotline. *Ochsner J* 2020;20(2):144-145 [FREE Full text] [doi: [10.31486/toj.20.0065](#)] [Medline: [32612467](#)]
18. Carson S, Peraza LR, Pucci M, Huynh J. Student hotline improves remote clinical skills and access to rural care. *PRiMER* 2020 Sep 21;4:22. [doi: [10.22454/primer.2020.581719](#)]
19. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. *J Telemed Telecare* 2016 Oct 16;24(1):4-12. [doi: [10.1177/1357633x16674087](#)]
20. Nishijima M, Ivanauskas TM, Sarti FM. Evolution and determinants of digital divide in Brazil (2005–2013). *Telecommunications Policy* 2017 Feb;41(1):12-24. [doi: [10.1016/j.telpol.2016.10.004](#)]

Edited by T Sanchez; submitted 05.10.20; peer-reviewed by M Das, M Wright; comments to author 16.11.20; revised version received 06.01.21; accepted 14.01.21; published 04.03.21.

Please cite as:

Cerqueira-Silva T, Carreiro R, Nunes V, Passos L, Canedo BF, Andrade S, Ramos PIP, Khouri R, Santos CBS, Nascimento JDS, Paste AA, Paiva Filho IDM, Santini-Oliveira M, Cruz Á, Barral-Netto M, Boaventura V

Bridging Learning in Medicine and Citizenship During the COVID-19 Pandemic: A Telehealth-Based Case Study

JMIR Public Health Surveill 2021;7(3):e24795

URL: <https://publichealth.jmir.org/2021/3/e24795>

doi: [10.2196/24795](https://doi.org/10.2196/24795)

PMID: [33630746](https://pubmed.ncbi.nlm.nih.gov/33630746/)

©Thiago Cerqueira-Silva, Roberto Carreiro, Victor Nunes, Louran Passos, Bernardo F Canedo, Sofia Andrade, Pablo Ivan P Ramos, Ricardo Khouri, Carolina Barbosa Souza Santos, Jedson Dos Santos Nascimento, Aurea Angélica Paste, Ivan De Mattos Paiva Filho, Marília Santini-Oliveira, Álvaro Cruz, Manoel Barral-Netto, Viviane Boaventura. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 04.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Early Improvement of Acute Respiratory Distress Syndrome in Patients With COVID-19 in the Intensive Care Unit: Retrospective Analysis

Zhu Zhan^{1,2*}, PhD; Xin Yang^{3*}, MBBS; Hu Du¹, MM; Chuanlai Zhang¹, MM; Yuyan Song⁴, MM; Xiaoyun Ran¹, MM; An Zhang¹, PhD; Mei Yang⁵, MM

¹Department of Intensive Care Unit, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China

²Chongqing Key Laboratory of Ultrasound Molecular Imaging, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China

³Department of Respiratory and Geriatrics, Chongqing Public Health Medical Center, Chongqing, China

⁴Department of Intensive Care Unit, Chongqing Public Health Medical Center, Chongqing, China

⁵Department of Intensive Care Unit, Chongqing Sixth People's Hospital, Chongqing, China

*these authors contributed equally

Corresponding Author:

An Zhang, PhD

Department of Intensive Care Unit

The Second Affiliated Hospital of Chongqing Medical University

76 Linjiang Rd

Yuzhong District

Chongqing, 600010

China

Phone: 86 63693449

Email: zhangan@hospital.cqmu.edu.cn

Abstract

Background: Since the start of the COVID-19 pandemic, there have been over 2 million deaths globally. Acute respiratory distress syndrome (ARDS) may be the main cause of death.

Objective: This study aimed to describe the clinical features, outcomes, and ARDS characteristics of patients with COVID-19 admitted to the intensive care unit (ICU) in Chongqing, China.

Methods: The epidemiology of COVID-19 from January 21, 2020, to March 15, 2020, in Chongqing, China, was analyzed retrospectively, and 75 ICU patients from two hospitals were included in this study. On day 1, 56 patients with ARDS were selected for subgroup analysis, and a modified Poisson regression was performed to identify predictors for the early improvement of ARDS (eiARDS).

Results: Chongqing reported a 5.3% case fatality rate for the 75 ICU patients. The median age of these patients was 57 (IQR 25-75) years, and no bias was present in the sex ratio. A total of 93% (n=70) of patients developed ARDS during ICU stay, and more than half had moderate ARDS. However, most patients (n=41, 55%) underwent high-flow nasal cannula oxygen therapy, but not mechanical ventilation. Nearly one-third of patients with ARDS improved (arterial blood oxygen partial pressure/oxygen concentration >300 mm Hg) in 1 week, which was defined as eiARDS. Patients with eiARDS had a higher survival rate and a shorter length of ICU stay than those without eiARDS. Age (<55 years) was the only variable independently associated with eiARDS, with a risk ratio of 2.67 (95% CI 1.17-6.08).

Conclusions: A new subphenotype of ARDS—eiARDS—in patients with COVID-19 was identified. As clinical outcomes differ, the stratified management of patients based on eiARDS or age is highly recommended.

(JMIR Public Health Surveill 2021;7(3):e24843) doi:[10.2196/24843](https://doi.org/10.2196/24843)

KEYWORDS

acute respiratory distress syndrome; ARDS; Chongqing; COVID-19; critically ill; intensive care unit; outcome; characteristic; mortality; epidemiology; improvement

Introduction

Background

In December 2019, Wuhan, Hubei Province, China, reported a cluster of pneumonia cases of unknown cause, later identified as COVID-19 [1]. This contagious disease is caused by SARS-CoV-2. COVID-19 was declared a pandemic by the World Health Organization (WHO) on March 11, 2020 [2]. As of January 17, 2021, 1 year has passed since the pandemic began, with more than 93 million cases and 2 million deaths reported worldwide [3].

The leading cause of COVID-19-related death may be severe respiratory failure caused by acute respiratory distress syndrome (ARDS) [4]. According to some autopsy results, the lesions primarily appear in the lungs, characterized by diffuse alveolar damage. Other organs, such as heart tissue, have no obvious histological changes [5-7]. Previous studies reported that 48.6% of patients with COVID-19 had ARDS, of whom 29% died, and the mortality rate increased with the severity of ARDS [8,9]. Therefore, the characteristics of ARDS in patients with COVID-19 need to be fully understood.

This study aimed to describe the epidemiology, clinical features, laboratory findings, treatments, and outcomes of intensive care unit (ICU) patients with COVID-19 in Chongqing, China, which neighbors Hubei Province. In addition, a subphenotype of ARDS—early improvement of ARDS (eiARDS)—which occurred in about one-third of patients, was identified. eiARDS predicted a favorable clinical outcome.

Methods

Study Design and Participants

This retrospective cohort study included two cohorts of ICU patients from the Chongqing Public Health Medical Center and the Chongqing Three Gorges Central Hospital (Chongqing, China), both of which were designated hospitals for the treatment of patients with COVID-19 in Chongqing. Patients admitted to the ICU between January 21, 2020 (the date the first patient was admitted), and March 15, 2020 (the date the last patient was discharged during the first wave), were enrolled in this study.

Patients with COVID-19 were confirmed by a positive real-time reverse transcription–polymerase chain reaction (RT-PCR) assay using nasal swab specimens per WHO guidance [10]. The severity of COVID-19 was judged according to the Fifth Revised Trial Version of the Novel Coronavirus Pneumonia Diagnosis and Treatment Guidance of China [11]. Patients meeting any of the following criteria were defined as having a severe course of disease: (1) respiratory distress with a respiratory rate of more than 30 breaths per minute, (2) oxygen saturation $\leq 93\%$ in the resting state, and (3) arterial blood oxygen partial pressure (PaO_2)/oxygen concentration (FiO_2) ≤ 300 mm Hg. Patients meeting one of the following criteria were defined as critically ill: (1) mechanical ventilation needed as a result of respiratory failure, (2) shock, and (3) intensive care needed owing to the failure of other organs.

Data Collection

The two designated hospitals were Grade A hospitals in China, and all case data were retrieved from their electronic case system. Epidemiological and demographic data, symptoms, underlying diseases, comorbidities, treatments, clinical course, and outcome data of the patients were recorded in a spreadsheet. Vital signs, arterial blood gas analysis, laboratory data, acute physiology and chronic health evaluation II, and the sequential organ failure assessment score were recorded on specified dates (day 0: admission to hospital; day 1: admission to ICU; day 3, day 7, and day 14) for each patient. If any question emerged regarding the case data, clarification was sought from the treating team physician. As data collection was completed, another doctor was responsible for checking and integrating the data. The proportion of pneumonia volume was calculated by the Pulmonary Infection–Assisted Diagnosis System (V1.7.1) based on computed tomography (CT) imaging.

Definition

ARDS was diagnosed according to the Berlin definition [12]. Liver injury was diagnosed according to the following criteria: alanine aminotransferase >3 upper limit of normal (ULN), aspartate aminotransferase >3 ULN, or total bilirubin >2 ULN, regardless of a chronic liver disease diagnosis [13]. Acute kidney injury was diagnosed on the basis of serum creatinine [14]. Cardiac injury was diagnosed if the serum concentration of hypersensitive cardiac troponin T was greater than the upper limit of the reference range (>14 pg/mL). Cessation of viral shedding was defined as two consecutive negative nasal swab PCR tests (with an interval of least 24 hours).

Statistical Analysis

SPSS 26.0 (IBM Corp) was used for statistical analysis. Normally distributed continuous variables were presented as mean (SD), and the independent Student *t* test was used for comparison between two groups. The continuous variables that did not meet the criteria for normal distribution were presented as the median (IQR), and the Mann-Whitney *U* test was used to compare differences between two groups. Categorical variables were summarized using frequencies and percentages, and the chi-square test or the Fisher exact test was used to compare two or more groups. A modified Poisson regression analysis (sandwich variance estimator) was performed to identify the predictors of eiARDS; variables with $P < .05$ in the univariate analysis were entered into the multivariate Poisson regression analysis. All tests were two sided, and $P < .05$ was considered statistically significant.

Ethics Approval and Consent to Participate

The study was approved by the Research Ethics Committee of the Second Affiliated Hospital of Chongqing Medical University, the Chongqing Public Health Medical Center, and the Chongqing Three Gorges Central Hospital. Written informed consent was waived by the ethics committees of the designated hospitals for emerging infectious diseases.

Availability of Data and Materials

Original data can be requested from the corresponding author.

Results

Clinical Characteristics of ICU Patients in Chongqing, China

From January 21 to March 15, 2020, Chongqing reported 576 new cases of COVID-19 and 6 deaths. In this study, 75 ICU patients from two hospitals were recruited, including 48 severe patients, 27 critically ill patients, and 4 deceased patients.

A comparison of clinical characteristics between the two groups is shown in Table 1. The median age of the 75 patients was 57 years, and no bias was found in the sex ratio. Smoking was more prevalent among critically ill patients than in severe patients. The most frequent chronic medical illnesses were diabetes and hypertension. The most common symptoms were cough, fever, and dyspnea. In addition, 2 patients experienced an asymptomatic period prior to hospitalization and were dyspneic without fever during their hospital stay.

ARDS developed in most patients ($n=70$, 93%), and more than half ($n=38$, 51%) had moderate ARDS (Table 1). However, most patients ($n=41$, 55%) were supported with high-flow nasal cannula (HFNC) oxygen therapy (21 patients also received ventilation during their stay in the ICU). Only 26 (35%) patients received noninvasive ventilation (7 patients also received

invasive ventilation during their stay in the ICU), and 7 (9%) patients received invasive ventilation. Other supportive treatments included prone-position ventilation in 7 patients, extracorporeal membrane oxygenation in 3 patients, renal replacement therapy in 3 patients, and vasoconstrictive agents in 7 patients. Although bacterial pneumonia was identified by microbiological culture of sputum or alveolar lavage fluid in only 4 patients, antibacterial agents were administered to 62 patients and antifungal agents to 12 patients. The liver was the most commonly injured extrapulmonary organ, followed by the heart and kidneys. Lymphopenia was a very noteworthy feature in these patients, with a lower incidence of leukopenia and thrombocytopenia. Antiviral agents were used in all patients; the combination of an Aluvia (lopinavir and ritonavir) tablet and interferon-alpha was the most commonly used antiviral formula ($n=69$, 92%), and 4 patients (5.3%) also used ribavirin. Glucocorticoids and intravenous immunoglobulin were more commonly used in critically ill patients due to anti-inflammation properties and ability to neutralize SARS-CoV-2, respectively. Thymopeptide (thymosin alpha 1 or thymopentin) was used in most patients ($n=63$, 84%) to improve antiviral immunity. Traditional Chinese medicine, which is made by decoction of more than a dozen kinds of herbs, was used in 87% ($n=65$) of patients owing to potential antiviral and anti-inflammation properties.

Table 1. Clinical characteristics of severe or critically ill patients with COVID-19 admitted to the intensive care unit in Chongqing, China (from January 21 to March 15, 2020).

Characteristic	Total (N=75)	Severe (n=48)	Critically ill (n=27) ^a	P value
Age (years), median (IQR)	57 (25-75)	56 (47-70)	63 (51-69)	.57
Sex				.64
Female	36 (48)	24 (50)	12 (44)	
Male	39 (52)	24 (50)	15 (56)	
Smoking	9 (12)	1 (2)	8 (30)	.002
Exposure				.052
Recent travel to Hubei Province	16 (21)	13 (27)	3 (11)	
Contact with patients from Hubei Province	24 (32)	18 (38)	6 (22)	
Contact with confirmed patients	17 (23)	7 (15)	10 (37)	
No definite epidemiological link	18 (24)	10 (21)	8 (30)	
Chronic medical illness				
Hypertension	14 (19)	8 (17)	6 (22)	.55
Diabetes	20 (27)	12 (25)	8 (30)	.66
Chronic cardiac disease	7 (9)	6 (13)	1 (4)	.40
Chronic obstructive pulmonary disease	4 (5)	4 (8)	0 (0)	.31
Malignancy	1 (1)	1 (2)	0 (0)	>.99
Symptoms				
Fever	51 (68)	28 (58)	23 (85)	.02
Cough	62 (83)	40 (83)	22 (81)	>.99
Expectoration	29 (39)	21 (44)	8 (30)	.23
Dyspnea	43 (57)	23 (48)	20 (74)	.03
Myalgia	20 (27)	12 (25)	8 (30)	.66
Headache	9 (12)	7 (15)	2 (7)	.58
Diarrhea	7 (9)	4 (8)	3 (11)	>.99
Comorbidities				
Acute respiratory distress syndrome^b	70 (93)	43 (90)	27 (100)	.01
None	5 (7)	5 (10)	0 (0)	
Mild	10 (13)	9 (19)	1 (4)	
Moderate	38 (51)	25 (52)	13 (48)	
Severe	22 (29)	9 (19)	13 (48)	
Pneumothorax	1 (1)	0 (0)	1 (4)	.77
Bacterial pneumonia ^c	4 (5)	3 (6)	1 (4)	>.99
Cardiac injury	14 (19)	7 (15)	7 (26)	.23
Liver injury	19 (25)	10 (21)	9 (33)	.23
Kidney injury	8 (11)	5 (10)	3 (11)	>.99
Shock	7 (9)	0 (0)	7 (26)	.001
Leukopenia	19 (25)	11 (23)	8 (30)	.52
Lymphopenia	71 (95)	46 (96)	25 (93)	.95
Thrombocytopenia	20 (27)	9 (19)	11 (41)	.04
Treatment				
High-flow nasal cannula	41 (55)	20 (42)	21 (78)	.003

Characteristic	Total (N=75)	Severe (n=48)	Critically ill (n=27) ^a	P value
Mechanical ventilation				
Noninvasive	26 (35)	0 (0)	26 (96)	<.001
Invasive	7 (9)	0 (0)	7 (26)	.001
Prone position ventilation	7 (9)	0 (0)	7 (26)	.001
Extracorporeal membrane oxygenation	3 (4)	0 (0)	3 (11)	.08
Renal replacement therapy	3 (4)	0 (0)	3 (11)	.08
Vasoconstrictive agents	7 (9)	0 (0)	7 (26)	.001
Antiviral agents	75 (100)	48 (100)	27 (100)	— ^d
Antibacterial agents	62 (83)	37 (77)	25 (93)	.17
Antifungal	12 (16)	4 (8)	8 (30)	.04
Glucocorticoids	46 (61)	22 (46)	24 (89)	<.001
Immunoglobulin	30 (40)	13 (27)	17 (63)	.002
Thymopeptides	63 (84)	37 (77)	26 (96)	.06
Traditional Chinese medicine	65 (87)	41 (85)	24 (89)	.94

^aFour patients who died in the ICU were included.

^bARDS stages were defined by the worst PaO₂/FiO₂ value.

^cBacterial pneumonia was confirmed by sputum or alveolar lavage fluid culture.

^dNot applicable.

Clinical Course and Outcomes

The clinical course and outcomes of patients with COVID-19 in Chongqing are shown in Table 2. Chongqing reported 6 deaths from COVID-19 up to March 15, 2020, with a mortality rate of 1.04% in all 576 patients. As 2 patients died in the emergency department, only 4 deceased patients with clinical

data were included in this study, with a 28-day case fatality rate of 5.3% and a 28-day mechanical ventilation dependency of 1.3% among ICU patients. The duration from any initial symptom to diagnosis confirmed by PCR test was 5 days, to ARDS was 7 days, to ICU admission was 8 days, and to death was 16 days. The length of ICU stay was 13 days and hospital stay was 22 days.

Table 2. Clinical course and outcomes of patients with COVID-19 admitted to the intensive care unit (ICU) in Chongqing, China (from January 21 to March 15, 2020).

Clinical course and outcomes	Value
Duration from initial symptom(s) to...^a (days), median (IQR)	
Diagnosis confirmed by PCR ^b test	5 (2-7)
Hospital admission	7 (4-8)
Acute respiratory distress syndrome	7 (6-10)
ICU admission	8 (6-11)
Ventilation	10 (7-14)
Cessation of viral shedding ^c	20 (16-26)
Death ^d	16 (15-28)
Length of ICU stay	13 (9-19)
Length of hospital stay	22 (16-34)
Outcomes (n=75), n (%)	
28-day mortality ^d	4 (5.3)
28-day mechanical ventilation dependency	1 (1.3)
Location of death (n=6), n (%)	
ICU	4 (66.7)
Emergency department	2 (33.3)

^aTwo patients without any symptoms until hospital admission were excluded from statistical analysis.

^bPCR: polymerase chain reaction.

^cTwo consecutive negative nasal swab PCR tests performed at an interval of least 24 hours).

^dFour patients who died in the ICU were included.

Early Improvement of ARDS

In order to clarify the characteristics of ARDS in patients with COVID-19, 56 patients with ARDS ($\text{PaO}_2/\text{FiO}_2 < 300$ mm Hg) on day 1 (first day of ICU admission) were included for a subgroup analysis. These patients were then divided into two groups based on the severity of illness on day 7: eiARDS patients with $\text{PaO}_2/\text{FiO}_2 \geq 300$ mm Hg and non-eiARDS patients with $\text{PaO}_2/\text{FiO}_2 < 300$ mm Hg (Figure 1A).

Of the 56 patients with ARDS, a total of 18 patients had eiARDS. No significant differences were found in $\text{PaO}_2/\text{FiO}_2$ on day 1 (Figure 1B), the proportion of pneumonia volume on day 1 (Figure 1C), and the rate of ventilator usage ($\chi^2=2.46$,

$P=.12$) between the two groups. Predictably and regrettably, all the four deceased patients did not have eiARDS. Moreover, patients with eiARDS spent a shorter duration in the ICU than those without (10.5 days, IQR 8-16 vs 18 days, IQR 13-22, respectively) ($P=.001$; Figure 1D).

A Poisson regression analysis was performed to determine the factors associated with eiARDS. Table 3 shows that two variables (age and white blood cell) with $P<.05$ in the univariate analysis were chosen for multivariable analysis. Age (<55 years) was the only variable independently associated with eiARDS, with a risk ratio of 2.67 (95% CI 1.17-6.08). This finding indicated that patients younger than 55 years were 2.67 times more likely to have eiARDS than older people.

Figure 1. Comparison of patients with and without early improvement of acute respiratory distress syndrome (eiARDS). (A) All patients with ARDS on day 1, divided into two groups (eiARDS and non-eiARDS) according to $\text{PaO}_2/\text{FiO}_2$ on day 7. (B) No significant difference between the two groups in the $\text{PaO}_2/\text{FiO}_2$ on day 1. (C) No significant difference between the two groups in the proportion of pneumonia volume on day 1. (D) The length of intensive care unit stay exhibited differences between the two groups.

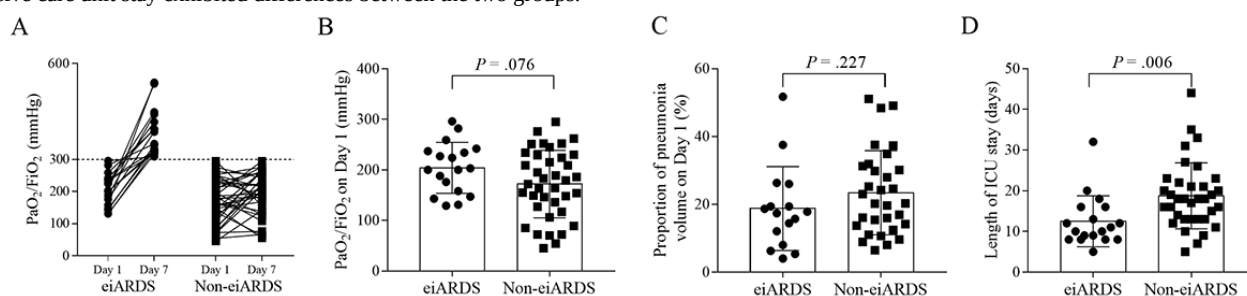


Table 3. Univariate and multivariate analysis of predictors for the early improvement of acute respiratory distress syndrome (eiARDS) in patients with COVID-19 in Chongqing, China (from January 21 to March 15, 2020).

Characteristic	eiARDS (n=18)	Non-eiARDS (n=38)	Relative risk (95% CI)	P value
Demographic and clinical characteristics				
Age (years)^a, n (%)				
≥55	6 (33.3)	26 (68.4)	1 (reference)	— ^b
<55 (univariate)	12 (66.7)	12 (31.6)	2.67 (1.17-6.08)	.02
<55 (multivariate)	—	—	2.36 (1.05-5.23)	.04
Male sex (vs female), n (%)	12 (66.7)	19 (50.0)	1.61 (0.71-3.70)	.26
Time of symptom onset to intensive care unit admission, median (days)	7.5	8	1.02 (0.92-1.14)	.69
Smoking, n (%)	4 (22.2)	4 (10.5)	1.71 (0.75-3.90)	.20
Hypertension, n (%)	2 (11.2)	7 (18.4)	0.65 (0.18-2.36)	.52
Diabetes, n (%)	4 (22.2)	9 (26.3)	0.86 (0.34-2.18)	.75
Chronic obstructive pulmonary disease, n (%)	1 (5.6)	3 (7.9)	0.77 (0.13-4.36)	.76
APACHE II ^c , median	6.5	9	0.90 (0.76-1.05)	.17
SOFA ^d , median	4	3	0.97 (0.67-1.42)	.89
Temperature, n (%)				
<37.3 °C	15 (83.3)	18 (47.4)	1 (reference)	—
≥37.3 °C (univariate)	3 (16.7)	20 (52.6)	0.33 (0.11-1.02)	.054
Heart rate per min, n (%)				
<100	13 (72.2)	33 (86.8)	1 (reference)	—
≥100	5 (27.8)	5 (13.2)	1.77 (0.82-3.83)	.15
Respiratory rate per min, n (%)				
<30	15 (83.3)	37 (97.4)	1 (reference)	—
≥30	3 (16.7)	1 (2.6)	1.17 (0.48-2.86)	.73
Systolic pressure, n (%)				
<140 mm Hg	16 (88.9)	35 (92.1)	1 (reference)	—
≥140 mm Hg	2 (11.1)	3 (7.9)	1.28 (0.41-4.02)	.68
Ventilation (vs non)	2 (11.1)	5 (13.2)	0.88 (0.25-3.02)	.83
Proportion of pneumonia volume, median	17.6	20.3	0.98 (0.94-1.02)	.28
Laboratory findings				
pH, n (%)				
7.35-7.45	4 (22.2)	12 (31.6)	1 (reference)	—
>7.45	14 (77.8)	26 (65.8)	1.44 (0.56-3.70)	.45
PaCO₂^e, n (%)				
34-45 mm Hg	9 (50.0)	21 (55.3)	1 (reference)	—
≤34 mm Hg	8 (44.4)	17 (44.7)	1.07 (0.48-2.35)	.87
PaO₂^f, n (%)				
≥60 mm Hg	14 (77.8)	23 (60.5)	1 (reference)	—
<60 mm Hg	4 (22.2)	15 (39.5)	0.56 (0.21-1.46)	.23
White blood cell count (×10⁹/L)^a, n (%)				
<4 (univariate)	7 (38.9)	6 (16.2)	2.21 (1.04-4.73)	.04

Characteristic	eiARDS (n=18)	Non-eiARDS (n=38)	Relative risk (95% CI)	P value
<4 (multivariate)	—	—	1.94 (0.97-3.86)	.06
4-10	9 (50.0)	28 (75.7)	1 (reference)	—
>10 (univariate)	2 (11.1)	3 (8.1)	1.64 (0.49-5.54)	.42
>10 (multivariate)	—	—	1.37 (0.48-3.89)	.55
Lymphocyte ($\times 10^9$), median	0.8	0.8	2.18 (0.79-6.02)	.13
Platelet ($\times 10^9/L$), n (%)				
≥100	16 (88.9)	35 (94.6)	1 (reference)	—
<100	2 (11.1)	2 (5.4)	1.59 (0.55-4.60)	.39
Potassium, n (%)				
3.5-4.5 mmol/L	14 (77.8)	25 (67.6)	1 (reference)	—
<3.5 mmol/L	3 (16.7)	11 (29.7)	0.61 (0.21-1.82)	.38
Sodium, n (%)				
135-145 mmol/L	10 (55.6)	20 (54.1)	1 (reference)	—
<135 mmol/L	8 (44.4)	16 (43.2)	1.00 (0.47-2.14)	>.99
Albumin, n (%)				
≥40 g/L	3 (17.6)	4 (10.8)	1 (reference)	—
<40 g/L	14 (82.4)	33 (89.2)	0.70 (0.27-1.82)	.46
Total bilirubin, n (%)				
≤17.1 $\mu\text{mol/L}$	10 (55.6)	27 (71.1)	1 (reference)	—
>17.1 $\mu\text{mol/L}$	8 (44.4)	11 (28.9)	1.56 (0.74-3.29)	.25
Alanine aminotransferase, n (%)				
≤40 U/L	13 (72.2)	23 (60.5)	1 (reference)	—
>40 U/L	5 (27.8)	15 (39.5)	0.69 (0.29-1.66)	.41
Creatine kinase, n (%)				
≤200 U/L	13 (81.2)	28 (75.7)	1 (reference)	—
>200 U/L	3 (18.8)	9 (24.3)	0.67 (0.27-2.32)	.79
High-sensitivity cardiac troponin T, n (%)				
≤0.014 ng/mL	9 (90.0)	19 (86.4)	1 (reference)	—
>0.014 ng/mL	1 (10.0)	3 (13.6)	0.78 (0.13-4.62)	.78
D-dimer, n (%)				
≤0.55 $\mu\text{g/L}$	9 (81.8)	15 (39.5)	1 (reference)	—
>0.55 $\mu\text{g/L}$	2 (18.2)	23 (60.5)	0.96 (0.40-2.33)	.93
Procalcitonin, n (%)				
≤0.046 ng/ml	8 (53.3)	6 (16.7)	1 (reference)	—
>0.046 ng/ml	7 (46.7)	30 (83.3)	0.49 (0.23-1.05)	.07
High-sensitivity C reaction protein (mg/L), median	44.6	95.2	0.99 (0.98-1.00)	.07

^aTwo variables (age, white blood cell count) were chosen for multivariable analysis.

^bNot applicable.

^cAPACHE: acute physiology and chronic health evaluation.

^dSOFA: sequential organ failure assessment.

^ePaCO₂: arterial partial pressure of carbon dioxide.

^fPaO₂: arterial partial pressure of oxygen.

Discussion

Principal Findings

In this study, the mortality rate associated with COVID-19 in Chongqing was 1.04%, and the 28-day case fatality rate of ICU patients was 5.3%. ARDS developed in 93% of ICU patients, and HFNC was the most commonly used type of oxygen therapy. About one-third of patients with ARDS improved in 1 week, which we defined as eiARDS. Patients younger than 55 years were more likely to exhibit eiARDS.

The mortality of COVID-19 varied widely across different periods and areas. In the early stage of the outbreak, Wuhan reported a mortality rate of 4.3% among hospitalized patients [15] and 61.5% among critically ill patients [16]. However, mortality in ICU patients gradually decreased to 32.5%-38.5% as time elapsed [17,18]. On the other hand, it was 26% in Lombardy, Italy [19], and 50% in Seattle, United States [20]. In the present study, a case fatality rate of only 5.3% was reported in ICU patients in Chongqing. The large differences in mortality were probably because medical resources could not be supplied timely, including health workers and hospital beds [21]. As a matter of fact, a substantial number of health workers from other provinces provided aid to Hubei Province, alongside an increase in acute care beds [21]. Similar to the model of Hubei Province, as the first cluster cases of COVID-19 were detected in Chongqing, four designated hospitals were arranged and prepared for only patients with COVID-19, and medical experts were invited from different hospitals in Chongqing. The centralized dispatch of medical resources was key for treating COVID-19 in China.

ARDS is the primary factor that increases mortality. According to the Berlin definition, stages of mild, moderate, and severe ARDS were associated with increased mortality (27%, 32%, and 45%, respectively) [12]. Studies showed that ARDS was one of the risk factors for death in patients with COVID-19 [16,22]. Much effort has been made to treat ARDS. However, only mechanical ventilation was shown to be effective [23]. Interestingly, in this study, although 93% of patients were affected by ARDS, only 35% received noninvasive ventilation and 9% received invasive ventilation. The most commonly used oxygen therapy was HFNC, which accounted for 55% (although patients may receive both HFNC and ventilation), meaning that HFNC was effective for COVID-19-induced ARDS. Similar conclusions were observed in a previous study [4]. The authors held that HFNC was suitable for patients with COVID-19 and mild ARDS, and even safe for patients with moderate and severe outcomes, which was clearly inconsistent with the stratified treatment strategies of ARDS caused by other factors [4].

In our study, nearly one-third of patients with ARDS recovered in 1 week, which we defined as eiARDS. However, eiARDS was found in only 18% of patients with mild ARDS caused by other factors, with 36% of patients persisting and 46% worsening during the first week after ARDS onset [24]. It is worth mentioning why so many patients with COVID-19 had eiARDS and why HFNC oxygen therapy was so effective for these patients. Gattinoni et al [25] proposed two types of patients with COVID-19 pneumonia: non-ARDS type 1 and ARDS type

2. Although both types of patients met the ARDS Berlin definition, severe hypoxemia in type 1 patients was associated with nearly normal respiratory system compliance, which led to ventilation/perfusion mismatch [25]. The aforementioned problems could be explained if we assume that patients with type 1 pneumonia would improve quickly compared with those with type 2. In addition, Gattinoni et al [25] proposed that the gas volume and percentage of nonaerated tissue could be clearly distinguished via CT scan between type 1 and type 2 pneumonia. However, no differences were found in the proportion of pneumonia volume between eiARDS and non-eiARDS patients in this study.

Regardless of the reasons why the proportion of eiARDS was high, paying attention to eiARDS itself is clinically meaningful. Early or rapid improvements in ARDS has always been associated with increased survival or better outcomes [24,26]. For patients with COVID-19, early improvement in oxygenation was associated with being discharged alive from the ICU [22]. Patients with eiARDS were found to have a higher survival rate and lower length of ICU stay compared with non-eiARDS patients. Dynamic observation of ARDS in the short term was worthwhile for the prognosis of COVID-19, and patients whose ARDS did not improve in 1 week should be given more attention. If most patients had mild ARDS at baseline, that would explain the higher percentage of patients with eiARDS. However, at baseline, 38.8% of patients had moderate ARDS, and no significant differences were found between patients with and without eiARDS. In other words, the initial $\text{PaO}_2/\text{FiO}_2$ value was independently associated with eiARDS. Indeed, multiple studies showed that older age (>65 years) was one of the risk factors for death in patients with COVID-19 [16,27-29] and establishing risk stratification through age (>60 years) might be helpful to clinicians [30]. Similar underlying mechanisms might be identified with regard to the effect of age on death and the development of ARDS. Nevertheless, age should be given immense attention during the management of patients with COVID-19.

Limitations

This study had several limitations. First, because of the retrospective study design, laboratory tests (except arterial blood gas analysis, which was performed daily) might not be performed for all patients at a specific time; the missing data were replaced by values obtained within the prior 2 days. Second, although the treatment strategies of the two hospitals were in accordance with the guidelines issued by the Chinese National Health Commission, some of the treatments were different, such as the composition of traditional Chinese medicine, leading to different clinical outcomes. Third, the sample size was relatively small, and some of the conclusions need to be verified using multiple care centers and larger sample sizes.

Conclusions

A new subphenotype of ARDS—eiARDS—in patients with COVID-19 was identified. As clinical outcomes differ, the stratified management of patients based on eiARDS or age is highly recommended.

Acknowledgments

The authors thank the patients described in this study, the health care personnel who cared for them, the staff members of the Health Commission of Chongqing City, and the staff members of the Chongqing Center for Disease Control and Prevention. We thank Dr Weijun Zheng (Zhejiang Chinese Medical University, China) for his professional statistical advice.

This work was supported by Chongqing Special Research Project for Novel Coronavirus Pneumonia Prevention and Control (#cstc2020jscx-fyzxX0012) and Emergency Research Project of COVID-19 of the Chongqing Health Commission (#2020NCPZX04).

Authors' Contributions

ZZ was responsible for conceptualization, methodology, formal analysis, and writing of the original draft. XY contributed to methodology, formal analysis, and writing of the original draft. HD was involved in the investigation and project administration. CZ and XR contributed to the investigation. YS was responsible for the investigation and resources. AZ was responsible for conceptualization, resources, supervision, and review and editing of the manuscript. MY contributed to resources and supervision. This manuscript was approved by all authors for publication.

Conflicts of Interest

None declared.

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020 Feb 20;382(8):727-733. [doi: [10.1056/nejmoa2001017](https://doi.org/10.1056/nejmoa2001017)]
2. Cucinotta D, Vanelli M. WHO Declares COVID-19 a Pandemic. *Acta Biomed* 2020 Mar 19;91(1):157-160 [FREE Full text] [doi: [10.23750/abm.v91i1.9397](https://doi.org/10.23750/abm.v91i1.9397)] [Medline: [32191675](https://pubmed.ncbi.nlm.nih.gov/32191675/)]
3. Weekly epidemiological update - 19 January 2021. World Health Organization. 2021 Jan 19. URL: <https://www.who.int/publications/m/item/weekly-epidemiological-update---19-january-2021> [accessed 2021-03-05]
4. Li X, Ma X. Acute respiratory failure in COVID-19: is it "typical" ARDS? *Crit Care* 2020 May 06;24(1):198 [FREE Full text] [doi: [10.1186/s13054-020-02911-9](https://doi.org/10.1186/s13054-020-02911-9)] [Medline: [32375845](https://pubmed.ncbi.nlm.nih.gov/32375845/)]
5. Xu Z, Shi L, Wang Y, Zhang J, Huang L, Zhang C, et al. Pathological findings of COVID-19 associated with acute respiratory distress syndrome. *The Lancet Respiratory Medicine* 2020 Apr;8(4):420-422 [FREE Full text] [doi: [10.1016/S2213-2660\(20\)30076-X](https://doi.org/10.1016/S2213-2660(20)30076-X)] [Medline: [32085846](https://pubmed.ncbi.nlm.nih.gov/32085846/)]
6. Duarte-Neto AN, Monteiro RAA, da Silva LFF, Malheiros DMAC, de Oliveira EP, Theodoro-Filho J, et al. Pulmonary and systemic involvement in COVID-19 patients assessed with ultrasound-guided minimally invasive autopsy. *Histopathology* 2020 Aug 24;77(2):186-197 [FREE Full text] [doi: [10.1111/his.14160](https://doi.org/10.1111/his.14160)] [Medline: [32443177](https://pubmed.ncbi.nlm.nih.gov/32443177/)]
7. Ackermann M, Verleden SE, Kuehnel M, Haverich A, Welte T, Laenger F, et al. Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. *N Engl J Med* 2020 Jul 09;383(2):120-128 [FREE Full text] [doi: [10.1056/NEJMoa2015432](https://doi.org/10.1056/NEJMoa2015432)] [Medline: [32437596](https://pubmed.ncbi.nlm.nih.gov/32437596/)]
8. Tang X, Du R, Wang R, Cao T, Guan L, Yang C, et al. Comparison of Hospitalized Patients With ARDS Caused by COVID-19 and H1N1. *Chest* 2020 Jul;158(1):195-205 [FREE Full text] [doi: [10.1016/j.chest.2020.03.032](https://doi.org/10.1016/j.chest.2020.03.032)] [Medline: [32224074](https://pubmed.ncbi.nlm.nih.gov/32224074/)]
9. Liu Y, Sun W, Li J, Chen L, Wang Y, Zhang L, et al. Clinical features and progression of acute respiratory distress syndrome in coronavirus disease 2019. *medRxiv*. Preprint posted online February 27, 2020. [doi: [10.1101/2020.02.17.20024166](https://doi.org/10.1101/2020.02.17.20024166)]
10. Clinical management of severe acute respiratory infection when novel coronavirus (nCoV) infection is suspected. World Health Organization. 2020 Jan 12. URL: <https://www.who.int/publications-detail-redirect/10665-332299> [accessed 2021-03-05]
11. The Fifth Revised Trial Version of the Novel Coronavirus Pneumonia Diagnosis and Treatment Guidance. National Health Commission of the People's Republic of China. 2020 Feb 8. URL: <http://www.nhc.gov.cn/yzygj/s7653p/202002/d4b895337e19445f8d728fcdf1e3e13a.shtml> [accessed 2021-03-05]
12. ARDS Definition Task Force, Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, et al. Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012 Jun 20;307(23):2526-2533. [doi: [10.1001/jama.2012.5669](https://doi.org/10.1001/jama.2012.5669)] [Medline: [22797452](https://pubmed.ncbi.nlm.nih.gov/22797452/)]
13. Chinese Digestion Association, Chinese Medical Doctor Association, Chinese Society of Hepatology, Chinese Medical Association. [The protocol for prevention, diagnosis and treatment of liver injury in coronavirus disease 2019]. *Zhonghua Gan Zang Bing Za Zhi* 2020 Mar 20;28(3):217-221. [doi: [10.3760/cma.j.cn501113-20200309-00095](https://doi.org/10.3760/cma.j.cn501113-20200309-00095)] [Medline: [32306655](https://pubmed.ncbi.nlm.nih.gov/32306655/)]
14. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO Clinical Practice Guideline for Acute Kidney Injury. *Kidney Inter Suppl* 2012;2(1):1-138 [FREE Full text]

15. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA* 2020 Mar 17;323(11):1061-1069 [[FREE Full text](#)] [doi: [10.1001/jama.2020.1585](https://doi.org/10.1001/jama.2020.1585)] [Medline: [32031570](https://pubmed.ncbi.nlm.nih.gov/32031570/)]
16. Yang X, Yu Y, Xu J, Shu H, Xia J, Liu H, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med* 2020 May 24;8(5):475-481 [[FREE Full text](#)] [doi: [10.1016/S2213-2600\(20\)30079-5](https://doi.org/10.1016/S2213-2600(20)30079-5)] [Medline: [32105632](https://pubmed.ncbi.nlm.nih.gov/32105632/)]
17. Yu Y, Xu D, Fu S, Zhang J, Yang X, Xu L, et al. Patients with COVID-19 in 19 ICUs in Wuhan, China: a cross-sectional study. *Crit Care* 2020 May 14;24(1):219 [[FREE Full text](#)] [doi: [10.1186/s13054-020-02939-x](https://doi.org/10.1186/s13054-020-02939-x)] [Medline: [32410714](https://pubmed.ncbi.nlm.nih.gov/32410714/)]
18. Li X, Xu S, Yu M, Wang K, Tao Y, Zhou Y, et al. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. *J Allergy Clin Immunol* 2020 Jul;146(1):110-118 [[FREE Full text](#)] [doi: [10.1016/j.jaci.2020.04.006](https://doi.org/10.1016/j.jaci.2020.04.006)] [Medline: [32294485](https://pubmed.ncbi.nlm.nih.gov/32294485/)]
19. Grasselli G, Zangrillo A, Zanella A, Antonelli M, Cabrini L, Castelli A, COVID-19 Lombardy ICU Network. Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy. *JAMA* 2020 Apr 28;323(16):1574-1581 [[FREE Full text](#)] [doi: [10.1001/jama.2020.5394](https://doi.org/10.1001/jama.2020.5394)] [Medline: [32250385](https://pubmed.ncbi.nlm.nih.gov/32250385/)]
20. Bhatraju PK, Ghassemieh BJ, Nichols M, Kim R, Jerome KR, Nalla AK, et al. Covid-19 in Critically Ill Patients in the Seattle Region — Case Series. *N Engl J Med* 2020 May 21;382(21):2012-2022. [doi: [10.1056/nejmoa2004500](https://doi.org/10.1056/nejmoa2004500)]
21. Zhang Z, Yao W, Wang Y, Long C, Fu X. Wuhan and Hubei COVID-19 mortality analysis reveals the critical role of timely supply of medical resources. *J Infect* 2020 Jul;81(1):147-178 [[FREE Full text](#)] [doi: [10.1016/j.jinf.2020.03.018](https://doi.org/10.1016/j.jinf.2020.03.018)] [Medline: [32209384](https://pubmed.ncbi.nlm.nih.gov/32209384/)]
22. Zangrillo A, Beretta L, Scandroglio AM, Monti G, Fominskiy E, Colombo S, COVID-BioB Study Group. Characteristics, treatment, outcomes and cause of death of invasively ventilated patients with COVID-19 ARDS in Milan, Italy. *Crit Care Resusc* 2020 Apr 23. [Medline: [32353223](https://pubmed.ncbi.nlm.nih.gov/32353223/)]
23. Fan E, Brodie D, Slutsky AS. Acute Respiratory Distress Syndrome: Advances in Diagnosis and Treatment. *JAMA* 2018 Feb 20;319(7):698-710. [doi: [10.1001/jama.2017.21907](https://doi.org/10.1001/jama.2017.21907)] [Medline: [29466596](https://pubmed.ncbi.nlm.nih.gov/29466596/)]
24. Pham T, Serpa Neto A, Pelosi P, Laffey JG, De Haro C, Lorente JA, LUNG SAFE Investigators*the European Society of Intensive Care Medicine Trials Group. Outcomes of Patients Presenting with Mild Acute Respiratory Distress Syndrome: Insights from the LUNG SAFE Study. *Anesthesiology* 2019 Feb;130(2):263-283 [[FREE Full text](#)] [doi: [10.1097/ALN.0000000000002508](https://doi.org/10.1097/ALN.0000000000002508)] [Medline: [30499850](https://pubmed.ncbi.nlm.nih.gov/30499850/)]
25. Gattinoni L, Chiumello D, Rossi S. COVID-19 pneumonia: ARDS or not? *Crit Care* 2020 Apr 16;24(1):154 [[FREE Full text](#)] [doi: [10.1186/s13054-020-02880-z](https://doi.org/10.1186/s13054-020-02880-z)] [Medline: [32299472](https://pubmed.ncbi.nlm.nih.gov/32299472/)]
26. Schenck EJ, Oromendia C, Torres LK, Berlin DA, Choi AMK, Siempos II. Rapidly Improving ARDS in Therapeutic Randomized Controlled Trials. *Chest* 2019 Mar;155(3):474-482 [[FREE Full text](#)] [doi: [10.1016/j.chest.2018.09.031](https://doi.org/10.1016/j.chest.2018.09.031)] [Medline: [30359616](https://pubmed.ncbi.nlm.nih.gov/30359616/)]
27. Du R, Liang L, Yang C, Wang W, Cao T, Li M, et al. Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: a prospective cohort study. *Eur Respir J* 2020 May;55(5) [[FREE Full text](#)] [doi: [10.1183/13993003.00524-2020](https://doi.org/10.1183/13993003.00524-2020)] [Medline: [32269088](https://pubmed.ncbi.nlm.nih.gov/32269088/)]
28. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Intern Med* 2020 Jul 01;180(7):934-943 [[FREE Full text](#)] [doi: [10.1001/jamainternmed.2020.0994](https://doi.org/10.1001/jamainternmed.2020.0994)] [Medline: [32167524](https://pubmed.ncbi.nlm.nih.gov/32167524/)]
29. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet* 2020 Mar;395(10229):1054-1062. [doi: [10.1016/s0140-6736\(20\)30566-3](https://doi.org/10.1016/s0140-6736(20)30566-3)]
30. Liu Y, Mao B, Liang S, Yang J, Lu H, Chai Y, Shanghai Clinical Treatment Experts Group for COVID-19. Association between age and clinical characteristics and outcomes of COVID-19. *Eur Respir J* 2020 May;55(5) [[FREE Full text](#)] [doi: [10.1183/13993003.01112-2020](https://doi.org/10.1183/13993003.01112-2020)] [Medline: [32312864](https://pubmed.ncbi.nlm.nih.gov/32312864/)]

Abbreviations

ARDS: acute respiratory distress syndrome
CT: computed tomography
eiARDS: early improvement of ARDS
FiO₂: oxygen concentration
HFNC: high-flow nasal cannula
ICU: intensive care unit
PaO₂: arterial blood oxygen partial pressure
PCR: polymerase chain reaction
RT-PCR: real-time reverse transcription–polymerase chain reaction
ULN: upper limit of normal
WHO: World Health Organization

Edited by T Sanchez; submitted 07.11.20; peer-reviewed by C Dandoy, A Hakkim; comments to author 02.02.21; revised version received 16.02.21; accepted 17.02.21; published 09.03.21.

Please cite as:

Zhan Z, Yang X, Du H, Zhang C, Song Y, Ran X, Zhang A, Yang M

Early Improvement of Acute Respiratory Distress Syndrome in Patients With COVID-19 in the Intensive Care Unit: Retrospective Analysis

JMIR Public Health Surveill 2021;7(3):e24843

URL: <https://publichealth.jmir.org/2021/3/e24843>

doi:[10.2196/24843](https://doi.org/10.2196/24843)

PMID:[33630743](https://pubmed.ncbi.nlm.nih.gov/33630743/)

©Zhu Zhan, Xin Yang, Hu Du, Chuanlai Zhang, Yuyan Song, Xiaoyun Ran, An Zhang, Mei Yang. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 09.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

The Nature and Extent of Online Marketing by Big Food and Big Alcohol During the COVID-19 Pandemic in Australia: Content Analysis Study

Florentine Martino¹, PhD; Ruby Brooks¹, BNut (Hons); Jennifer Browne¹, PhD; Nicholas Carah², PhD; Christina Zorbas¹, PhD; Kirstan Corben³, MHP; Emma Saleeba³, MPH; Jane Martin⁴, MPH; Anna Peeters¹, PhD; Kathryn Backholer¹, PhD

¹Deakin University, Geelong, Australia, Global Obesity Centre, Institute for Health Transformation

²School of Communication and Arts, The University of Queensland, Brisbane, Australia

³Victorian Health Promotion Foundation, Melbourne, Australia

⁴Obesity Policy Coalition, Melbourne, Australia

Corresponding Author:

Kathryn Backholer, PhD

Deakin University

Global Obesity Centre

Institute for Health Transformation

1 Gheringhap Street

Geelong, 3220

Australia

Phone: 61 3 924 43836

Email: kathryn.backholer@deakin.edu.au

Abstract

Background: Emerging evidence demonstrates that obesity is associated with a higher risk of COVID-19 morbidity and mortality. Excessive alcohol consumption and “comfort eating” as coping mechanisms during times of high stress have been shown to further exacerbate mental and physical ill-health. Global examples suggest that unhealthy food and alcohol brands and companies are using the COVID-19 pandemic to further market their products. However, there has been no systematic, in-depth analysis of how “Big Food” and “Big Alcohol” are capitalizing on the COVID-19 pandemic to market their products and brands.

Objective: We aimed to quantify the extent and nature of online marketing by alcohol and unhealthy food and beverage companies during the COVID-19 pandemic in Australia.

Methods: We conducted a content analysis of all COVID-19-related social media posts made by leading alcohol and unhealthy food and beverage brands (n=42) and their parent companies (n=12) over a 4-month period (February to May 2020) during the COVID-19 pandemic in Australia.

Results: Nearly 80% of included brands and all parent companies posted content related to COVID-19 during the 4-month period. Quick service restaurants (QSRs), food and alcohol delivery companies, alcohol brands, and bottle shops were the most active in posting COVID-19-related content. The most common themes for COVID-19-related marketing were *isolation activities* and *community support*. Promotion of hygiene and home delivery was also common, particularly for QSRs and alcohol and food delivery companies. Parent companies were more likely to post about corporate social responsibility (CSR) initiatives, such as donations of money and products, and to offer health advice.

Conclusions: This is the first study to show that Big Food and Big Alcohol are incessantly marketing their products and brands on social media platforms using themes related to COVID-19, such as *isolation activities* and *community support*. Parent companies are frequently posting about CSR initiatives, such as donations of money and products, thereby creating a fertile environment to loosen current regulation or resist further industry regulation. “COVID-washing” by large alcohol brands, food and beverage brands, and their parent companies is both common and concerning. The need for comprehensive regulations to restrict unhealthy food and alcohol marketing, as recommended by the World Health Organization, is particularly acute in the COVID-19 context and is urgently required to “build back better” in a post-COVID-19 world.

KEYWORDS

alcohol; food and beverage; COVID-19; marketing; social media

Introduction

In March 2020, the World Health Organization declared a global COVID-19 pandemic, which is caused by the virus SARS-CoV-2 [1]. People over the age of 60 years and those with underlying noncommunicable diseases, such as cancer, obesity, cardiovascular disease, and diabetes, are at higher risk of developing severe illness from COVID-19 [2]. These conditions impair the immune system and the ability to fight off viruses [3]. Consumption of alcohol and ultraprocessed foods are leading risk factors for these underlying conditions [4]. Additionally, heavy alcohol consumption, both short and long term, is known for its direct immunosuppressive effects [5].

The COVID-19 pandemic is also having serious consequences for mental health. A recent poll demonstrated that the pandemic has negatively impacted the mental health of 48% of parents and 36% of children in Australia [6]. Well-known coping mechanisms for dealing with stress and negative emotions include excessive consumption of alcohol [7,8] and overeating of unhealthy, highly palatable foods [9]. Indeed, evidence shows that population consumption of unhealthy foods and excessive alcohol intake increases after mass traumas, such as global financial crises, terrorism, and natural disasters [7,10,11]. This may, in turn, exacerbate mental and physical ill-health. For example, alcohol use is associated with increased reporting of interpersonal and domestic violence incidents [12]. Similarly, unhealthy diets are associated with poor mental health outcomes for adults and children, such as depression [13] as well as anxiety and mood disorders [14]. Building healthy populations resilient to the ongoing threat of COVID-19 and potential future public health threats will require a concerted effort to reduce obesity, unhealthy dietary intake, and excessive alcohol consumption across the population [15].

Marketing of unhealthy foods and beverages and alcohol influences attitudes, preferences, expectations, and consumption of these products over the life course [16]. With the rise of social media and other digital platforms, marketing is omnipresent, increasingly targeted, immersive, and engaging [17]. Consumption of alcohol and unhealthy foods and beverages is frequently portrayed by the alcohol and food industries as fun and harmless [18,19]. Global examples of marketing by these industries during the COVID-19 pandemic suggest that unhealthy food and alcohol brands and companies are capitalizing on the pandemic to promote their products, market corporate social responsibility (CSR), pursue partnerships, and shape policy environments [20]. The alignment of brands and companies with a social or health issue, known as *cause marketing*, is a subset of CSR business practices [21]. Cause marketing is used to build goodwill and to enhance public perceptions of the brand [22]. Social media platforms are ideal platforms to promote CSR activities, as interaction with consumers is high, maximizing opportunities to expand audience

reach [23]. Regardless of whether or not cause marketing has positive social impacts, it is counterproductive if companies promote products that contribute to the problems they purport to solve. For example, alcohol brands have been found to “pinkwash” their products by aligning themselves with the pink ribbon of the breast cancer awareness campaign, which is paradoxical, as alcohol is an established risk factor for breast cancer [24].

No study to date has systematically and comprehensively examined if “Big Alcohol” and “Big Food” are capitalizing on the COVID-19 pandemic to increase sales by “COVID-washing” their online posts. This study aimed to examine the nature and extent of COVID-19-related online posts by leading alcohol and unhealthy food and beverage brands and their parent companies in Australia over 4 months (February to May 2020) during the COVID-19 pandemic.

Methods

Study Design

Similar to previous studies that investigated marketing by Big Food and Big Alcohol on social media [25,26], we conducted a content analysis of the public accounts on social media platforms of leading alcohol brands, unhealthy (ie, energy-dense and nutrient-poor) food and beverage brands, and their parent companies. Content analysis is “a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use” [27]. All COVID-19-related posts were retrospectively extracted in June 2020 for a 4-month time period during the first wave of the COVID-19 pandemic in Australia, with data collected between February 1 and May 31, 2020. During these 4 months, various actions to limit the spread of COVID-19 were implemented in Australia, including public health orders to stay at home. As this research only involved analysis of publicly available social media accounts, institutional ethics approval was not required.

Selection of Brands and Companies

We selected the top five brands in Australia and their parent companies across the following categories: (1) confectionery, (2) snacks, (3) soft drinks, (4) quick service restaurants (QSRs), (5) food delivery services, (6) beer, (7) wine, (8) spirits, (9) ready-to-drink (RTD) alcoholic beverages, (10) alcohol retailers, and (11) alcohol delivery services (see [Multimedia Appendix 1](#)). These Australian brands were selected based on Australian market share (ie, sales) [28].

Some of the top five of RTD alcoholic beverage and spirit brands overlapped (eg, Jim Beam had an RTD alcoholic beverage and a spirit in the top five), and some of the top five bottle shops and alcohol delivery companies overlapped (eg, Dan Murphy’s was in top five of both categories). We combined these categories when presenting results. Brands that did not sell ultraprocessed foods and beverages according to the NOVA

classification [29], such as dairy brands, were excluded and replaced by the brand with the next largest market share.

Selection of Social Media Platforms

We selected brands' and companies' official public accounts on the following social media platforms, where available: Facebook, Instagram, YouTube, and Twitter. These four social media platforms were chosen for the following reasons:

1. Facebook, Instagram, and YouTube are the most frequently used platforms by Australians, and Twitter is an important platform for public relations communication from major brands and corporations.
2. These platforms are commonly used by large brands and companies for both advertising and public relations.
3. Facebook, including its subsidiary Instagram, and Google, including its subsidiary YouTube, constitute an advertising duopoly. The Australian Consumer and Competition Commission estimates that Facebook and Google receive two-thirds of the online advertising spend in Australia [30].

We sampled the social media accounts that were most likely to target Australian audiences. Therefore, only the Australian accounts of brands were included. Parent companies, on the other hand, are often global entities with multiple brands in their

portfolio and often did not have a local social media presence in Australia. If this was the case, international accounts were selected, where available. Occasionally, brands and parent companies used the same accounts (eg, Arnott's, Lindt, and Uber Eats), in which case the data were only captured in the *brands* analysis. Finally, we excluded social media accounts that had not been active during the prior 12 months (ie, since February 1, 2019).

Data Collection

For each social media platform, we collected data on the number of followers, the total number of posts, and details of COVID-19-related posts that brands and companies shared using their public social media accounts. The date of these posts, the product marketed, and number of likes, views, and shares—at least one week after sharing the post—were recorded in a standardized template.

COVID-19-related posts were included if they directly or indirectly referred to the COVID-19 pandemic or issues pertaining to COVID-19 (see Figure 1), such as lockdown and social distancing. Indirect marketing included posts where the only relationship to COVID-19 was the addition of hashtags (eg, #workfromhome, #quarantinecooking, and #quarantini).

Figure 1. Examples of posts directly (left) and indirectly (right) related to the COVID-19 context.



Only official brand- and company-generated posts on their social media accounts or channels were included in our analysis. For Facebook, we collected posts and reposts, including photos, videos, and events; for Instagram, we collected posts, pinned stories, and hashtag promotions in account bios; for YouTube,

we collected videos; and finally, for Twitter, we collected tweets and retweets. We did not include user-generated content. For all four social media platforms, we collected the number of followers or channel subscribers (see Table 1).

Table 1. Social media followers by platform and Australian brand account during the first week of May 2020.

Brands	Followers on social media platforms, n			
	Facebook	Instagram	YouTube	Twitter
Confectionery				
Cadbury	16,704,000	24,500	3250	15,100
Allen's	378,000	11,500	513	<100
Darrell Lea	128,000	14,600	<100	1000
Lindt	7,196,000	14,000	1030	N/A ^a
Snacks				
Arnott's	93,700	13,900	3350	2800
Peters	42,200	N/A	267	N/A
Magnum	12,412,000	N/A	3400	N/A
Smith's	161,900	600	3400	N/A
Soft drinks				
Coca-Cola	1,059,000	45,000	7450	6200
Pepsi Max	37,511,000	1200	2990	800
Schweppes	85,000	4100	367	1200
Kirks	130,000	200	<100	N/A
Quick service restaurants				
McDonald's	80,537,000	156,000	22,900	32,800
KFC	55,241,000	154,000	9510	29,600
Hungry Jack's	664,000	61,200	4600	11,800
Domino's Pizza	1,126,000	97,400	4120	37,700
Subway	809,600	18,100	2460	9300
Food delivery services				
Uber Eats	2,767,000	49,700	N/A	N/A
Deliveroo	751,000	21,100	<100	600
DoorDash	N/A	N/A	N/A	N/A
Menulog	118,000	6800	1350	3700
Beer				
XXXX Gold	426,000	13,200	31,500	300
Corona Extra	15,119,000	34,800	15,800	700
Carlton Premium Dry	191,000	17,600	2140	900
Victoria Bitter	154,000	33,000	1100	1900
Great Northern Original Lager	121,000	31,900	800	N/A
Spirits and ready-to-drink alcoholic beverages				
Smirnoff Red	14,168,000	6800	440	300
Johnnie Walker Red Label	14,901,000	N/A	1410	N/A
Jim Beam	3,205,000	N/A	600	300
Jack Daniel's	17,304,000	9000	1060	5700
Bundaberg	372,000	22,200	1640	1700
Canadian Club	283,000	12,900	200	700
Woodstock & Cola	60,100	2000	<100	N/A
Wine				

Brands	Followers on social media platforms, n			
	Facebook	Instagram	YouTube	Twitter
De Bortoli	20,000	10,000	306	8200
McWilliam's	14,800	12,700	<100	2700
Stanley Wines	N/A	N/A	N/A	N/A
Jacob's Creek	853,000	27,200	1430	5000
Berri Estates	<100	N/A	N/A	N/A
Bottle shops and alcohol delivery services				
BWS	269,000	11,600	N/A	N/A
Liquorland	152,000	3400	<100	1400
Dan Murphy's	477,000	42,500	6850	9300
Cellarbrations	60,200	1000	<100	200
IGA Plus Liquor	13,600	N/A	N/A	N/A
Shop My Local	<100	<100	N/A	N/A
Jimmy Brings	16,400	5179	Not displayed ^b	200

^aN/A: not applicable; the brand did not have an Australian social media account or had not posted since February 2019 and was excluded from the analysis.

^bThis brand did have an active platform, YouTube in this case; however, the number of subscribers was not displayed. This account was included in the analysis.

Data Analysis

We developed a coding framework based on an initial analysis of COVID-19-related posts by a subset of brands in our sample. Three investigators identified COVID-19-related themes used in social media posts by the top brand and company in each category. These initial themes were discussed with the research team in order to develop a coding framework that could be applied across all sampled social media accounts. The framework continued to develop iteratively as additional themes were identified in the data (see [Table 2](#)).

Three researchers extracted and coded the social media posts. Every individual post was independently extracted and coded by one researcher (RB or JB), and a subsample (20%) of all posts by the top brands in every product category was

cross-checked for consistency of data extraction and coding by a third researcher (FM). Discrepancies in data extraction and coding were discussed between researchers and, where necessary, through consultation with a fourth researcher (KB) until consensus was reached. Because there were very few discrepancies in coding, we did not continue this process for the remaining brands. Similarly, any unclear posts and how these should be coded were discussed between the three coders and the senior member of the research team (KB) throughout the analysis.

Once all the COVID-19-related posts had been identified and coded according to the framework, we calculated the frequency with which each COVID-19-related theme was used by each brand and company and across product categories (see [Table 2](#)).

Table 2. COVID-19-related marketing themes.

COVID-19 theme	Description	Example
Trading or event updates	Practical updates around trading hours, opening or closing of stores, and events during the COVID-19 pandemic, excluding delivery	"We're still open! As you've just heard from our PM, no restrictions will apply to off-licence venues. So take your time, visit us in-store or online 7 days a week. #danmurphys" (Dan Murphy's—Twitter, local brand)
Home delivery or take away in lockdown period	Mentions of home delivery or take away options during the COVID-19 pandemic (eg, no need to leave the house and staying at home during difficult times)	"Working from home? You can still enjoy your Macca's faves and have them dropped on your doorstep with contactless McDelivery. We'll deliver to you." (McDonald's—Instagram, local brand)
Hygiene or zero contact	References to reducing chances of virus spread through hygiene practices when preparing food or handling food and drinks, social distancing by employees and customers, and extra cleaning and disinfecting (eg, contactless, zero contact, and keeping community safe)	"We won't let anything get in the way of you and your sub, not even a little <-- social distancing --> Come see us in-restaurant. We're still open for takeaway -- just make sure you stay about five Subway Footlong® subs apart" (Subway—Facebook, local brand)
Community support or feeling	References to standing together during these challenging, unprecedented, or unexpected times and/or the brand or company being there to support consumers and we are all in this together	"Connect with your mates online and we'll get through this together, with a GOLD in hand" (XXXX Gold—Facebook, local brand)
Applauding health staff or essential workers	Referencing or thanking essential workers, health workers, frontline workers, etc, including discounts for these workers	"The Philippines continues to be under strict lock down due to the coronavirus, making it difficult for many communities to provide for their families. Our frontline healthcare workers are overworked as we overcome this pandemic. To help share comfort and nutrition, #TeamMDLZ #Philippines has shared P12MM worth of snack products to communities and frontline healthcare workers, working with 40+ organizations. #StrongerTogether" (Mondelez—Facebook, international parent company)
Donations	References to large-scale product or financial donation	"Beam Suntory and Southern Glazer's Wine & Spirits \$1 million donation will provide resources and financial aid to workers affected by mandated closures amidst the COVID-19 pandemic" (Beam Suntory—Facebook, international parent company)
Isolation activities	Suggestions for things to do while in isolation that include brand use or promotion	"#StayHome and get your bake on this week with this tantalising TeeVee Snacks Caramel Slice! A deliciously indulgent snack that will really hit the sweet spot. https://bit.ly/3bqG1xd " (Arnott's—Facebook, local brand)
Consumption helps with coping with COVID-19	References to consumption making consumers feel better or to consumers deserving the advertised product	"Jump on our site to organise drinks for your staff in 3 simple steps. Then jump on Zoom and crack a cold one together - a bit of normalcy is great for team morale!" (Jimmy Brings—Instagram, local brand)
Supporting local business or trading partners	Suggests consumers should support local businesses, or announcements that the brand or company supports local businesses or their trading partners	"This is a very challenging time for many of our customers. As part of our ongoing support, Coca-Cola Amatil has established a free 24-hour customer support and counselling service for our customers who are struggling with the unprecedented impact of COVID-19. The Coca-Cola Amatil Customer Support Line is run by Assure, a trusted Amatil partner. The 24/7 support service offers confidential counselling and financial coaching and is available to all Amatil customers completely free of charge, in the strictest confidence. Contact your Coca-Cola Amatil representative for all the details." (Coca-Cola Amatil—Facebook, local parent company)
Health advice, including mental health	Posts include health, including mental health, advice with reference to COVID-19	"Touching multiple surfaces, then our faces, is one of the most common causes of infection. Find out how bleach can clean and protect your home from our expert Suresh Nadakatti." (Unilever—Twitter, international parent company)
Production of sanitizing products	References to the brand or company making hand sanitizer	"#ICYMI: Brown-Forman and others in the spirits industry convert distilling operations to produce sanitizer to combat the COVID-19 health crisis. https://bddy.me/2y3FNEb " (Brown Forman—Twitter, international parent company)

COVID-19 theme	Description	Example
Maintaining an essential product supply chain	References to supply chains being maintained, ensuring consumers' needs are met	"Our nine factories across Australia and New Zealand are hard at work making product. Our dedicated teams are doing their best to bring your favourite Nestlé food and beverage products to a store near you. Look after yourself and each other out there! ♥ Nestlé" (Nestlé—Facebook, international parent company)

Results

Food and Alcohol Brands

After excluding brands and services without Australian accounts in Australia, 42 brands were included in the analysis. [Table 3](#) shows that most brands had low or no activity on YouTube and Twitter (<20 posts over 4 months), with a few exceptions such as Jimmy Brings on YouTube (24 posts; 0% COVID-19 related) and Domino's on Twitter (24 posts; 38% [9/24] COVID-19 related). The number of account followers (see [Table 1](#)) shows that brands generally have the greatest social media followings on Facebook and Instagram. For example, McDonald's Australia had 80,537,000 followers on Facebook, 156,000 on Instagram, 23,000 on YouTube, and 33,000 on Twitter.

Most brands (33/42, 79%) posted content that related to the COVID-19 pandemic. A total of 916 COVID-19-related posts out of 2796 posts (32.8%) were identified. Engagement by followers in terms of likes, shares, and views was substantial

and varied greatly between brands and type of content (see [Table 3](#)), with more active accounts generating more interactions, in general, as did videos regardless of platform. The product categories with the highest proportions of total posts related to COVID-19 were bottle shops and alcohol delivery (231/619, 37.3%), QSRs (353/915, 38.6%), food delivery (52/142, 36.6%), beer (86/226, 38.1%), spirits and RTD alcoholic beverages (34/121, 28.1%), and wine (83/261, 31.8%) (see [Table 3](#)). COVID-19-related marketing activity was very low or nonexistent for brands within the snacks and soft drink categories. Arnott's was the most active in its category (ie, snacks), with 18 COVID-19-related posts on Facebook and Instagram (18/92, 20%) across the 4-month period.

Specific brands that were most actively posting COVID-19-related content were Domino's and Dan Murphy's, with both brands posting more than 100 COVID-19-related posts across the four social media platforms (Domino's: 231/623, 37.1% of total posts; Dan Murphy's: 119/253, 47.0% of total posts) during the 4-month period.

Table 3. COVID-19-related marketing on social media platforms by the top Australian unhealthy food, beverage, and alcohol brands over 4 months (February to May 2020).

Brands	COVID-19-related posts on Australian brands' social media accounts, n/N (%)					Shares, views, and likes per COVID-19 post, mean (SD), range
	Facebook	Instagram	YouTube	Twitter	Total	
Confectionery						
Cadbury ^a	10/48 (21)	3/14 (21)	N/A ^b	No posts ^c	13/62 (21)	590 (772), 14-2569
Allen's	1/7 (14)	N/A	0/1 (0)	N/A	1/8 (14)	43 ^d
Darrell Lea	7/50 (14)	6/44 (14)	2/7 (29)	N/A	15/101 (14.9)	6528 (19,255), 126-76,000
Lindt	11/86 (13)	16/125 (12.8)	No posts	N/A	27/211 (12.8)	242 (269), 24-1114
Total	29/191 (15.2)	25/183 (13.7)	2/8 (25)	— ^e	56/382 (14.7)	Not calculated ^f
Snacks						
Arnott's	10/48 (21)	8/44 (18)	No posts	No posts	18/92 (20)	554 (637), 45-1874
Peters	0/4 (0)	N/A	N/A	N/A	0/4 (0)	—
Magnum	0/3 (0)	N/A	No posts	N/A	0/3 (0)	—
Smith's	0/13 (0)	N/A	No posts	N/A	0/13 (0)	—
Total	10/68 (15)	8/44 (1)	—	—	18/112 (16.1)	Not calculated
Soft drinks						
Coca-Cola	1/6 (17)	0/4 (0)	0/1 (0)	No posts	1/11 (9)	247 ^d
Pepsi Max	2/4 (50)	N/A	0/1 (0)	N/A	2/5 (40)	29 (7), 24-34
Schweppes	0/1 (0)	0/1 (0)	No posts	N/A	0/2 (0)	—
Kirks ^g	N/A	N/A	N/A	N/A	N/A	—
Total	3/11 (27)	0/5 (0)	0/2 (0)	—	3/18 (17)	Not calculated
Quick service restaurants						
McDonald's	12/23 (52)	9/20 (45)	1/13 (8)	3/3 (100)	25/59 (42)	29,387 (131,637), 33-660,977
KFC	14/24 (58)	6/14 (43)	0/2 (0)	No posts	20/40 (50)	4061 (6136), 6-22,706
Hungry Jack's	17/49 (35)	9/31 (29)	1/7 (14)	4/9 (44)	31/96 (32)	3944 (16,266), 6-91,384
Domino's	147/365 (40.3)	70/227 (30.8)	5/7 (71)	9/24 (38)	231/623 (30.1)	1028 (3649), 2-300,015
Subway	29/61 (48)	16/33 (48)	1/3 (33)	No posts	46/97 (47)	19,010 (65,422), 25-355,044
Total	219/522 (42.0)	110/325 (33.8)	8/32 (25)	16/36 (44)	353/915 (38.6)	Not calculated
Food delivery services						
Uber Eats	3/7 (43)	12/18 (67)	N/A	N/A	15/25 (60)	588 (557), 19-1911
Deliveroo	0/3 (0)	7/22 (32)	No posts	0/3 (0)	7/28 (25)	819 (419), 155-1506
DoorDash ^g	N/A	N/A	N/A	N/A	N/A	—
Menulog	11/26 (42)	11/45 (24)	2/5 (40)	6/13 (46)	30/89 (34)	238 (768), 0-4223
Total	14/36 (39)	30/85 (35)	2/5 (40)	6/13 (46)	52/142 (36.6)	Not calculated
Beer						
XXXX Gold	7/16 (44)	4/9 (44)	1/1 (100)	N/A	12/26 (46)	1073 (682), 265-2376
Corona Extra	1/1 (100)	0/23 (0)	No posts	N/A	1/24 (4)	41 ^d
Carlton Premium Dry	14/23 (61)	N/A	N/A	N/A	14/23 (61)	1448 (3395), 38-12,821
Victoria Bitter	24/62 (39)	13/52 (25)	N/A	N/A	37/114 (32.5)	1198 (660), 221-2781

Brands	COVID-19-related posts on Australian brands' social media accounts, n/N (%)					Shares, views, and likes per COVID-19 post, mean (SD), range
	Facebook	Instagram	YouTube	Twitter	Total	
Great Northern Original Lager	6/8 (75)	14/29 (48)	2/2 (100)	N/A	22/39 (56)	19,461 (64,807), 10-282,560
Total	52/110 (47.3)	31/113 (27.4)	3/3 (100)	—	86/226 (38.1)	Not calculated
Spirits and ready-to-drink alcoholic beverages (combined)						
Smirnoff Red	0/4 (0)	No posts	N/A	N/A	0/4 (0)	—
Johnnie Walker Red Label	No posts	N/A	No posts	N/A	No posts	—
Jim Beam	0/14 (0)	0/4 (0)	0/2 (0)	N/A	0/20 (0)	—
Jack Daniel's	7/19 (37)	6/17 (35)	N/A	N/A	13/36 (36)	506 (482), 41-1546
Bundaberg	6/13 (46)	6/11 (55)	No posts	N/A	12/24 (50)	1121 (1317), 265-5000
Canadian Club	3/3 (100)	6/8 (75)	0/1 (0)	N/A	9/12 (75)	3044 (6427), 87-20,105
Woodstock & Cola	0/9 (0)	0/16 (0)	N/A	N/A	0/25 (0)	—
Total	16/62 (26)	18/56 (32)	0/3 (0)	—	34/121 (28.1)	Not calculated
Wine						
De Bortoli	21/61 (34)	24/56 (43)	0/11 (0)	2/13 (15)	47/141 (33.3)	161 (154), 3-695
McWilliam's	7/34 (21)	7/32 (22)	N/A	N/A	14/66 (21)	39 (25), 9-84
Stanley Wines ^g	N/A	N/A	N/A	N/A	N/A	—
Jacob's Creek	10/25 (40)	9/20 (45)	0/1 (0)	N/A	19/46 (41)	144 (164), 1-481
Berri Estates	3/8 (38)	N/A	N/A	N/A	3/8 (38)	23 (23), 2-48
Total	41/128 (32.0)	40/108 (37.0)	0/12 (0)	2/13 (15)	83/261 (31.8)	Not calculated
Bottle shops and alcohol delivery services (combined)						
BWS	34/77 (44)	34/55 (68)	N/A	N/A	68/132 (51.5)	213 (376), 0-1657
Liquorland	3/15 (20)	3/24 (13)	N/A	N/A	6/39 (15)	85 (119), 6-317
Dan Murphy's	75/100 (75.0)	36/135 (26.7)	4/14 (29)	4/4 (100)	119/253 (47.0)	13,751 (142,233), 2-1,552,065
Cellarbrations	0/10 (0)	N/A	N/A	N/A	0/10 (0)	—
IGA Plus Liquor	1/11 (9)	N/A	N/A	N/A	1/11 (9)	39 ^d
Shop My Local	1/31 (3)	0/21 (0)	N/A	N/A	1/52 (2)	3 ^d
Jimmy Brings	14/39 (36)	19/50 (38)	0/24 (0)	3/9 (33)	36/122 (29.5)	148 (360), 1-1792
Total	128/283 (45.2)	92/285 (32.3)	4/38 (11)	7/13 (54)	231/619 (37.3)	Not calculated

^aCadbury represents two of the top five brands in the confectionery category: Cadbury and Cadbury Dairy Milk. Cadbury Dairy Milk was also in the top five brands within the snack category.

^bN/A: not applicable; brand did not have an Australian social media account for this platform or had not posted since February 2019 and was excluded from analysis.

^cBrand did have an Australian social media account for this platform, but there were no posts of any kind during this period and, therefore, no COVID-19-related posts; percentage cannot be calculated because division by zero is undefined.

^dThere was only 1 COVID-19-related post, so SD and range could not be calculated.

^eCould not be calculated because there were no COVID-19-related posts or no posts of any kind on these platforms, the brands did not use this social media platform, or the brands did not have an Australian social media account.

^fGrand means of shares, views, and likes per COVID-19 post within brand categories were not calculated, as there was a large variation between the brands in terms of number of followers.

^gThis brand did not have any Australian accounts and was excluded from the analysis.

The analysis of COVID-19-related posts (n=916) identified a broad range of themes used, with often more than one theme used per post, in particular when posting video content. [Table 4](#) shows that, across all brands, the themes most consistently used were (1) *isolation activities*, which included suggestions

and activities for people to do at home while in isolation (eg, cocktail recipes by BWS, online trivia nights organized by Dan Murphy's, and virtual parties by Domino's), and (2) *community support*.

Table 4. Proportion of COVID-19-related themes of social media posts by brand category.

COVID-19 theme	Posts containing each theme per category, n (%) ^a								
	Confectionery (n=56) ^b	Snacks (n=18)	Soft drinks (n=3)	QSRs ^c (n=353)	Food delivery (n=52)	Beer (n=86)	Wine (n=83)	Spirits and RTD ^d alcoholic beverages (n=34)	Bottle shops and alcohol delivery (n=231)
Trading and events updates	7 (13)	0 (0)	0 (0)	22 (6.2)	7 (14)	1 (1)	15 (18)	4 (12)	70 (30.3)
Home delivery and take away	2 (4)	0 (0)	0 (0)	197 (55.8)	35 (67)	1 (1)	30 (36)	0 (0)	73 (31.6)
Hygiene and contact free	2 (4)	0 (0)	0 (0)	143 (40.5)	17 (33)	0 (0)	12 (15)	4 (12)	26 (11.3)
Community support and feeling	21 (38)	4 (22)	3 (100)	56 (15.9)	19 (37)	12 (14)	24 (29)	10 (29)	7 (3.0)
Applaud health staff and essential workers	3 (5)	0 (0)	3 (100)	43 (12.2)	5 (10)	0 (0)	4 (5)	0 (0)	3 (1.3)
Donations	5 (9)	0 (0)	0 (0)	34 (9.6)	4 (8)	2 (2)	9 (11)	6 (18)	3 (1.3)
Isolation activities	25 (45)	16 (89)	2 (67)	63 (17.8)	3 (6)	48 (56)	44 (53)	14 (41)	150 (64.9)
Consumption helps coping	5 (9)	2 (11)	0 (0)	26 (7.4)	4 (8)	5 (6)	9 (11)	0 (0)	21 (9.1)
Supporting local business and trading partners	0 (0)	0 (0)	0 (0)	11 (3.1)	19 (37)	2 (2)	1 (1)	0 (0)	0 (0)
Other	0 (0)	0 (0)	0 (0)	6 (1.7)	0 (0)	11 (13)	2 (2)	2 (6)	6 (2.6)
No clear theme ^e	10 (18)	0 (0)	0 (0)	33 (9)	2 (4)	14 (16)	0 (0)	8 (24)	12 (5.2)

^aPosts could be coded for multiple themes; therefore, the columns do not add up to 100%.

^bAll n values in this row represent total COVID-19-related posts.

^cQSR: quick service restaurant.

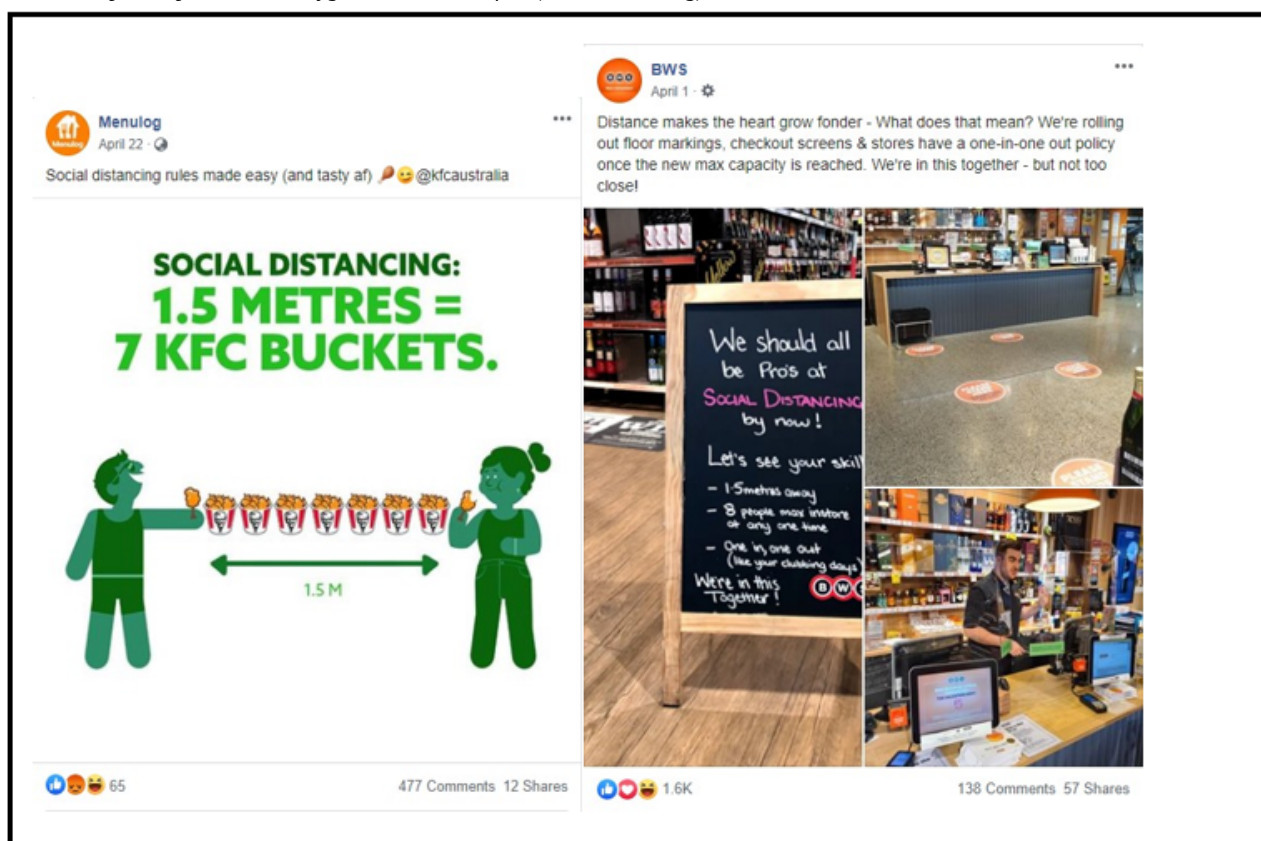
^dRTD: ready-to-drink.

^eIndirect link to COVID-19 context (eg, using hashtags or referring to working from home).

The theme *isolation activities* was particularly prominent among alcohol brands as well as bottle shops and alcohol delivery services (combined), with 40% to 64% of all COVID-19-related posts from alcohol categories using this theme. The proportion of COVID-19-related posts that used the theme *trading and events updates* was greatest for bottle shops and alcohol delivery services (70/231, 30.3%). *Home delivery and take away*-themed posts were frequently used by bottle shops and alcohol delivery services (73/231, 31.6%) and food delivery services (35/52,

67%), but they were also used by QSRs (197/353, 55.8%) and wine brands (30/83, 36%).

Over a third of all posts from brands representing establishments where food is handled by staff used the *hygiene and contact free* theme, such as QSRs (143/353, 40.5%) and food delivery services (17/52, 33%). This often included a mention of safe food handling practices and hygiene standards, an emphasis on contact-free delivery or pickup, and physical distancing requirements in stores (see [Table 2](#) and [Figure 2](#)).

Figure 2. Examples of posts with the *hygiene and contact free (social distancing)* theme.

Brands also used the theme *applauding health staff and essential workers* for their efforts during the pandemic (QSRs: 43/353, 12.2%) or making *donations* to charity or other organizations (spirits and RTD alcoholic beverages: 6/34, 18%; wine: 9/83, 11%; QSRs: 34/353, 9.6%). Approximately 5% to 10% of COVID-19-related posts across almost every category referred to the products helping people cope with isolation by creating a sense of normality or of the consumer deserving a treat in difficult times (see Table 2).

Food and Alcohol Parent Companies

A total of 16 out of 28 parent companies (57%) were excluded from the analysis, either because the brand and parent company used the same social media account (10/16, 63%) or because the parent company did not have a social media account (6/16, 38%). All 12 parent companies included in the analysis posted content that related to COVID-19. International parent companies such as Yum! Brands (n=194 posts), Unilever (n=133 posts), Mondelez (n=76 posts), and local Australian parent company Coca-Cola Amatil (n=70 posts) were most prolific in posting COVID-19-related content during the 4-month period

under analysis. Of all COVID-19-related posts by parent companies across all social media platforms (n=783 posts), 54.7% (428/783) were on Twitter and 31.2% (244/783) were on Facebook (data not shown).

Table 5 illustrates the most frequently used themes for COVID-19-related posts by parent companies. For almost all product categories, the theme *community support* was highly prevalent, often expressed using hashtags such as #InThisTogether and #StrongerTogether. The *donations* theme was also frequently used across all categories. This could refer to financial donations to support relief organizations and hospitality staff funds or product donations, such as sanitizing products and food. For example, Mondelez posted that they donated US \$20 million worth of chocolate and snacks globally, mainly to health care workers and food banks. The donations theme was often combined with other frequently used themes, namely *applauding health staff and essential workers*. For example, a Mondelez Facebook post showed a photo of health workers with Oreos and Tang—orange, powdered drink mix—in their hands (see Table 2).

Table 5. COVID-19–related themes of social media posts by unhealthy food, beverage, and alcohol parent companies.

COVID-19 theme	Posts containing each theme per category, n (%) ^{a,b}						
	Confectionery (2 companies) (n=109) ^c	Snacks (1 company) (n=133)	Soft drinks (1 company) (n=70)	QSRs ^d (1 company) (n=194)	Beer (2 companies) (n=23)	Wine (2 companies) (n=7)	Spirits and RTD ^e alcoholic beverages (3 companies) (n=103)
Trading and events updates	0 (0)	0 (0)	4 (6)	0 (0)	4 (17)	1 (14)	6 (5.8)
Home delivery and take away	0 (0)	0 (0)	0 (0)	27 (13.9)	0 (0)	0 (0)	0 (0)
Hygiene and contact free	1 (0.9)	8 (6.0)	5 (7)	39 (20.1)	1 (4)	0 (0)	2 (1.9)
Community support and feeling	70 (64.2)	86 (64.7)	30 (43)	114 (58.8)	17 (74)	1 (14)	20 (19.4)
Applaud health staff and essential workers	43 (39.4)	33 (24.8)	13 (19)	88 (45.4)	1 (4)	0 (0)	19 (18.4)
Donations	55 (50.5)	60 (45.1)	27 (39)	128 (66.0)	6 (26)	2 (29)	67 (65.0)
Isolation activities	11 (10.1)	6 (4.5)	5 (7)	10 (5.2)	5 (22)	0 (0)	15 (14.6)
Consumption helps coping	7 (6.4)	2 (1.5)	2 (3)	1 (0.5)	2 (9)	0 (0)	0 (0)
Supporting local business and trading partners	0 (0)	2 (1.5)	5 (7)	0 (0)	2 (9)	2 (29)	4 (3.9)
Production of sanitizing products	3 (2.8)	43 (32.3)	17 (24)	0 (0)	1 (4)	0 (0)	24 (23.3)
Health advice, including mental health	15 (13.8)	33 (24.8)	2 (3)	0 (0)	4 (17)	4 (57)	19 (18.4)
Maintaining essential supply chain	25 (22.9)	2 (1.5)	3 (4)	0 (0)	1 (4)	0 (0)	0 (0)

^aPosts could be coded for multiple themes; therefore, the columns do not add up to 100%.

^bThe categories *bottle shops and alcohol delivery services* and *food delivery services* were excluded from this table, as there were no parent companies or no parent companies with social media accounts.

^cAll n values in this row represent total COVID-19-related posts.

^dQSR: quick service restaurant.

^eRTD: ready-to-drink.

Posts relating to production and supply of hand sanitizer were also frequently made by parent companies that were in the business of distilling (eg, alcohol companies) or manufacturing of sanitizing products, such as soaps and bleach. These companies also often provided *health advice* in terms of how to clean and sanitize to prevent spread of COVID-19 (see [Table 2](#)).

The prevalence of COVID-19-related posts that provided mental health advice was highest among alcohol companies, mainly through reposting videos from DrinkWise—the Australian alcohol industry’s social aspect organization—featuring a celebrity doctor warning about the use of alcohol as a coping mechanism for stress during the COVID-19 pandemic.

Some companies posted specifically to support their business or trading partners (see [Table 2](#)). Finally, 23% of confectionery companies’ posts (Mondelez and Nestlé) aimed to reassure consumers that the company was doing everything it could to maintain the “essential” food supply chain, often shown in behind-the-scenes videos of production and supply chains (see [Table 2](#)).

Discussion

Principal Findings

This is the first comprehensive and systematic analysis of the nature and extent of COVID-19-related online marketing by major alcohol and unhealthy food and beverage brands and their parent companies. We demonstrated that during a 4-month

period (February to May 2020) of the first wave of the COVID-19 pandemic in Australia, 79% of the top 42 alcohol and unhealthy food and beverage brands as well as all 12 parent companies posted content related to COVID-19. Approximately one-third of social media posts by brands were COVID-19 related, and the number of likes, views, and shares indicate that engagement was substantial. Australian QSRs, alcohol brands, bottle shops, and food and alcohol delivery services were the most active in posting COVID-19-related marketing during this time. While brands in the snack, confectionery, and soft drink categories were relatively less active with regard to the number of COVID-19-related posts, their parent companies were often more active. COVID-19-related social media posts most commonly related to *isolation activities*, particularly for alcohol brands, with approximately half of all posts using this theme. Two other common COVID-19-related themes for both brands and companies were *community support*, positioning themselves as “in this together” with consumers, and *applauding health staff*. Parent companies also frequently posted about CSR activities, such as donations of money, food, or hand sanitizer.

Our results support and extend evidence from the marketing literature that businesses often employ cause marketing on social media platforms to improve public perception of their brand over the long term [23]. We found that parent companies, in particular, frequently posted about their CSR initiatives, such as financial donations to support relief efforts; manufacturing and/or donation of hand sanitizer by alcohol companies; donating soaps and sanitizing products, such as by the snack company Unilever; and donating chocolates, such as by the confectionery companies Mondelez and Nestlé. Communicating CSR initiatives has been shown to increase positive public attitudes to alcohol and unhealthy food brands and legitimizes their consumption [31]. Such marketing not only increases sales, but creates a fertile environment to loosen current regulations or resist further regulation of the industry, as these companies are seen as “part of the solution” [31]. For example, the industry body Alcohol Beverages Australia recently called on governments to minimize “regulatory and tax burdens” in response to the impacts of COVID-19 policies on sales [32]. Providing “health advice” in relation to avoiding contraction of COVID-19 (eg, how to wash hands) and messages related to responsible consumption of alcohol was another strategy frequently used by parent companies. While these may be viewed as health promotion messages, there is evidence to suggest that this type of messaging, when propagated by the alcohol industry, actually reinforces the idea that high levels of alcohol consumption are normal [33].

Another common reason for cause marketing is to increase immediate sales and business returns by aligning with trending topics. Our results indicate that in order to effectively engage consumers, many brands and services referred to COVID-19 in their social media posts to make themselves part of the zeitgeist, for example, by providing fun isolation activities, health and hygiene advice, and claiming “we’re all in this together.” Additionally, we found that food and alcohol delivery services, QSRs, and bottle shops responded to the rapid changes in consumer practices during the COVID-19 pandemic (eg, buying take away instead of going to restaurants). Their posts often

used COVID-19 themes related to health and safety procedures, such as contact-free pickup and delivery of food and alcohol and hygienic preparation of meals. The success of this marketing and the accompanying technological advancements in online sales is illustrated by bank card data showing that Australians spent 46.2% more in bottle shops in the week ending on May 22, 2020, compared to the same week in 2019, while spending in pubs and hotels plummeted (63%) [34]. Given that the social media accounts included in our analysis each had thousands, or often hundreds of thousands, of followers, large numbers of people are likely exposed to these posts, making social media posts a very efficient form of (unpaid) advertising.

“COVID-washing” of social media posts by Big Food and Big Alcohol is propagated to ultimately further increase sales and consumption of their unhealthy products and is in direct contrast with the goal of “building back better” from COVID-19 [35]. Alcohol consumption and unhealthy food and beverage consumption are both major risk factors for excess weight gain and obesity, which in turn are risk factors for comorbidities and mortality from COVID-19 [4]. Reducing population levels of alcohol and unhealthy food consumption will not only protect against the ongoing threat from COVID-19 and other pandemics in the future but is also vital for maintaining physical and mental health during pandemics and associated times of restricted movement. Regulating the marketing of these unhealthy products is an important step toward achieving this public health goal.

Limitations

Our study did not capture posts from smaller brands and companies. However, this is the first study to comprehensively describe the nature and extent of “COVID-washing” on social media platforms. A strength of our analysis is that we systematically captured all posts on social media platforms from the top brands (n=42) and parent companies (n=12), by market share, in Australia across a 4-month period during the COVID-19 pandemic. While examining all food and alcohol posts from social media platforms would provide a more accurate indication of COVID-19 marketing breadth, it would miss the depth that we were able to capture by focusing on the largest brands and companies. Another limitation of this study is that it only included publicly available marketing on four social media platforms and has missed posts on other social media platforms, including those where young people, who are more vulnerable to the effects of marketing, are active, such as Snapchat and TikTok. Our initial scoping of posts by the brands and companies included in our study suggested that the brands and companies under analysis did not have public accounts on these social media platforms; however, further work is required to better understand how these social media platforms are used by alcohol and food companies to engage consumers, including during times of crisis. A further limitation is that we were unable to determine from our analysis what the reach and actual exposure levels to these posts were. This information is only available to the marketers. However, the number of followers of active brand accounts and the level of engagement with COVID-19-related posts suggest that reach was quite high. Finally, this study only captured a portion of the marketing material created by these brands. We included public posts on

official brand and company accounts, but not user-generated content, which is now an integral part of brands' and companies' marketing campaigns (eg, the use of influencers). We also did not capture paid advertisements, as companies are unlikely to post ads on their public pages. Thus, the practices we have identified may well be much more extensive and culturally embedded than what we describe here. It is important that future studies are designed to monitor both covert and overt marketing techniques.

Conclusions

"COVID-washing" by Big Food and Big Alcohol is both common and concerning. Brands and companies have

strategically designed their marketing campaigns to positively align with the COVID-19 pandemic to increase brand loyalty and sales and potentially to influence and oppose regulation. This is problematic, as consumption of alcohol and unhealthy foods are directly or indirectly associated with poorer COVID-19-related outcomes; they also negatively impact physical and mental health beyond the pandemic. The implementation of comprehensive regulations to restrict unhealthy food and alcohol marketing, as recommended by the World Health Organization [36,37], is particularly acute in the COVID-19 context and is urgently required to "build back better" in a post-COVID-19 world.

Acknowledgments

This study would not have been possible without the support of the Victorian Health Promotion Foundation (VicHealth), which funded this study. KB is also supported by a Heart Foundation Future Leader Fellowship (102047).

Authors' Contributions

FM, NC, CZ, KC, ES, JM, AP, and KB conceived and designed the study. FM, RB, and JB collected and coded the data. FM performed the analysis. FM, RB, JB, and KB interpreted the data and prepared the original manuscript. All authors reviewed and approved the final version of the manuscript. Two of the study authors (KC and ES) are employees at VicHealth and contributed to the paper as stated above.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Inclusion of brands and parent companies in the audit.

[DOCX File, 22 KB - [publichealth_v7i3e25202_app1.docx](#)]

References

1. Rolling updates on coronavirus disease (COVID-19). World Health Organization. 2020 Jul 30. URL: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen> [accessed 2020-11-20]
2. Coronavirus disease (COVID-19). World Health Organization. 2020 Oct 12. URL: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19> [accessed 2020-11-20]
3. Yang J, Tian C, Chen Y, Zhu C, Chi H, Li J. Obesity aggravates COVID-19: An updated systematic review and meta-analysis. *J Med Virol* 2020 Nov 17;1-13 [FREE Full text] [doi: [10.1002/jmv.26677](https://doi.org/10.1002/jmv.26677)] [Medline: [33200825](https://pubmed.ncbi.nlm.nih.gov/33200825/)]
4. GBD 2013 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* 2015;386(10010):2287-2323. [doi: [10.1016/S0140-6736\(15\)00128-2](https://doi.org/10.1016/S0140-6736(15)00128-2)] [Medline: [26364544](https://pubmed.ncbi.nlm.nih.gov/26364544/)]
5. Sarkar D, Jung MK, Wang HJ. Alcohol and the immune system. *Alcohol Res* 2015;37(2):153-155 [FREE Full text]
6. Rhodes A. National Child Health Poll. COVID-19 Pandemic: Effects on the Lives of Australian Children and Families. Melbourne, Australia: The Royal Children's Hospital; 2020 Jul. URL: <https://www.rchpoll.org.au/wp-content/uploads/2020/07/nchp-poll18-report-covid.pdf> [accessed 2020-11-16]
7. Keyes KM, Hatzenbuehler ML, Hasin DS. Stressful life experiences, alcohol consumption, and alcohol use disorders: The epidemiologic evidence for four main types of stressors. *Psychopharmacology (Berl)* 2011 Nov;218(1):1-17 [FREE Full text] [doi: [10.1007/s00213-011-2236-1](https://doi.org/10.1007/s00213-011-2236-1)] [Medline: [21373787](https://pubmed.ncbi.nlm.nih.gov/21373787/)]
8. Kushner M, Abrams K, Borchardt C. The relationship between anxiety disorders and alcohol use disorders: A review of major perspectives and findings. *Clin Psychol Rev* 2000 Mar;20(2):149-171. [doi: [10.1016/s0272-7358\(99\)00027-6](https://doi.org/10.1016/s0272-7358(99)00027-6)] [Medline: [10721495](https://pubmed.ncbi.nlm.nih.gov/10721495/)]
9. Pool E, Delplanque S, Coppin G, Sander D. Is comfort food really comforting? Mechanisms underlying stress-induced eating. *Food Res Int* 2015 Oct 01;76:207-215. [doi: [10.1016/j.foodres.2014.12.034](https://doi.org/10.1016/j.foodres.2014.12.034)]

10. de Goeij MCM, Suhrcke M, Toffolutti V, van de Mheen D, Schoenmakers TM, Kunst AE. How economic crises affect alcohol consumption and alcohol-related health problems: A realist systematic review. *Soc Sci Med* 2015 Apr;131:131-146. [doi: [10.1016/j.socscimed.2015.02.025](https://doi.org/10.1016/j.socscimed.2015.02.025)] [Medline: [25771482](https://pubmed.ncbi.nlm.nih.gov/25771482/)]
11. Kemp E, Kennett-Hensel P, Williams K. The calm before the storm: Examining emotion regulation consumption in the face of an impending disaster. *Psychol Mark* 2014 Oct 09;31(11):933-945. [doi: [10.1002/mar.20744](https://doi.org/10.1002/mar.20744)]
12. Curtis A, Vandenberg B, Mayshak R, Coomber K, Hyder S, Walker A, et al. Alcohol use in family, domestic and other violence: Findings from a cross-sectional survey of the Australian population. *Drug Alcohol Rev* 2019 May;38(4):349-358. [doi: [10.1111/dar.12925](https://doi.org/10.1111/dar.12925)] [Medline: [30942525](https://pubmed.ncbi.nlm.nih.gov/30942525/)]
13. Lai J, Hiles S, Bisquera A, Hure A, McEvoy M, Attia J. A systematic review and meta-analysis of dietary patterns and depression in community-dwelling adults. *Am J Clin Nutr* 2014 Jan;99(1):181-197 [FREE Full text] [doi: [10.3945/ajcn.113.069880](https://doi.org/10.3945/ajcn.113.069880)] [Medline: [24196402](https://pubmed.ncbi.nlm.nih.gov/24196402/)]
14. O'Neil A, Quirk SE, Housden S, Brennan SL, Williams LJ, Pasco JA, et al. Relationship between diet and mental health in children and adolescents: A systematic review. *Am J Public Health* 2014 Oct;104(10):e31-e42. [doi: [10.2105/AJPH.2014.302110](https://doi.org/10.2105/AJPH.2014.302110)] [Medline: [25208008](https://pubmed.ncbi.nlm.nih.gov/25208008/)]
15. Hamer M, Kivimäki M, Gale C, Batty G. Lifestyle risk factors, inflammatory mechanisms, and COVID-19 hospitalization: A community-based cohort study of 387,109 adults in UK. *Brain Behav Immun* 2020 Jul;87:184-187 [FREE Full text] [doi: [10.1016/j.bbi.2020.05.059](https://doi.org/10.1016/j.bbi.2020.05.059)] [Medline: [32454138](https://pubmed.ncbi.nlm.nih.gov/32454138/)]
16. Armstrong G, Adam S, Denize S, Kotler P. Principles of Marketing. 7th edition. Melbourne, Australia: Pearson Australia; 2018.
17. Carah N. Alcohol corporations and marketing in social media. In: Lyons AC, McCreanor T, Goodwin I, Moewaka Barnes H, editors. *Youth Drinking Cultures in a Digital World: Alcohol, Social Media and Cultures of Intoxication*. New York, NY: Routledge; 2017:115-131.
18. Cairns G, Angus K, Hastings G, Caraher M. Systematic reviews of the evidence on the nature, extent and effects of food marketing to children. A retrospective summary. *Appetite* 2013 Mar;62:209-215. [doi: [10.1016/j.appet.2012.04.017](https://doi.org/10.1016/j.appet.2012.04.017)] [Medline: [22561190](https://pubmed.ncbi.nlm.nih.gov/22561190/)]
19. Pettigrew S, Jongenelis M, Jongenelis G, Pierce H, Stafford J, Keric D. Get them laughing to get them drinking: An analysis of alcohol advertising themes across multiple media in Australia. *J Stud Alcohol Drugs* 2020 May;81(3):311-319. [doi: [10.15288/jsad.2020.81.311](https://doi.org/10.15288/jsad.2020.81.311)]
20. Collin J, Raston R, Hill S. Signalling Virtue, Promoting Harm: Unhealthy Commodity Industries and COVID-19. Edinburgh, Scotland: NCD Alliance, SPECTRUM; 2020. URL: https://ncdalliance.org/sites/default/files/resource_files/Signalling%20Virtue%2C%20Promoting%20Harm_Sept2020_FINAL.v.pdf [accessed 2020-11-30]
21. Rego MM. The Impact of Cause-Related Marketing on Global Consumers: A Meta-Analysis [doctoral dissertation]. Storrs, CT: University of Connecticut; 2017. URL: <https://opencommons.uconn.edu/cgi/viewcontent.cgi?article=7910&context=dissertations> [accessed 2021-03-02]
22. Chernev A, Blair S. Doing well by doing good: The benevolent halo of corporate social responsibility. *J Consum Res* 2015 Apr 01;41(6):1412-1425 [FREE Full text] [doi: [10.1086/680089](https://doi.org/10.1086/680089)]
23. Araujo T, Kollat J. Communicating effectively about CSR on Twitter. *Internet Res* 2018 Apr 04;28(2):419-431 [FREE Full text] [doi: [10.1108/intr-04-2017-0172](https://doi.org/10.1108/intr-04-2017-0172)]
24. Mart S, Giesbrecht N. Red flags on pinkwashed drinks: Contradictions and dangers in marketing alcohol to prevent cancer. *Addiction* 2015 Oct;110(10):1541-1548. [doi: [10.1111/add.13035](https://doi.org/10.1111/add.13035)] [Medline: [26350708](https://pubmed.ncbi.nlm.nih.gov/26350708/)]
25. Hendriks H, Wilmsen D, van Dalen W, Gebhardt WA. Picture me drinking: Alcohol-related posts by Instagram influencers popular among adolescents and young adults. *Front Psychol* 2019;10:2991 [FREE Full text] [doi: [10.3389/fpsyg.2019.02991](https://doi.org/10.3389/fpsyg.2019.02991)] [Medline: [32038379](https://pubmed.ncbi.nlm.nih.gov/32038379/)]
26. Bragg M, Pageot YK, Amico A, Miller AN, Gasbarre A, Rummo PE, et al. Fast food, beverage, and snack brands on social media in the United States: An examination of marketing techniques utilized in 2000 brand posts. *Pediatr Obes* 2020 May;15(5):e12606. [doi: [10.1111/ijpo.12606](https://doi.org/10.1111/ijpo.12606)] [Medline: [31875654](https://pubmed.ncbi.nlm.nih.gov/31875654/)]
27. Krippendorff K. Content Analysis: An Introduction to its Methodology. 2nd edition. Thousand Oaks, CA: SAGE Publications; 2004.
28. Brand Shares Australia 2019. Euromonitor International. URL: <https://www.euromonitor.com/australia> [accessed 2021-05-08]
29. Monteiro CA, Cannon G, Levy R, Moubarac J, Jaime P, Martins AP, et al. NOVA. The star shines bright. *World Nutr* 2016;7(1-3):28-38 [FREE Full text]
30. Australian Competition and Consumer Commission. Digital Platforms Inquiry. Canberra, Australia: Commonwealth of Australia; 2019 Jun. URL: <https://www.accc.gov.au/system/files/Digital%20platforms%20inquiry%20-%20final%20report.pdf> [accessed 2020-09-10]
31. Mialon M, McCambridge J. Alcohol industry corporate social responsibility initiatives and harmful drinking: A systematic review. *Eur J Public Health* 2018 Aug 01;28(4):664-673 [FREE Full text] [doi: [10.1093/eurpub/cky065](https://doi.org/10.1093/eurpub/cky065)] [Medline: [29697783](https://pubmed.ncbi.nlm.nih.gov/29697783/)]
32. Worst month on record for Australian beer, wine, spirits producers: ABA covid report. Alcohol Beverages Australia. 2020. URL: <https://www.alcoholbeveragesaustralia.org.au/worst-month-on-record-for-australian-beer-wine-spirits-producers-aba-covid-report/> [accessed 2020-09-23]

33. Pettigrew S, Biagioni N, Daube M, Stafford J, Jones S, Chikritzhs T. Reverse engineering a 'responsible drinking' campaign to assess strategic intent. *Addiction* 2016 Jun;111(6):1107-1113. [doi: [10.1111/add.13296](https://doi.org/10.1111/add.13296)] [Medline: [27157907](https://pubmed.ncbi.nlm.nih.gov/27157907/)]
34. CommBank card spending data shows continued improvement. Commonwealth Bank of Australia. 2020 May 26. URL: <https://www.commbank.com.au/guidance/business/commbank-card-spending-data-shows-continued-improvement-202005.html> [accessed 2020-09-22]
35. Building back better: A sustainable, resilient recovery after COVID-19. OECD. 2020 Jun 05. URL: <http://www.oecd.org/coronavirus/policy-responses/building-back-better-a-sustainable-resilient-recovery-after-covid-19-52b869f5/> [accessed 2020-09-23]
36. Set of Recommendations on the Marketing of Foods and Non-Alcoholic Beverages to Children. Geneva, Switzerland: World Health Organization; 2010. URL: https://apps.who.int/iris/bitstream/handle/10665/44416/9789241500210_eng.pdf [accessed 2020-09-22]
37. Global Status Report on Alcohol and Health. Geneva, Switzerland: World Health Organization; 2014. URL: https://apps.who.int/iris/bitstream/handle/10665/112736/9789240692763_eng.pdf?sequence=1 [accessed 2020-09-22]

Abbreviations

CSR: corporate social responsibility

QSR: quick service restaurant

RTD: ready-to-drink

VicHealth: Victorian Health Promotion Foundation

Edited by T Sanchez; submitted 21.10.20; peer-reviewed by L Horwitz, G Hu; comments to author 06.11.20; revised version received 29.11.20; accepted 16.12.20; published 12.03.21.

Please cite as:

Martino F, Brooks R, Browne J, Carah N, Zorbas C, Corben K, Saleeba E, Martin J, Peeters A, Backholer K

The Nature and Extent of Online Marketing by Big Food and Big Alcohol During the COVID-19 Pandemic in Australia: Content Analysis Study

JMIR Public Health Surveill 2021;7(3):e25202

URL: <https://publichealth.jmir.org/2021/3/e25202>

doi:[10.2196/25202](https://doi.org/10.2196/25202)

PMID:[33709935](https://pubmed.ncbi.nlm.nih.gov/33709935/)

©Florentine Martino, Ruby Brooks, Jennifer Browne, Nicholas Carah, Christina Zorbas, Kirstan Corben, Emma Saleeba, Jane Martin, Anna Peeters, Kathryn Backholer. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 12.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

The Influence of Social Distancing on COVID-19 Mortality in US Counties: Cross-sectional Study

Phoebe Tran¹, MSc, MPhil; Lam Tran², MSc; Liem Tran³, PhD

¹Chronic Disease Epidemiology Department, Yale School of Public Health, Yale University, New Haven, CT, United States

²Department of Biostatistics, Michigan School of Public Health, University of Michigan, Ann Arbor, MI, United States

³Department of Geography, University of Tennessee, Knoxville, TN, United States

Corresponding Author:

Liem Tran, PhD

Department of Geography

University of Tennessee

1000 Phillip Fulmer Way

BGB 306

Knoxville, TN

United States

Phone: 1 865 974 6034

Email: ltran1@utk.edu

Abstract

Background: Previous studies on the impact of social distancing on COVID-19 mortality in the United States have predominantly examined this relationship at the national level and have not separated COVID-19 deaths in nursing homes from total COVID-19 deaths. This approach may obscure differences in social distancing behaviors by county in addition to the actual effectiveness of social distancing in preventing COVID-19 deaths.

Objective: This study aimed to determine the influence of county-level social distancing behavior on COVID-19 mortality (deaths per 100,000 people) across US counties over the period of the implementation of stay-at-home orders in most US states (March-May 2020).

Methods: Using social distancing data from tracked mobile phones in all US counties, we estimated the relationship between social distancing (average proportion of mobile phone usage outside of home between March and May 2020) and COVID-19 mortality (when the state in which the county is located reported its first confirmed case of COVID-19 and up to May 31, 2020) with a mixed-effects negative binomial model while distinguishing COVID-19 deaths in nursing homes from total COVID-19 deaths and accounting for social distancing- and COVID-19-related factors (including the period between the report of the first confirmed case of COVID-19 and May 31, 2020; population density; social vulnerability; and hospital resource availability). Results from the mixed-effects negative binomial model were then used to generate marginal effects at the mean, which helped separate the influence of social distancing on COVID-19 deaths from other covariates while calculating COVID-19 deaths per 100,000 people.

Results: We observed that a 1% increase in average mobile phone usage outside of home between March and May 2020 led to a significant increase in COVID-19 mortality by a factor of 1.18 ($P < .001$), while every 1% increase in the average proportion of mobile phone usage outside of home in February 2020 was found to significantly decrease COVID-19 mortality by a factor of 0.90 ($P < .001$).

Conclusions: As stay-at-home orders have been lifted in many US states, continued adherence to other social distancing measures, such as avoiding large gatherings and maintaining physical distance in public, are key to preventing additional COVID-19 deaths in counties across the country.

(JMIR Public Health Surveill 2021;7(3):e21606) doi:[10.2196/21606](https://doi.org/10.2196/21606)

KEYWORDS

COVID-19; marginal effects; mortality; negative binomial model; social distancing

Introduction

With the rapid spread of COVID-19 across the United States in early 2020, many states began enacting social distancing measures (stay-at-home orders, maintaining physical distance in public, and avoiding large gatherings) from March 2020 [1]. Enforcement of social distancing measures varied widely throughout the United States with some states including Arkansas, Iowa, Nebraska, North Dakota, Oklahoma, South Dakota, Utah, and Wyoming opting to not implement stay-at-home orders altogether [1]. Even with social distancing measures in place in most parts of the country, over 108,000 COVID-19 deaths were reported by the end of May 2020 [2]. As stay-at-home orders have been lifted in most US states and numbers of COVID-19 cases continue to rise, it is critical to assess the role of social distancing in preventing COVID-19 deaths in the United States [1].

Although some studies have assessed the influence of social distancing on COVID-19 mortality in the United States with actual rather than simulated data, these studies—including those performed by Medline et al [3] and Siedner et al [4]—have examined this relationship at the national level but not at a more granular geographic scale. Furthermore, these 2 studies [3,4] do not consider mortality directly but rather consider tangentially related mortality measures (change in the mortality growth rate and the period between the report of the first confirmed COVID-19 case and the peak number of deaths), and the study by Medline et al [3] does not include all US states. Additionally, Medline et al [3] did not separate COVID-19 deaths in nursing homes—a major (~30%) source of all COVID-19 fatalities—from total COVID-19 deaths, which could have potentially led to biased findings that are not representative of the ongoing pandemic [3,4]. These approaches may obscure the actual effectiveness of social distancing and differences in social distancing behaviors by county.

Consequently, we sought to bridge the gap in studies on COVID-19-related mortality in the United States by investigating the association between social distancing and COVID-19 mortality at the county level across the United States while simultaneously separating COVID-19 deaths in nursing homes from total COVID-19 deaths and accounting for social distancing and COVID-19-related factors. To accomplish this, we modeled the relationship between the average proportion of mobile phone usage outside of home during the time when stay-at-home orders were in place nationwide (a proxy variable for social distancing) and COVID-19 mortality by using a mixed-effects negative binomial model and marginal effects at the mean.

Methods

Assessment of COVID-19 Mortality

Data on deidentified confirmed COVID-19 deaths in all 3142 US counties were obtained from the Coronavirus Resource Center of the Center for Systems Science and Engineering at Johns Hopkins University [5]. This data set contains comprehensive information on COVID-19 death counts provided by the US Centers for Disease Control and Prevention (CDC)

and state health departments [5]. Additionally, we retrieved deidentified data on COVID-19 deaths in US nursing homes from the COVID-19 nursing home data of the Centers for Medicare and Medicaid Services [6]. We then used ArcGIS Pro 2.5 [7] to identify all COVID-19 deaths in nursing homes at facilities within a county's geographic boundaries for all US counties. We then defined a county's COVID-19 mortality (deaths per 100,000 people) from the time a county reported its first confirmed case of COVID-19 and up to May 31, 2020, as follows:



We selected May 31, 2020, as the study endpoint as most social distancing orders by state health departments in the United States were enforced in late March to early April until mid- and late May of 2020 [1]. COVID-19 deaths in nursing homes were excluded from a county's total number of confirmed COVID-19 deaths because social distancing measured through mobile phone tracking is not an accurate measure of social distancing behavior among nursing home residents [8].

Social Distancing Metrics

County-level social distancing metrics used in this study were derived from the SafeGraph COVID-19 Consortium [8], which provides social distancing data curated through anonymous global positioning system-based tracking of mobile phones [8]. We used the average proportion of mobile phone usage outside of home between March and May 2020 as our social distancing measure as this timespan encompasses the period of the enforcement of stay-at-home orders in most parts of the United States [1,8]. The average proportion of mobile phone usage outside of home in February 2020 was used to establish a baseline on which to compare the social distancing measure with that since February 2020 when the number of COVID-19 cases began to rapidly rise in the United States [9].

Covariates

We included covariates that have been reported to be associated with social distancing and COVID-19 from previous studies in these areas [10-13]. These covariates were the number of days between the report of the first confirmed case of COVID-19 and May 31, 2020 (determined for each US county from Johns Hopkins University's COVID-19 data set [5]); population density; social vulnerability; and hospital resource availability. With data on population and county size from the United States Census Bureau, we calculated a county's population density by dividing the number of individuals in a county by that county's area in square miles [14]. For population density, we used "100 persons per square mile" as the unit because the US population density at the county level varies across 5 orders of magnitude (eg, mean=270, median=40, minimum<1, and maximum>72,000), and multiples of hundreds of the average population density allow for easier interpretation of the results [15]. Social vulnerability was assessed using 15 measures that constitute the CDC's social vulnerability index: proportion of the population below poverty line, unemployment rate, per-capita income, proportion of the population >25 years of age with no high school diploma, proportion of the population >65 years of age, proportion of the population <17 years of age,

proportion of the civilian noninstitutionalized population with a disability, proportion of the population that is a single-parent household with children <18 years of age, proportion of the population that is a minority, proportion of the population >5 years of age who speak English “less than well,” proportion of housing that is a structure with >10 units, proportion of housing that is a mobile home, proportion of occupied housing units with more people than rooms, proportion of households with no vehicles, and the proportion of the population in institutionalized group quarters [16]. Data on hospital resource availability were obtained from the Environmental Systems Research Institute’s COVID-19 Resource Center [17]. Specifically, we used the number of staffed hospital beds within a county as our measure of hospital resource availability because this variable had complete data for the largest number of counties, being highly correlated with other measures of hospital resource availability such as the number of adult beds in the intensive care unit, number of licensed beds, and average ventilator use [17,18].

Statistical Modeling

In order to assess the influence of social distancing on COVID-19 mortality, we developed a mixed-effects negative binomial model. Our choice of a mixed-effects negative binomial model was motivated by 2 factors observed during the data processing phase of our study: (1) considerable variation in the average state-level mortality rates, suggesting the necessity of a state-level random intercept and (2) the variance in county-level mortality rates in each state tending to be markedly greater than the average state-level mortality rate, suggesting the use of a negative binomial model. A county’s COVID-19 mortality up to May 31, 2020, was specified as the model outcome, while the social distancing exposure in the model was average proportion of mobile phone usage outside of home between March and May 2020. The aforementioned covariates associated with social distancing and COVID-19 were also included in the model for adjustment purposes. As larger counties are more likely to have a greater COVID-19 mortality than smaller counties, owing to their population size, the model included a population size offset to account for this tendency. Potential correlations resulting from counties within the same state and with comparable behavioral factors, health care systems, and COVID-19 testing policies were dealt with by including a random intercept by state in the model. Results from the negative binomial model are reported as incidence rate ratios, which were calculated by exponentiating the model’s coefficients. Using the same mixed-effects negative binomial model, we also performed an analysis that included COVID-19 deaths in nursing homes, and the results displayed considerable differences between the model that included COVID-19 deaths in nursing homes and the main model that excluded these deaths (Multimedia Appendix 1).

Based on the results of the mixed-effects negative binomial model, we examined the impact of social distancing on a county’s COVID-19 mortality apart from the influence of other factors through marginal effects at the mean. The marginal effects at the mean set all covariates besides the social distancing variable (average proportion of mobile phone usage outside of home between March and May 2020) to their average value [19]. Compared to traditional regression models, marginal effects at the mean facilitated clearer isolation from other factors involved in the effect of social distancing on COVID-19 deaths within a county [19]. Marginal effects analyses were performed for values of the average proportion of mobile phone usage outside of home of 25%-41% between March and May 2020 because this range contains the mean (31.73%) of the social distancing variable as well as both extremes of social distancing behavior [8]. All statistical analyses were carried out using Stata (version 15, StataCorp), and results were considered significant when Cronbach $\alpha=.05$ [20].

Data Availability

The COVID-19 and social distancing data used in this study are publicly available and can be found on the Center for Systems Science and Engineering at Johns Hopkins University’s website [5] and SafeGraph’s website [8].

Results

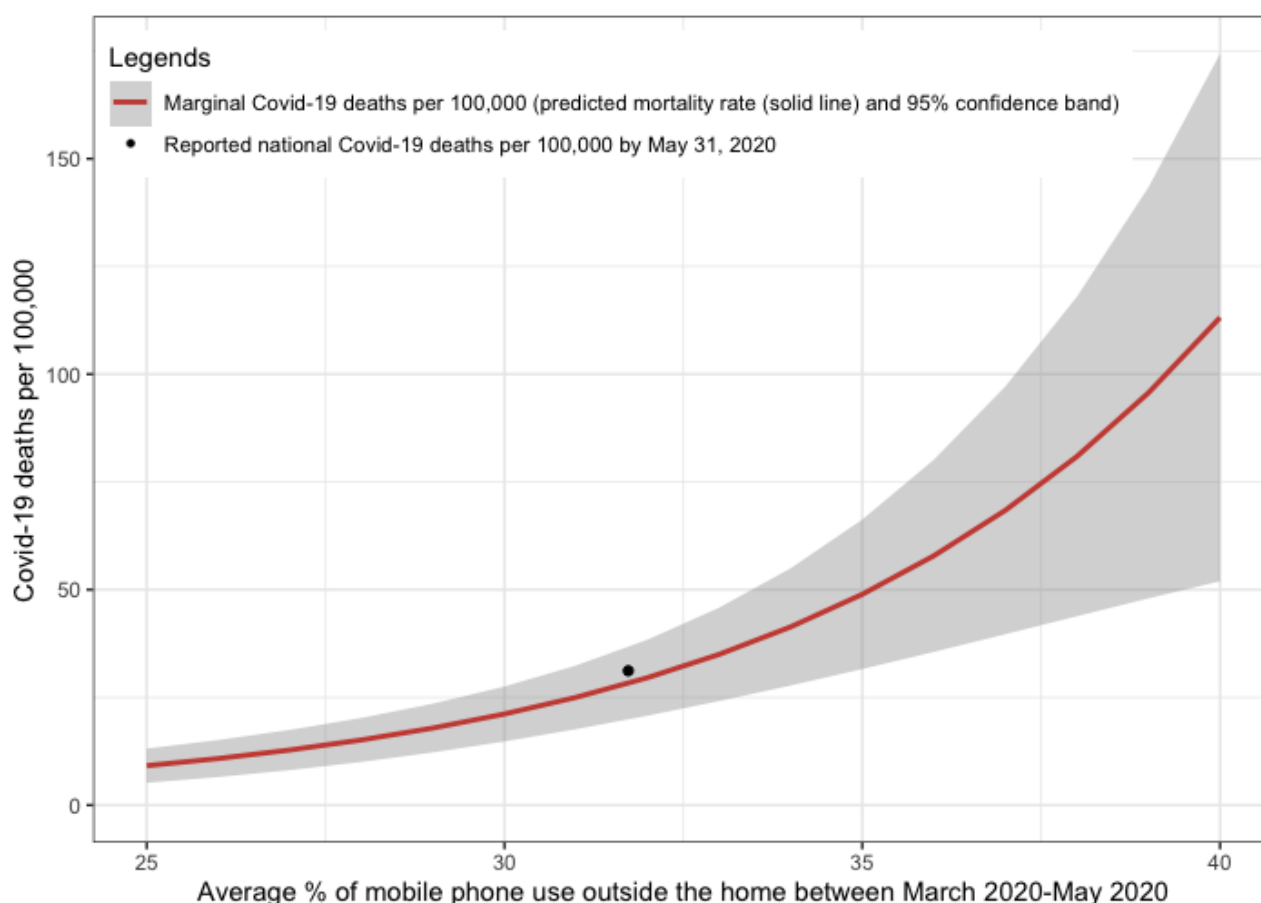
On including COVID-19 deaths in nursing homes, the total number of COVID-19 deaths during the study period was 102,958, while the mean county-level COVID-19 mortality was 13 deaths per 100,000 people. Modeling results for the social distancing variable and other covariates of interest are presented in Table 1. With respect to our social distancing variable, for every 1% increase in the average proportion of mobile phone usage outside of home between March and May 2020, COVID-19 mortality—in terms of deaths per 100,000 people—was expected to significantly increase by a factor of 1.18 ($P<.001$). This is in contrast with the variable in February, where for every 1% increase in the average proportion of mobile phone usage outside of home in February, COVID-19 mortality significantly decreased by a factor of 0.90 ($P<.001$). Coefficient estimates for population density and the number of days between the report of the first confirmed case of COVID-19 and May 31, 2020, were also significant. For every extra 100 individuals per square mile, the expected COVID-19 mortality was projected to increase by a factor of 1.02 (1.0002^{100}), whereas for every extra day between the first confirmed case of COVID-19 and May 31, 2020, the expected COVID-19 mortality increased by a factor of 1.03.

Table 1. Mixed-effects negative binomial model examining the impact of social distancing on COVID-19 mortality (deaths per 100,000 people) in US counties between the period of the report of the first case of COVID-19 and May 31, 2020.

Variables	Incidence rate ratio (95% CI)	P value
Average proportion of mobile phone usage outside of home between March and May 2020	1.18 (1.12-1.24)	<.001
Average proportion of mobile phone usage outside of home in February 2020	0.90 (0.86-0.94)	<.001
Population density (100 persons per square mile)	1.02 (1.01-1.04)	<.001
Days between the report of the first confirmed case of COVID-19 and May 31, 2020	1.03 (1.02-1.04)	<.001

We plotted the expected nationwide COVID-19 mortality rates in Figure 1 for a range of proportions of mobile phone usage outside of home, corresponding to different stringencies of lockdown measures and adherence to social distancing. The red line represents the marginal predicted nationwide deaths per

100,000 people, with the shaded band indicating the 95% CI around it. The black dot represents the reported nationwide COVID-19 mortality rate by May 31, 2020, and falls well within the predicted interval.

Figure 1. Predicted COVID-19 deaths per 100,000 people in the United States vs the average proportion of mobile phone usage outside of the nursing home between March and May 2020.

Discussion

Our study examined the influence of social distancing on COVID-19 mortality across US counties while simultaneously accounting for COVID-19 deaths in nursing homes and covariates associated with social distancing and COVID-19. Social distancing, as assessed from the average proportion of mobile phone usage outside of home between March and May 2020 but not in February 2020, was found to correspond with a significant increase in COVID-19 mortality in a county. In addition, by estimating nationwide COVID-19 mortality rates

by May 31, 2020, for a range of proportions of mobile phone usage outside of home, we found that COVID-19 mortality by this date could have been even higher had there been poorer adherence to social distancing.

While our finding of increased mobile phone usage outside of home during the period of enforcement of stay-at-home orders in many parts of the United States being associated with higher COVID-19 mortality is somewhat expected, it is of some interest that a significant negative association between the average proportion of mobile phone use outside of home in February 2020 and COVID-19 mortality was observed. In fact, high

proportions of mobile phone usage outside of home were recorded in February 2020 more often in suburban and rural counties. As the first appearance of COVID-19 tended to be later in these less densely populated regions, COVID-19 mortality in February 2020 in suburban and rural counties was lower than that in urban regions where the virus had time to spread, which resulted in the negative association observed in February 2020.

Given the current prevalence and numerous ramifications of COVID-19, a few studies carried out in the United States [3,4] have examined the impact of social distancing on COVID-19 at the national level. Concurrent with Medline et al [3] and Siedner et al [4], we also determined that social distancing was a critical factor in reducing COVID-19 mortality. However, it is difficult to directly compare our findings with those of Medline et al [3] and Siedner et al [4] because the method of assessing social distancing and COVID-19 mortality varies between these 2 studies and our study. For instance, Medline et al [3] assessed social distancing on the basis of the period between the report of the first COVID-19 case and the implementation of social distancing measures and estimated COVID-19 mortality on the basis of the period between the report of the first COVID-19 case and the peak of COVID-19 deaths. They reported a positive association between social distancing and COVID-19 mortality where each additional day in implementing a stay-at-home order additionally contributed to the peak of COVID-19 deaths [3]. Siedner et al [4] assessed social distancing by comparing the number of COVID-19 cases 14 days before and 3 days after a stay-at-home order was implemented, and they considered the change in the COVID-19 mortality growth rate as the outcome. Thus, Siedner et al [4] reported that 1 week after stay-at-home orders were implemented, the COVID-19 mortality growth rate significantly decreased day by day by 2%. By including all US counties and separating COVID-19 deaths in nursing homes from total COVID-19 deaths, our results complement those of previous studies on social distancing and COVID-19 mortality in the United States while simultaneously providing additional information regarding this relationship at a more granular level.

Although our study provides novel estimates of the impact of social distancing on COVID-19 deaths, it still has limitations that need to be considered. Our study used a proxy variable for social distancing (average proportion of mobile phone usage outside of home during March-May 2020), which would be unable to track social distancing when individuals leave their

phones at home or turn off their location services, and may not encompass some aspects of social distancing such as maintaining physical distance in public spaces and avoiding large group gatherings [8]. Although our metric for social distancing is incomplete, it is arguably one of the more quantitative approaches to track social distancing that may be prone to less bias relative to other metrics such as the dates of the implementation of stay-at-home orders or attendance at cultural, sporting, or religious events [2,21]. Additionally, approximately 95% of the adult population in the United States owns a mobile phone, with no significant deviations in ownership by race, age, education, income, or urban or rural residence [22]. Consequently, there is no evidence that any measurement error would systematically bias the social distancing coefficient either upwards or downwards and any random measurement error that does exist would not lead to any significant changes in the estimated covariate values. Finally, as with many studies on COVID-19 mortality in the United States, our study is limited by the quality of currently available data on COVID-19 mortality owing to a combination of factors including insufficient testing kits to properly diagnose COVID-19, suspected COVID-19 deaths that were instead attributed to other respiratory illnesses such as pneumonia and influenza, and delays in COVID-19 deaths being reported to the CDC [23,24].

The impact of social distancing on the COVID-19 pandemic has been particularly challenging to model at a national level. Different jurisdictions have imposed inconsistent policies with varying levels of restrictions, and substantial heterogeneity exists in the actual adherence to these social distancing guidelines. Furthermore, there is considerable variability in the availability and quality of COVID-19 mortality data. In this study, we considered daily county-level mobile phone usage outside of home and aggregated it into a national level data set to quantify the influence of social distancing on COVID-19 mortality. In doing so, we managed to account for within-county heterogeneity in the implementation of and adherence to social distancing measures. By including both mobile data usage outside of home from February 2020 (before social distancing measures) and during March-May 2020 (during and after social distancing measures), our study examined the entire span of stay-at-home orders in US states and provides a comprehensive overview of the role of social distancing in reducing COVID-19 deaths. Future studies are required to further examine the relationship between social distancing and COVID-19 given the relaxation of social distancing measures and the potential burden of continued COVID-19 mortality.

Acknowledgments

We would like to thank the Center for Systems Science and Engineering at Johns Hopkins University and SafeGraph for making their data available and free to use for research purposes.

Authors' Contributions

Pheobe Tran conceived and designed the study. Lam Tran analyzed and interpreted the data. Liem Tran helped design the study. All authors drafted the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Results from the mixed-effects negative binomial model with nursing home COVID-19 deaths included.

[DOCX File, 13 KB - [publichealth_v7i3e21606_app1.docx](#)]

References

1. Tolbert J, Kates J, Levitt L. Lifting Social Distancing Measures in America: State Actions and Metrics. Kaiser Family Foundation. 2020 May 04. URL: <https://www.kff.org/coronavirus-policy-watch/lifting-social-distancing-measures-in-america-state-actions-metrics/> [accessed 2021-03-15]
2. Social Distancing: Keep a Safe Distance to Slow the Spread. Centers for Disease Control and Prevention. 2020 Nov 17. URL: <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html> [accessed 2021-03-15]
3. Medline A, Hayes L, Valdez K, Hayashi A, Vahedi F, Capell W, et al. Evaluating the impact of stay-at-home orders on the time to reach the peak burden of Covid-19 cases and deaths: does timing matter? BMC Public Health 2020 Nov 23;20(1):1750 [FREE Full text] [doi: [10.1186/s12889-020-09817-9](https://doi.org/10.1186/s12889-020-09817-9)] [Medline: [33225945](https://pubmed.ncbi.nlm.nih.gov/33225945/)]
4. Siedner MJ, Harling G, Reynolds Z, Gilbert RF, Haneuse S, Venkataramani AS, et al. Social distancing to slow the US COVID-19 epidemic: Longitudinal pretest-posttest comparison group study. PLoS Med 2020 Aug;17(8):e1003244 [FREE Full text] [doi: [10.1371/journal.pmed.1003244](https://doi.org/10.1371/journal.pmed.1003244)] [Medline: [32780772](https://pubmed.ncbi.nlm.nih.gov/32780772/)]
5. Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Johns Hopkins University: Coronavirus Resource Center. URL: <https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6> [accessed 2021-03-15]
6. COVID-19 Nursing Home Data. Centers for Medicare and Medicaid Services. URL: <https://data.cms.gov/stories/s/COVID-19-Nursing-Home-Data/bkwz-xpvg> [accessed 2021-03-15]
7. What's new in ArcGIS Pro 2.5. Environmental Systems Research Institute. URL: <https://pro.arcgis.com/en/pro-app/2.6/get-started/what-s-new-in-arcgis-pro-2-5.htm> [accessed 2021-03-15]
8. SafeGraph COVID-19 Data Consortium. SafeGraph. URL: <https://www.safegraph.com/covid-19-data-consortium> [accessed 2021-03-15]
9. Schuchat A, CDC COVID-19 Response Team. Public Health Response to the Initiation and Spread of Pandemic COVID-19 in the United States, February 24-April 21, 2020. MMWR Morb Mortal Wkly Rep 2020 May 08;69(18):551-556 [FREE Full text] [doi: [10.15585/mmwr.mm6918e2](https://doi.org/10.15585/mmwr.mm6918e2)] [Medline: [32379733](https://pubmed.ncbi.nlm.nih.gov/32379733/)]
10. Price G, van Holm E. The Effect of Social Distancing On The Spread of Novel Coronavirus: Estimates From Linked State-Level Infection And American Time Use Survey Data. ResearchGate. 2020. URL: <https://tinyurl.com/w85nc28y> [accessed 2021-03-15]
11. Ngonghala C, Iboi E, Eikenberry S, Scotch M, MacIntyre C, Bonds M, et al. Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus. Math Biosci 2020 Jul;325:108364 [FREE Full text] [doi: [10.1016/j.mbs.2020.108364](https://doi.org/10.1016/j.mbs.2020.108364)] [Medline: [32360770](https://pubmed.ncbi.nlm.nih.gov/32360770/)]
12. Nayak A, Islam S, Mehta A, Ko Y, Patel S, Goyal A. Impact of Social Vulnerability on COVID-19 Incidence and Outcomes in the United States. medRxiv. Preprint posted online April 17, 2020. [doi: [10.1101/2020.04.10.20060962](https://doi.org/10.1101/2020.04.10.20060962)]
13. Matrajt L, Leung T. Evaluating the effectiveness of social distancing interventions against COVID-19. medRxiv. Preprint posted online March 31, 2020. [doi: [10.1101/2020.03.27.20044891](https://doi.org/10.1101/2020.03.27.20044891)]
14. County Population Totals: 2010-2019. United States Census Bureau. URL: <https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html> [accessed 2021-03-15]
15. Cohen D. Understanding Population Density. United States Census Bureau. 2015. URL: <https://www.census.gov/newsroom/blogs/random-samplings/2015/03/understanding-population-density.html> [accessed 2021-03-15]
16. CDC SVI 2018 Documentation. Centers for Disease Control and Prevention. 2020. URL: https://svi.cdc.gov/Documents/Data/2018_SVI_Data/SVI2018Documentation.pdf [accessed 2021-03-15]
17. Definitive Healthcare: USA Hospital Beds. Environmental Systems Research Institute. URL: https://coronavirus-resources.esri.com/datasets/1044bb19da8d4dbfb6a96eb1b4ebf629_0?geometry=154.511%2C-16.820%2C-179.473%2C72.123 [accessed 2021-03-15]
18. Phillip PJ, Mullner R, Andes S. Toward a better understanding of hospital occupancy rates. Health Care Financ Rev 1984;5(4):53-61 [FREE Full text] [Medline: [10310946](https://pubmed.ncbi.nlm.nih.gov/10310946/)]
19. Williams R. Using the Margins Command to Estimate and Interpret Adjusted Predictions and Marginal Effects. The Stata Journal 2018 Nov 19;12(2):308-331. [doi: [10.1177/1536867x1201200209](https://doi.org/10.1177/1536867x1201200209)]
20. StataCorp. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC. StataCorp. 2017. URL: <https://www.stata.com/support/faqs/resources/citing-software-documentation-faqs/> [accessed 2021-03-15]
21. Rotondi V, Andriano L, Dowd J, Mills M. Early evidence that social distancing and public health interventions flatten the COVID-19 curve in Italy. OSF Preprints. Preprint posted online April 26, 2020. [doi: [10.31219/osf.io/wah4e](https://doi.org/10.31219/osf.io/wah4e)]

22. Lavrakas PJ. Encyclopedia of Survey Research Methods. Thousand Oaks, CA: SAGE; 2008.
23. COVID-19 Testing Overview. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/testing.html> [accessed 2021-03-15]
24. Excess Deaths Associated with COVID-19: Provisional Death Counts for Coronavirus Disease (COVID-19). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm [accessed 2021-03-15]

Abbreviations

CDC: Centers for Disease Control and Prevention

Edited by T Sanchez; submitted 22.06.20; peer-reviewed by S Newbold, J Long, A Bhatia; comments to author 14.12.20; revised version received 14.01.21; accepted 20.01.21; published 18.03.21.

Please cite as:

Tran P, Tran L, Tran L

The Influence of Social Distancing on COVID-19 Mortality in US Counties: Cross-sectional Study

JMIR Public Health Surveill 2021;7(3):e21606

URL: <https://publichealth.jmir.org/2021/3/e21606>

doi:[10.2196/21606](https://doi.org/10.2196/21606)

PMID:[33497348](https://pubmed.ncbi.nlm.nih.gov/33497348/)

©Phoebe Tran, Lam Tran, Liem Tran. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 18.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Automated Travel History Extraction From Clinical Notes for Informing the Detection of Emergent Infectious Disease Events: Algorithm Development and Validation

Kelly S Peterson^{1,2}, MS; Julia Lewis^{1,2}, DO; Olga V Patterson^{1,2}, PhD; Alec B Chapman^{1,2}, MS; Daniel W Denhalter^{1,3}, MSPH; Patricia A Lye⁴, MT(ASCP); Vanessa W Stevens^{1,2}, PhD; Shantini D Gamage^{4,5}, MPH, PhD; Gary A Roselle^{4,5,6}, MD; Katherine S Wallace^{7,8}, PharmD; Makoto Jones^{1,2}, MSCI, MD

¹VA Salt Lake City Health Care System, US Department of Veterans Affairs, Salt Lake City, UT, United States

²Division of Epidemiology, Department of Internal Medicine, University of Utah, Salt Lake City, UT, United States

³Department of Rocky Mountain Cancer Data Systems, University of Utah, Salt Lake City, UT, United States

⁴National Infectious Diseases Service, Specialty Care Services, US Department of Veterans Affairs, Cincinnati, OH, United States

⁵Division of Infectious Diseases, Department of Internal Medicine, University of Cincinnati College of Medicine, Cincinnati, OH, United States

⁶Cincinnati VA Medical Center, US Department of Veterans Affairs, Cincinnati, OH, United States

⁷Office of Biosurveillance, Veterans Affairs Central Office, US Department of Veterans Affairs, Washington, DC, United States

⁸National Biosurveillance Integration Center, Countering Weapons of Mass Destruction, Department of Homeland Security, Washington, DC, United States

Corresponding Author:

Kelly S Peterson, MS

VA Salt Lake City Health Care System

US Department of Veterans Affairs

500 Foothill Dr

Salt Lake City, UT, 84148

United States

Phone: 1 801 582 1565

Email: kelly.peterson2@va.gov

Abstract

Background: Patient travel history can be crucial in evaluating evolving infectious disease events. Such information can be challenging to acquire in electronic health records, as it is often available only in unstructured text.

Objective: This study aims to assess the feasibility of annotating and automatically extracting travel history mentions from unstructured clinical documents in the Department of Veterans Affairs across disparate health care facilities and among millions of patients. Information about travel exposure augments existing surveillance applications for increased preparedness in responding quickly to public health threats.

Methods: Clinical documents related to arboviral disease were annotated following selection using a semiautomated bootstrapping process. Using annotated instances as training data, models were developed to extract from unstructured clinical text any mention of affirmed travel locations outside of the continental United States. Automated text processing models were evaluated, involving machine learning and neural language models for extraction accuracy.

Results: Among 4584 annotated instances, 2659 (58%) contained an affirmed mention of travel history, while 347 (7.6%) were negated. Interannotator agreement resulted in a document-level Cohen kappa of 0.776. Automated text processing accuracy (F1 85.6, 95% CI 82.5-87.9) and computational burden were acceptable such that the system can provide a rapid screen for public health events.

Conclusions: Automated extraction of patient travel history from clinical documents is feasible for enhanced passive surveillance public health systems. Without such a system, it would usually be necessary to manually review charts to identify recent travel or lack of travel, use an electronic health record that enforces travel history documentation, or ignore this potential source of information altogether. The development of this tool was initially motivated by emergent arboviral diseases. More recently, this system was used in the early phases of response to COVID-19 in the United States, although its utility was limited to a relatively

brief window due to the rapid domestic spread of the virus. Such systems may aid future efforts to prevent and contain the spread of infectious diseases.

(*JMIR Public Health Surveill* 2021;7(3):e26719) doi:[10.2196/26719](https://doi.org/10.2196/26719)

KEYWORDS

natural language processing; machine learning; travel history; COVID-19; Zika; infectious disease surveillance; surveillance applications; biosurveillance; electronic health record

Introduction

Epidemiologic clues are critical to understand how infectious diseases spread. When working up a rapidly unfolding event, a history of travel to an endemic region can be a valuable piece of evidence for public health authorities and biosurveillance experts who must often work quickly in tracing linkages to manage outbreaks. When information about travel to endemic areas is present, it may confirm an existing understanding or flag an epidemiologist to gather additional information. Alternatively, an absence of travel records for patients who are infected may suggest the possibility of local transmission. Even on an individual patient level, the importance of evaluating travel history has been previously recognized as crucial, in addition to symptoms, in establishing an appropriate differential diagnosis and optimizing testing [1,2].

Unfortunately, information about patient travel in the electronic health record (EHR) is still not typically recorded in a structured format but in unstructured clinical documents [3], especially when screening questions have not been mandated for emergency department triage. Instead, much of this information can be recorded in notes by specialists or others who suspect a travel-related disease. Although important, the need for significant annotated data and an effective document selection process are likely reasons why automatic extraction of travel history from clinical documents has not been explored extensively [4]. It has been noted that no standard tools exist for identifying critical information. Several gaps remain in gathering crucial data including accurate and timely patient travel history [5]. Reliable methods are needed for automated identification of such information from documents.

Machine learning has been applied increasingly to information extraction tasks from text. Recent developments allow scaling to large amounts of data while achieving accurate results. Such methods have proven useful to identify hate speech in social media [6] and for accelerating information gathering from legal documents [7]. These advancements have shown to be effective in several aspects of health care ranging from automated processing of radiology reports [8] to detection and relation of adverse drug events [9].

Previous work has identified geographic locations in text such as newswire articles and social media posts [10,11]. Biomedical literature has been previously used to extract geolocation information for infected hosts to enable virus spread modeling [12]. Although other means of identifying location exist, such as mobile apps, challenges remain to use these resources in health care, including privacy concerns and integration with clinical indicators of disease such as laboratory results recorded

in the EHR. Using EHR documentation to identify patient travel is rarely reported potentially due to the complexity of the task.

In health care, a comprehensive understanding of locations visited by a patient is important for understanding key epidemiologic links. This information is especially crucial for respiratory diseases, which can be spread rapidly through international travel [1]. Most notably, early efforts in managing the spread of COVID-19 have relied upon information about travel to prevent the spread of the virus [13-15]. Several recent studies have leveraged publicly available data resources such as social media to conduct surveillance of the virus including travel history, symptoms, and concerns [16-18].

In this study, we created a reference standard for detection of patient travel history from the EHR through manual chart abstraction, developed an automatic text extraction pipeline, and deployed the system in preparedness for public health threats. The goal of the system in 2017 and 2018 was not only to evaluate the travel in the context of Zika and other emerging arboviral diseases at the time but also with the intent to expand the system to support the surveillance of other emerging infectious diseases. To that effect, this capability was used during a brief window to monitor the initial spread of COVID-19 infections in the United States.

Methods

Data Set

The Veterans Affairs (VA) Corporate Data Warehouse contains clinical data from 170 medical centers and over 1000 outpatient sites across the United States [19]. This study was developed with Corporate Data Warehouse data from 2015 to 2017 across more than 6.4 million patients with over 694 million clinical notes. This study was conducted in VA to provide tools to facilitate biosurveillance and health care operations. After the initial development, accuracy was evaluated on a sample of 57 patients whose documents were authored up to April 2018. This system currently continues to run for operational biosurveillance data insights.

Annotation Guideline

The annotation guideline was developed in collaboration with epidemiologists and natural language processing (NLP) researchers. The scope of annotation was to collect past travel history and differentiate it from future or hypothetical travel discussion. Any reason for travel was allowed, including military deployment or medical tourism. No restrictions were made on the level of geographic specificity. Mentions of international travel vary in the level of geographic specificity. Although some travel locations are mentioned precisely (eg, "Mexico City"),

others include landmarks (eg, “Great Barrier Reef”) or free-form description (eg, “beach in the Mexican Riviera”). This annotation scope was simplified into one sentence to enable rapid annotation: “Does this text indicate that the patient has been somewhere outside the United States?”

Annotation Corpus Selection

To train an automated text pipeline, we annotated a corpus for travel history mentions. Some examples are illustrated in Figure

1. Since clinician documentation of patient travel history in the EHR may vary, we used a strategy to identify likely candidate statements. The cohort for these potential travel mentions were all patients who received laboratory testing for the arboviral diseases Zika, dengue, or chikungunya from January 2015 to April 2017, a date range that corresponded with concerns about Zika virus transmission. For this cohort, we included clinical notes authored between 1 month before and 6 months after lab specimens were taken.

Figure 1. Illustration of this study’s contributions. Included here are synthetic text examples, annotation output format, and model extraction of positive location mentions now deployed for public health response.

Text Example	Annotations	Model Extractions
Pt recently returned from a trip to Germany and Italy.	(Germany, Affirmed) (Italy, Affirmed)	Germany Italy
Travel to Miami? [] Y [x] N	(Miami, Negated)	None
Pt plans to travel to Japan in the Fall.	No annotation	None

To further reduce the annotation corpus for relatively rare events, a semiautomated process was implemented similar to previous work [20]. The first step involved seed terms and phrases likely related to travel. These terms were comprised of cities and countries from the GeoNames gazetteer [21]. This also included phrases likely related to travel history, such as “recently returned from X” or “traveled on a cruise to X.” This set also included curated patterns for travel history questions common in the EHR. Some questions asked patients about relatively recent concerns in 2017 such as Zika or dengue. Meanwhile, others included questions about past travel related to Ebola or Middle East respiratory syndrome. These question patterns were not used as inclusion criteria for the annotation corpus since this remained limited to arboviral testing patients.

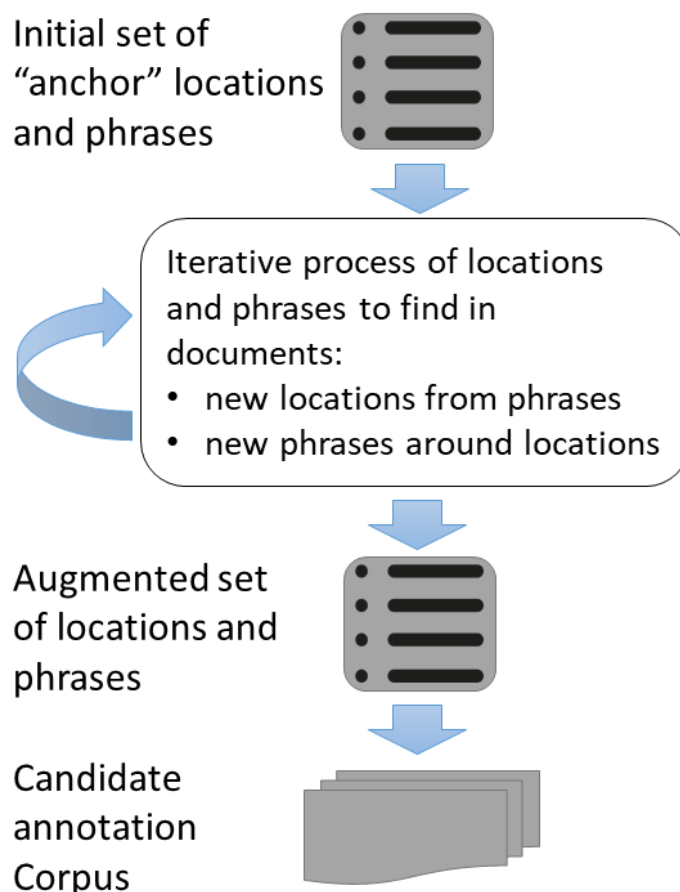
Next, these initial terms and phrases were used to retrieve documents from the patient cohort, which received arbovirus lab testing. An iterative bootstrapping approach was then used to collect additional documents. This approach was chosen to efficiently identify patterns from a large corpus of notes. At

each iteration, new travel location terms and travel-related phrases expanded the initial seed terms, as these were extracted in a semiautomated fashion illustrated in Figure 2.

New travel locations were found when co-occurring with detected travel phrases. Likewise, travel location matches were used to extract new travel-related phrases by inspecting n-grams collocated with locations such as “10 day vacation to X...” and “symptoms after a 12 hour flight from X...” Where possible, these phrases were generalized to regular expressions. Locations of varying specificity and lexical variation were gathered, such as “the big island,” “porto rica,” and “deep Mexico.” This process was repeated several times until a set of candidate documents had been gathered for annotation.

For rapid annotation, short snippets of three sentences were collected where one of the sentences contained a location, phrase, or question template. Document retrieval and sentence retrieval rules were implemented using the Leo framework [22]. These snippets were then deduplicated by performing exact lexical comparison with no preprocessing.

Figure 2. Bootstrapping process to identify previously unknown locations and phrases to enrich a document set for travel history mentions.



Annotation Methods

Annotation was performed by three annotators and consisted of span level annotation on short snippets in ChartReview [23]. In each reviewed snippet, annotators highlighted the span of a travel location mention and specified whether the location was affirmed to be a location that the patient visited or if travel was negated (eg, the patient did not visit the location or denied such visits). Examples of text and annotation data are presented in Figure 1.

Annotation agreement was measured on a set of 100 snippets annotated by two annotators. Agreement was measured at a span level of locations and at a record level of travel history in a snippet.

Location agreement was calculated for all annotated location text spans and required an exact match of text offset and negation status. Any difference in status was counted as a disagreement and any difference in text span was considered as a separate annotation element. Record agreement combined any annotated location status so that each snippet would be assigned a class of either no travel mentioned, negated locations, positive locations, or mixed.

Extraction Methods

A text processing pipeline was built to test feasibility of travel history extraction for operational use. This pipeline used models trained from the reference standard created from arboviral illness cases. The objective of this pipeline is to label affirmed travel location mentions. Although negated locations were gathered in annotation, these were not used in model training. Figure 1 provides an example of this consideration.

Given a requirement for rapid processing of documents with available computing resources, the model was constructed using conditional random fields (CRF) implemented in CRFSuite in Python [24,25]. Classes of contextual features around tokens were evaluated.

Annotated documents were split into 80% training, 10% validation, and 10% test to be held out from any training or evaluation until the final experiments. Feature types and model hyperparameters (eg, token window size) were evaluated via random search and cross validation using scikit-learn [26]. The set of features in this model include tokens, lemmas, character n-grams, part of speech tags, token shape (ie, capitalization, digits, and punctuation), gazetteer match of cities and countries from GeoNames [21], and word embedding clusters [27,28]. An example illustrating feature encoding is presented in Figure 3 and Table 1. Part of speech tags (eg, VBD) are defined as per the Penn Treebank tag values [29].

Figure 3. Illustration of feature encoding window for the token "Senegal" in an example sentence.**Table 1.** Feature encoding examples for the token "Senegal" in Figure 3.

Feature class	Encoding examples ^a
Tokens	-2_token: 'traveled'; -1_token: 'to'; 0_token: 'senegal'; 1_token: '2'; 2_token: 'weeks'
Lemmas	-2_lemma: 'travel'; -1_lemma: 'to'; 0_lemma: 'seneg'; 1_lemma: '2'; 2_lemma: 'week'
Token shape	-2_islower: 1; -1_islower: 1; 0_iscapital: 1; 1_isdigit: 1 2_islower: 1
Part of speech [29]	-2_pos: 'VBD'; 1_pos: 'TO'; 0_pos: 'JJ'; 1_pos: 'CD'; 2_pos: 'NNS'; -2_-1_pos: 'VBD_TO'; -1_0_pos: 'TO_JJ'; [...]
Character n-grams	[...]; 0_char_bigram_se: 1; 0_char_bigram_en: 1; [...]; 0_char_trigram_sen: 1; 0_char_trigram_ene: 1; [...]
Gazetteer match	-2_gaz: 0; -1_gaz: 0; 0_gaz: 1; 1_gaz: 0; 2_gaz: 0
Word embedding clusters	-2_cluster: '6101'; -1_cluster: '8804'; 0_cluster: '5470'; [...]; -2_-1_cluster: '6101_8804'; -1_0_cluster: '8804_5470'; [...]

^aSingle quoted values are encoded as categorical features rather than ordinal. When present, a prefix '-2_' indicates features for the token 2 to the left, '0_' for the current token, etc.

Model performance was evaluated with classification metrics of precision, recall, and F1 as well as bootstrapped sampling methods at a 95% CI [30,31]. All reported findings here used strict matching such that if the annotator marked a multi-word visited location such as "Western Africa," both tokens must be detected.

As a comparison, we selected two general purpose location named-entity recognition (NER) models, which were evaluated for geoparsing. Since these systems were developed for geolocation detection and not travel history, the intent of this comparison is to provide motivation that the task of travel history is distinct from geoparsing. In one comparison experiment, we applied the NER model from the Stanford NLP library to the annotated travel corpus [10]. In the other experiment, we used spaCy [32] to label the location entity. Although these are general-purpose models for location extraction, they have been used in several system comparisons for parsing geographic locations from text [33]. Among the potential libraries reviewed in [33], these were selected given convenient usage in the Python programming language. As this was a simple experiment to provide motivation of this being a distinct task, these systems were not retrained or augmented with gazetteer entries, as they were applied to the annotated data set with no modifications. Spans labeled as locations by each baseline model were used in this simplified evaluation of identifying patient travel location. Two comparisons were made since these existing systems do not allow for an ideal comparison. One was limited to locations annotated for past affirmed travel and the other for any annotated location, which would include negated, future, or hypothetical travel.

Postdeployment Evaluation Methods

After an initial model had been deployed for operational purposes, an analyst with clinical experience performed a chart review to estimate system performance. This evaluation was performed to account for a potential selection bias due to the keyword-based selection of records for the original reference standard and to evaluate the system on more recent data not available during the training phase.

To conduct this evaluation, 57 unique patients were randomly selected among patients who received laboratory testing for Zika, dengue, or chikungunya prior to April 2018. For each patient, clinical documents were reviewed surrounding the dates of the specimen acquisition and test completion. This review resulted in the manual gathering of past affirmed patient locations. These locations were compared against automated locations identified in operational reporting.

Ethics Approval

This analysis was performed under project approval from the University of Utah Institutional Review Board and the VA Salt Lake City Health Care System Research and Development Office.

Results

Annotation Corpus Selection Results

The process for selecting candidate travel history mentions was initiated on 250,133 clinical documents among 2274 unique patients who received arbovirus testing. The mean age of these patients was 59 years, and 1960 (86.2%) patients were male. A total of 6482 snippets were identified as potential travel mentions, and 3894 (60.1%) were acquired from locations, 2576 (39.7%) from phrases, and 12 (0.2%) from template questions.

After removing duplicates, 4584 unique snippets remained to be annotated. Since there were over 1800 duplicates in the candidates, some of the most frequent are explained here. Many of the duplicate snippets consisted of several travel question templates, which were identical in hundreds of snippets across multiple patients, such as “Has the patient traveled to an Ebola affected area? [] Y [X] N” or “Has the patient traveled to West Africa?” Other duplicates included verbatim travel history records that had been pasted into the same patient’s medical record multiple times across separate documents. Other frequent duplicate snippets matched location strings yet were involved in unrelated common clinical templates such as “Sleep Disorders: Berlin and Epworth Questionnaires.”

Annotation Results

Of the 4584 unique snippets annotated, 2659 (58.0%) were annotated as affirmed travel history and 347 (7.6%) were marked as negated. The remaining 1578 (34.4%) snippets were annotated as not containing any mention of travel history.

Both location and record agreement were calculated using Cohen kappa [34]. In the set of 100 double annotated snippets, span agreement was measured as $\kappa=0.706$ and record agreement as $\kappa=0.776$.

Documentation of patient travel history in these annotations vary in a spectrum from semistructured questionnaires (eg, “Have you visited a region known for Zika transmission?” “Has the patient recently returned from Brazil, Mexico or Miami, Florida? [] Y [X] N”) to coarse grain descriptions (eg, “went to Europe”) to detailed traces of travel activity (eg, “returned to the United States on July 2 from Guatemala by way of Mexico City (July 1)”).

Instances of annotator disagreement were examined to identify the categories of differences and explore potential refinement for the annotation guideline. Several disagreements stemmed from differing attribution of past affirmed travel as opposed to future or hypothetical travel. For example, in one synthetic example, one annotator marked the mention “...travelling to

visit sister in Hungary in May” as future travel, while another marked this as past affirmed travel. These errors were the most frequent cause for disagreement.

Additional challenges in this task occurred as this study was conducted using Department of VA records, where military deployment is common but actual patient exposure may remain ambiguous. For example, a template such as “Service Era: Vietnam” does not necessarily imply that the patient was in Vietnam. Statements like these on their own did not lead to much disagreement, yet we observed that annotators made different inferences in more complicated examples. In short snippets, longer distance statements could lead to differing interpretations. For example, in a snippet such as “Service: Persian Gulf War...<another long sentence>...but after coming home and getting married...,” one annotator inferred that the patient had been in the Persian Gulf region while the other did not.

Geographic specificity also led to some annotation disagreement. Although the annotation gave no requirement for specificity, there were instances where a sentence like “Pt traveled outside the United States from early May to August” was marked by one annotator as a past affirmed travel location “outside the United States.” Although the value of such travel history could be debated for the purposes of infectious disease surveillance, additional examples in the annotation guideline could have provided clarity in these situations.

The annotation identified 561 distinct locations across 8127 location spans. The most frequently annotated affirmed and negated locations are listed in Table 2. These locations exhibit a range of geographic specificity. Some mentions are specific cities or countries, while others are entire continents or regions (eg, Africa, West Africa) and others are highly specific such as the name of a beach resort or an oceanic cruise itinerary. Some location mentions are ambiguous, such as “Jamaica,” which could refer to the country or the neighborhood in New York City.

Table 2. Most frequently affirmed and negated annotated locations and their percentage of the 8127 total location text spans annotated.

Locations	Annotations, n (%)
Affirmed	
Iraq	471 (5.8)
Mexico	374 (4.6)
Vietnam	251 (3.1)
Costa Rica	251 (3.1)
Dominican Republic	244 (3.0)
Afghanistan	236 (2.9)
Jamaica	187 (2.3)
Puerto Rico	179 (2.2)
Negated	
Liberia	341 (4.2)
Guinea	341 (4.2)
Sierra Leone	333 (4.1)
Democratic Republic of Congo	130 (1.6)
West Africa	114 (1.4)
Mali	114 (1.4)
Nigeria	89 (1.1)
Western Africa	81 (1.0)

Extraction Results

The best performing model to label affirmed travel history mentions was evaluated on the test set as 88.0% precision, 83.3% recall, and 85.6% F1 measure. A comparison to two existing systems in detecting past affirmed travel is presented

in [Table 3](#). Another example of these systems for any annotated travel location mentions, whether affirmed, past, negated, or future, is provided in [Table 4](#). Although neither comparison is ideal, these differences highlight that identifying travel history is distinct from typical geolocation recognition.

Table 3. Comparison of our model to two other general-purpose baselines in the task of identifying past affirmed travel locations.

Model	Precision	Recall	F1 score
Stanford location NER ^a	40.2	76.7	52.8
spaCy NER	36.0	62.0	45.5
Proposed model	88.0	83.3	85.6

^aNER: named-entity recognition.

Table 4. Comparison of the two other general-purpose baselines in the task of identifying any annotated travel location, whether affirmed, negated, or future.

Model	Precision	Recall	F1 score
Stanford Location NER ^a	62.3	81.7	70.7
spaCy NER	52.5	63.4	57.5

^aNER: named-entity recognition.

Feature importance of our proposed model was measured as an ablation experiment quantifying contribution of each feature

set with respect to a baseline set of surface-level features. The results of this feature ablation are presented in [Table 5](#).

Table 5. Ablation results of model features.

Feature set	Precision (95% CI)	Recall (95% CI)	F1 score (95% CI)
Tokens, lemmas, token shape	85.0 (80.8-87.0)	69.1 (64.5-74.1)	76.3 (72.7-79.5)
+ Part of speech	87.4 (83.3-90.5)	72.2 (67.6-76.9)	79.1 (75.8-82.1)
+ Character n-grams	85.9 (82.1-89.4)	77.2 (72.9-81.1)	81.3 (78.0-84.3)
+ Gazetteer match	84.1 (79.9-87.1)	75.2 (70.9-79.7)	79.4 (76.0-82.6)
+ Word embedding clusters	87.6 (83.7-90.9)	74.9 (70.7-79.3)	80.8 (77.3-83.7)
All features	88.0 (84.2-90.9)	83.3 (78.9-86.6)	85.6 (82.5-87.9)

Since hyperparameter values such as token window size and the size of character n-grams have an impact on model performance, some of these were varied to identify the optimal parameters and quantify sensitivity to value changes. These

values did not appear to cause much variance in precision but did improve recall. The best performing model integrated features within a window size of 2 tokens and character n-grams of size 2 and 3. This analysis is presented in [Table 6](#).

Table 6. Results of varying certain feature extraction hyperparameters.

Parameter difference	Precision (95% CI)	Recall (95% CI)	F1 score (95% CI)
Window size=1	87.8 (84.3-91.2)	80.5 (76.7-84.3)	84.1 (80.9-86.4)
Window size=3	87.4 (83.9-90.9)	79.2 (74.8-83.6)	83.1 (79.7-86.1)
Character n-gram size=2	87.9 (84.5-91.1)	82.7 (79.0-86.4)	85.2 (82.6-88.1)
Character n-gram size=3	87.9 (84.0-91.0)	81.2 (77.4-84.8)	84.4 (81.4-87.0)
Character n-gram size=4	87.9 (84.1-91.0)	81.0 (77.1-84.6)	84.3 (81.6-86.9)
Optimal hyperparameters (window size=2, character n-gram sizes=2 and 3)	88.0 (84.2-90.9)	83.3 (78.9-86.6)	85.6 (82.5-87.9)

Since this pipeline has been deployed for running daily operations, we have measured the total processing time for querying source documents, processing them, and storing the results back to the database. In February 2020, responding to the spread of COVID-19, over 978,000 documents were processed with a median length of 2414 characters. Throughput for the system is an average of 2.01 documents per second using a single machine core (ie, Intel Xeon E5-4650 @ 2.10 GHz).

Postdeployment Evaluation Results

Evaluation of the system on a previously unseen sample of patients and documents following deployment resulted in an estimated system performance of 71.9% precision, 78.1% recall, and 74.9% F1 measure. In consultation with infectious disease analysts, the measured performance was determined to be not sufficient, so the newly created annotations were used to perform an error analysis of model predictions. Several of the false positives were instances in which the initial model predicted future or negated travel. Many of the false negatives were locations whose spelling did not match canonical gazetteer entries. Besides these errors, the annotated corpus was reviewed, and some annotations were adjudicated and modified. Several of these adjudicated annotations occurred in statements of future or hypothetical travel, indicating that the guideline may benefit from refinement. Additional feature types were evaluated including word embedding clusters and character n-grams. Ablation experiments were performed to determine the best feature set. Before these improvements, the model's performance on the test split of annotated documents was 82.9% precision, 76.4% recall, and 79.5% F1 measure. The improved model

exhibited a 6.1% increase in F1 and is reported in the preceding section.

Discussion

Principal Findings

In this study, we have explored and demonstrated the feasibility of extracting patient travel history from clinical documentation of a large national cohort. Relying on semiautomated methods for lexicon expansion, we manually labeled a data set to train an extraction model. Our findings demonstrate that training an accurate model to extract travel mentions is feasible in an automated system. Both labeled sets and the modeling approaches were chosen to minimize development time and computational resources necessary to continue surveillance in day-to-day operations. The baseline comparison presented here is a simplified evaluation, but it demonstrates that general-purpose geoparsing solutions alone result in lower precision. This is due to these systems labeling all locations in text, while the task in this study was to identify past affirmed travel for public health response. Our system has been deployed for operational use.

Since graphics processing unit acceleration is not currently available in the computing environments where system development was conducted and models later deployed, more costly model options such as recurrent neural networks and contextual language models (ie, Bidirectional Encoder Representations from Transformers, Generative Pre-trained Transformer 2, and Universal Language Model Fine-tuning) have not yet been evaluated [35-37]. Rather, we present a linear

model that leverages information from neural language models that can be rapidly trained on widely available hardware. Our system has been deployed on a central processing unit to support health care operations churning through large volumes of clinical documents daily. Since deployment in November 2017 through July 2020, more than 18 million documents have been processed as part of a daily data pipeline.

Our proposed system exhibits promising performance of location mention retrievals in text. Future work may attempt to disambiguate location text or resolve mentions to a geolocation. Additionally, annotation of timing or duration of travel could provide another layer of travel history. We have not yet measured how frequently such information is recorded in EHR, but we hope to answer this question in a future study.

Further, although negated location mentions were annotated for the reference standard, these were not used at the time of training. The extraction model may be improved by distinguishing between affirmed and negated locations, and public health efforts could be assisted by such information when conducting efforts to rule out exposures.

Originally deployed in November 2017, this system has been used in VA to rapidly respond to public health events. Given that this system processes millions of notes each week, it can be integrated with other biosurveillance metrics to give a sense to decision makers about the extent that travel history may be playing a role in disease dynamics. These higher-level metrics include the absolute counts and percentages of emergency department visits with respect to syndromic categories such as respiratory illness, gastrointestinal conditions, and diseases.

After this system was deployed, it was initially used by VA biosurveillance analysts to rapidly screen new cases of Zika in the nation. Rapid visibility of automated travel history extractions allowed analysts to scan new cases and perform additional chart review daily if there was no travel history in the patient profile. For example, if a new case of Zika occurred for a patient living in a midwestern state without an automated travel profile, the next step could be to perform a chart review to determine whether the case might be reflective of local transmission. Since chart review can be a time-consuming process, the fact that many patients already had a travel profile to an endemic area meant that less time was consumed in manual review. Cases that did have a profile could be easily verified by performing keyword searches stemming from locations visited.

This system was already running in early 2020 such that it could be leveraged in response to the spread of COVID-19 in the United States. This capability was useful in early stages of transmission, as it was able to identify mentions of travel to endemic areas.

One way that this extraction was useful was by setting up broad surveillance and then using travel as additional criteria. Specifically, case review could be prioritized using case definitions for respiratory and influenza-like illness we had developed for broad syndromic surveillance. These definitions required both relevant International Classification of Diseases, Tenth Revision (ICD-10) diagnostic coding and concepts for

signs and symptoms extracted from chief complaint text as concept unique identifiers (CUI) in the Unified Medical Language System [38]. ICD-10 codes were drawn from resources provided by the International Society for Disease Surveillance [39]. CUI values for these broad syndromic definitions were collected from prior work [40]. Combining multiple aspects of clinical data with travel information permitted rapid selection of cases to review in January 2020. During February 2020, as testing for COVID-19 became more standard, combining orders or results for such testing and travel regions permitted more specific case review.

This study was limited as the selection of annotation corpus was performed by location keywords and travel phrases. This could entail a bias in the location types and variants used in annotation and model training. We attempted to address these limitations of the annotation corpus by performing manual chart review and random sampling. This review was limited by resources, so future work may ensure that the annotation corpus is free of potential bias by performing further review by random sampling or other strategies leveraging clinical evidence. Although acceptable, the relatively low level of annotator agreement may be a potential limitation of our system. Additionally, preliminary data from corpus selection indicates that past affirmed travel mentions in clinical documents are likely to be imbalanced. Given this, experiments in ongoing work may achieve better model performance by exploring class imbalance strategies such as undersampling or oversampling.

The utility of this system for COVID-19 specifically was limited to a relatively brief window in the early phases of transmission when containment was possible and travel was a pertinent risk factor. The Centers for Disease Control and Prevention (CDC) guidance for Persons Under Investigation on February 12, 2020, included explicit mention for travel to Wuhan or Hubei Province [41]. By March 4, the CDC removed these criteria and instead encouraged clinicians to use best judgment for virus testing [42]. In some surveillance efforts, travel history was deemed to be less important in risk assessment once community acquisition increased [43].

When planned in 2017, considerations for this capability were primarily concerned with importation of infectious disease, so the scope was focused on travel outside the United States. Presently, the deployed model does extract some location mentions within the continental United States, but the frequency of these extractions and remaining false negatives have not yet been estimated. Such a capability could be useful for aiding containment and tracking of international spread for future epidemics.

Future work would benefit from refining the annotation guideline and training to prevent disagreement like the examples presented here. Furthermore, model training would benefit from additional annotation sampled from other time periods and infectious diseases to assess true validity and improve performance for continued surveillance preparedness. As no data augmentation methods were applied in this study, there are several approaches using keyword or pattern replacement that could improve model performance with available resources [44-46].

Extracting travel history from clinical notes involves the same challenges as any efforts based on secondary use of clinical data. Patient travel history is not always recorded in clinical notes, so this is still an incomplete representation of actual events. Perl and Price [15] emphasized the importance of travel history documentation in the EHR, particularly to enable timely response to emerging pathogens like COVID-19 to protect both patients and frontline health care providers. We note that recording travel history in a structured format would yield more accurate results if faithfully adhered to, but we shall continue to monitor, evaluate, and improve this system as we leverage it to attain a more complete picture of disease for global health. It remains impossible to predict future transition phases in response to COVID-19, but we believe that such a capability could provide value in regions engaged in the phases of prevention or containment.

Conclusion

We have proposed methods for automated extraction of patient travel history from clinical documents and have demonstrated enhanced capabilities for improving public health systems. This system has been combined with surveillance metrics to inform decision makers on the role of travel history in disease dynamics. It has also been used to reduce a manual chart review to verify whether an infectious disease case may have been imported or locally acquired. These methods have now been deployed to ongoing operations to support enhanced understanding and monitoring. Validity of the system has been sufficient to reduce analyst burden, while computational requirements remain low, allowing thousands of documents to be processed daily. Such capabilities have been leveraged in infectious disease responses, such as Zika and the importation of COVID-19 to the United States, amid dynamically evolving situations. Such systems may aid future efforts to prevent and contain the spread of COVID-19 and other infectious diseases.

Acknowledgments

This project was funded by the Department of Homeland Security, National Biosurveillance Integration Center. This project was supported with facilities at the VA Salt Lake City Informatics, Decision Enhancement, and Analytic Sciences Center in Salt Lake City, Utah and data resources from VA Informatics and Computing. The views expressed are those of the authors and do not necessarily reflect the position or policy of the Department of VA or the United States government. We thank the editor and anonymous reviewers for their feedback in ameliorating the reporting of this study.

Conflicts of Interest

None declared.

References

1. Duong TN, Waldman SE. Importance of a travel history in evaluation of respiratory infections. *Curr Emerg Hosp Med Rep* 2016;4(3):141-152 [FREE Full text] [doi: [10.1007/s40138-016-0109-y](https://doi.org/10.1007/s40138-016-0109-y)] [Medline: [32226655](https://pubmed.ncbi.nlm.nih.gov/32226655/)]
2. Alexander C, Cottom L, Smith K, Perrow K, Coyne M, Jones B. Schistosomiasis in Scottish travellers: public health importance of laboratory testing and the need for enhanced surveillance. *J Public Health (Oxf)* 2018 Mar 01;40(1):138-145. [doi: [10.1093/pubmed/idx024](https://doi.org/10.1093/pubmed/idx024)] [Medline: [28335010](https://pubmed.ncbi.nlm.nih.gov/28335010/)]
3. Chapman WW, Gundlapalli AV, South BR, Dowling JN. Natural language processing for biosurveillance: detection and characterization from textual clinical reports. In: Castillo-Chavez C, Chen H, Lober WB, Thurmond M, Zeng D, editors. *Infectious Disease Informatics and Biosurveillance: Research, Systems and Case Studies*. Berlin, Germany: Springer; 2011:279-310.
4. Peterson K, Denhalter D, Patterson OV, Jones M. Can electronic clinical notes identify travelers with Zika? *Open Forum Infect Dis* 2018;5(suppl_1):S168. [doi: [10.1093/ofid/ofy210.457](https://doi.org/10.1093/ofid/ofy210.457)]
5. Morris T, Cicchinelli M, McGarvey S, Johnson J, Conn L, Loonsk J. PHIN preparedness: outbreak management. 2005 Presented at: 2005 AMIA Annual Symposium; 2005; Washington, DC.
6. Salminen J, Almerikhi H, Milenković M, Jung SG, An J, Kwak H, et al. Anatomy of online hate: developing a taxonomy and machine learning models for identifying and classifying hate in online news media. 2018 Presented at: International AAAI Conference On Web And Social Media; 2018; Stanford, CA.
7. García-Constantino M, Atkinson K, Bollegala D, Chapman K, Coenen F, Roberts C, et al. CLIEL: context-based information extraction from commercial law documents. In: *Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law*. 2017 Presented at: ICAIL '17; June 2017; London, UK p. 79-87. [doi: [10.1145/3086512.3086520](https://doi.org/10.1145/3086512.3086520)]
8. Steinkamp JM, Chambers C, Lalevic D, Zafar HM, Cook TS. Toward complete structured information extraction from radiology reports using machine learning. *J Digit Imaging* 2019 Aug;32(4):554-564 [FREE Full text] [doi: [10.1007/s10278-019-00234-y](https://doi.org/10.1007/s10278-019-00234-y)] [Medline: [31218554](https://pubmed.ncbi.nlm.nih.gov/31218554/)]
9. Christopoulou F, Tran T, Sahu S, Miwa M, Ananiadou S. Adverse drug events and medication relation extraction in electronic health records with ensemble deep learning methods. *J Am Med Inform Assoc* 2020 Jan 01;27(1):39-46 [FREE Full text] [doi: [10.1093/jamia/ocz101](https://doi.org/10.1093/jamia/ocz101)] [Medline: [31390003](https://pubmed.ncbi.nlm.nih.gov/31390003/)]

10. Finkel JR, Grenager T, Manning C. Incorporating non-local information into information extraction systems by Gibbs sampling. In: Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics. 2005 Presented at: ACL '05; June 2005; Ann Arbor, MI p. 363-370. [doi: [10.3115/1219840.1219885](https://doi.org/10.3115/1219840.1219885)]
11. Liu X, Zhang S, Wei F, Zhou M. Recognizing named entities in tweets. 2011 Presented at: 49th Annual Meeting of the Association for Computational Linguistics; June 19-24, 2011; Portland, OR p. 359-367.
12. Magge A, Weissenbacher D, Sarker A, Scotch M, Gonzalez-Hernandez G. Deep neural networks and distant supervision for geographic location mention extraction. *Bioinformatics* 2018 Jul 01;34(13):i565-i573 [[FREE Full text](#)] [doi: [10.1093/bioinformatics/bty273](https://doi.org/10.1093/bioinformatics/bty273)] [Medline: [29950020](https://pubmed.ncbi.nlm.nih.gov/29950020/)]
13. Niehus R, De Salazar PM, Taylor AR, Lipsitch M. Estimating underdetection of internationally imported COVID-19 cases. *medRxiv*. Preprint posted online February 14, 2020. [[FREE Full text](#)]
14. Niehus R, De Salazar PM, Taylor AR, Lipsitch M. Quantifying bias of COVID-19 prevalence and severity estimates in Wuhan, China that depend on reported cases in international travelers. *medRxiv*. Preprint posted online February 18, 2020. [[FREE Full text](#)] [doi: [10.1101/2020.02.13.20022707](https://doi.org/10.1101/2020.02.13.20022707)]
15. Perl TM, Price CS. Managing emerging infectious diseases: should travel be the fifth vital sign? *Ann Intern Med* 2020 Apr 21;172(8):560-561 [[FREE Full text](#)] [doi: [10.7326/M20-0643](https://doi.org/10.7326/M20-0643)] [Medline: [32120386](https://pubmed.ncbi.nlm.nih.gov/32120386/)]
16. 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 Jun 08;6(2):e19509 [[FREE Full text](#)] [doi: [10.2196/19509](https://doi.org/10.2196/19509)] [Medline: [32490846](https://pubmed.ncbi.nlm.nih.gov/32490846/)]
17. Dawood FS, Ricks P, Njie GJ, Daugherty M, Davis W, Fuller JA, et al. Observations of the global epidemiology of COVID-19 from the pre-pandemic period using web-based surveillance: a cross-sectional analysis. *Lancet Infect Dis* 2020 Nov;20(11):1255-1262 [[FREE Full text](#)] [doi: [10.1016/S1473-3099\(20\)30581-8](https://doi.org/10.1016/S1473-3099(20)30581-8)] [Medline: [32738203](https://pubmed.ncbi.nlm.nih.gov/32738203/)]
18. Xu B, Kraemer MUG, Open COVID-19 Data Curation Group. Open access epidemiological data from the COVID-19 outbreak. *Lancet Infect Dis* 2020 May;20(5):534 [[FREE Full text](#)] [doi: [10.1016/S1473-3099\(20\)30119-5](https://doi.org/10.1016/S1473-3099(20)30119-5)] [Medline: [32087115](https://pubmed.ncbi.nlm.nih.gov/32087115/)]
19. Veterans Health Administration. US Department of Veterans Affairs. URL: <https://www.va.gov/health/> [accessed 2021-03-15]
20. Alba P, Patterson O, Viernes B, Denhalter D, Bailey N, Wilson A, et al. The super annotator: a method of semi-automated rare event identification for large clinical data sets. 2016 Presented at: AMIA 2016 Annual Symposium; November 12-16, 2016; Chicago, IL.
21. GeoNames. 2020. URL: <http://www.geonames.org/> [accessed 2021-03-15]
22. Cornia R, Patterson O, Ginter T, DuVall S. Rapid NLP development with Leo. 2014 Presented at: AMIA 2014 Annual Symposium; November 15-19, 2014; Washington, DC.
23. DuVall S, Forbush T, Cornia R, Ginter T, Adams B, Palmer M, et al. Reducing the manual burden of medical record review through informatics: 772. *Pharmacoepidemiology Drug Safety* 2014;772.
24. Lafferty J, McCallum A, Pereira FCN. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: Proceedings of the Eighteenth International Conference on Machine Learning. 2001 Presented at: ICML '01; June 2001; Williamstown, MA p. 282-289 URL: https://repository.upenn.edu/cis_papers/159/
25. Okazaki N. CRFsuite: a fast implementation of conditional random fields (CRFs). Naoaki Okazaki. 2007. URL: <http://www.chokkan.org/software/crfsuite/> [accessed 2021-03-15]
26. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Machine Learning Res* 2011;12:2825-2830.
27. Guo J, Che W, Wang H, Liu T. Revisiting embedding features for simple semi-supervised learning. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. 2014 Presented at: EMNLP '14; October 2014; Doha, Qatar p. 110-120 URL: <http://www.aclweb.org/anthology/D/D14/D14-1012.pdf> [doi: [10.3115/v1/d14-1012](https://doi.org/10.3115/v1/d14-1012)]
28. Chapman AB, Peterson KS, Alba PR, DuVall SL, Patterson OV. Detecting adverse drug events with rapidly trained classification models. *Drug Saf* 2019 Jan;42(1):147-156 [[FREE Full text](#)] [doi: [10.1007/s40264-018-0763-y](https://doi.org/10.1007/s40264-018-0763-y)] [Medline: [30649737](https://pubmed.ncbi.nlm.nih.gov/30649737/)]
29. Marcus M, Santorini B, Marcinkiewicz MA. Building a large annotated corpus of English: The Penn Treebank. *Penn Libraries* 1993.
30. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York, NY: Chapman and Hall/CRC; 1994.
31. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. *J Am Stat Assoc* 1983 Jun;78(382):316-331. [doi: [10.1080/01621459.1983.10477973](https://doi.org/10.1080/01621459.1983.10477973)]
32. spaCy. URL: <https://spacy.io/> [accessed 2021-03-15]
33. Gritta M, Pilehvar MT, Collier N. A pragmatic guide to geoparsing evaluation: toponyms, Named Entity Recognition and pragmatics. *Lang Resour Eval* 2020;54(3):683-712 [[FREE Full text](#)] [doi: [10.1007/s10579-019-09475-3](https://doi.org/10.1007/s10579-019-09475-3)] [Medline: [32802011](https://pubmed.ncbi.nlm.nih.gov/32802011/)]
34. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychological Meas* 2016 Jul 02;20(1):37-46. [doi: [10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104)]
35. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. *arXiv*. Preprint posted online October 11, 2018. [[FREE Full text](#)]

36. Howard J, Ruder S. Universal language model fine-tuning for text classification. arXiv. Preprint posted online January 18, 2018. [FREE Full text] [doi: [10.18653/v1/p18-1031](https://doi.org/10.18653/v1/p18-1031)]
37. Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I. Language models are unsupervised multitask learners. OpenAI Blog 2019;1(8):9.
38. Lindberg DAB, Humphreys BL, McCray AT. The Unified Medical Language System. Methods Inf Med 2018 Feb 06;32(04):281-291. [doi: [10.1055/s-0038-1634945](https://doi.org/10.1055/s-0038-1634945)]
39. Hicks P, Pavlin JA, Baer A, Swenson D, Kite-Powell A, Jayatilleke AU, et al. Syndromic surveillance. Online J Public Health Inform 2015;7(1):e28.
40. Chapman WW, Dowling JN, Baer A, Buckeridge DL, Cochrane D, Conway MA, et al. Developing syndrome definitions based on consensus and current use. J Am Med Inform Assoc 2010;17(5):595-601 [FREE Full text] [doi: [10.1136/jamia.2010.003210](https://doi.org/10.1136/jamia.2010.003210)] [Medline: [20819870](https://pubmed.ncbi.nlm.nih.gov/20819870/)]
41. Evaluating and reporting persons under investigation (PUI). Centers for Disease Control and Prevention. URL: <https://web.archive.org/web/20200213123640/https://www.cdc.gov/coronavirus/2019-nCoV/hcp/clinical-criteria.html> [accessed 2020-02-13]
42. Evaluating and reporting persons under investigation (PUI). Centers for Disease Control and Prevention. URL: <https://web.archive.org/web/20200305002956/https://www.cdc.gov/coronavirus/2019-nCoV/hcp/clinical-criteria.html> [accessed 2020-03-05]
43. Krausz M, Westenberg JN, Vigo D, Spence RT, Ramsey D. Emergency response to COVID-19 in Canada: platform development and implementation for eHealth in crisis management. JMIR Public Health Surveill 2020 May 15;6(2):e18995 [FREE Full text] [doi: [10.2196/18995](https://doi.org/10.2196/18995)] [Medline: [32401218](https://pubmed.ncbi.nlm.nih.gov/32401218/)]
44. Kobayashi S. Contextual augmentation: data augmentation by words with paradigmatic relations. 2018 Presented at: 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2; June 2018; New Orleans, LA p. 452-457. [doi: [10.18653/v1/n18-2072](https://doi.org/10.18653/v1/n18-2072)]
45. Wang X, Pham H, Dai Z, Neubig G. SwitchOut: an efficient data augmentation algorithm for neural machine translation. 2018 Presented at: 2018 Conference on Empirical Methods in Natural Language Processing; October 2018; Brussels, Belgium. [doi: [10.18653/v1/d18-1100](https://doi.org/10.18653/v1/d18-1100)]
46. Wei J, Zou K. EDA: easy data augmentation techniques for boosting performance on text classification tasks. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. 2019 Presented at: EMNLP-IJCNLP '19; November 2019; Hong Kong, China. [doi: [10.18653/v1/d19-1670](https://doi.org/10.18653/v1/d19-1670)]

Abbreviations

CDC: Centers for Disease Control and Prevention
CRF: conditional random fields
CUI: concept unique identifiers
EHR: electronic health record
ICD-10: International Classification of Diseases, Tenth Revision
NER: named-entity recognition
NLP: natural language processing
VA: Veterans Affairs

Edited by T Sanchez; submitted 23.12.20; peer-reviewed by M Rodrigues, Z Ren, O Serban, S Das Bhattacharjee; comments to author 18.01.21; revised version received 05.02.21; accepted 12.02.21; published 24.03.21.

Please cite as:

Peterson KS, Lewis J, Patterson OV, Chapman AB, Denhalter DW, Lye PA, Stevens VW, Gamage SD, Roselle GA, Wallace KS, Jones M

Automated Travel History Extraction From Clinical Notes for Informing the Detection of Emergent Infectious Disease Events: Algorithm Development and Validation

JMIR Public Health Surveill 2021;7(3):e26719

URL: <https://publichealth.jmir.org/2021/3/e26719>

doi:[10.2196/26719](https://doi.org/10.2196/26719)

PMID:[33759790](https://pubmed.ncbi.nlm.nih.gov/33759790/)

Surveillance (<http://publichealth.jmir.org>), 24.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

App Use and Usability of a Barcode-Based Digital Platform to Augment COVID-19 Contact Tracing: Postpilot Survey and Paradata Analysis

Thomas Foster Scherr¹, PhD; Jenna Maria DeSousa^{1*}, BSc; Carson Paige Moore^{1*}, BSc; Austin Hardcastle¹, BEng; David Wilson Wright¹, PhD

Department of Chemistry, Vanderbilt University, Nashville, TN, United States

*these authors contributed equally

Corresponding Author:

Thomas Foster Scherr, PhD

Department of Chemistry

Vanderbilt University

1234 Stevenson Center Lane

Nashville, TN, 37212

United States

Phone: 1 615 322 5516

Email: thomas.f.scherr@vanderbilt.edu

Abstract

Background: The COVID-19 pandemic has drastically changed life in the United States, as the country has recorded over 23 million cases and 383,000 deaths to date. In the leadup to widespread vaccine deployment, testing and surveillance are critical for detecting and stopping possible routes of transmission. Contact tracing has become an important surveillance measure to control COVID-19 in the United States, and mobile health interventions have found increased prominence in this space.

Objective: The aim of this study was to investigate the use and usability of MyCOVIDKey, a mobile-based web app to assist COVID-19 contact tracing efforts, during the 6-week pilot period.

Methods: A 6-week study was conducted on the Vanderbilt University campus in Nashville, Tennessee. The study participants, consisting primarily of graduate students, postdoctoral researchers, and faculty in the Chemistry Department at Vanderbilt University, were asked to use the MyCOVIDKey web app during the course of the study period. Paradata were collected as users engaged with the MyCOVIDKey web app. At the end of the study, all participants were asked to report on their user experience in a survey, and the results were analyzed in the context of the user paradata.

Results: During the pilot period, 45 users enrolled in MyCOVIDKey. An analysis of their enrollment suggests that initial recruiting efforts were effective; however, participant recruitment and engagement efforts at the midpoint of the study were less effective. App use paralleled the number of users, indicating that incentives were useful for recruiting new users to sign up but did not result in users attempting to artificially inflate their use as a result of prize offers. Times to completion of key tasks were low, indicating that the main features of the app could be used quickly. Of the 45 users, 30 provided feedback through a postpilot survey, with 26 (58%) completing it in its entirety. The MyCOVIDKey app as a whole was rated 70.0 on the System Usability Scale, indicating that it performed above the accepted threshold for usability. When the key-in and self-assessment features were examined on their own, it was found that they individually crossed the same thresholds for acceptable usability but that the key-in feature had a higher margin for improvement.

Conclusions: The MyCOVIDKey app was found overall to be a useful tool for COVID-19 contact tracing in a university setting. Most users suggested simple-to-implement improvements, such as replacing the web app framework with a native app format or changing the placement of the scanner within the app workflow. After these updates, this tool could be readily deployed and easily adapted to other settings across the country. The need for digital contact tracing tools is becoming increasingly apparent, particularly as COVID-19 case numbers continue to increase while more businesses begin to reopen.

(JMIR Public Health Surveill 2021;7(3):e25859) doi:[10.2196/25859](https://doi.org/10.2196/25859)

KEYWORDS

contact tracing; COVID-19; mobile health; usability; app; usage; tracking; monitoring; survey; pilot

Introduction

The COVID-19 pandemic quickly evolved from localized transmission to broad and sustained community transmission across the United States [1]. Most states initially enacted stay-at-home orders to curb the spread of the virus, with schools and businesses shifting to virtual operations early in the pandemic [2-5]. As local restrictions have lifted, many schools and workplaces have implemented new changes that enable a safe return to work. These adjustments include masking and social distancing requirements as well the implementation of daily health checks [6-8].

Relative to many other infectious diseases, COVID-19 has a high degree of asymptomatic transmission [9,10]. An infected person may not know that they are infected, and if they are not engaging in good public health measures (eg, regular handwashing, mask wearing, social distancing), they may come into direct contact with other individuals and spread the virus. As such, there is a clear ceiling on the usefulness of symptom monitoring alone. It is widely recognized that broad testing and effective contact tracing are necessary to counter instances of unknowing transmission, particularly because widespread vaccine distribution has been slow to take hold.

Across the United States, contact tracing efforts have been implemented to varying extents [11-17]. When an individual is confirmed or suspected to be positive for COVID-19, contact tracers will interview that person and identify any close contacts that they have had during their infectious window. After building a list of potential contacts for each index case, tracers reach out to each of the contacts and let them know of their potential exposure, either helping them to locate nearby testing options or providing counseling on effective self-isolation.

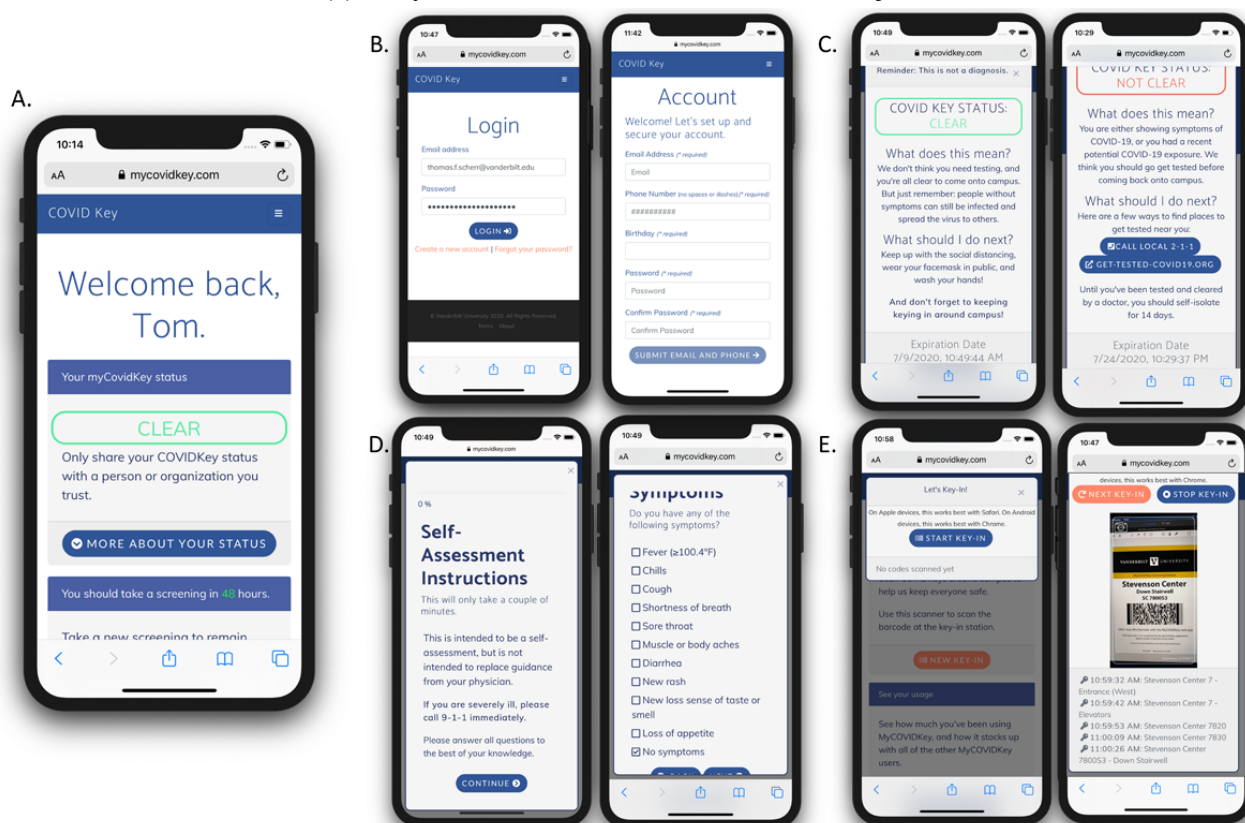
In states and counties where there has been a rapid rise in cases, the need for contact tracing has often outpaced the ability to implement a rigorous surveillance system. This has presented an opportunity for developers of digital health tools, which were already increasing in use prior to the pandemic, to redirect their efforts to build contact tracing platforms [18-22]. Several digital contact tracing platforms have been described in the academic

literature [23-25], and more are available through for-profit technology companies [20]. Although these tools use a variety of technologies, two of the most popular strategies are (1) continuous location monitoring and (2) observing Bluetooth interactions between devices. Due to the size and commercial motives of the developers of these platforms, they have been subject to intense scrutiny over potential privacy concerns regarding data ownership and use—even before they have been released. These apps, while simple and useful for contact tracing efforts, are viewed with skepticism by many people, who may not wish to share such granular personal data.

Due to the rapid emergence of the pandemic and the digital contact tracing tools that soon emerged in response, few formal studies have been performed to understand user priorities and improve usability. Modeling has been applied to examine contact tracing app acceptance rates [26], and frameworks have been proposed to evaluate the potential scalability of these apps [27]. Recently, a survey study identified the importance of enhancing perceived benefits and self-efficacy and also identified the perceived barrier of privacy concerns [28]. Although these findings are useful in the initial design of contact tracing platforms, none of these studies investigated specific existing apps.

We previously described an alternative digital contact tracing tool, MyCOVIDKey, that is designed to supplement existing contact tracing infrastructure [29]. Our primary motivation was to develop a tool that would be less invasive while retaining efficacy. The software is a mobile-friendly web app that is based around recurring self-assessments and barcode-based location “key-ins” in which users scan a bar code specific to a particular location (Figure 1). Users are assigned a status of CLEAR or NOT CLEAR and are then provided personalized recommendations based on their risk and their location. A thorough detailing of the development, implementation, and utility of the app for contact tracing is shown elsewhere, but briefly: over the duration of the pilot study, 45 unique accounts were created, 227 self-assessments were performed, and users performed 1410 key-ins at 48 unique locations (out of a possible 71 locations) [29].

Figure 1. The main screens of the MyCOVIDKey web app: (A) the landing page, which presents a user's status after a valid login and allows them to access self-assessments and key-ins; (B) the login and account creation pages; (C) the screens for CLEAR (left) and NOT CLEAR (right) statuses; (D) the brief COVID-19 risk assessment; and (E) the key-in feature, which enables users to scan location-specific bar codes.



At the conclusion of the previously described MyCOVIDKey pilot study, we analyzed aggregate and individual app use data, and we also asked our users to provide feedback on their experience with the app. In this manuscript, we provide an analysis of these use statistics and the user feedback; we also describe the subsequent adjustments we are making to improve the app. We present this information so that public health officials who are preparing to implement digital contact tracing tools and software developers who are building them can learn from our users and their experience.

Methods

Institutional Review Board Approval

This study was reviewed and approved by the Vanderbilt University Institutional Review Board (#200976; June 1, 2020).

Pilot Study Design

The MyCOVIDKey pilot study ran from June 17 to July 29, 2020, and was centered around a series of interconnected science and engineering buildings on Vanderbilt University's campus. During this phase of the COVID-19 pandemic, most work was being conducted remotely, with the exception of research that required the researcher's physical presence on campus.

Anyone over 18 years of age with an internet-connected mobile device was eligible to participate. To recruit participants, potential users were informed of the ongoing pilot by flyers posted around the participating buildings and by two recruiting emails that were distributed through the department mailing list

to faculty, staff, and students. At the end of the first week, an email was sent to department-wide email lists and the flyers were updated in Stevenson Center to announce the introduction of a weekly raffle for a US \$20 Amazon gift card. A second email and an update to the posted recruiting flyers were deployed near the end of the fourth week to the same email lists, detailing an increase in the weekly raffle prize from US \$20 to \$45 and the addition of a second prize of a \$15 gift card.

To incentivize participation, weekly raffles were performed in which the number of entries for each user correlated to that user's app use. Briefly, users were awarded 10 points for each self-assessment performed and 1 point for each key-in; there was a cap of 30 points per week for each category, limiting users to a maximum of 60 entries into the raffle (3 self-assessments and 30 key-ins per week), and the number of points reset at the start of each week. Users could view their individual statistics and accumulation of raffle points within the web app. A modal pop-up window that displayed the user's most frequently keyed-in locations and weekly points obtained toward the raffle could be viewed by clicking the "See Your Stats" button on the home page. Users could view box plots generated for both average daily and all-time key-ins versus average weekly and all-time scans, allowing them to compare their use to the aggregate and anonymous data of the MyCOVIDKey user base. A pop-up window was available to describe the point system toward the raffle.

Paradata Collection and Analysis

The term *paradata* in this manuscript refers to collected data that indicate how users engaged with the app. This information includes timestamped user-performed actions and events, and it describes the process by which users interact with the MyCOVIDKey site. An in-house paradata library was built to collect data on app use. The paradata library was built using JavaScript and enabled “behind-the-scenes” data collection using AJAX, the commonly used asynchronous HTTP request library. Each time a page was loaded or a button was clicked, the following information was sent asynchronously (without blocking the user experience on the front end) to our database: timestamp, user action, user ID (if the user was authenticated with a valid username and password), the user’s current PHP session ID, the page that the action occurred on, the user’s IP address, and the user’s device and browser information. When users were authenticated, their paradata could be associated with other user feedback and actions (ie, results of screenings and key-ins).

Postpilot Survey

Near the conclusion of the study, participants were asked to voluntarily provide feedback on their experience with the MyCOVIDKey app. All individuals who consented and participated in the MyCOVIDKey on-campus pilot study and who provided a verified email address were invited to participate in the postpilot survey. A custom survey was hosted on Research Electronic Data Capture (REDCap), a secure research electronic database, and an individual, nonpublic link was provided by email to all registered MyCOVIDKey users at the end of the 6-week trial period [30,31]. Data entered on the survey webpage were stored directly on the REDCap server. The survey totaled 59 questions across eight sections, and users could refer to the previous pages to review or change their answers until they submitted the survey. The eight survey sections included demographics (5 questions), COVID-19 testing history (3 questions), system usability scales (10 questions per feature, 30 total questions), impressions of MyCOVIDKey (12 questions), impressions of digital contact tracing tools and features (6 questions), and open-response questions specific to MyCOVIDKey (3 questions). Demographic data included age, gender identity, race, and on-campus role (eg, student, postdoctoral researcher, faculty, or staff). Usability was measured using a System Usability Scale (SUS) [32]. The SUS consisted of 10 statements, such as “I found MyCOVIDKey unnecessarily complex,” which were then ranked using a 5-point Likert scale in which respondents were asked to what degree they agreed with the statements (strongly disagree to strongly agree). The SUS was used to assess the perceived usability of the MyCOVIDKey app as a whole and the key-in and the self-assessment features of the app independently. The SUS

was scored by converting each answer to a score from 0-4, summing the total responses for each question, and then multiplying the total by 2.5. This produces a score from 0-100; 68 is considered a benchmark score for usability, and scores lower than this value are considered to indicate below-average usability. Impressions of MyCOVIDKey were also measured on a 5-point Likert scale, where participants responded to phrases such as “I found it easy to take screenings every two days” or “Using MyCOVIDKey positively impacted my feeling of safety on campus.” All phrases were positively coded to ensure consistent composite scores for all questions. To determine impressions of general contact tracing and digital contact tracing tools, a binary yes/no system was implemented to determine general user impressions regarding the importance (eg, “Do you think contact tracing is important?”), effectiveness, security, ease of use, and time and effort costs of contact tracing interventions. The final section of the survey encouraged users to fill in free-response questions relating to their personal MyCOVIDKey experiences as well as suggestions for the development team to improve usability. To encourage participation, participants who completed the survey were entered into a raffle for a US \$50 Amazon gift card.

Data Analysis

At the conclusion of the study, all paradata were exported from the MySQL database. Similarly, all survey responses were exported from REDCap. Distributions were analyzed using box plots, violin plots, and custom Likert-style plots. Linear regressions were performed to analyze user sign-up data. All analytical and quantitative statistical analysis was performed with statistical packages in Python (ie, StatsModels, NumPy, SciPy). All data visualizations were generated in Python using common numerical plotting packages (ie, matplotlib, Seaborn).

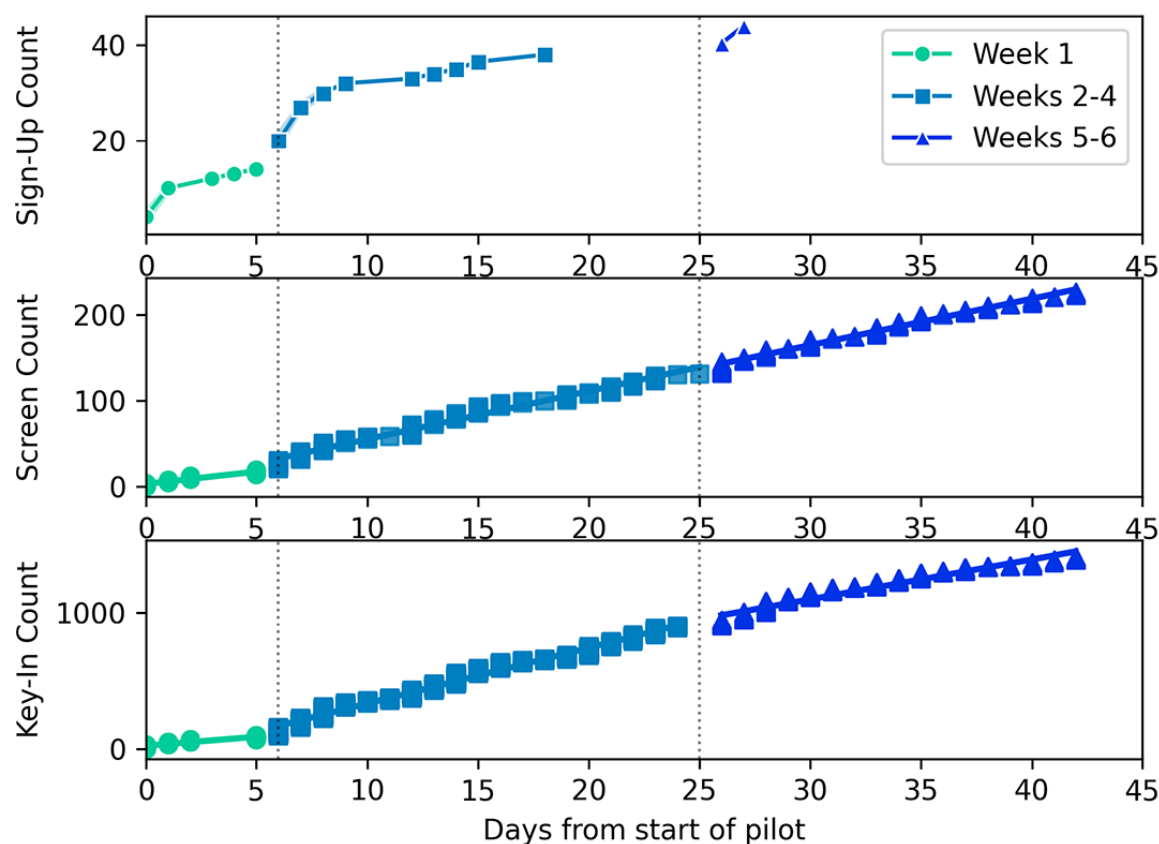
Results

Paradata Analysis

User-Aggregated Paradata

In the first week, organic growth of the user base quickly plateaued (Figure 2). The MyCOVIDKey user base grew organically to 14 users in the first week of use through signup flyers posted throughout the participating buildings. We observed that the first recruiting email had the most substantial impact on new user signups, while a second recruiting email was less effective (Figure 2 (top)). An additional 6 user accounts were created on the day that the first recruiting email was sent, and the user base reached 32 accounts by day 9 of the pilot and 38 users by day 18, after which it remained constant for 8 more days. The second recruiting email had a more limited effect, adding only 7 users over the course of the next week.

Figure 2. The user signups, screenings, and key-ins for the MyCOVIDKey app over three time periods during the pilot study: week 1, weeks 2-4, and weeks 5-6. The vertical dashed lines at days 6 and 25 represent days on which recruiting emails were distributed.



The self-screen and key-in counts relative to the launch date of the web app are reported in Figure 2 (center and bottom) and summarized in Table 1. In the first week, 2.85 self-assessments per day were completed across all users. This number increased to 5.58 screenings completed per day after the first promotional

email and remained approximately constant (5.39 per day) after the second promotional email.

Key-ins saw a similar uptick in use after the first promotional email, increasing from 13.1 key-ins per day to 40.8 key-ins per day. After the second email, the number of key-ins per day decreased to a rate of 29.2 per day.

Table 1. Linear regression parameters for screenings and key-ins over each of the three study periods.

Slope and study week	Screens per day	r^2 (SE)
Screening slope		
1	2.85	0.944 (0.228)
2-4	5.58	0.989 (0.078)
5-6	5.39	0.991 (0.077)
Key-in slope		
1	13.1	0.929 (0.531)
2-4	40.8	0.993 (0.170)
5-6	29.2	0.970 (0.325)

Paradata to Determine Incentive-Driven App Use

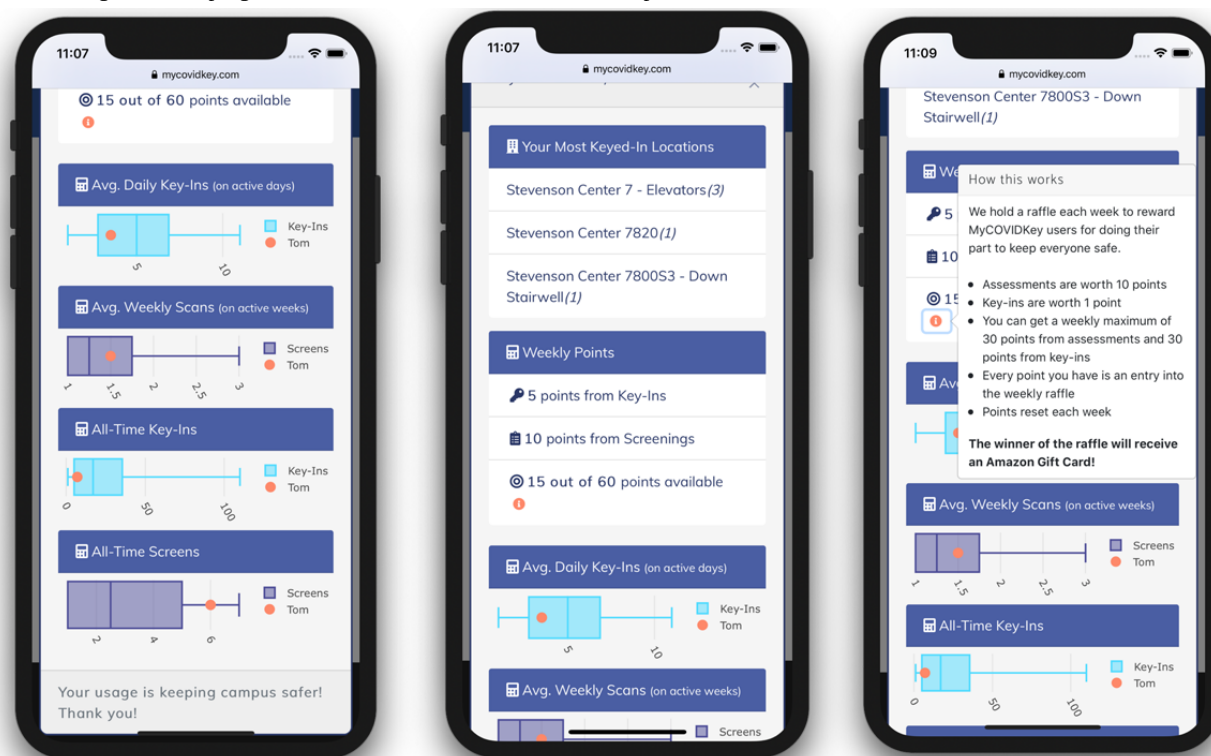
We evaluated the user-generated paradata to see if there were any users that would be considered “high-score seekers”—users who are primarily interested in reaching the maximum number

of points (Figure 3). Specifically, we compared the number of scans and key-ins for each user with the number of times that user clicked on the “See Your Stats” modal button to view their number of entries in the weekly raffle. The paradata showed that the number of views of the statistics modal for each user

generally correlated with increased use of the other app features (key-ins, [Multimedia Appendix 1](#) top; self-assessments, [Multimedia Appendix 1](#) bottom). This trend was also generally seen when comparing the number of logins for each user (color of markers). There are two obvious outliers on different ends of the analysis: (1) a user who viewed their statistics

substantially more than twice the number of times of the next highest user (ie, used the screening feature more often than average) but who did not key in frequently; and (2) a user who had nearly double the number of key-ins as the next highest user, performed relatively few self-assessments, and viewed their stats only a few times.

Figure 3. Screenshots of the user statistics modal in MyCOVIDKey. Users were presented with statistics comparing their app use to that of the rest of the userbase along with their progress toward the maximum number of raffle points allowed each week.



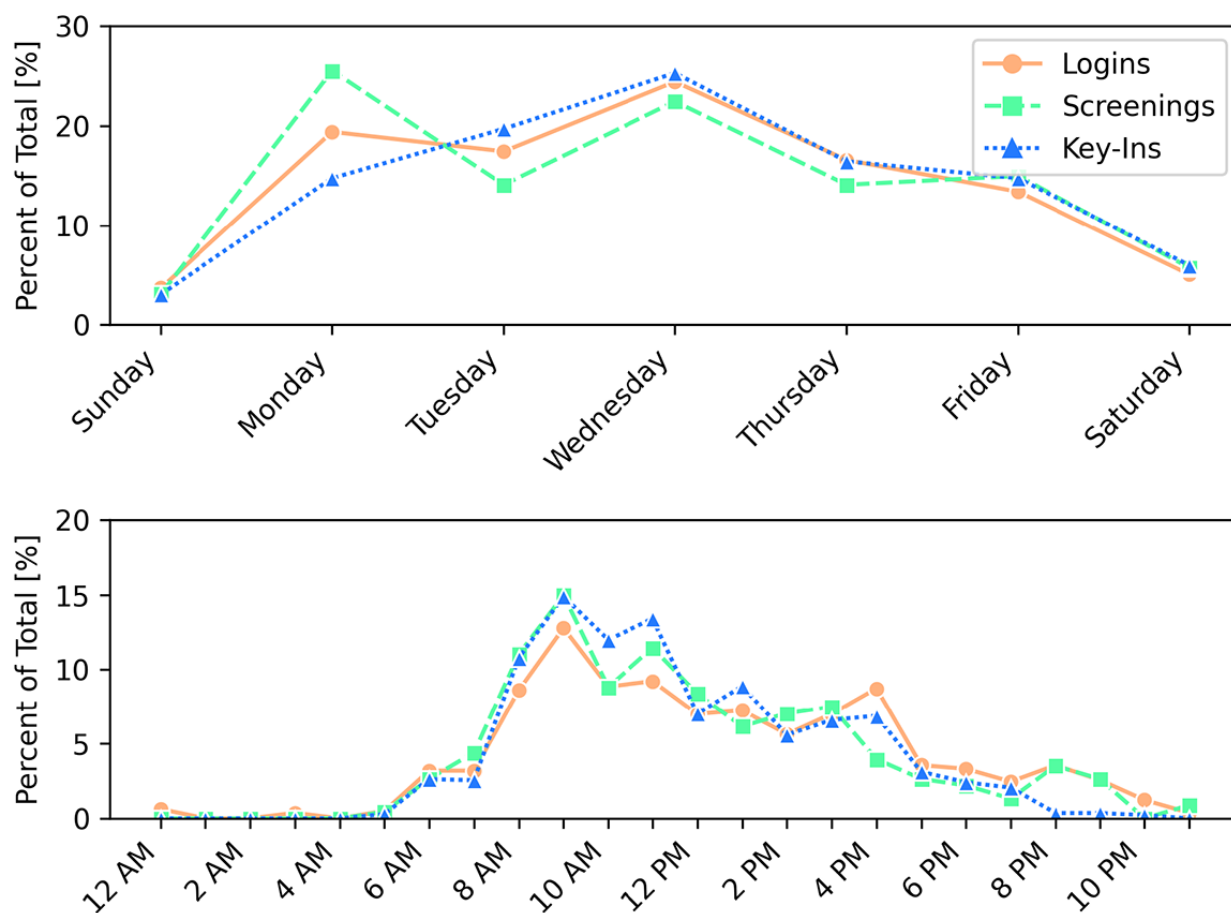
Event Statistics

Throughout the 6-week pilot study of MyCOVIDKey, paradata were collected and analyzed to better understand the use and usability of the contact tracing platform. From the paradata, 45 users logged a total of 1270 unique sessions and used 114 distinct browser/mobile device combinations. The time required to create a MyCOVIDKey account was measured from the first presentation of the account creation page to the time that the completed user registration was recorded in our database. The entire account creation process took users an average of 2.30 minutes (SD 2.07) to complete. Once users created an account, they were asked to complete an initial self-assessment. Each of the self-assessments had an expiration lifetime of 48 hours, at which point the user was required to complete another assessment prior to accessing other features in the app. On average, the time to complete a screening was 18.22 seconds (SD 20.04) ([Multimedia Appendix 2](#), top), with an average time of 3.83 days (SD 4.23) between screenings ([Multimedia Appendix 2](#), center). This was expected, as each screening remained valid for 48 hours. Measuring from the time that a user launched the modal to scan a barcode to the time that the pop-up window was closed, key-in events had a mean duration of 75.30 seconds (SD 97.89) ([Multimedia Appendix 2](#), bottom). Removing any instances in which the modal was presented and

the user did not scan a barcode at a location, these key-in events included an average of 3.17 key-ins (SD 4.59) per time that the modal was launched. This indicates that most users scanned barcodes at multiple locations within the same session. These disparities were then compared for each individual user, and the results showed that individual users mostly mirrored the aggregate distributions ([Multimedia Appendix 3](#)).

During the study, the Stevenson Center Complex operated on a limited access basis; graduate students and faculty maintained staggered schedules, which varied from hourly shifts to alternating days for each research group. Thus, several people were entering and exiting the building at various points throughout the day. Every login, key-in, and screening was grouped by the day of the week and time of day that each was performed, and the distribution of percentages is shown in [Figure 4](#). Both logins and key-ins followed a similar trend in which they each exhibited an increase until a midweek peak. Notably, activity was minimal during the weekend, as expected. The highest percentage of screenings (58/227, 25.6%), on the other hand, was performed on Monday. Once a user finishes a screening, the CLEAR or NOT CLEAR status remains for a period of 48 hours. Therefore, if a screening was completed on Monday, the user would not have to take another until Wednesday, resulting in a decrease of screenings on Tuesdays and another increase two days later on Wednesday.

Figure 4. Use of the MyCOVIDKey app by (top) time of day and (bottom) day of week. Time of day is user local time (CST). Total values for reference: 814 logins, 227 screenings, and 1410 key-ins.



Similarly, each login, key-in, and screening was sorted by the time of day at which it was accessed (Figure 4). After peaking at 9 AM, logins remained relatively consistent between 10 AM and 4 PM, followed by a sharp decrease through the late afternoon and evening. Consistent logins were expected, as the server automatically logged users out for security reasons after 20 minutes of inactivity. After that, a new login was required to key in at a new location or complete a screening. More than half of all logins (61.5%), key-ins (72.4%), and screenings (68.28%) occurred before 1 PM, and most of them occurred from 9 AM to 10 AM (logins, screenings, and key-ins: 12.78%, 14.98%, and 14.89%, respectively).

User Demographics and COVID-19 Testing

Of the 45 MyCOVIDKey users during the pilot period, 26 (58%) completed the postpilot survey; 4 users (9%) started the survey but did not complete it. Our survey respondents were primarily White (24/30, 80%), female (20/30, 67%), and aged 20-30 years (22/30, 73%) (Table 2). Three quarters of our users (23/30, 77%) identified themselves as graduate students. The high proportion of graduate students enrolled during the pilot period was expected, as Vanderbilt University's reopening policies emphasized remote work, which was more readily achievable for faculty and administrative staff than for graduate students involved in laboratory research.

Table 2. Survey respondents' demographic characteristics and campus roles (N=45).

Characteristic	Value, n (%)
Survey status	
Sent	45 (100)
Complete	26 (58)
Incomplete	4 (9)
Not started	15 (33)
Gender	
Female	20 (67)
Male	10 (33)
Race	
White	24 (80)
Asian	2 (7)
Hispanic, Latino, or of Spanish Origin and ^a White	2 (7)
Black or African American	1 (3)
Asian and ^a White	1 (3)
Age (years)	
20-30	22 (73)
31-40	6 (20)
41-50	1 (3)
>60	1 (3)
Campus role	
Graduate student	23 (77)
Faculty member	6 (20)
Postdoctoral researcher	1 (3)

^aUsers selected multiple checkboxes.

Although the self-assessment feature provided users with a symptom selection that could indicate potential infection with SARS-CoV-2, users were not asked to provide any information about their experiences or results from COVID-19 diagnostic testing within the MyCOVIDKey app. In the postpilot survey, we asked users if they had received diagnostic testing, and if

so, to provide the results of this diagnostic testing ([Table 3](#)). More than one-third (11/30, 37%) of our users were tested for SARS-CoV-2 throughout the pilot, with 30% (9/30) of the users being tested only one time and 7% (2/30) of the users being tested twice. One user (1/30, 3%) indicated that they had tested positive for SARS-CoV-2 during the pilot period.

Table 3. Survey responses related to COVID-19 testing during the pilot period (n=30).

Variable	Value, n (%)
Not tested	19 (63)
Tested	11 (37)
Number of tests	
1	9 (30)
2	2 (7)
Test result	
Positive	1 (3)
Negative	10 (33)

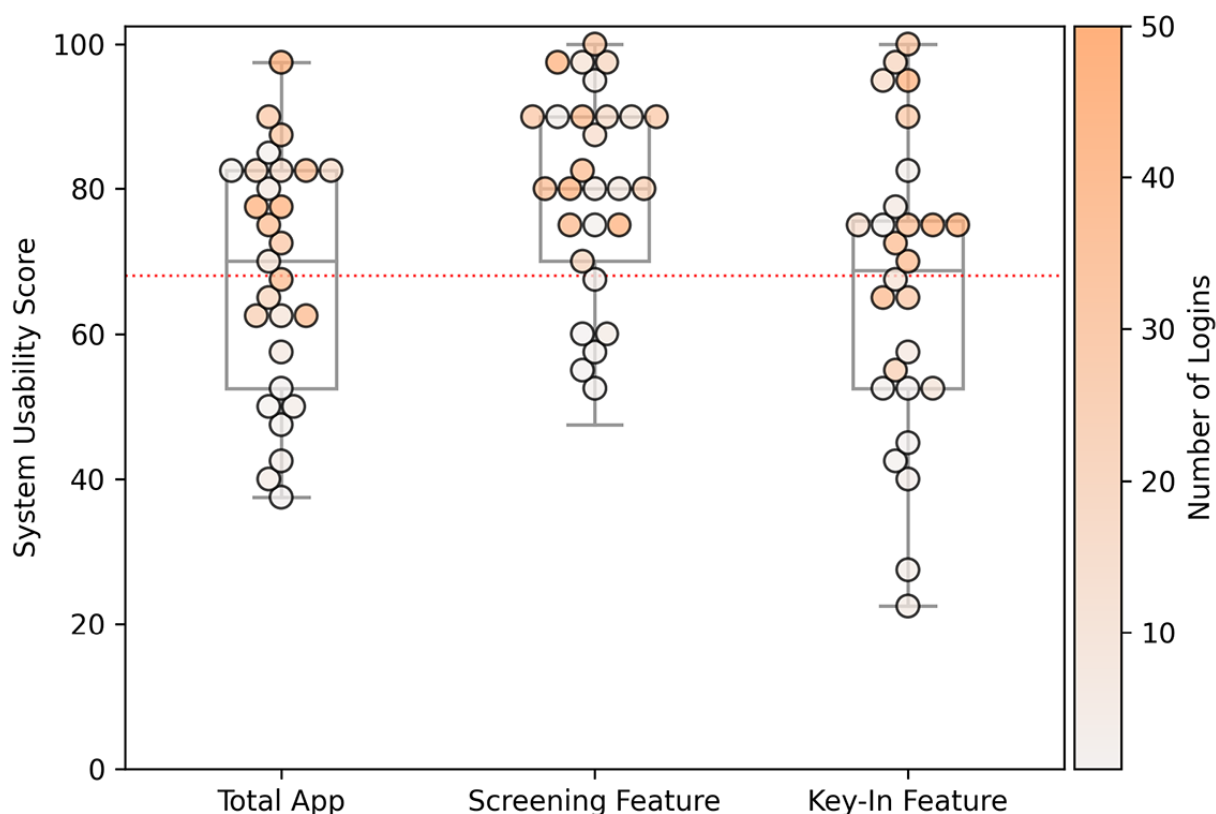
Postpilot Survey

System Usability Scores

Users were also asked to provide their impressions on the MyCOVIDKey app as a whole as well as the self-assessment and key-in features individually (Figure 5). Using an SUS score of 68 as the threshold of acceptable usability (red dotted line in Figure 5) [32], the app as a whole can be considered to have adequate usability (SUS 70). The screening feature easily passed this metric, with a median score of 80.0 and a bottom quartile score of 70.0. The key-in feature appeared to be more polarizing in its usability; the range of its SUS scores was much larger,

with a minimum and maximum score of 22.5 and 100, respectively. Although it still surpassed the threshold for usability, the key-in feature had the lowest top quartile score (75.6), median score (68.75), bottom quartile score (52.5), and minimum score. Each individual user's SUS score was compared to that user's number of logins (Multimedia Appendix 4, top), self-assessments (Multimedia Appendix 4, center), and key-ins (Multimedia Appendix 4, bottom). For the app as a whole and for the individual features (self-assessments and key-ins), there were positive correlations between more frequent use and higher SUS scores.

Figure 5. System Usability Scale scores for the total MyCOVIDKey app, the screening feature, and the key-in feature. The threshold for acceptable usability of 68 is represented with a dashed horizontal line. The markers represent the scores provided by individual users, and the intensity of the color correlates to the number of logins for that particular user. The maximum color intensity indicates users with more than 50 logins.



User Preferences for Contact Tracing Apps

In addition to the SUS-related questions, this study focused on understanding user perceptions (Figure 6) of MyCOVIDKey using a series of questions on the Likert scale (strongly disagree: 1, to strongly agree: 5). Users strongly agreed with the statements that taking screenings was simple and easy to do every two days; however, a majority of users disagreed that the number of screenings should be increased. In general, users thought that the coverage of MyCOVIDKey key-in stations around the buildings used for the pilot study was appropriate.

There was a positive shift in user perception of keying in over the course of the pilot, with 43% (12/28) of users agreeing or strongly agreeing that keying in felt natural by the end of the study compared to 25% (7/28) at the beginning. A large portion of users (12/28, 43%), were mostly ambivalent on whether MyCOVIDKey made them feel safer around campus or at locations that they were visiting, with slightly more respondents disagreeing with those statements than agreeing. Over 70% of users (20/28, 71%) agreed or strongly agreed that they felt their health information was kept private while they used MyCOVIDKey.

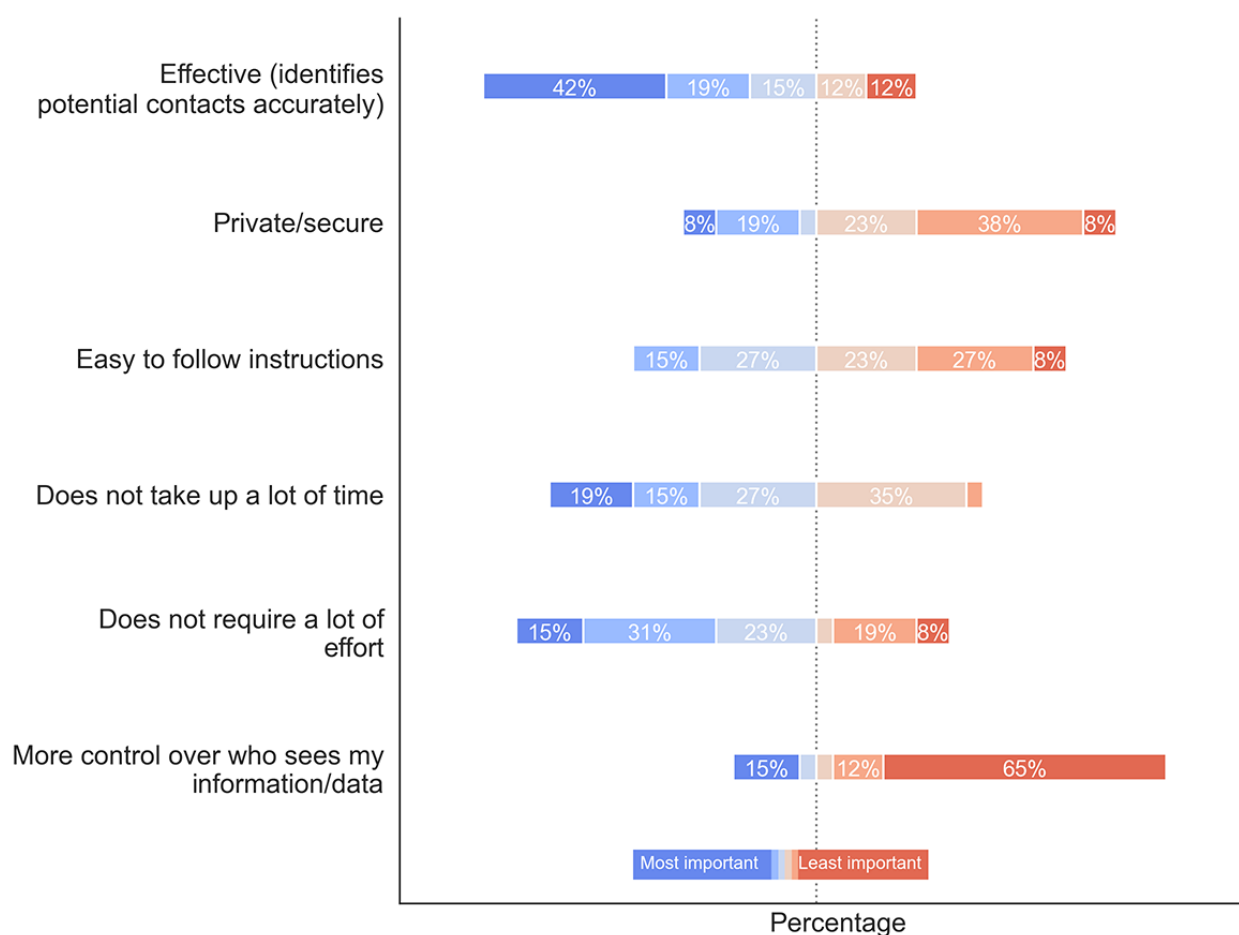
Figure 6. The distributions of responses to specific questions about MyCOVIDKey (n=28).

When these distributions were separated into groups based on the SUS score that the user provided ([Multimedia Appendix 5](#)) and the number of times the user logged in ([Multimedia Appendix 6](#)), the distributions across the Likert scale were mostly similar. Users who used the app less frequently mostly self-identified as less frequent users ([Multimedia Appendix 6](#)), and this distribution was also skewed toward less positive ratings of the app ([Multimedia Appendix 5](#)). There are several other notable differences; users who logged into the app more frequently and gave it a higher SUS score had a more skewed distribution than those who logged in less frequently and gave lower SUS scores. For instance, more users who gave SUS scores below the threshold of usability disagreed or strongly disagreed with the statements that taking screenings was a simple process, it was easy to take a screening every two days, the coverage of MyCOVIDKey key-in stations was appropriate, and that it felt natural keying in to places at the end of the study. This analysis held true when the users were separated as logging in more or less than the median number of logins. Users who

rated the app higher and users who logged in more frequently had more positive impressions of safety.

Although users were not explicitly asked to compare MyCOVIDKey to other contact tracing approaches, we recognize that contact tracing is a new phenomenon to the general public; thus, we sought to understand which features users prioritized ([Figure 7](#)). The study population strongly indicated that contact tracing effectiveness (ie, the platform accurately identifies potential contacts) was the most important value proposition, as 42% of users (11/26) ranked it as the most important trait of a contact tracing tool. Users next valued minimization of effort and time, where 46% (12/26) and 34% (9/26) of users ranked these traits as either the most or second most important, respectively. Conversely, the MyCOVIDKey users appeared to deprioritize privacy, given the trade-off between the other characteristics. Surprisingly, 65% (17/26) of the survey respondents indicated that having more control over who sees their information or data was the least important feature.

Figure 7. A diagram depicting how MyCOVIDKey users ranked the importance of the features of the app. The questions asked the users (n=26) to identify the most important (blue) and least important (red) features of the app.



User preferences were largely the same across users who gave MyCOVIDKey a lower SUS score ([Multimedia Appendix 7](#)) and across infrequent users ([Multimedia Appendix 8](#)). Some interesting deviations from this pattern are that more users who rated the app below the usability threshold ranked privacy as less important and ranked minimal effort as their most important preference. In contrast, users who logged into MyCOVIDKey more frequently more often ranked minimizing effort as their lowest priority.

Direct User Response to MyCOVIDKey

Although the majority of the poststudy survey questions enabled users to select from a predefined set of answers, users also provided open-ended responses regarding their opinions on the best and worst parts of MyCOVIDKey and on how the app could be improved. The open-ended responses are aggregated in [Table 4](#).

Table 4. Aggregated user responses to open-ended questions.

Category and response	Value, n (%)
Best part of MyCOVIDKey	
Total responses	18 (100)
Useful tool, good for contact tracing	6 (33)
Accessible/good locations for barcodes	3 (17)
Easy to use/simple	3 (17)
Tracks statistics/game-like	2 (11)
Minimal time required	2 (11)
Scanning worked well	2 (11)
Self-assessments	2 (11)
Gift card incentives	1 (6)
Worst part of MyCOVIDKey	
Total responses	19 (100)
Web browser	10 (53)
Key-in did not always work as expected	6 (32)
Difficult to use when carrying things	3 (16)
Automatic logout	3 (16)
Effort required	2 (11)
Unsure who sees information	1 (5)
Frequent self-assessments	1 (5)
Number of steps to get to key-in window	1 (5)
Unclear instructions	1 (5)
Unsure if others are using frequently/effectively?	1 (5)
Expectation of receiving notification if someone tested positive for SARS-CoV-2	1 (5)
Ways to improve MyCOVIDKey	
Total responses	20 (100)
Make it a native app	12 (60)
Open directly to scanner, faster scanning	4 (20)
Integration with other location services, self-report locations visited	4 (20)
Offline mode	2 (10)
More key-in locations	2 (10)
More transparency on how information is collected and shared	1 (5)
Clearer instructions for use	1 (5)
More use statistics	1 (5)
Option to self-report positive test result	1 (5)

In the “best part of MyCOVIDKey” open-ended responses, users indicated that they believed the tool had purpose and was a suitable option for contact tracing. Other responses noted that the platform was simple, was accessible at expected locations, worked well, and had recurring self-assessments. There were also two mentions of statistics/game-like mentality, and one mention of the gift card incentives. From the “worst part of MyCOVIDKey” and the “improve MyCOVIDKey” responses, there is a clear directive to build the platform as a native app (10 of 19 responses identified the web browser in the “worst”

section, and 12 of 20 responses in the “improve” section asked for it specifically to be made into an app). This appears to be an umbrella response for many users, as users also noted room for improvement with regard to the automatic logout feature, an open direct to scanner feature, offline mode, and integration with other location services. A few users indicated that it was difficult to use MyCOVIDKey when carrying things, which is a now-obvious problem for chemistry graduate students working on experiments in multiple laboratories. These users, and others, noted that the ability to integrate with other location services

or to self-report locations they had visited would be a helpful remedy. Although 2 of 18 users shared that scanning worked well, some users (6/19) did have unexpected issues with the key-in feature.

Discussion

Contact tracing is poised to play a large role in the strategic preparedness and response plan during the remainder of the COVID-19 pandemic in the United States. With evidence of broad asymptomatic transmission, identifying individuals who may have been exposed and providing them with proper testing and isolation is an essential means to slow the spread of disease. As case numbers and deaths continue to increase rapidly, resource-intensive manual contact tracing can be augmented with digital tools that can efficiently identify those at risk of exposure. We developed MyCOVIDKey to provide an alternative to other digital contact tracing solutions that use constant GPS and Bluetooth monitoring and pose potential privacy concerns. In this work, we focused on analysis of user-generated paradata and a postpilot survey to understand user impressions and develop a roadmap for improvements.

Usability of MyCOVIDKey

SUS scores were used to evaluate the platform as a whole as well as its two main features. An SUS presents a simple and objective tool to evaluate basic usability and identify areas for improvement. The median SUS score for the MyCOVIDKey app overall was above the threshold for acceptable usability, albeit close to it. Although the key-in feature was above this threshold, it was closer to the cutoff, and the survey respondents expressed more grossly divergent opinions about this feature. In contrast, the recurring self-assessments scored well, with more than 75% of the feature-specific SUS scores deeming its usability to be acceptable. In general, users who more frequently participated in the study through app use reported higher SUS scores on average for each feature of MyCOVIDKey.

In addition to the SUS questions, specific sections of the survey were focused on understanding user impressions of the two main features of the app. Survey respondents indicated that the self-assessments were easy to use, simple, and noninvasive. This was reinforced by the paradata, which showed that the majority of assessments could be completed in less than 1 minute. The responses to the survey showed that users had a more polarized opinion of the key-in feature, which also aligns well with the SUS scores and paradata.

User-generated paradata represent an important tool to understand individual and aggregate behavior within an app. Although paradata are commonly tracked and analyzed in consumer apps, their use has received considerably less attention in health care-related apps. Despite survey respondents' criticisms of the ease of use of the key-in feature of MyCOVIDKey, we note that it was frequently used, and the paradata indicate that users were able to perform this task relatively quickly. However, the negative survey responses were strong, and the paradata may reflect the fact that a core group of users dominated the use of the app and perhaps were less critical of the feature.

User Priorities

From our survey responses, several clear user priorities were identified. Our users indicated that their three most important characteristics of a digital contact tracing tool were (1) effectiveness at accurately tracing contacts, (2) not requiring much time, and (3) not requiring much effort. When given the opportunity to rank their preferences, users ranked control over who sees information and security as the two least important characteristics of a digital contact tracing tool. This result was surprising, considering the discourse surrounding mobile contact tracing apps; however, it was also rarely mentioned in the open response section of the survey. Additionally, this result contrasts with those in previously published studies describing attitudes toward contact tracing digital apps [26,28]. Our users, regardless of how they scored the usability of the app or how often they logged in, indicated that they felt their information was securely maintained in MyCOVIDKey. In that sense, these responses may reflect users' opinions specifically on MyCOVIDKey in the context of the pilot study, particularly one that took place at a research institution and primarily enrolled graduate students and faculty, and may not be representative of the broader population. Furthermore, as the Department of Chemistry is a relatively small, self-contained environment, participants may have felt more comfortable regarding privacy concerns knowing that the study was occurring within their community. Indeed, concerns over privacy have been linked to larger, for-profit corporations and technology companies, and data security is typically a larger consumer concern in the event of any data breach.

Incentivized Participation in Digital Contact Tracing

It is generally understood that digital contact tracing platforms must reach a critical user volume to be effective. Employers or educational institutions can require that their employees or students use these platforms as a condition of their employment or access to facilities; however, this requirement may be met with resentment and have a negative impact on user perception and cooperation with contact tracing teams. Regardless, we were unable to require user participation in our pilot study. In place of a mandate, after a week of moderate enrollment, we deployed a weekly raffle to encourage uptake and continued use of MyCOVIDKey. The number of entries for any given user was based upon the number of key-ins and self-assessments that the user performed that week, with a cap on each to minimize the effect of high-score seekers. Flyers advertising the pilot study were modified to announce the raffle, and a recruiting email was sent out to departmental email lists to promote the study and the raffle. In the three days after this change, the number of accounts created more than doubled, and we saw at least 2-fold increases in the rates of key-ins and self-assessments. This approximately proportional increase suggests that the increases in use rates was tied to the increase in users and was not simply due to increased use by previous users due to the raffle. Because the raffle was announced at the same time as the first marketing email, we cannot decouple the effect of one from that of the other.

After 2 weeks, the raffle prize was increased, and we announced a runner-up raffle prize to (1) avoid a drop-off in use after the

July 4 holiday and (2) attempt to further increase our user enrollment. The number of new user accounts created after this change was minimal, and there was a surprising decrease in use rates (both key-ins and self-assessments per day). There are several possible explanations for this, including user signup saturation, individual work schedules, and pandemic fatigue. Regardless of the reasoning, this result does suggest that the initial incentive of the raffle was sufficient. Although 2 users noted the presence of an incentive or the gamification of the scoring system as features they enjoyed about MyCOVIDKey, the analysis of the paradata did not suggest that this was the only motivational factor in their use of the app.

It is clear that the first recruiting email, with its incentive for use, had a positive impact on user signups. The sharp increases in the screening rate and key-in rate were mostly a result of the influx of users, after accounting for the number of users active during each period of the study. The second recruiting email, even with increased incentives, resulted in modest new account signups and decreased use rates for self-assessments and key-ins.

With the number of potential users (those who were working in our study buildings throughout the pilot period) remaining constant, the lower number of signups can likely be explained by the theory of innovation diffusion. We had likely captured the early adopters and early majority, and by weeks 5-6 we were beginning to approach the late majority of users. Interestingly, the second recruiting email and the more valuable raffle prize coincided with a slight decrease in screenings per day and a more noticeable drop in key-in rate.

Strengths and Limitations of This Study

This study is one of the first rigorous pilot evaluations performed on a digital contact tracing app. It is one of the first formal studies to investigate the usability of a COVID-19 contact tracing app with the intent of making iterative improvements. The usability analysis combines both quantitative and qualitative user feedback. This study is also one of the first studies to compare user-generated paradata to user survey responses in mobile health (mHealth) apps. This valuable tool provides unique insight into the difference between perception of usability and actual use patterns.

The primary limitations of this study include a relatively modest sample size and a narrow user demographic. These weaknesses were mostly circumstantial: social distancing requirements and “safer-at-home” orders that were in place during the study limited the number of people on campus, and our prospective participant pool was limited mainly to students and faculty within Stevenson Center, our study location. This resulted in the selection of a young and technology-savvy cohort. This does limit the overall generalizability of the study, as this demographic is potentially more comfortable with technology or currently available mHealth interventions than the broader population. Another limitation of this study is that the study was stopped on July 29, and not all users used the app for 6 full weeks prior to completing the survey.

Guidance for MyCOVIDKey and for Other Digital Contact Tracing Apps

Based on the analysis of the paradata and user feedback received from the survey, we established a set of directions for improving the MyCOVIDKey platform. The MyCOVIDKey users, while a narrow demographic, showed strong preferences for a platform that was effective at identifying potential contacts while also minimizing the effort and time required for use. It is our hope that other developers can learn from the feedback that we received.

In general, the recurring self-assessment was favorably received by our users. It received high usability scores and positive feedback in survey responses; also, the paradata indicated that the task could be accomplished quickly. The users indicated a preference for not increasing the frequency of the required self-assessments; however, this may not be a barrier based on the approval that the feature received. As such, our focus on the self-assessment will be the addition of questions related to diagnostic testing and results. The goal of the MyCOVIDKey pilot was to rapidly deploy a solution for beta testing and to identify improvements prior to an anticipated larger rollout. Because this could be accomplished without sharing personal testing results, we made the explicit decision not to ask our users for this information. However, the benefits of integrating diagnostic results, when available, are obvious. This information has the easily recognizable utility of confirming the positivity of users that indicated symptoms in the app and, also of great importance, of removing persons of interest who test negative for SARS-CoV-2 from contact tracing queues. These questions were unnecessary during the pilot evaluation but will be critical in any broader release.

As the most frequently performed user action inside the app, we will focus substantial effort on improving the key-in feature of MyCOVIDKey. Users indicated that this was something that they wanted to be able to perform faster, with some users indicating that they would like to be able to do it directly from the home screen. Additionally, some users expressed concerns over network connectivity and how that hindered their ability to use the key-in feature. All these concerns can be readily addressed by converting MyCOVIDKey from a mobile-friendly web app to a native phone app. Although mobile-friendly web apps can often blur the lines between native and web apps, it is clear that in this instance, users have preferences that can be better met by a native app. Indeed, from our survey responses, many users explicitly stated their preference for a native app instead of a browser-based platform.

Conclusions

Digital platforms are uniquely positioned to play a large role in contact tracing efforts during the COVID-19 pandemic. In response, we have developed MyCOVIDKey, a web-based contact tracing app, and evaluated it over the course of a 6-week pilot study. In this work, we analyzed aggregate and individual use data and compared it to user feedback from a postpilot survey. We were able to obtain quantitative data to understand how and when MyCOVIDKey was used as well as how users felt about the different components of the app. Although the app and its individual features received acceptable usability

scores, this work clearly shows that users prioritize contact tracing effectiveness along with minimal time and effort requirements. This feedback provides us with a clear blueprint for how to improve our app prior to an expanded rollout as well as guidelines for other digital contact tracing efforts moving forward.

Acknowledgments

The authors would like to acknowledge support from the Amazon Web Services Diagnostic Development Initiative. This work used REDCap, which is supported by the National Center for Advancing Translational Sciences at the National Institutes for Health (UL1TR000445). The authors would like to acknowledge Mr Jason Bradley of Vanderbilt University Information Technology for his assistance in server configuration and deployment.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The number of times each user launched the modal window to view their statistics compared to the number of (top) self-assessments they performed, (bottom) key-ins they performed, and (color and size) the number of times they logged in.

[[PDF File \(Adobe PDF File\), 77 KB - publichealth_v7i3e25859_app1.pdf](#)]

Multimedia Appendix 2

The probability density distribution and quartile distributions of the duration of (top) key-in events, (middle) time between screenings, and (bottom) duration of self-assessments.

[[PDF File \(Adobe PDF File\), 102 KB - publichealth_v7i3e25859_app2.pdf](#)]

Multimedia Appendix 3

The distributions of the (top) duration of scan events, (middle) duration of self-assessments, and (bottom) time between completed self-assessments for each individual user.

[[PDF File \(Adobe PDF File\), 111 KB - publichealth_v7i3e25859_app3.pdf](#)]

Multimedia Appendix 4

System Usability Scale scores for the entire app and the number of logins for each user (top). Each user's System Usability Scale score for the self-assessment feature of the app and the number of self-assessments that they completed (middle). Each user's System Usability Scale score for the key-in feature of the app and the number of times that the user keyed in to a location on campus (bottom).

[[PDF File \(Adobe PDF File\), 84 KB - publichealth_v7i3e25859_app4.pdf](#)]

Multimedia Appendix 5

The distribution of responses to MyCOVIDKey-specific questions, separated into responses from users who gave the overall app System Usability Scale scores above (green, above) and below (red, bottom) the usability threshold of 68. A score of 1 reflects a response of strongly disagree, and a score of 5 reflects a response of strongly agree. Markers reflect individual responses (jitter has been artificially added to enhance visualization; only discrete integer values could be selected).

[[PDF File \(Adobe PDF File\), 158 KB - publichealth_v7i3e25859_app5.pdf](#)]

Multimedia Appendix 6

The distribution of responses to MyCOVIDKey-specific questions, separated into responses from users who logged in more (green, above) and less (red, bottom) than the median number of logins. A score of 1 reflects a response of strongly disagree, and a score of 5 reflects a response of strongly agree. Markers reflect individual responses (jitter has been artificially added to enhance visualization; only discrete integer values could be selected).

[[PDF File \(Adobe PDF File\), 156 KB - publichealth_v7i3e25859_app6.pdf](#)]

Multimedia Appendix 7

The distribution of responses to user preferences in a digital contact tracing tool, separated into responses from users who gave the overall app system usability scores above (green, above) and below (red, bottom) the usability threshold of 68. A score of 1 reflects a feature that is most important to that user, and a score of 5 reflects a feature that is least important to that user. Markers represent individual responses (jitter has been artificially added to enhance visualization; only discrete integer values could be selected).

[[PDF File \(Adobe PDF File\), 93 KB - publichealth_v7i3e25859_app7.pdf](#)]

Multimedia Appendix 8

The distribution of responses to user preferences in a digital contact tracing tool, separated into responses from users who logged in more (green, above) and less (red, bottom) than the median number of logins. A score of 1 reflects a feature that is most important to that user, and a score of 5 reflects a feature that is least important to that user. Markers represent individual responses (jitter has been artificially added to enhance visualization; only discrete integer values could be selected).

[[PDF File \(Adobe PDF File\), 93 KB - publichealth_v7i3e25859_app8.pdf](#)]

References

1. Healthcare facilities: managing operations during the COVID-19 pandemic. US Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/hcp/guidance-hcf.html> [accessed 2020-04-13]
2. Nussbaumer-Streit B, Mayr V, Dobrescu A, Chapman A, Persad E, Klerings I, et al. Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review. *Cochrane Database Syst Rev* 2020 Apr 08;4:CD013574. [doi: [10.1002/14651858.CD013574](https://doi.org/10.1002/14651858.CD013574)] [Medline: [32267544](https://pubmed.ncbi.nlm.nih.gov/32267544/)]
3. Engle S, Stromme J, Zhou A. Staying at home: mobility effects of COVID-19. *SSRN Journal*. Preprint posted online on April 15, 2020. [doi: [10.2139/ssrn.3565703](https://doi.org/10.2139/ssrn.3565703)]
4. Czeisler, Tynan MA, Howard ME, Honeycutt S, Fulmer EB, Kidder DP, et al. Public attitudes, behaviors, and beliefs related to COVID-19, stay-at-home orders, nonessential business closures, and public health guidance - United States, New York City, and Los Angeles, May 5-12, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Jun 19;69(24):751-758 [FREE Full text] [doi: [10.15585/mmwr.mm6924e1](https://doi.org/10.15585/mmwr.mm6924e1)] [Medline: [32555138](https://pubmed.ncbi.nlm.nih.gov/32555138/)]
5. Gostin LO, Wiley LF. Governmental public health powers during the COVID-19 pandemic: stay-at-home orders, business closures, and travel restrictions. *JAMA* 2020 Jun 02;323(21):2137-2138 [FREE Full text] [doi: [10.1001/jama.2020.5460](https://doi.org/10.1001/jama.2020.5460)] [Medline: [32239184](https://pubmed.ncbi.nlm.nih.gov/32239184/)]
6. Larochelle MR. "Is it safe for me to go to work?" Risk stratification for workers during the Covid-19 pandemic. *N Engl J Med* 2020 Jul 30;383(5):e28. [doi: [10.1056/nejmp2013413](https://doi.org/10.1056/nejmp2013413)]
7. Testing, screening, and outbreak response for institutions of higher education (IHEs). US Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/community/colleges-universities/ihe-testing.html> [accessed 2020-08-10]
8. Shaw WS, Main CJ, Findley PA, Collie A, Kristman VL, Gross DP. Opening the workplace after COVID-19: what lessons can be learned from return-to-work research? *J Occup Rehabil* 2020 Sep 19;30(3):299-302 [FREE Full text] [doi: [10.1007/s10926-020-09908-9](https://doi.org/10.1007/s10926-020-09908-9)] [Medline: [32562129](https://pubmed.ncbi.nlm.nih.gov/32562129/)]
9. Jones J, Sullivan PS, Sanchez TH, Guest JL, Hall EW, Luisi N, et al. Similarities and differences in COVID-19 awareness, concern, and symptoms by race and ethnicity in the United States: cross-sectional survey. *J Med Internet Res* 2020 Jul 10;22(7):e20001 [FREE Full text] [doi: [10.2196/20001](https://doi.org/10.2196/20001)] [Medline: [32614778](https://pubmed.ncbi.nlm.nih.gov/32614778/)]
10. Bai Y, Yao L, Wei T, Tian F, Jin D, Chen L, et al. Presumed asymptomatic carrier transmission of COVID-19. *JAMA* 2020 Apr 14;323(14):1406-1407 [FREE Full text] [doi: [10.1001/jama.2020.2565](https://doi.org/10.1001/jama.2020.2565)] [Medline: [32083643](https://pubmed.ncbi.nlm.nih.gov/32083643/)]
11. Contact tracing for COVID-19. US Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/contact-tracing.html> [accessed 2021-03-16]
12. Baraniuk C. Covid-19 contact tracing: a briefing. *BMJ* 2020 May 13;369:m1859. [doi: [10.1136/bmj.m1859](https://doi.org/10.1136/bmj.m1859)] [Medline: [32404304](https://pubmed.ncbi.nlm.nih.gov/32404304/)]
13. Beidas RS, Buttenheim AM, Feuerstein-Simon R, Kilaru AS, Asch DA, Volpp KG, et al. Optimizing and Implementing Contact Tracing through Behavioral Economics. *NEJM Catal* 2020 Jun 23 [FREE Full text]
14. Sun K, Viboud C. Impact of contact tracing on SARS-CoV-2 transmission. *Lancet Infect Dis* 2020 Aug;20(8):876-877 [FREE Full text] [doi: [10.1016/S1473-3099\(20\)30357-1](https://doi.org/10.1016/S1473-3099(20)30357-1)]
15. Maytin L, Maytin J, Agarwal P, Krenitsky A, Krenitsky J, Epstein R. Attitudes and perceptions toward COVID-19 digital surveillance: survey of young adults in the United States. *JMIR Form Res* 2021 Jan 08;5(1):e23000 [FREE Full text] [doi: [10.2196/23000](https://doi.org/10.2196/23000)] [Medline: [33347420](https://pubmed.ncbi.nlm.nih.gov/33347420/)]
16. Zeng K, Bernardo SN, Havins WE. The use of digital tools to mitigate the COVID-19 pandemic: comparative retrospective study of six countries. *JMIR Public Health Surveill* 2020 Dec 23;6(4):e24598 [FREE Full text] [doi: [10.2196/24598](https://doi.org/10.2196/24598)] [Medline: [33302255](https://pubmed.ncbi.nlm.nih.gov/33302255/)]
17. Kondylakis H, Katehakis DG, Kouroubali A, Logothetidis F, Triantafyllidis A, Kalamaras I, et al. COVID-19 mobile apps: a systematic review of the literature. *J Med Internet Res* 2020 Dec 09;22(12):e23170 [FREE Full text] [doi: [10.2196/23170](https://doi.org/10.2196/23170)] [Medline: [33197234](https://pubmed.ncbi.nlm.nih.gov/33197234/)]
18. Wang S, Ding S, Xiong L. A new system for surveillance and digital contact tracing for COVID-19: spatiotemporal reporting over network and GPS. *JMIR mHealth uHealth* 2020 Jun 10;8(6):e19457 [FREE Full text] [doi: [10.2196/19457](https://doi.org/10.2196/19457)] [Medline: [32499212](https://pubmed.ncbi.nlm.nih.gov/32499212/)]

19. Yasaka TM, Lehrich BM, Sahyouni R. Peer-to-peer contact tracing: development of a privacy-preserving smartphone app. *JMIR mHealth uHealth* 2020 Apr 7;8(4):e18936. [doi: [10.2196/18936](https://doi.org/10.2196/18936)]
20. Privacy-preserving contact tracing. Apple | Google. URL: <https://www.apple.com/covid19/contacttracing> [accessed 2020-08-10]
21. Digital contact tracing tools. US Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/digital-contact-tracing-tools.html> [accessed 2021-03-16]
22. Altmann S, Milsom L, Zillesen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: Cross-country survey evidence. *medRxiv*. Preprint posted online on May 08, 2020. [doi: [10.1101/2020.05.05.20091587](https://doi.org/10.1101/2020.05.05.20091587)]
23. Nakamoto I, Wang S, Guo Y, Zhuang W. A QR code-based contact tracing framework for sustainable containment of COVID-19: evaluation of an approach to assist the return to normal activity. *JMIR mHealth uHealth* 2020 Sep 07;8(9):e22321 [FREE Full text] [doi: [10.2196/22321](https://doi.org/10.2196/22321)] [Medline: [32841151](https://pubmed.ncbi.nlm.nih.gov/32841151/)]
24. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* 2020 May 08;368(6491) [FREE Full text] [doi: [10.1126/science.abb6936](https://doi.org/10.1126/science.abb6936)] [Medline: [32234805](https://pubmed.ncbi.nlm.nih.gov/32234805/)]
25. Anglemeyer A, Moore T, Parker L, Chambers T, Grady A, Chiu K, et al. Digital contact tracing technologies in epidemics: a rapid review. *Cochrane Database Syst Rev* 2020 Aug 18;8:CD013699. [doi: [10.1002/14651858.CD013699](https://doi.org/10.1002/14651858.CD013699)] [Medline: [33502000](https://pubmed.ncbi.nlm.nih.gov/33502000/)]
26. Trang S, Trenz M, Weiger WH, Tarafdar M, Cheung CM. One app to trace them all? Examining app specifications for mass acceptance of contact-tracing apps. *Eur J Inf Syst* 2020 Jul 27;29(4):415-428. [doi: [10.1080/0960085x.2020.1784046](https://doi.org/10.1080/0960085x.2020.1784046)]
27. Dar AB, Lone AH, Zahoor S, Khan AA, Naaz R. Applicability of mobile contact tracing in fighting pandemic (COVID-19): issues, challenges and solutions. *Comput Sci Rev* 2020 Nov;38:100307 [FREE Full text] [doi: [10.1016/j.cosrev.2020.100307](https://doi.org/10.1016/j.cosrev.2020.100307)] [Medline: [32989380](https://pubmed.ncbi.nlm.nih.gov/32989380/)]
28. Walrave M, Waeterloos C, Ponnet K. Adoption of a contact tracing app for containing COVID-19: a health belief model approach. *JMIR Public Health Surveill* 2020 Sep 01;6(3):e20572 [FREE Full text] [doi: [10.2196/20572](https://doi.org/10.2196/20572)] [Medline: [32755882](https://pubmed.ncbi.nlm.nih.gov/32755882/)]
29. Scherr T, Hardcastle AN, Moore C, DeSousa JM, Wright D. Understanding on-campus interactions with a semi-automated, barcode-based platform to augment COVID-19 contact tracing: application development and usability study. *JMIR Mhealth Uhealth* 2021 Feb 25;a [FREE Full text] [doi: [10.2196/24275](https://doi.org/10.2196/24275)] [Medline: [33690142](https://pubmed.ncbi.nlm.nih.gov/33690142/)]
30. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
31. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform* 2019 Jul;95:103208 [FREE Full text] [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
32. Brooke J. SUS: a 'quick and dirty' usability scale. In: *Usability Evaluation in Industry*. Boca Raton, FL: CRC Press; 1996:189-194.

Abbreviations

mHealth: mobile health

REDCap: Research Electronic Data Capture

SUS: System Usability Scale

Edited by G Eysenbach; submitted 20.11.20; peer-reviewed by H Mehdizadeh, A Fioravanti; comments to author 09.12.20; revised version received 15.01.21; accepted 16.01.21; published 26.03.21.

Please cite as:

Scherr TF, DeSousa JM, Moore CP, Hardcastle A, Wright DW

App Use and Usability of a Barcode-Based Digital Platform to Augment COVID-19 Contact Tracing: Postpilot Survey and Paradata Analysis

JMIR Public Health Surveill 2021;7(3):e25859

URL: <https://publichealth.jmir.org/2021/3/e25859>

doi:[10.2196/25859](https://doi.org/10.2196/25859)

PMID:[33630745](https://pubmed.ncbi.nlm.nih.gov/33630745/)

distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Rapid Surveillance Report

Analyzing Cross-country Pandemic Connectedness During COVID-19 Using a Spatial-Temporal Database: Network Analysis

Amanda MY Chu¹, PhD; Jacky NL Chan², BMSIM; Jenny TY Tsang¹, PhD; Agnes Tiwari^{3,4}, PhD; Mike KP So², PhD

¹Department of Social Sciences, The Education University of Hong Kong, Hong Kong, China (Hong Kong)

²Department of Information Systems, Business Statistics and Operations Management, The Hong Kong University of Science and Technology, Hong Kong, China (Hong Kong)

³Hong Kong Sanatorium & Hospital, Hong Kong, China (Hong Kong)

⁴Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:

Mike KP So, PhD

Department of Information Systems, Business Statistics and Operations Management

The Hong Kong University of Science and Technology

Clear Water Bay

Hong Kong

China (Hong Kong)

Phone: 852 23587726

Email: immkps@ust.hk

Abstract

Communicable diseases including COVID-19 pose a major threat to public health worldwide. To curb the spread of communicable diseases effectively, timely surveillance and prediction of the risk of pandemics are essential. The aim of this study is to analyze free and publicly available data to construct useful travel data records for network statistics other than common descriptive statistics. This study describes analytical findings of time-series plots and spatial-temporal maps to illustrate or visualize pandemic connectedness. We analyzed data retrieved from the web-based Collaborative Arrangement for the Prevention and Management of Public Health Events in Civil Aviation dashboard, which contains up-to-date and comprehensive meta-information on civil flights from 193 national governments in accordance with the airport, country, city, latitude, and the longitude of flight origin and the destination. We used the database to visualize pandemic connectedness through the workflow of travel data collection, network construction, data aggregation, travel statistics calculation, and visualization with time-series plots and spatial-temporal maps. We observed similar patterns in the time-series plots of worldwide daily flights from January to early-March of 2019 and 2020. A sharp reduction in the number of daily flights recorded in mid-March 2020 was likely related to large-scale air travel restrictions owing to the COVID-19 pandemic. The levels of connectedness between places are strong indicators of the risk of a pandemic. Since the initial reports of COVID-19 cases worldwide, a high network density and reciprocity in early-March 2020 served as early signals of the COVID-19 pandemic and were associated with the rapid increase in COVID-19 cases in mid-March 2020. The spatial-temporal map of connectedness in Europe on March 13, 2020, shows the highest level of connectedness among European countries, which reflected severe outbreaks of COVID-19 in late March and early April of 2020. As a quality control measure, we used the aggregated numbers of international flights from April to October 2020 to compare the number of international flights officially reported by the International Civil Aviation Organization with the data collected from the Collaborative Arrangement for the Prevention and Management of Public Health Events in Civil Aviation dashboard, and we observed high consistency between the 2 data sets. The flexible design of the database provides users access to network connectedness at different periods, places, and spatial levels through various network statistics calculation methods in accordance with their needs. The analysis can facilitate early recognition of the risk of a current communicable disease pandemic and newly emerging communicable diseases in the future.

(*JMIR Public Health Surveill* 2021;7(3):e27317) doi:[10.2196/27317](https://doi.org/10.2196/27317)

KEYWORDS

air traffic; coronavirus; COVID-19; human mobility; network analysis; travel restrictions

Introduction

Communicable diseases remain a major public health threat worldwide. The COVID-19 pandemic is a stark reminder of the ongoing challenge posed by communicable diseases on human health [1]. Timely surveillance and estimation of the risk of a pandemic are crucial for curbing the spread of communicable diseases. Without efficacious medications and vaccines, the implementation of nonpharmaceutical interventions, such as air travel restrictions and social distancing measures, is vital for controlling communicable diseases [2]. The conventional method of estimating the risk of a pandemic based only on the number of confirmed cases provides limited information about pandemic trends. Network analysis is a powerful tool to estimate the risk of a pandemic through network connectedness [3] when air travel is a common route of communicable disease transmission [4]. Network connectedness analysis using air travel data records can help visualize the effect of air travel restrictions on pandemic connectedness.

The risk of in-flight communicable disease transmission has been a global health concern well before the emergence of COVID-19 [5]. A number of in-flight communicable disease transmissions have been documented, including influenza [6], severe acute respiratory syndrome [7], multidrug-resistant tuberculosis [8], measles [9], meningococcal infections [10], norovirus [11], shigellosis [12], and cholera [13]. Studies on the transmission of influenza [14] and severe acute respiratory syndrome [15] on aircraft further indicated that air travel can serve as a channel for the rapid spread of newly emerging communicable diseases. A study on COVID-19 control in Latin America suggested that countries serving as air transportation hubs are more prone to disease transmission. The practicable use of travel data for the prediction of the risk of the COVID-19 pandemic has been previously reported [16].

We previously analyzed travel data retrieved from the web-based Collaborative Arrangement for the Prevention and Management of Public Health Events in Civil Aviation (CAPSCA) dashboard [17], which contains up-to-date and comprehensive meta-information on civil flights from 193 national governments in accordance with the airport, country, city, latitude, and the longitude of flight origin and the destination. Unlike official travel data sources, such as the Federal Aviation Administration (FAA) and the International Air Transport Association (IATA), the travel data of the CAPSCA are free and publicly available.

The use of an Automated Dependent Surveillance Broadcast system as one of the travel data collection elements enables the CAPSCA to provide up-to-date travel locations, among other travel data (latitude and longitude of flight origin and the destination alongside a timestamp) [18,19]. The CAPSCA provides civil flight data for both passenger and cargo flights. While it is common knowledge that cargo flight crews may spread the disease through air travel, cargo is not always considered a health risk [20]. Travel data from both passenger and cargo flights are therefore more comprehensive for pandemic connectedness analysis.

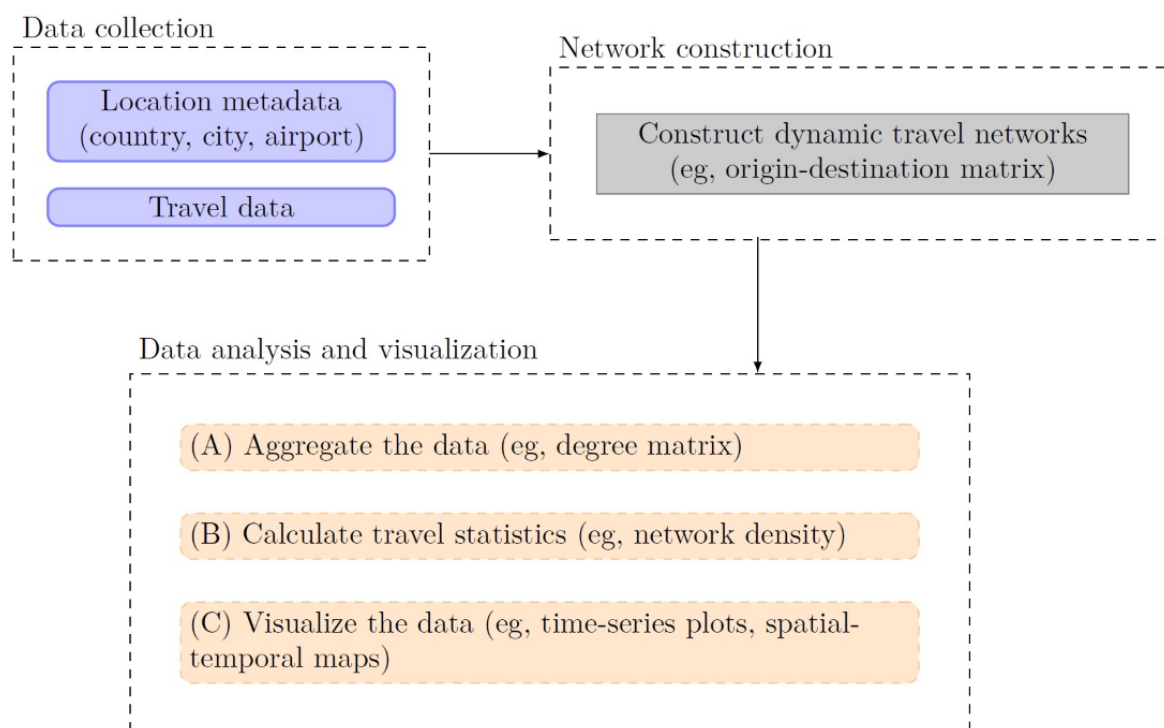
Flexible analysis of travel data can be performed through in-database processing. Specifically, users can analyze travel data through various network statistics calculations, including network analysis [3], network density [21], and reciprocity [22], which are powerful tools to estimate the risk of a pandemic through network connectedness. Simple analyses with time-series plots and spatial-temporal maps would facilitate clear visualization of the analytical results. Time-series plots show changes in network density and reciprocity, which are likely to be early signs of alterations in the risk of a pandemic [21]. Spatial-temporal maps during network analysis illustrate the connectedness among places and reflect changes in the risk of a pandemic before unprocessed data on the number of confirmed cases are obtained [16]. Moreover, diverse tools can be used to analyze pandemic connectedness.

The CAPSCA dashboard allows us to create a spatial-temporal database by integrating the travel data from different airports to illustrate connectedness at the city, country, or regional levels in accordance with the users' preference. The database can facilitate research and policymaking at local and global levels and provide a spatial outlook of the evolution of the pandemic network for predicting and assessing the risk of a communicable disease pandemic.

Methods

Workflow of Data Collection and Analysis

The database enables the visualization of pandemic connectedness through the workflow of travel data collection, network construction, data aggregation, travel statistics calculation, and visualization with time-series plots and spatial-temporal maps. The workflow of data collection and analysis is summarized in Figure 1.

Figure 1. Workflow of data collection and analysis.

Data Collection

Travel data and meta-information were retrieved from the CAPSCA dashboard through two separate procedures:

1. Extracting the airport meta-information from the rendered JavaScript object “airportData” through a HyperText Transfer Protocol request.
2. Downloading and extracting the flight numbers of civil flights, including both passenger and cargo flights, from the JavaScript Object Notation responses through multiple Asynchronous JavaScript and Extensible Markup Language requests.

The collected travel data were filtered for valid International Civil Aviation Organization (ICAO)–formatted airport codes. Raw JavaScript Object Notation responses contain formatted airport codes from various data sources such as the ICAO, IATA, and FAA. Each type of formatted airport code has its specific format. Valid ICAO codes start only with a letter and have 4 letters or digits, valid IATA codes consist of 3 letters, while valid FAA codes are 3–5–character alphanumeric codes. Each format type helps develop a filtering system to extract travel data with specific codes.

Data Records

A live version of the data record, which is maintained up-to-date with the latest data, can be downloaded from our travel database project repository [23]. The data records consist of 2 major parts: aggregated raw input and calculated or computed records.

The aggregated raw inputs are location metadata that contain data at multiple levels—country, city, airport, and geolocation (latitude and longitude)—and travel data, which contain daily

information regarding flight origin and the destination, starting from January 2019. These data encompass >200 countries and regions worldwide.

The data records (details) are structured into the following 3 comma-separated value (CSV) files:

1. [ICAO_airport_meta.csv] Table of the location metadata (ICAO-CAPSCA airport meta). The fields of the table are as follows:
 - a. *countryName* is the name of the country
 - b. *countryCode* is the International Organization for Standardization (ISO)-3166 alpha-3 code of the country
 - c. *airportName* is the name of the airport
 - d. *airportCode* is the ICAO code of the airport
 - e. *cityName* is the name of the city
 - f. *latitude* is the geolocation (latitude) of the airport
 - g. *longitude* is the geolocation (longitude) of the airport
2. [flight_2019-01-01_2020-12-03.csv] Table of travel data (daily flight numbers from origin to destination). The fields of the table are as follows:
 - a. *date* is the record date
 - b. *num_flight* is the number of flights from the origin airport to the destination airport
 - c. *orig_airportCode* is the ICAO code of the origin airport
 - d. *orig_airportName* is the name of the origin airport
 - e. *orig_countryCode* is the ISO-3166 alpha-3 country code of the origin airport
 - f. *orig_countryName* is the country name of the origin airport
 - g. *orig_cityName* is the city name of the origin airport
 - h. *orig_latitude* is the geolocation (latitude) of the destination airport

- i. *orig_longitude* is the geolocation (longitude) of the destination airport
 - j. *dest_airportCode* is the ICAO code of the destination airport
 - k. *dest_airportName* is the name of the destination airport
 - l. *dest_countryCode* is the ISO-3166 alpha-3 country code of the destination airport
 - m. *dest_countryName* is the name of the country where the destination airport is located
 - n. *dest_cityName* is the name of the city where the destination airport is located
 - o. *dest_latitude* is the geolocation (latitude) of the destination airport
 - p. *dest_longitude* is the geolocation (longitude) of the destination airport
3. [network_statistics.csv] Table of the calculated network statistics. The fields of the table are as follows:
 - a. *date* is the reference date of the network statistics at time t
 - b. V_t is the number of vertices (V_t) at time t
 - c. E_t is the number of edges (E_t) at time t
 - d. D_t is the edge density (D_t) at time t
 - e. R_t is the reciprocity (R_t) at time t

Dynamic Network Construction and Data Aggregation

The travel data can be used to construct the travel network structure [24]. The basic network components involve nodes (vertices) and links (edges). The nodes represent the target entity (location), such as airport, city, or country. As the travel data

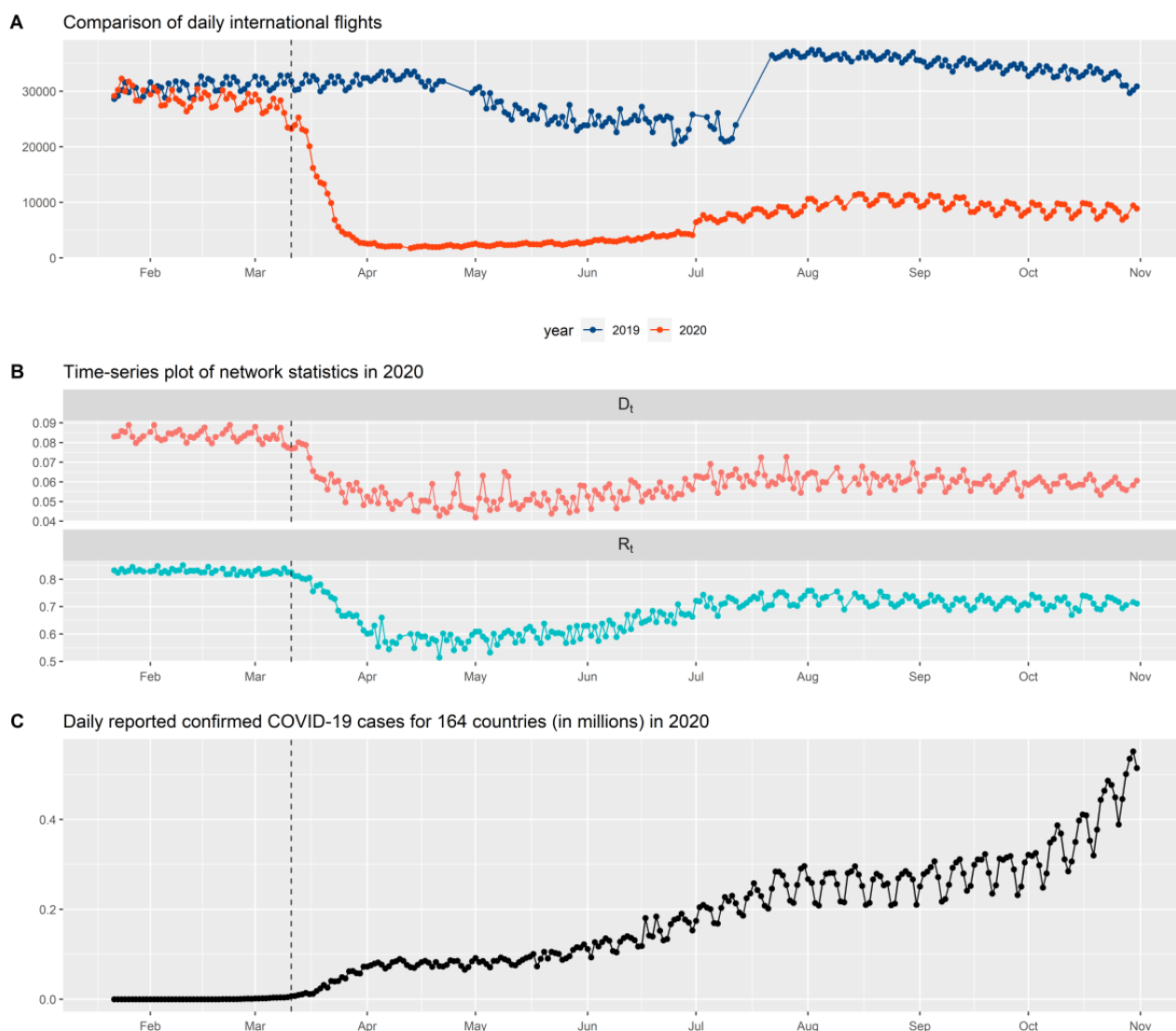
contain detailed airport-to-airport records, they can be transformed by merging data from the airports to form nodes of cities, countries, regions, or groups of any geolocations in accordance with the users' preference. A link represents a relationship (connection) between 2 target entities. The relationship can be binary or numeric (eg, flight frequency), indicating the existence or strength of a travel connection, respectively.

For example, if we focus on global analysis, we aggregate the airport data at the country level and input the country data as a new set of nodes to form travel subnetworks, which are represented by the country-country–origin-destination matrix with entries being the flight frequencies between 2 countries.

Data Analysis: Travel Network Statistics

We can further aggregate the travel data to obtain overall worldwide flight information. Figure 2A shows the time-series plots of worldwide daily flights in 2019 and 2020. By comparing the 2 time-series curves in Figure 2A, we observed similar patterns of worldwide daily flights from January to early-March in 2019 and 2020. A sharp decline in the number of daily flights recorded in mid-March 2020 was likely related to large-scale air travel restrictions owing to the COVID-19 pandemic. In addition, the subnetwork can be further used to generate the degree matrix, where its diagonal entries contain the number of edges connected to different nodes (number of connected countries of each node). The combined use of the origin-destination and degree matrices can yield spatial-temporal maps (Figure 3).

Figure 2. Time-series plots of (A) daily international flights, (B) global network statistics, and (C) daily reported number of confirmed COVID-19 cases.



If V_t is the number of vertices of the dynamic network at time t , and E_t is the number of edges of the dynamic network at time t , network statistics such as D_t (network density [21]) and R_t (reciprocity [22]) can be determined (Figure 2A). For example, in the dynamic networks in Figure 3, countries are represented by vertices, and travel connections are represented by edges. Network density D_t is based on an undirected network structure, and is defined as follows:

$$D_t = \frac{E_t}{V_t(V_t - 1)}$$

which refers to the ratio of the number of connections with respect to the maximum possible connections among countries.

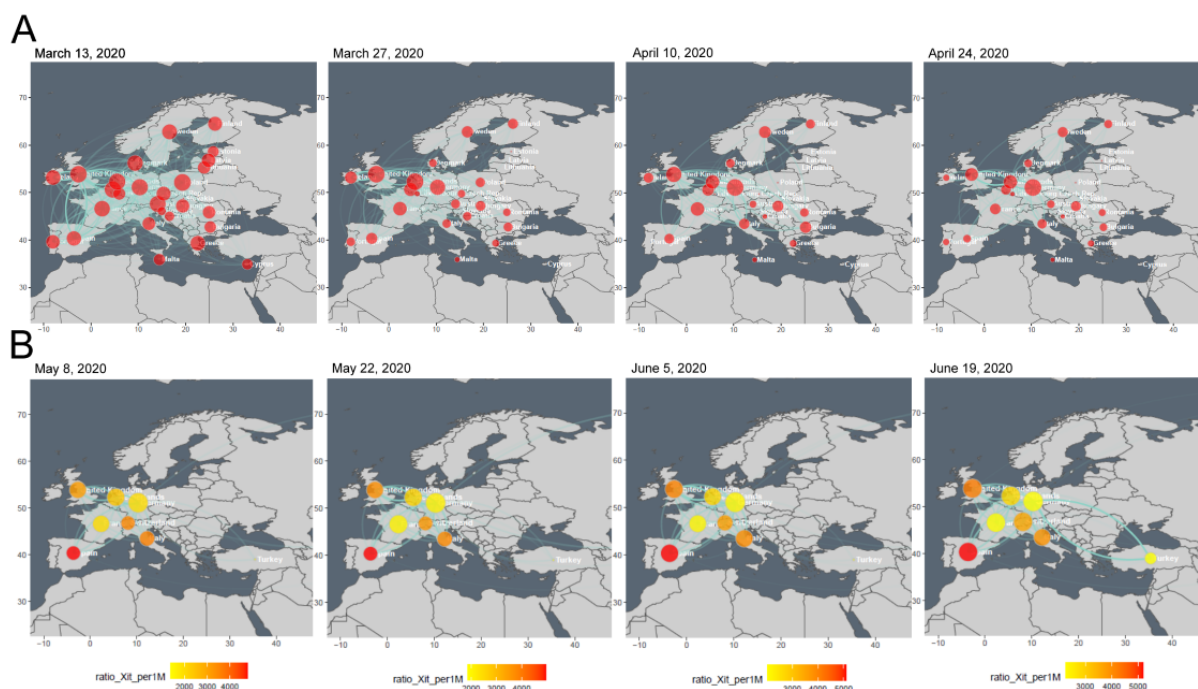
This equation illustrates how dense the connections in the dynamic network are at time t .

Reciprocity R_t is based on a directed network structure and is defined as follows:

$$R_t = \frac{E_t^{<->}}{L_t}$$

where R_t is the ratio of the number of links pointing in both directions, $E_t^{<->}$ (mutual links), to the total number of links, L_t . In other words, the value of R_t represents the average possibility that a link is reciprocated.

Figure 3. Spatial-temporal maps of connectedness in Europe at (A) the country level between March 13 and April 24, 2020, and (B) air transportation hub level between May 8 and June 19, 2020.



Results

Data Visualization

Network statistics data were visualized using time-series plots. In addition, data records of the origin-destination and degree matrices can be visualized using spatial-temporal maps.

Figure 2 displays the time-series plots of worldwide daily flights (Figure 2A), global network statistics (Figure 2B), and daily reported confirmed COVID-19 cases (Figure 2C). We found that changes in network density can serve as an early signal of the risk of a pandemic.

The time-series plots of network density and reciprocity shown in Figure 2B demonstrate the practical use of network statistics to predict the risk of a global pandemic. With initial reports of COVID-19 cases worldwide, a high network density and reciprocity in early-March 2020 were early signals of the COVID-19 pandemic and were associated with the rapid increase in COVID-19 cases reported by the World Health Organization (WHO) in mid-March 2020 (Figure 2C). The sharp reduction in the network density and reciprocity in mid-March 2020 suggests a reduction in the risk of the pandemic, which was associated with a steady number of daily confirmed cases from mid-March to May 2020. The gradual increase in network density and reciprocity from mid-May 2020 suggests an increase in the pandemic risk, which was associated with a rapid surge in COVID-19 cases from mid-May to August 2020. The in-database processing system allows users to analyze network density and reciprocity at local or global levels during selected periods in order to identify possible trends in the evolution of the risk of pandemics.

Figure 3 shows some spatial-temporal maps to illustrate the network connectedness among countries under different spatial

levels. The levels of connectedness among places are strong indicators of the risk of a pandemic. The spatial-temporal maps showing the connectedness among places are constructed on the basis of the origin-destination and degree matrices, which provide directional data and summarized node data (degree of vertices), respectively. The maps display the connectedness among places in the form of connections and bubbles. The connections (links) illustrate the connectedness among places. The bubbles (vertices), which contain aggregated information presented by the node size, further enhance the visualization of connectedness, especially when the network density is high. The in-database processing system allows users to generate spatial-temporal maps at multiple spatial levels, such as airports, cities, countries, or any regions of interest, with various temporal settings, such as different periods and time-intervals. The flexibility of the database facilitates data analysis in accordance with the users' preferences.

For example, to analyze the connectedness in Europe during the first wave of the COVID-19 pandemic in early 2020, spatial-temporal maps of Europe at 14-day intervals between March and April 2020 were generated (Figure 3A). Maps displaying connectedness at the country level are plotted, with each bubble representing a country. The size of the colored bubbles represents the number of vertices or countries. The color intensity of the bubbles indicates the number of daily confirmed COVID-19 cases per 1 million population of the country. The light-green connections (links) represent the number of daily international flights. Thicker the green line, higher the connectedness between 2 countries. The map for March 13, 2020, shows the highest level of connectedness among European countries, reflecting a severe outbreak of COVID-19 in late March and early April 2020. The significant

reduction in connectedness from March 13 to April 24 indicates that the first wave of the pandemic subsided in May 2020.

Figure 3B shows a different set of spatial-temporal maps to analyze connectedness in Europe before the onset of the second wave of the COVID-19 pandemic. Spatial-temporal maps of Europe at 14-day intervals between May and June 2020 were generated. As countries serving as air transportation hubs were particularly prone to the spread of COVID-19 [16], spatial-temporal maps at the country level for these hubs have been plotted to investigate network connectedness between the countries harboring these hubs during the growth of the pandemic. Each bubble represents a country with airports defined as among the top 10 air transportation hubs in Europe by the Official Airline Guide MegaHub Index of 2019 [25]. The color indicates the accumulated total number of confirmed COVID-19 cases per 1 million population of a particular country at time t (Figure 3B, ratio_Xit_per1M). The significant increase in connectedness in mid-June 2020 indicates an increase in the risk of the pandemic in Europe, especially among hub countries.

The aforementioned data were visualized using igraph in the R software package (The R Foundation) or software facilitating network visualization, such as Gephi (The Gephi Consortium).

Technical Validation

CAPSCA is a voluntary cross-sectorial and multi-organizational collaborative program managed by the ICAO with support from the WHO. The quality of the data on the dashboard should be guaranteed. As a quality control measure, we used the aggregated number of international flights from April to October 2020 to compare the number of COVID-19 cases at hub airports officially reported by the ICAO [26] with the data collected from the CAPSCA dashboard. We analyzed the correlation between the 2 data sets for 7 different regions during this period: Asia Pacific, East and South Africa, Europe and North Atlantic, Middle East, North America and the Caribbean, South America, and West and Central Africa. The results are summarized in Table 1. We found that most coefficients of correlation were >0.99 , except for the Asia Pacific region ($r=0.96$). This difference probably results from data synchronization (lagging by 3 days) by the ICAO after collecting the daily raw counts.

Table 1. Correlation between the number of daily international flights officially reported by the International Civil Aviation Organization and those obtained from the Collaborative Arrangement for the Prevention and Management of Public Health Events in Civil Aviation dashboard for 7 geographic regions from April to October 2020.

ICAO ^a -identified geographic regions	Coefficient of correlation
Asia Pacific	0.96
East and South Africa	1.00
Europe and North Atlantic	1.00
Middle East	1.00
North America and the Caribbean	0.99
South America	1.00
West and Central Africa	1.00

^aICAO: International Civil Aviation Organization.

Discussion

Principal Findings

Ongoing systematic surveillance is important to help detect early outbreaks and to evaluate the effectiveness of public health measures and programs [27,28]. In this study, we attempted to conduct flexible analysis of freely available travel data collected from the CAPSCA dashboard to identify certain patterns and early signals of the COVID-19 pandemic, which may help policymakers take appropriate action. By linking the database to daily numbers of confirmed COVID-19 cases, we can develop a user-friendly platform for timely and flexible visualization of network connectedness to facilitate surveillance and early recognition of the risk of a pandemic, including the high network density and reciprocity in early March 2020 through time-series analysis and the high level of connectedness among European countries on March 13, 2020, through spatial-temporal mapping. The surveillance and findings are important for curbing the spread of communicable diseases and balancing disease control and economic recovery.

Our study provides a clear workflow for data collection and analysis and the suggested software for analysis. Because air transportation is highly relevant to the dissemination of communicable diseases, the database and the analysis can also be applied to investigate the risk of other communicable disease pandemics occurring currently or emerging in the future. In addition, researchers may replicate our workflow for assessing pandemic connectedness by using the same database or other databases.

For data quality assurance, we conducted correlation analysis and validated the database. We compared the number of aggregated international flights officially reported by the ICAO for 7 regions from April to October 2020 with corresponding data we collected from the CAPSCA dashboard. We found only minor differences probably owing to data synchronization performed by the ICAO after collecting the daily raw counts of COVID-19 cases. We expect that our findings may help researchers explore and validate freely available health-related databases to conduct ongoing and systematic analysis and interpretation to identify early warning signals, such that

necessary action can be taken to prevent and control the spread of communicable diseases.

Conclusions

In this study, we demonstrated a workflow to analyze freely available travel data retrieved from the CAPSCA dashboard, together with data on confirmed COVID-19 cases reported by the WHO, for systematic surveillance. Flexible analysis of the

travel data can be performed through in-database processing, which allows us to visualize and analyze the risk of a pandemic and pandemic connectedness by using different tools such as time-series plots and spatial-temporal maps. This analysis facilitates early recognition of the risk of pandemics of current communicable diseases and newly emerged communicable diseases in the future.

Acknowledgments

This study was partially supported by the Hong Kong University of Science and Technology research grant “Big Data Analytics on Social Research” (grant# CEF20BM04).

Authors' Contributions

AMYC and MKPS conceived and designed the study. JNLC collected the data. JNLC, AMYC, and MKPS processed the data. AMYC, JTYT, and JNLC drafted the manuscript. AT and MKPS finalized the manuscript. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. COVID-19 a Reminder of the Challenge of Emerging Infectious Diseases. National Institutes of Health. 2020 Feb 28. URL: <https://www.nih.gov/news-events/news-releases/covid-19-reminder-challenge-emerging-infectious-diseases> [accessed 2020-12-04]
2. Guidelines for the implementation of non-pharmaceutical interventions against COVID-19. European Centre for Disease Prevention and Control. 2020 Sep 24. URL: <https://www.ecdc.europa.eu/en/publications-data/covid-19-guidelines-non-pharmaceutical-interventions#no-link> [accessed 2020-12-04]
3. So MK, Tiwari A, Chu AM, Tsang JT, Chan JN. Visualizing COVID-19 pandemic risk through network connectedness. *Int J Infect Dis* 2020 Jul;96:558-561 [FREE Full text] [doi: [10.1016/j.ijid.2020.05.011](https://doi.org/10.1016/j.ijid.2020.05.011)] [Medline: [32437929](https://pubmed.ncbi.nlm.nih.gov/32437929/)]
4. Mangili A, Gendreau MA. Transmission of infectious diseases during commercial air travel. *Lancet* 2005;365(9463):989-996 [FREE Full text] [doi: [10.1016/S0140-6736\(05\)71089-8](https://doi.org/10.1016/S0140-6736(05)71089-8)] [Medline: [15767002](https://pubmed.ncbi.nlm.nih.gov/15767002/)]
5. Weiss H, Hertzberg VS, Dupont C, Espinoza JL, Levy S, Nelson K, FlyHealthy Research Team. The Airplane Cabin Microbiome. *Microb Ecol* 2019 Jan;77(1):87-95 [FREE Full text] [doi: [10.1007/s00248-018-1191-3](https://doi.org/10.1007/s00248-018-1191-3)] [Medline: [29876609](https://pubmed.ncbi.nlm.nih.gov/29876609/)]
6. Baker M, Thornley CN, Mills C, Roberts S, Perera S, Peters J, et al. Transmission of pandemic A/H1N1 2009 influenza on passenger aircraft: retrospective cohort study. *BMJ* 2010 May 21;340:c2424 [FREE Full text] [doi: [10.1136/bmj.c2424](https://doi.org/10.1136/bmj.c2424)] [Medline: [20495017](https://pubmed.ncbi.nlm.nih.gov/20495017/)]
7. Olsen SJ, Chang H, Cheung TY, Tang AF, Fisk TL, Ooi SP, et al. Transmission of the severe acute respiratory syndrome on aircraft. *N Engl J Med* 2003 Dec 18;349(25):2416-2422. [doi: [10.1056/NEJMoa031349](https://doi.org/10.1056/NEJMoa031349)] [Medline: [14681507](https://pubmed.ncbi.nlm.nih.gov/14681507/)]
8. Kenyon TA, Valway SE, Ihle WW, Onorato IM, Castro KG. Transmission of multidrug-resistant *Mycobacterium tuberculosis* during a long airplane flight. *N Engl J Med* 1996 Apr 11;334(15):933-938. [doi: [10.1056/NEJM199604113341501](https://doi.org/10.1056/NEJM199604113341501)] [Medline: [8596593](https://pubmed.ncbi.nlm.nih.gov/8596593/)]
9. Centers for Disease Control (CDC). Interstate importation of measles following transmission in an airport--California, Washington, 1982. *MMWR Morb Mortal Wkly Rep* 1983 Apr 29;32(16):210, 215-210, 216. [Medline: [6406807](https://pubmed.ncbi.nlm.nih.gov/6406807/)]
10. O'Connor BA, Chant KG, Binotto E, Maidment CA, Maywood P, McAnulty JM. Meningococcal disease--probable transmission during an international flight. *Commun Dis Intell Q Rep* 2005;29(3):312-314 [FREE Full text] [Medline: [16220872](https://pubmed.ncbi.nlm.nih.gov/16220872/)]
11. Kirking H, Cortes J, Burrell S, Hall A, Cohen N, Lipman H, et al. Likely transmission of norovirus on an airplane, October 2008. *Clin Infect Dis* 2010 May 01;50(9):1216-1221. [doi: [10.1086/651597](https://doi.org/10.1086/651597)] [Medline: [20353365](https://pubmed.ncbi.nlm.nih.gov/20353365/)]
12. Hedberg CW, Levine WC, White KE, Carlson RH, Winsor DK, Cameron DN, et al. An international foodborne outbreak of shigellosis associated with a commercial airline. *JAMA* 1992 Dec 09;268(22):3208-3212. [Medline: [1433760](https://pubmed.ncbi.nlm.nih.gov/1433760/)]
13. Eberhart-Phillips J, Besser RE, Tormey MP, Koo D, Feikin D, Araneta MR, et al. An outbreak of cholera from food served on an international aircraft. *Epidemiol Infect* 1996 Feb;116(1):9-13 [FREE Full text] [doi: [10.1017/s0950268800058891](https://doi.org/10.1017/s0950268800058891)] [Medline: [8626007](https://pubmed.ncbi.nlm.nih.gov/8626007/)]
14. Shankar AG, Janmohamed K, Olowokure B, Smith GE, Hogan AH, De Souza V, et al. Contact tracing for influenza A(H1N1)pdm09 virus-infected passenger on international flight. *Emerg Infect Dis* 2014 Jan;20(1):118-120 [FREE Full text] [doi: [10.3201/eid2001.120101](https://doi.org/10.3201/eid2001.120101)] [Medline: [24377724](https://pubmed.ncbi.nlm.nih.gov/24377724/)]

15. Desenclos J, van der Werf S, Bonmarin I, Levy-Bruhl D, Yazdanpanah Y, Hoen B, et al. Introduction of SARS in France, March-April, 2003. *Emerg Infect Dis* 2004 Feb;10(2):195-200 [FREE Full text] [doi: [10.3201/eid1002.030351](https://doi.org/10.3201/eid1002.030351)] [Medline: [15030682](https://pubmed.ncbi.nlm.nih.gov/15030682/)]
16. Chu A, Tsang J, Chan J, Tiwari A, So M. Analysis of travel restrictions for COVID-19 control in Latin America through network connectedness. *J Travel Med* 2020 Dec 23;27(8):taaa176 [FREE Full text] [doi: [10.1093/jtm/taaa176](https://doi.org/10.1093/jtm/taaa176)] [Medline: [32970124](https://pubmed.ncbi.nlm.nih.gov/32970124/)]
17. Infectious Disease App. International Civil Aviation Organization. URL: <http://quips.anbdata.com/project/dev/5c1c21b205c09f70bfe60eeeb46316af89506e9.html> [accessed 2020-12-04]
18. Miller S, Moat HS, Preis T. Using aircraft location data to estimate current economic activity. *Sci Rep* 2020 May 05;10(1):7576 [FREE Full text] [doi: [10.1038/s41598-020-63734-w](https://doi.org/10.1038/s41598-020-63734-w)] [Medline: [32371997](https://pubmed.ncbi.nlm.nih.gov/32371997/)]
19. International Civil Aviation Organization. ADS - B Data as a source for analytical solutions of traffic behaviour in airspace. ICAO Big Data Project. 2016 Sep 06. URL: <https://www.icao.int/NACC/Documents/Meetings/2016/PARAST25/PARAST25-P1.pdf> [accessed 2020-12-04]
20. Graham J. Positive COVID-19 test in Alaska associated with cargo flight. *Air Cargo Week*. 2020 Mar 13. URL: <https://www.aircargoweek.com/positive-covid-19-test-in-alaska-associated-with-cargo-flight/> [accessed 2020-12-04]
21. Chu A, Tiwari A, So M. Detecting early signals of COVID-19 global pandemic from network density. *J Travel Med* 2020 Aug 20;27(5):taaa084 [FREE Full text] [doi: [10.1093/jtm/taaa084](https://doi.org/10.1093/jtm/taaa084)] [Medline: [32463088](https://pubmed.ncbi.nlm.nih.gov/32463088/)]
22. Garlaschelli D, Loffredo MI. Patterns of link reciprocity in directed networks. *Phys Rev Lett* 2004 Dec 31;93(26 Pt 1):268701. [doi: [10.1103/PhysRevLett.93.268701](https://doi.org/10.1103/PhysRevLett.93.268701)] [Medline: [15698035](https://pubmed.ncbi.nlm.nih.gov/15698035/)]
23. Chu A, Chan J, Tsang J, Tiwari A, So M. Analyzing cross-country pandemic connectedness in COVID-19: Network analysis using a spatial-temporal database. *JMIR Public Health Surveill* 2021 Mar 04 [FREE Full text] [doi: [10.2196/27317](https://doi.org/10.2196/27317)] [Medline: [33711799](https://pubmed.ncbi.nlm.nih.gov/33711799/)]
24. Bertin J. *Semiology of graphics: diagrams, networks, maps*. Madison, WI: University of Wisconsin Press; 1983.
25. Megahubs Index 2019. OAG. URL: <https://www.oag.com/oag-megahubs-2019> [accessed 2020-12-04]
26. Global COVID-19 Airport Status. International Civil Aviation Organization. URL: <https://www.icao.int/safety/pages/covid-19-airport-status.aspx> [accessed 2020-12-04]
27. Thacker SB, Berkelman RL. Public health surveillance in the United States. *Epidemiol Rev* 1988;10:164-190. [doi: [10.1093/oxfordjournals.epirev.a036021](https://doi.org/10.1093/oxfordjournals.epirev.a036021)] [Medline: [3066626](https://pubmed.ncbi.nlm.nih.gov/3066626/)]
28. Herbuela VRDM, Karita T, Carvajal TM, Ho HT, Lorena JMO, Regalado RA, et al. Early Detection of Dengue Fever Outbreaks Using a Surveillance App (Mozzify): Cross-sectional Mixed Methods Usability Study. *JMIR Public Health Surveill* 2021 Mar 01;7(3):e19034 [FREE Full text] [doi: [10.2196/19034](https://doi.org/10.2196/19034)] [Medline: [33646128](https://pubmed.ncbi.nlm.nih.gov/33646128/)]

Abbreviations

CAPSCA: Collaborative Arrangement for the Prevention and Management of Public Health Events in Civil Aviation

CSV: comma-separated value

FAA: Federal Aviation Administration

IATA: International Air Transport Association

ICAO: International Civil Aviation Organization

ISO: International Organization for Standardization

WHO: World Health Organization

Edited by T Sanchez; submitted 21.01.21; peer-reviewed by B Lam, A Chong; comments to author 11.02.21; revised version received 03.03.21; accepted 04.03.21; published 29.03.21.

Please cite as:

Chu AMY, Chan JNL, Tsang JTY, Tiwari A, So MKP

Analyzing Cross-country Pandemic Connectedness During COVID-19 Using a Spatial-Temporal Database: Network Analysis

JMIR Public Health Surveill 2021;7(3):e27317

URL: <https://publichealth.jmir.org/2021/3/e27317>

doi:[10.2196/27317](https://doi.org/10.2196/27317)

PMID:[33711799](https://pubmed.ncbi.nlm.nih.gov/33711799/)

©Amanda MY Chu, Jacky NL Chan, Jenny TY Tsang, Agnes Tiwari, Mike KP So. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 29.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution,

and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Early Detection of Dengue Fever Outbreaks Using a Surveillance App (Mozzify): Cross-sectional Mixed Methods Usability Study

Von Ralph Dane Marquez Herbuela^{1,2}, PhD; Tomonori Karita³, PhD; Thaddeus Marzo Carvajal^{1,4}, PhD; Howell Tsai Ho⁵, PhD; John Michael Olea Lorena⁶, PhD; Rachele Arce Regalado⁷, MA; Girly Dirilo Sobrepeña⁸, MD; Kozo Watanabe^{1,4}, PhD

¹Center for Marine Environmental Studies, Ehime University, Matsuyama, Japan

²Department of Civil and Environmental Engineering, Graduate School of Science and Engineering, Ehime University, Matsuyama, Japan

³Department of Special Needs Education, Graduate School of Education, Ehime University, Matsuyama, Japan

⁴Biological Control Research Unit, Center for Natural Science and Environmental Research, De La Salle University, Manila, Philippines

⁵College of Arts, Sciences and Education, Trinity University of Asia, Quezon City, Philippines

⁶St. Luke's College of Nursing, Trinity University of Asia, Quezon City, Philippines

⁷Guidance Counseling and Testing Department, University of Santo Tomas–Angelicum College, Quezon City, Philippines

⁸Pediatrics Department, Novaliches District Hospital, Quezon City, Philippines

Corresponding Author:

Kozo Watanabe, PhD

Center for Marine Environmental Studies

Ehime University

3 Bunkyo-cho

Matsuyama

Japan

Phone: 81 89 927 9847

Email: watanabe.kozo.mj@ehime-u.ac.jp

Related Article:

This is a corrected version. See correction statement: <https://publichealth.jmir.org/2021/4/e29795/>

Abstract

Background: While early detection and effective control of epidemics depend on appropriate surveillance methods, the Philippines bases its dengue fever surveillance system on a passive surveillance method (notifications from barangay/village health centers, municipal or city health offices, hospitals, and clinics). There is no available mHealth (mobile health) app for dengue fever that includes all the appropriate surveillance methods in early detection of disease outbreaks in the country.

Objective: This study aimed to evaluate the usability of the Mozzify app in terms of objective quality (engagement, functionality, aesthetics, information) and app subjective and app-specific qualities and compare total app mean score ratings by sociodemographic profile and self and family dengue fever history to see what factors are associated with high app mean score rating among school-based young adult samples and health care professionals. Individual interviews and focus group discussions were also conducted among participants to develop themes from their comments and suggestions to help structure further improvement and future development of the app.

Methods: User experience sessions were conducted among participants, and the Mobile Application Rating Scale (MARS) professional and user versions (uMARS) were administered followed by individual interviews and focus group discussions. Descriptive statistical analysis of the MARS and uMARS score ratings was performed. The total app mean score ratings by sociodemographic and dengue fever history using nonparametric mean difference analyses were also conducted. Thematic synthesis was used to develop themes from the comments and suggestions raised in individual interviews and focus group discussions.

Results: Mozzify obtained an overall >4 (out of 5) mean score ratings in the MARS and uMARS app objective quality (4.45), subjective (4.17), and specific (4.55) scales among 948 participants (79 health care professionals and 869 school-based samples). Mean difference analyses revealed that total app mean score ratings were not significantly different across ages and gender among

health care professionals and across age, income categories, and self and family dengue fever history but not gender ($P < .001$) among the school-based samples. Thematic syntheses revealed 7 major themes: multilanguage options and including other diseases; Android version availability; improvements on the app's content, design, and engagement; inclusion of users from low-income and rural areas; Wi-Fi connection and app size concerns; data credibility and issues regarding user security and privacy.

Conclusions: With its acceptable performance as perceived by health care professionals and school-based young adults, Mozzify has the potential to be used as a strategic health intervention system for early detection of disease outbreaks in the Philippines. It can be used by health care professionals of any age and gender and by school-based samples of any age, socioeconomic status, and dengue fever history. The study also highlights the feasibility of school-based young adults to use health-related apps for disease prevention.

(*JMIR Public Health Surveill* 2021;7(3):e19034) doi:[10.2196/19034](https://doi.org/10.2196/19034)

KEYWORDS

dengue fever; mHealth; public health surveillance; health communication; behavior modification; dengue outbreak

Introduction

In 1953, the first dengue hemorrhagic fever outbreak was reported in the Philippines [1]. Since then, this mosquito-borne viral infection that causes acute, potentially severe flu-like illness has been a leading public health burden in all regions of the country [2]. While early detection and effective control of epidemics depend on appropriate surveillance methods [3], the Philippines relies on a passive surveillance method that mainly depends on notifications from barangay/village health centers, municipal or city health offices, hospitals and clinics, and quarantine sections [4-6]. This limits reports of cases that are clinically diagnosed without laboratory confirmation [4,7] which is only a portion (14.3%) of the dengue cases [8,9]. This leaves patients with undifferentiated febrile illness or viral syndrome underreported and, thus, limits the capability to predict or control epidemics [8].

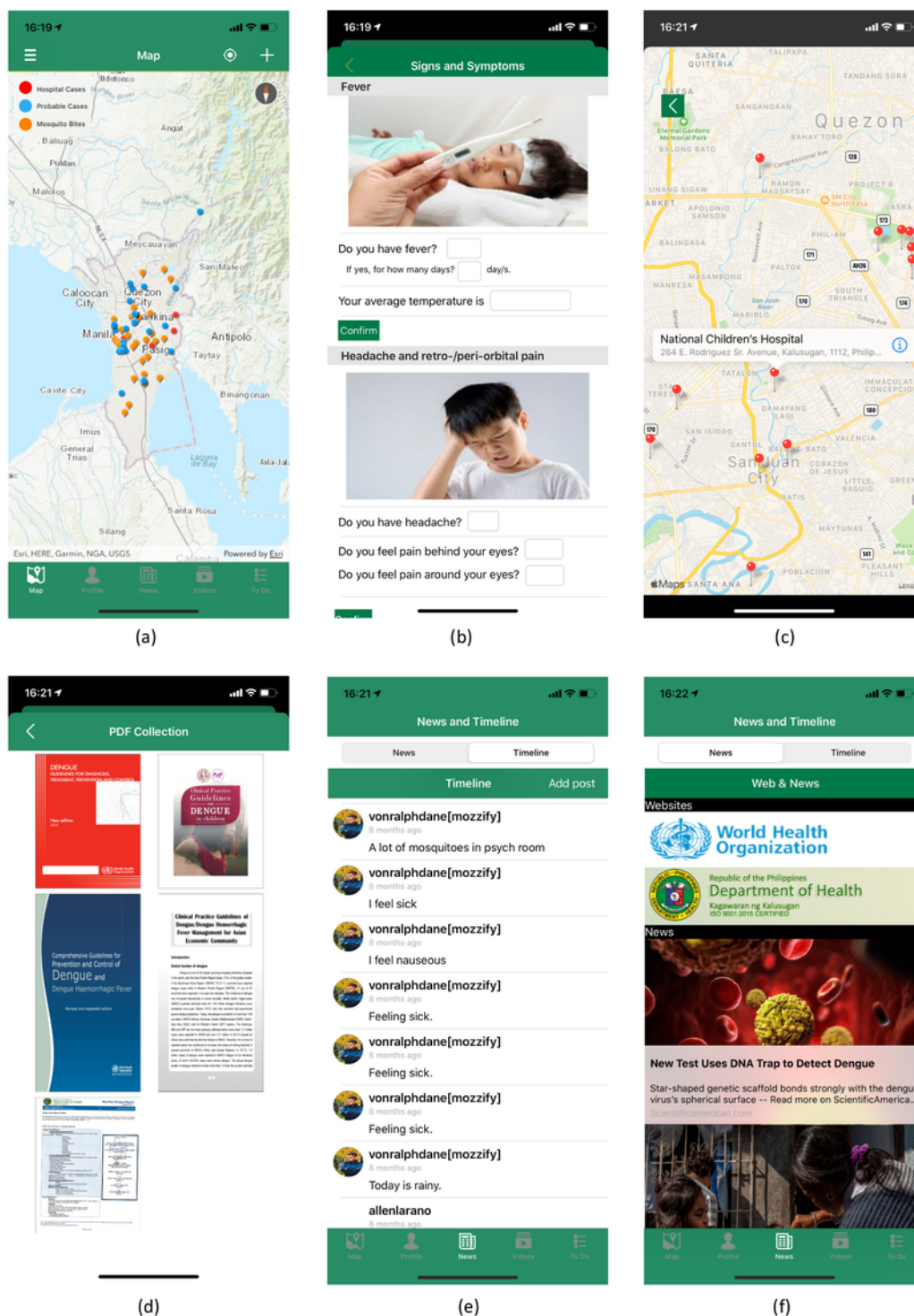
Public health surveillance aims to monitor dengue transmission accurately, triggering the necessary effective preventive measures and programs to prevent the occurrence and spread of diseases [8]. Recently, the use of mHealth (mobile health) technology, specifically mobile apps, has been gaining prominence as a potential surveillance system that would meet the need for real-time disease surveillance and timely identification of epidemics [3,10]. Mohanty et al [3] found 26 apps relevant to epidemic surveillance of diseases such as influenza, H1N1, Ebola, Zika, and dengue. Most of the apps were free of charge and provide real-time tracking and interactive maps, and the majority were on the Android platform only and focus on a single disease (mostly influenza). Some were country and language-specific and had narrow applicability; only a few were tailored to health professionals [3]. Thus, there is a pressing need to develop an app that addresses these limitations.

Mozzify is a noncommercial app that features real-time reporting and mapping of dengue cases, comprehensive health communication, and an evidence-based behavior modification system tailored for members of the general public and health care professionals [11] (Figure 1). It is an integrated mHealth app that combines appropriate surveillance methods in the early detection of disease outbreaks: indicator-based surveillance (IBS), event-based surveillance (EBS), and behavior

modification [3]. The app includes health care professionals in reporting laboratory-confirmed dengue fever cases, which is the provision of IBS [3,9,11]. It also uses ArcGIS's spatial analysis feature (Environmental Systems Research Institute) to identify hotspots, which is also an IBS method (sentinel surveillance) [3,11,12]. The reporting and mapping of patients with probable or suspected (with clinical symptoms) dengue fever through its interactive symptoms checker (syndromic surveillance, also an IBS method) is another app feature [3,9,11,13]. Mozzify also includes media reports and news, social media (timeline/chat forum), and links to websites of international and local health agencies to detect and monitor outbreaks which is the provision of the EBS method [3,11,14]. Most importantly, the app promotes behavior modification to address the poor translation of awareness or knowledge of the different preventive practices against dengue fever into practice [11].

Mozzify was pilot tested in Japan and obtained overall mean star ratings of ≥ 4.0 (out of 5) based on the Mobile Application Rating Scale (MARS) [11] among 5 health experts (4.4), 23 health-related researchers (4.1), and 22 nonclinical end users (4.1). It received ≥ 4.0 mean score ratings in objective and subjective quality (4.2) and ≥ 4.4 in specific quality from the participants [11]. The pilot study concluded that Mozzify may be a promising integrated strategic health intervention system for dengue case reporting and mapping, increasing awareness, improving knowledge, and changing attitudes about dengue [11,15]. However, the study had a small sample size and was not conducted in a country where dengue fever is endemic, which limited the generalizability of its findings. Thus, the next step was to evaluate the app's usability in the Philippines where there were approximately 106,630 dengue fever cases and 456 associated deaths reported from January 1 to June 29, 2019 (morbidity week 1 to week 26), which were higher than the alert and epidemic thresholds [16]. This led to the declarations of a National Dengue Alert on July 15, 2019, followed by a National Dengue Epidemic on August 6, 2019, by the Department of Health (DOH) when the cases reached 271,480, 113% higher than the same period in 2018 [17-20]. To the best of our knowledge, there is no other available integrated mHealth app for dengue fever that includes all the appropriate surveillance methods in early detection of disease outbreaks.

Figure 1. Screenshots of the Mozzify app: (a) Dengue fever cases and mosquito bite reporting and mapping, (b) symptoms checker, (c) hospital directions, (d) within-app local and international PDF collection on Dengue fever, (e) chat forum (timeline), (f) local and international health agency websites and news.



There were approximately 74 million (56.7% of the total population) mobile phone users in the Philippines as of 2019 [21,22]. According to a survey conducted in 2018 among those who own a mobile phone or use the internet (at least occasionally) in the Philippines, 93% are young adults (aged 18 to 29 years), of which 86% are pursuing higher education (those who are in university level and above) [23]. The survey

also reported that the second most common activity of mobile phone users in the Philippines is to seek information about health and medicine [23]. Given these data, school-based young adults can be encouraged to use mHealth apps. However, a prior study that involved this population reported that among those who had smartphones, only 14% had mHealth apps [24]. Despite this low use rate, 66.4% of those who had used mHealth apps

reported that these apps were useful, and 92.8% were satisfied with the functions of the mHealth apps that they owned [24]. Thus, this population, who comprise 30% of reported cases of dengue fever during the outbreak in 2019, can provide important information on the usability of the Mozzify app in the Philippines [20]. Further, the inclusion of health care professionals who have first-hand experience in the management of patients and are knowledgeable about the disease and its control and prevention will enable the app to address the actual needs of the users for whom it was developed.

Therefore, the objectives of this study were threefold: (1) test and evaluate the usability of the Mozzify app among health care professionals and school-based young adult samples in Metro Manila, Philippines, in terms of engagement, functionality, aesthetics, information, app subjective and app-specific qualities; (2) compare the total app mean score ratings by sociodemographic profile and self and family dengue fever history to identify factors associated with a high app mean score rating; and (3) conduct individual interviews and focus group discussions among the participants and analyze their comments and suggestions to help structure further improvement of the app. This app will be the first to be tested and introduced in the Philippines, specifically Metro Manila during an outbreak. Hence, we hypothesized that the mean score ratings among the participants will be high due to the app's relevance to the current outbreak and the lack of a surveillance system in the country. We also hypothesized that the total app mean score ratings will not be significantly different across ages, gender, income categories, and self and family dengue fever history, indicating that the app can be used as a potential and appropriate surveillance method in early detection of dengue fever outbreaks.

Methods

Study Design and Sampling Method

This cross-sectional mixed-methods usability study was conducted from August to September 2019 in Metro Manila, Philippines, one of the regions that was greatly affected by the dengue fever outbreak with 18,136 reported cases from January 1 to August 31, 2019 [20]. School-based samples were recruited based on inclusion criteria (university students aged between 18 to 29 years who owned a mobile phone) using the respondent-driven sampling technique by group from a university in Quezon City, the largest city with the highest population among the cities in Metro Manila as of 2015 [24,25]. Health care professionals were recruited from academic, clinical, research, or community settings based on the authors' local networks and through the use of a recruitment poster that was sent or posted via email and social networking sites.

Ethical Considerations

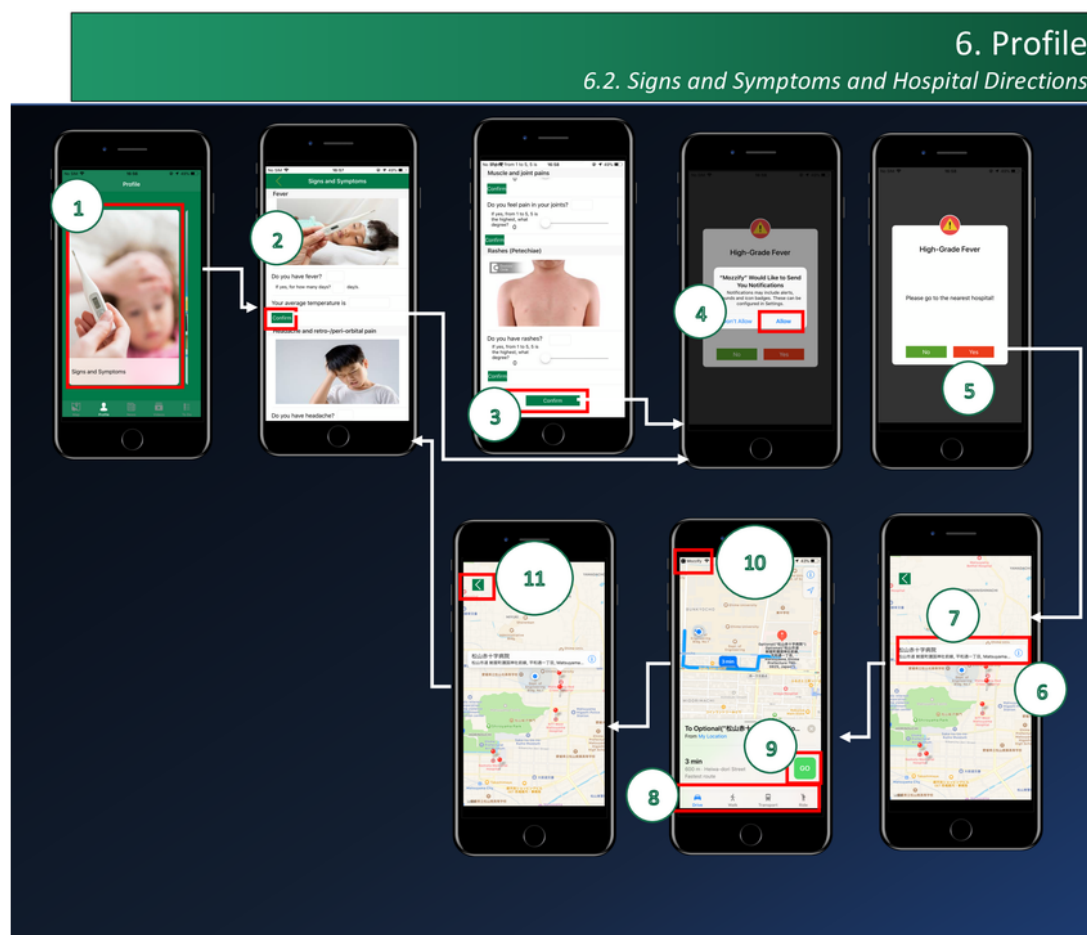
This study was approved and conducted in accordance with international ethical guidelines (2019-28-Herbuela-VPAA-Mozzify-v1) [26,27]. When found eligible, participants were asked to read and sign (for online Google Forms version, by clicking "Yes, I agree") an informed consent sheet and were told their participation in the study was voluntary and they may stop their participation at any time. All forms containing personal information were coded and stored in a password-protected database and computer.

Mozzify App

The app has 3 main components: real-time dengue fever case reporting and mapping system, health communication, and behavior modification. Through the use of ArcGIS, users can report a probable case or mosquito bite incidence by pinning their current location (or any other location). The map counts the pins in each barangay or village and colors them depending on the number of dengue cases. The app lets the user check their symptoms and alerts and navigates to the nearest hospitals that cater to dengue fever patients if it is found they require prompt medical assessment by a physician (determination is made after user answers 26 symptom-related questions) [11]. The app shows different evidence-based local and international guidelines in PDF, videos, and news on the control, prevention, diagnosis, and treatment of dengue fever and websites of the World Health Organization (WHO) and DOH [28,29]. It also has a chat forum (timeline) where users can post events, concerns, and questions about dengue fever. Mozzify promotes behavior modification using an alert system to encourage the practice of preventive measures against dengue based on WHO's communication for behavioral impact strategy [15,30,31].

Usability Testing

Usability testing was done to investigate the user-Mozzify interaction in terms of its functions, engagement, design, information, subjective and specific qualities, effectiveness, efficiency, and satisfaction of the users to produce recommendations for the design (or redesign), structure, or further iterations [32-35]. All participants were asked to install the app on their mobile phones; otherwise, they were asked to use the available mobile phones with the installed Mozzify app. The majority of health care professionals were invited for the individual user experience sessions, while school-based samples were invited for a series of 1-hour investigator-facilitated group user experience sessions. In the first 15 minutes, the investigator used the screen projection method to show the mobile app and its functions to the participants and let them perform the tasks based on a predetermined checklist matched with the functions of the app. However, in cases when participants could not attend the sessions, they were asked to use the app's user guide (Figure 2) or explore the app as they would do in other apps.

Figure 2. Sample screenshot of the app user guide.

1. Click **Signs and Symptoms** window.

2. Answer all the questions and click the **Confirm** button in each symptom. If your condition requires immediate clinical assessment, the app will alert you to go to a nearest hospital.

3. Finish answering all the questions in all the symptoms and click **Confirm**.

4. Click **Allow** button to send you notifications.

5. Click **Yes** button in the alert to show

hospitals around you.

6. Click the **red** pin to show the hospital information.

7. Click to show directions.

8. Choose among the **options**.

9. Click **GO** button to start route.

10. Click the **Mozzify** in the upper leftmost part of the screen to go back.

11. Click **back arrow** button to go back.



Mozzify User Guide, June 2019 version 1.0

All rights reserved. No part of this publication may be produced or transmitted in any form or by any means, including photocopying and recording, without seeking the permission of the developer. For inquiries, email herbuelavonralphdanc@gmail.com

Measures of Interest and Outcomes

Sociodemographic Profile

Participants were asked to provide their sociodemographic information such as age and gender to allow investigation of mean score differences within its categories. Age was clustered according to the minimum and maximum (aged 18 to 19 years for school-based samples and 19 to 53 years for health care professionals) ages of the participants with an equal interval

size of 4 (eg, 18-21, 22-25). The school-based samples were asked to provide their family monthly income, which was categorized (adjusted to 10,000-peso interval size) using the clusters from the indicative range of monthly family incomes (for a family of 5) based on a study on the profile and determinants of the middle-income class in the Philippines in 2015 and 2017 [36]. Information on self and family history of dengue fever was collected from school-based samples to investigate whether a previous diagnosis or history of dengue

fever in a family would be associated with higher total mean score app ratings.

MARS

The second phase of testing included the use of a standardized self-administered questionnaire ([Multimedia Appendix 1](#)). Two versions of the MARS, the professional 23-item MARS and the 20-item user version (uMARS), were used to evaluate the usability of Mozzify among health care professionals and the school-based samples either online or on paper [32,37]. Both versions use a 5-point scale (1=inadequate to 5=excellent) to assess the app in terms of objective quality (engagement, functionality, aesthetics, and information), subjective quality (4 items on recommending the app to others, using the app for a short or long time, and overall star rating of the app), and the app-specific quality (perceived impact of the app on the user's awareness and knowledge of and attitudes toward dengue fever, help-seeking behavior, intentions, and actual change of behavior in practicing preventive measures) [32,37]. Both versions had excellent internal consistency level Cronbach alpha=.90 [32,37].

Individual Interviews and Focus Group Discussions

After answering the questionnaire, health care professionals were invited to individual interviews and school-based samples invited to focus group discussions to provide more comprehensive qualitative feedback. The sessions were completed by one trained investigator who was guided with preselected questions (semistructured), and all the participants were asked to write their comments and suggestions to ensure consistency in the analysis.

Statistical and Data Analysis

Statistical analysis was conducted using SPSS Statistics version 25 (IBM Corporation). All items on the MARS and uMARS were scored based on responses on the 5-point Likert scale (5 points=excellent, 4 points=good, etc) [37]. The mean score of each subscale in the app objective quality scale, each item of the app subjective quality scale (reported as individual items and by mean score), and the app-specific scale were calculated separately for the school-based samples and health care

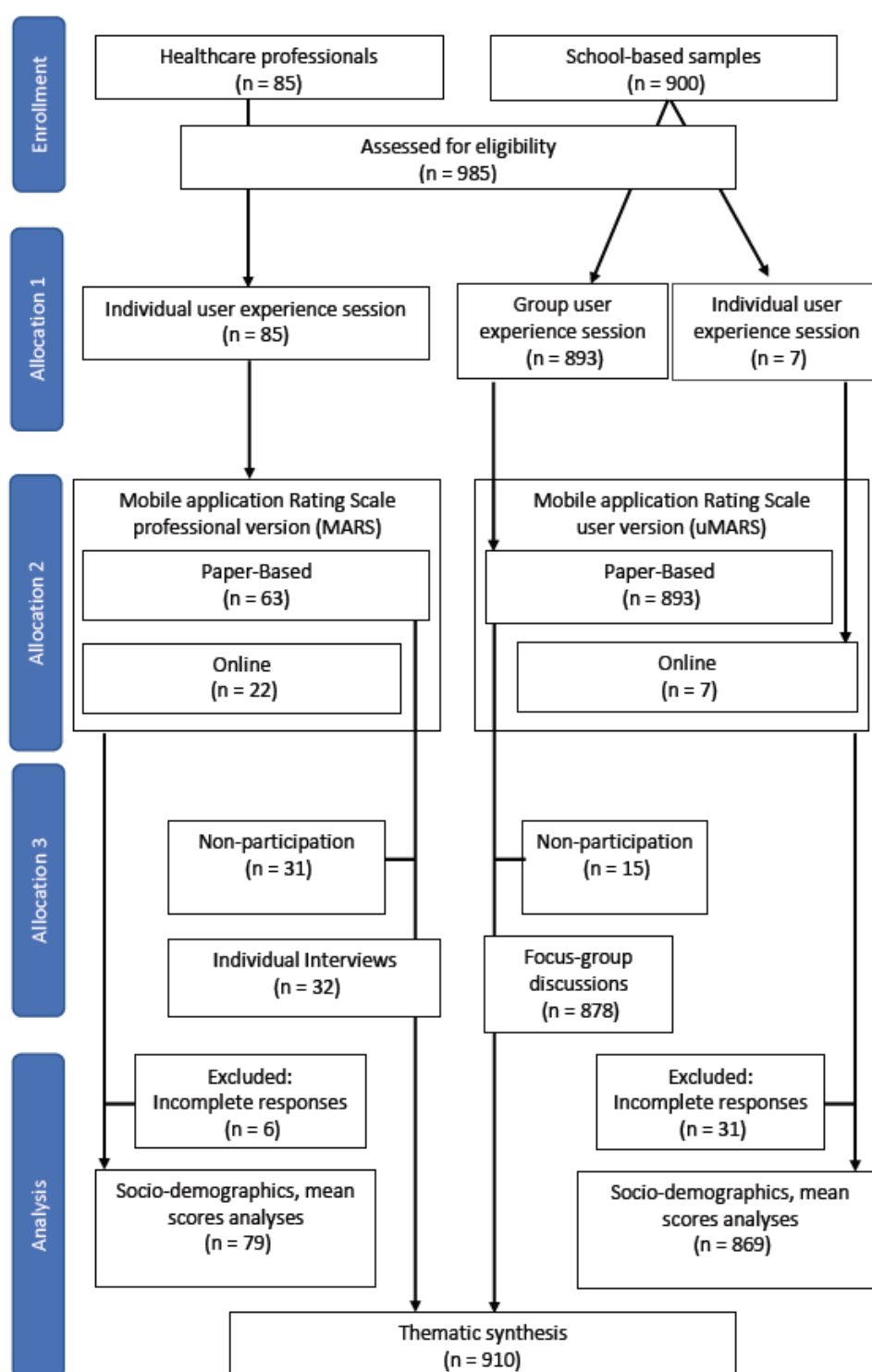
professionals. The mean scores of each scale were then combined to get the overall mean score among the participants, which was also used to investigate the differences in means within each category of age, gender, income, and dengue fever history (self and family). Nonparametric Mann-Whitney *U* tests and Kruskal-Wallis tests were used for dichotomous and multicategorical variables, respectively, as normality tests revealed that the data were not normally distributed ($P<.001$).

Comments and suggestions in the individual interviews and group discussions were analyzed using thematic synthesis [38]. The transcripts of the comments and suggestions of the participants were encoded verbatim to an Excel file (Microsoft Corporation) as a database. Topic and key concept or content unit was identified by examining (word-by-word) each transcript [39]. The identified content units were used to easily add to a bank of codes for themes, and new ones were developed as required. Similar or overlapping content units were coded (with numbers that corresponded to a theme) and quantified using quantitative content analysis where each content unit was scored 1 [40]. The similar or overlapping content units were grouped into minor themes and then grouped to major themes. The frequencies of content units in each theme were added to identify its relevance.

Results

Participant and Data Flow

We were able to recruit 985 eligible participants (health care professionals, $n=85$; school-based young adult samples, $n=900$). All of the health care professionals and school-based samples attended the user experience sessions. Of the health care professionals, 37.6% (32/85) were able to attend the individual interviews, while 98.3% (878/893) of school-based samples attended the focus group discussions. The Consolidated Standards of Reporting Trials (CONSORT) diagram ([Figure 3](#)) shows the detailed participant flow from user experience sessions, excluding (incomplete) responses, to the number of responses included in the analysis.

Figure 3. Consolidated Standards of Reporting Trials diagram of participant flow from enrollment, allocation, and analysis.

Sociodemographic and Dengue Fever History Profile

The school-based samples were mostly female (549/869, 63.2%) aged 18 to 21 years (698/869, 89.6%) with a mean age of 19.5 years; 28.8% (131/455) come from a family with a monthly

income of $\geq 50,000$ pesos. Three-quarters (417/555, 75.1%) had no dengue fever history, and 55.7% (306/550) had no family members with a history of dengue fever, as shown in Table 1. Health care professionals had a mean age of 31.8 years with a nearly equal distribution by gender (35/79, 50.7% females).

Table 1. Sociodemographic and dengue fever history profile of health care professionals and school-based samples.

Sociodemographic profile	Health care professionals (n=79), n (%)	School-based samples (n=869), n (%)
Gender		
Male	35 (49.3)	234 (26.9)
Female	36 (50.7)	549 (63.2)
Age in years		
18-21	3 (5.17)	698 (89.6)
22-25	14 (24.1)	68 (8.73)
26-29	12 (20.7)	13 (1.67)
30-33	12 (20.7)	— ^a
34-37	7 (12.1)	—
38-41	4 (6.90)	—
42-45	4 (6.90)	—
46-49	3 (5.17)	—
50-53	2 (3.45)	—
Income (₱^b)		
≤10,000	—	46 (10.1)
11,000-20,000	—	80 (17.6)
21,000-30,000	—	90 (19.8)
31,000-40,000	—	58 (12.8)
41,000-50,000	—	50 (11.01)
≥51,000	—	131 (28.8)
Self DF^c history		
None	—	417 (75.3)
Had DF	—	138 (24.9)
Family DF history		
None	—	306 (55.7)
≥1 member had DF	—	244 (44.4)

^aN/A: not applicable.

^b₱: Philippine peso (US \$1=₱50.5; 5-year average rate).

^cDF: dengue fever.

App Objective, Subjective, and Specific Quality

Mozzify obtained >4 (out of 5) mean score ratings in all app objective quality scales among the participants as shown in Table 2. Mean score ratings of >4.50 were found in the information subscale (mean 4.56 [SD 0.6]) among school-based young adult samples, followed by functionality (mean 4.26 [SD 0.51]) and engagement (mean 4.26 [SD 0.45]) among the health care professionals. In the app subjective scale, the item recommending the app to others obtained the highest mean

score ratings of 4.58 (SD 0.70) among the school-based sample and 4.62 (SD 0.63) with the health care professionals. In the app-specific scale, improving awareness, knowledge, and help-seeking behavior were the items with the highest mean score ratings (4.59, 4.52, and 4.59, respectively) among health care professionals while improving awareness and help-seeking behavior items got the highest mean score ratings of 4.63 and 4.61, respectively among school-based samples. Conversely, the item on using the app for the next 12 months under app subjective quality saw the lowest mean score ratings of 3.64.

Table 2. Mean scores of app objective, subjective, and specific quality ratings based on the Mobile Application Rating Scale from health care professionals and school-based samples.

MARS ^a /uMARS ^b subscales	Health care professionals (n=79), mean (SD)	School-based samples (n=869), mean (SD)	Combined (n=948), mean (SD)
App objective quality	4.22 (0.59)	4.47 (0.69)	4.45 (0.67)
1. Engagement	4.26 (0.45)	4.39 (0.75)	4.38 (0.74)
2. Functionality	4.26 (0.51)	4.46 (0.69)	4.45 (0.68)
3. Aesthetics	4.17 (0.49)	4.47 (0.67)	4.44 (0.67)
4. Information	4.19 (0.62)	4.56 (0.63)	4.53 (0.64)
App subjective quality	3.71 (0.57)	4.17 (0.88)	4.13 (0.88)
5. Recommending the app to others	4.62 (0.63)	4.58 (0.70)	4.59 (0.69)
6. Using the app for the next 12 months	3.35 (0.79)	3.67 (0.92)	3.64 (0.91)
7. Overall (star) rating of the app	4.42 (0.67)	4.25 (0.74)	4.27 (0.73)
App specific quality	4.48 (0.52)	4.55 (0.67)	4.54 (0.67)
8. Awareness (DF ^c symptoms, hospital locations, hotspots, prevention, treatment)	4.59 (0.63)	4.61 (0.66)	4.61 (0.66)
9. Knowledge (DF symptoms, prevention, treatment)	4.52 (0.64)	4.58 (0.66)	4.58 (0.66)
10. Attitude (severity, susceptibility, preventive practices against DF)	4.37 (0.64)	4.42 (0.72)	4.41 (0.72)
11. Intention-to-change (preventive practices against DF)	4.43 (0.61)	4.49 (0.67)	4.48 (0.67)
12. Help-seeking (for clinical assessment for presence of DF symptoms)	4.59 (0.63)	4.63 (0.63)	4.63 (0.63)
13. Behavior change (preventive practices against DF)	4.35 (0.56)	4.57 (0.67)	4.55 (0.66)

^aMARS: Mobile Application Rating Scale.^buMARS: Mobile Application Rating Scale–User Version.^cDF: dengue fever.

Mean Score Differences by Sociodemographics and Dengue Fever History

App mean score ratings were not significantly different across age and gender among health care professionals and age, income

brackets, and self and family dengue fever history, except gender among school-based samples as shown in [Table 3](#). Females had significantly ($P<.001$) higher app mean score ratings (mean 4.50 [SD 0.41]) than males (mean 4.34 [SD 0.52]).

Table 3. Total mean score comparison among health care professionals and school-based samples by sociodemographic and dengue fever history.

Sociodemographic profile	Health care professionals (n=79)		School-based samples (n=869)	
	Mean (SD)	P value	Mean (SD)	P value
Gender^a				
Male	4.31 (0.32)	.34	4.34 (0.52)	<.001
Female	4.22 (0.40)	— ^b	4.50 (0.41)	—
Age in years^c				
18-21	4.69 (0.02)	.67	4.45 (0.46)	.97
22-25	4.23 (0.18)	—	4.47 (0.42)	—
26-29	4.39 (0.40)	—	4.43 (0.52)	—
30-33	4.26 (0.34)	—	—	—
34-37	4.40 (0.40)	—	—	—
38-41	4.16 (0.74)	—	—	—
42-45	4.35 (0.27)	—	—	—
46-49	4.46 (0.48)	—	—	—
50-53	4.19 (0.08)	—	—	—
Income (₱)^{d,c}				
≤10,000	—	—	4.47 (0.58)	.18
11,000-20,000	—	—	4.45 (0.41)	—
21,000-30,000	—	—	4.42 (0.50)	—
31,000-40,000	—	—	4.38 (0.54)	—
41,000-50,000	—	—	4.33 (0.42)	—
≥51,000	—	—	4.48 (0.39)	—
Self DF^e history^{b,c}				
None	—	—	4.46 (0.44)	.61
Had DF	—	—	4.42 (0.48)	—
Family DF history^{b,c}				
None	—	—	4.44 (0.46)	.29
≥1 member had DF	—	—	4.47 (0.43)	—

^aMann-Whitney *U* test.^bN/A: not applicable.^cKruskal-Wallis test.^d₱: Philippine peso (US \$1=₱50.5; 5-year average rate).^eDF: dengue fever.

Individual Interviews and Focus Group Discussions

Thematic Synthesis

Of the 910 attendees (32 health care professionals and 878 school-based samples), only 424 transcripts were collected from 26 (out of 32) health care professionals and 398 (out of 878) school-based young adult samples. From 424 transcripts analyzed, a total of 443 content units (47 from the health care professionals and 369 from the school-based samples) were identified and grouped into 16 minor themes that were then

grouped into 7 major themes as shown in Table 4: positive comments regarding the app's concept, design, information, and features (320/443, 72.2%); suggestions on adding features like multilanguage options and including other diseases (48/443, 10.8%); Android version availability (36/443, 8.13%); inclusion of users from low-income and rural areas (14/443, 3.16%); improvements on the app's content, design, and engagement (14/443, 3.16%); Wi-Fi connection and app size concerns (6/443, 1.35%); and data credibility and issues regarding user security and privacy (5/443, 1.13%).

Table 4. The major themes that emerged from individual interviews and focus group discussions among health care professionals and school-based samples.

Themes	Content units		
	Health care professionals (n=47)	School-based samples (n=396)	Combined (n=443)
1. Positive comments (concept, design, information, features)	27 (57.4)	293 (73.9)	320 (72.2)
2. Suggestions (add features: multilanguage, include other diseases, alerts)	11 (23.4)	37 (9.34)	48 (10.8)
3. Android version availability	6 (12.8)	30 (7.58)	36 (8.13)
4. Inclusivity (rural areas, low-income families)	2 (4.33)	12 (3.03)	14 (3.16)
5. Improvement (content and design, engagement)	— ^a	14 (3.54)	14 (3.16)
6. Wi-Fi connection and app size	—	6 (1.52)	6 (1.35)
7. Data credibility and user security and privacy	1 (2.1)	4 (1.01)	5 (1.13)

^aN/A: not available.

Positive Comments

Almost three quarters (320/443, 72.2%) of the topics discussed in individual and group discussions included positive comments regarding the app. Health care professionals perceived that the app is concise and relevant, interesting, well-executed, and informative. They also perceived that the app will be of help to clinicians in improving health care awareness and services to the people, especially those patients with dengue fever. The school-based samples also had positive comments about the app: user-friendly, useful, educational, innovative, well-designed, interactive, and unique. Some commented about the capability of the app to improve public awareness of dengue fever.

This mobile app is what we need today to raise medical awareness about vector-borne diseases. This is timely because the country is facing a big dilemma on dengue. This could help doctors in locating possible dengue cases so that proper medical attention will be given. Looking forward to using this app and hoping to be approved by the Department of Health because it really helps in our country to give awareness and knowledge about dengue, and in future the high increase of dengue cases will be decreased.

Suggestions

Participants also suggested including some additional features and information that they perceived should be in the app. Suggestions including adding a feature that would allow users to easily access medical services professionals, multilanguage options for users, more information (through videos, graphics, and PDF files) about vector mosquitoes, information about the dengue vaccine, how to prevent or avoid dengue, vector-human interaction, and including other diseases (eg, HIV, tuberculosis, other mosquito-borne diseases). Participants also suggested connecting the app to social media like Facebook and Twitter, adding advertisements and games, and developing a desktop version for hospitals and clinics.

Maybe link, a feature that can immediately notify the nearby hospital for assistance; add emergency

hotlines of nearby hospitals/clinics. Have multiple languages, so a particular country can use their own languages to better understand the use of the app. This app should also include other diseases caused by mosquitoes. Maybe this app can also be applied not only for dengue cases or human health-related but also for environmental issues and disease outbreaks in the agriculture sector. Include videos or mini games that would make the app more interactive.

Android Version

A total of 36 (36/443, 8.13%) topics were raised regarding the need for an Android version of the app.

The app should be available to all and not restricted to iOS to fully utilize its intention to help raise awareness about these diseases caused by mosquitoes.

Inclusivity

Another major theme is the need for inclusivity (14/443, 3.02%). Participants commented that the app should bridge the gap between low-income and those who are in rural areas and receiving appropriate medical services through the use of the app.

This is a novel approach to mitigate the barriers between rural areas and health care. The people living in the line of poverty might not be able to access the app immediately and since most people who are experiencing poverty are the ones most vulnerable to dengue; Needs to be more accessible to people in different social classes.

Improvement

This theme emerged based on the comments of members of the school-based sample regarding improving the app's design, color, and mapping system and making the app more interactive (14/396, 3.54%).

Enhance the graphics to be more aesthetically pleasing so more would be attracted to use it often. It should be more colorful to attract millennial

students or even kids/teens. The user interface seems pretty bare, so maybe make it a bit more appealing.

Wi-Fi Connection and App Size

Concerns regarding limited access to Wi-Fi connections and the huge size of the app (6/396, 1.52%) were raised by the members of the school-based samples.

It is a useful app but what about for people who don't have strong Wi-Fi connections. I am not sure either if the PDFs and videos that supposedly can play without internet are required since I assume this will take a lot of space within the phone's storage.

Data Credibility and Issues Regarding User Security and Privacy

The last but probably one of the most important themes emerging were concerns regarding data security and privacy issues (5/443, 1.13%).

How can we prevent the abuse of misinformation and protect the credibility of the app? Concern regarding validation and data privacy act. Accessibility with all which is not advisable due to discrimination.

Discussion

Principal Findings

Mozzify obtained high mean score ratings (>4 out of 5) in the MARS app objective, subjective, and specific quality scales. Mean difference analyses revealed that total app mean score ratings were not significantly different across ages among health care professionals and across income categories and self and family dengue fever history but not gender among school-based samples. Thematic synthesis of the topics discussed in the individual interviews and focus group discussions revealed 7 major themes.

This study extends the pilot study conducted in Japan by including a significantly large sample size in the Philippines where dengue fever is endemic [14]. To our knowledge, Mozzify is the first mobile app for dengue fever that has been tested and introduced in the Philippines, especially in Metro Manila. Our study also provides the first report on the perception of health care professionals and school-based samples regarding the app's usability in terms of its functions, engagement, design, information, and subjective and specific qualities.

The high mean score ratings (>4 out of 5) in the MARS app objective quality scale, specifically in the information and functionality subscales, indicate that it contains high-quality information from a credible source and excellent functioning, easy-to-learn navigation flow, logic, and gestural design. The high mean score ratings of the item recommending the app to others in the subjective quality scale indicate that the participants had a high desire or intention to recommend the app to others. Most importantly, the results in the app-specific scale indicate that the app has the potential to improve user awareness and knowledge of dengue fever symptoms, hotspots, prevention, and management; encourage help-seeking behavior among users through its unique and interactive symptoms checker and

hospital navigation system; improve users' intentions; and encourage users to develop or practice preventive measures against dengue fever.

The total app mean score ratings, which were not significantly different across ages, indicate that Mozzify is relevant and can be used as a potential surveillance method for dengue fever as perceived by health care professionals and school-based samples of any age (above 18 years). This perception is also true among school-based samples of any socioeconomic status with or without a self or family history of dengue fever. Surprisingly, among the school-based samples, females had a significantly ($P<.001$) higher total app mean score rating than males. Aside from the fact that there were more females (549/869, 63.2%) than males (234/869, 26.9%) among the school-based samples, the nearly equal distribution of males and females whose mean ratings were not significantly different ($P=.34$) among health care professionals suggests that the app mean score ratings between female and male school-based samples would not also be statistically different if they would have had equal sample sizes.

The interview and discussion sessions revealed that while the participants had positive perceptions of the app, they also suggested adding features like multilanguage options and including other diseases. As mentioned earlier, most of the available mHealth apps for epidemic surveillance focus on a single disease (mostly influenza), and only a few were tailored for health professionals [3]. While the latter has been included from conception to evaluation of Mozzify, including other mosquito-borne diseases (eg, malaria, Zika, chikungunya, and Japanese encephalitis) in the app has been considered, hence the name Mozzify, which was termed after the word mosquito. To our knowledge, one similar app called FeverDX has included different vector-borne diseases and, like Mozzify, has received a high mean score rating of 4 (out of 5) among health care professionals [41].

Enhancing the app's engagement ability also emerged as one of the concerns of the participants. The relatively low mean score rating on using the app in the next 12 months can be associated with the relatively low overall mean score rating of 4.38 in the engagement subscale under the app objective quality scale. These emphasize the need to sustain the engagement of the users with the app for longer use and the possibility of including games, a potential tool to increase engagement among users that can be integrated with health behavior change concepts [42]. Moreover, the participants also suggested that the app should have an Android version and an added alert system. While some of the epidemic surveillance apps run on the Android platform [3], Mozzify is only available in iOS. An Android version has been developed, and an alert system has been designed in the map that would warn users when they enter or when they are near a barangay/village or area with high dengue fever case incidence and mosquito abundance. Currently, a version with multilanguage options is being developed in preparation for its use in other countries where dengue is prevalent.

The interview and discussion sessions also introduced important factors that should be considered in the further improvement of

the app. First, a stable Wi-Fi connection is indeed necessary for most of the features of the app to function. ArcGIS has an offline version of the online map; however, it does not allow users to input mosquito bite reports and dengue fever cases. This has been addressed by directly installing the videos and PDFs of online-sourced evidence-based local and international guidelines in the app. However, it increases the app's size which is another concern. Second, and probably the most important, are issues regarding the credibility of user data inputs and security and privacy. To ensure the credibility of data inputs from users, a mechanism to include email or phone number verification among users when signing up for the app that would allow the system to track suspicious accounts has been designed. Also, a plan to connect with and train health care workers to confirm inputted data or users within their barangays or neighborhood has been in place. However, there is also a need to limit the access of the online map that shows cases among health care providers and local government officials to aid in proper planning and mitigation with the affected areas.

Limitations

This study has several limitations. First, participants resided in urbanized areas and thus, the findings may not represent those who are in rural areas. Second, participants only used and assessed the app in a very short span of time, limiting the evaluation to the app's design and usability. Third, the investigator-led user experience sessions conducted may bias the results of this study as they may not reflect a real-world

scenario. Fourth, the experiences and perceptions of the school-based young adult samples in this study may not represent that of the majority of young adults. Fifth, the participants were mostly students and health care professionals in an academic and research-based setting, which could confound their true assessment of the app since they are more involved in academic- or research-approach assessment and may be more inclined to give positive responses. These limitations should be addressed by conducting a longitudinal (eg, 8-month use of the app) implementation study in urban and rural communities that may give insights on the app's actual use in a real-world scenario (eg, number of reported cases, pre and post levels of perceived change in awareness, knowledge on, and attitudes toward dengue fever, help-seeking behavior, intentions, and actual change of behavior in practicing preventive measures against dengue fever).

Conclusions

Despite these limitations, Mozzify, with its acceptable performance as perceived by health care professionals and school-based young adults, has the potential to be used as a strategic health intervention system in early detection of disease outbreaks in the Philippines. Based on our results, it could also be used by health care professionals and school-based samples of any age group, gender, and socioeconomic status. The study also emphasizes the participation of young adults in the usability testing of Mozzify, which highlights the feasibility of this population to use health-related apps for disease prevention.

Acknowledgments

This study was supported by the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (17H01624, 19H01144), JSPS Core-to-Core Program B Asia-Africa Science Platforms, and Endowed Chair Program of the Sumitomo Electric Industries Group Corporate Social Responsibility Foundation, which had no role in the design, data collection, statistical analysis, and writing of this manuscript. The authors would like to thank all the students and professors from Trinity University of Asia, Quezon City, Philippines, and health care professionals who participated in this study.

Authors' Contributions

VRDMH conceptualized, designed, and developed the Mozzify app; designed the study; wrote the protocol; conducted the data gathering; analyzed the data; and wrote this manuscript. TK helped design the framework and features of the mobile app. TMC, HTH, JMOL, and GDS supervised the data gathering and analysis procedures. RAR supervised the data processing and analysis. KW supervised the data gathering and provided guidance and comments on the analysis and the initial drafts of this manuscript. All authors contributed to and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Mobile Application Rating Scale form.

[[PDF File \(Adobe PDF File\), 175 KB](#) - [publichealth_v7i3e19034_app1.pdf](#)]

References

1. Gubler DJ. Dengue and dengue hemorrhagic fever. Clin Microbiol Rev 1998 Jul;11(3):480-496 [[FREE Full text](#)] [Medline: [9665979](#)]
2. Dengue and severe dengue. World Health Organization. 2018. URL: <https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue> [accessed 2018-05-12]

3. Mohanty B, Chughtai A, Rabhi F. Use of mobile apps for epidemic surveillance and response: availability and gaps. *Global Biosecurity* 2019 Sep 11;1(2):37 [FREE Full text] [doi: [10.31646/gbio.39](https://doi.org/10.31646/gbio.39)]
4. Edillo FE, Halasa YA, Largo FM, Erasmo JNV, Amino NB, Alera MTP, et al. Economic cost and burden of dengue in the Philippines. *Am J Trop Med Hyg* 2015 Feb;92(2):360-366 [FREE Full text] [doi: [10.4269/ajtmh.14-0139](https://doi.org/10.4269/ajtmh.14-0139)] [Medline: [25510723](https://pubmed.ncbi.nlm.nih.gov/25510723/)]
5. Edillo F, Madarieta S. Trends of dengue infections (1997-2008) in Cebu Province, Philippines. *Dengue Bull* 2012;36:37-49 [FREE Full text]
6. National Epidemiology Center of the Department of Health. Manual of Procedures for the Philippine Integrated Disease Surveillance and Response. 1st Edition. Manila: Department of Health; 2008. URL: <http://docshare02.docshare.tips/files/10846/108466912.pdf> [accessed 2021-02-18]
7. Shepard DS, Undurraga EA, Halasa YA. Economic and disease burden of dengue in Southeast Asia. *PLoS Negl Trop Dis* 2013;7(2):e2055 [FREE Full text] [doi: [10.1371/journal.pntd.0002055](https://doi.org/10.1371/journal.pntd.0002055)] [Medline: [23437406](https://pubmed.ncbi.nlm.nih.gov/23437406/)]
8. Ooi E, Gubler DJ. Dengue in Southeast Asia: epidemiological characteristics and strategic challenges in disease prevention. *Cad Saude Publica* 2009;25 Suppl 1:S115-S124 [FREE Full text] [doi: [10.1590/s0102-311x2009001300011](https://doi.org/10.1590/s0102-311x2009001300011)] [Medline: [19287856](https://pubmed.ncbi.nlm.nih.gov/19287856/)]
9. Yan SJ, Chughtai AA, Macintyre CR. Utility and potential of rapid epidemic intelligence from internet-based sources. *Int J Infect Dis* 2017 Oct;63:77-87 [FREE Full text] [doi: [10.1016/j.ijid.2017.07.020](https://doi.org/10.1016/j.ijid.2017.07.020)] [Medline: [28765076](https://pubmed.ncbi.nlm.nih.gov/28765076/)]
10. El-Khatib Z, Shah M, Zallappa SN, Nabeth P, Guerra J, Manengu CT, et al. SMS-based smartphone application for disease surveillance has doubled completeness and timeliness in a limited-resource setting: evaluation of a 15-week pilot program in Central African Republic (CAR). *Confl Health* 2018;12:42 [FREE Full text] [doi: [10.1186/s13031-018-0177-6](https://doi.org/10.1186/s13031-018-0177-6)] [Medline: [30386418](https://pubmed.ncbi.nlm.nih.gov/30386418/)]
11. Herbuela VRDM, Karita T, Francisco ME, Watanabe K. An integrated mHealth app for dengue reporting and mapping, health communication, and behavior modification: development and assessment of Mozzify. *JMIR Form Res* 2020 Jan 08;4(1):e16424 [FREE Full text] [doi: [10.2196/16424](https://doi.org/10.2196/16424)] [Medline: [31913128](https://pubmed.ncbi.nlm.nih.gov/31913128/)]
12. The National Center for Biotechnology Information (NCBI). MeSH Database: Sentinel Surveillance. The National Center for Biotechnology Information (NCBI). URL: <https://www.ncbi.nlm.nih.gov/mesh?Db=mesh&Cmd=DetailsSearch&Term=%22Sentinel+Surveillance%22%5BMeSH+Terms%5D> [accessed 2021-02-18]
13. Flamand C, Quenel P, Ardillon V, Carvalho L, Bringay S, Teisseire M. The epidemiologic surveillance of dengue-fever in French Guiana: when achievements trigger higher goals. *Stud Health Technol Inform* 2011;169:629-633. [Medline: [21893824](https://pubmed.ncbi.nlm.nih.gov/21893824/)]
14. Global Health Protection and Security. Centers For Disease Control and Prevention. URL: <https://www.cdc.gov/globalhealth/healthprotection/gddopscenter/how.html> [accessed 2019-05-30]
15. Herbuela VRDM, de Guzman FS, Sobrepeña GD, Claudio ABF, Tomas ACV, Arriola-Delos Reyes CM, et al. Knowledge, attitude, and practices regarding dengue fever among pediatric and adult in-patients in Metro Manila, Philippines. *Int J Environ Res Public Health* 2019 Nov 26;16(23):4705 [FREE Full text] [doi: [10.3390/ijerph16234705](https://doi.org/10.3390/ijerph16234705)] [Medline: [31779171](https://pubmed.ncbi.nlm.nih.gov/31779171/)]
16. Department of Health. Monthly Dengue Cases Report No 6, January 1 to June 29, 2019. Epidemiology Bureau, Public Health Surveillance Division. 2019. URL: <https://doh.gov.ph/sites/default/files/statistics/Dengue%20Monthly%20Report%20No.%206.pdf> [accessed 2019-09-23]
17. Situation report 1: dengue outbreak. Representative Office for the Philippines, World Health Organization. 2019 Jul 16. URL: https://reliefweb.int/sites/reliefweb.int/files/resources/WHO%20PHL%20SitRep1_Dengue%20Outbreak_16Jul2019_original.pdf [accessed 2019-07-31]
18. Situation report 4: dengue outbreak. Representative Office for the Philippines, World Health Organization. 2019 Aug 13. URL: https://reliefweb.int/sites/reliefweb.int/files/resources/SitRep4_Dengue%20Outbreak_13Aug2019.pdf [accessed 2019-07-31]
19. Situation report 6. Representative Office for the Philippines, World Health Organization. 2019 Sep 02. URL: https://reliefweb.int/sites/reliefweb.int/files/resources/SitRep6_Dengue%20Outbreak_2September2019.pdf [accessed 2019-08-31]
20. Monthly dengue cases report No. 8, January 1 to August 31, 2019. Epidemiology Bureau, Public Health Surveillance Division. 2019. URL: <https://www.doh.gov.ph/sites/default/files/statistics/2019%20Dengue%20Monthly%20Report%20No.%208.pdf> [accessed 2019-10-23]
21. Sanchez M. Smartphone penetration as share of population Philippines 2017-2025. Statista. URL: <https://www.statista.com/statistics/625427/smartphone-user-penetration-in-philippines/> [accessed 2019-02-18]
22. Sanchez M. Smartphone users in the Philippines 2015-2025. Statista. 2020. URL: <https://www.statista.com/statistics/467186/forecast-of-smartphone-users-in-the-philippines/> [accessed 2021-02-18]
23. Silver L, Huang C. Emerging economies, smartphone and social media users have broader social networks. Washington: Pew Internet and American Life Project; 2019 Aug 22. URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2019/08/PI-PG_2019-08-22_social-networks-emerging-economies_FINAL.pdf [accessed 2021-02-18]
24. Do TTT, Le MD, Van Nguyen T, Tran BX, Le HT, Nguyen HD, et al. Receptiveness and preferences of health-related smartphone applications among Vietnamese youth and young adults. *BMC Public Health* 2018 Jun 19;18(1):764 [FREE Full text] [doi: [10.1186/s12889-018-5641-0](https://doi.org/10.1186/s12889-018-5641-0)] [Medline: [29921258](https://pubmed.ncbi.nlm.nih.gov/29921258/)]

25. National Capital Region 2015 Census of Population. Philippine Statistics Office. URL: <http://rssoncr.psa.gov.ph/sites/default/files/QC.pdf> [accessed 2021-02-18]
26. 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](https://doi.org/10.1001/jama.2013.281053)] [Medline: [24141714](https://pubmed.ncbi.nlm.nih.gov/24141714/)]
27. European Medicines Agency. ICH Topic E6 (R1) Guideline for good clinical practice step 5 note for guidance on good clinical practice (CPMP/ICH/135/95). URL: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e6-r1-guideline-good-clinical-practice_en.pdf [accessed 2021-02-18]
28. World Health Organization. Fact sheets: dengue and severe dengue. 2020 Jun 23. URL: <http://www.who.int/mediacentre/factsheets/fs117/en> [accessed 2021-02-18]
29. Search results: dengue. Department of Health. URL: <https://doh.gov.ph/search/node/Dengue> [accessed 2021-02-23]
30. Kumaran E, Doum D, Keo V, Sokha L, Sam B, Chan V, et al. Dengue knowledge, attitudes and practices and their impact on community-based vector control in rural Cambodia. PLoS Negl Trop Dis 2018 Feb;12(2):e0006268 [FREE Full text] [doi: [10.1371/journal.pntd.0006268](https://doi.org/10.1371/journal.pntd.0006268)] [Medline: [29451879](https://pubmed.ncbi.nlm.nih.gov/29451879/)]
31. Parks W, Lloyd L. Planning social mobilization and communication for dengue fever prevention and control: a step-by-step guide. Geneva: World Health Organization; 2004. URL: http://www.who.int/immunization/hpv/communicate/planning_social_mobilization_and_communication_for_dengue_fever_prevention_and_control_who_cds_wmc_2004.pdf [accessed 2021-02-18]
32. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
33. Delikostidis I. Methods and Techniques for Field-Based Usability Testing of Mobile Geo-Applications [Thesis]. Enschede: International Institute for Geo-Information Science & Earth Observation (ITC); 2007.
34. Khan S, Timmings C, Moore JE, Marquez C, Pyka K, Gheihman G, et al. The development of an online decision support tool for organizational readiness for change. Implementation Sci 2014 May 10;9(1):56. [doi: [10.1186/1748-5908-9-56](https://doi.org/10.1186/1748-5908-9-56)]
35. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inf 2019 Jun;126:95-104. [doi: [10.1016/j.ijmedinf.2019.03.018](https://doi.org/10.1016/j.ijmedinf.2019.03.018)]
36. Albert J, Santos A, Vizmanos J. Profile and determinants of the middle-income class in the Philippines. Quezon City: Philippine Institute for Development Studies; 2018. URL: <https://pidswebs.pids.gov.ph/CDN/PUBLICATIONS/pidsdps1820.pdf> [accessed 2019-10-25]
37. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016;4(2):e72 [FREE Full text] [doi: [10.2196/mhealth.5849](https://doi.org/10.2196/mhealth.5849)] [Medline: [27287964](https://pubmed.ncbi.nlm.nih.gov/27287964/)]
38. Urquhart C. Systematic reviewing, meta-analysis and meta-synthesis for evidence-based library and information science. 2010. URL: <http://informationr.net/ir/15-3/colis7/colis708.html> [accessed 2021-02-18]
39. Campbell R, Pound P, Pope C, Britten N, Pill R, Morgan M, et al. Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. Soc Sci Med 2003 Feb;56(4):671-684. [Medline: [12560003](https://pubmed.ncbi.nlm.nih.gov/12560003/)]
40. Morgan DL. Qualitative content analysis: a guide to paths not taken. Qual Health Res 1993 Feb;3(1):112-121. [Medline: [8457790](https://pubmed.ncbi.nlm.nih.gov/8457790/)]
41. Rodríguez S, Sanz AM, Llano G, Navarro A, Parra-Lara LG, Krystosik AR, et al. Acceptability and usability of a mobile application for management and surveillance of vector-borne diseases in Colombia: an implementation study. PLoS One 2020;15(5):e0233269 [FREE Full text] [doi: [10.1371/journal.pone.0233269](https://doi.org/10.1371/journal.pone.0233269)] [Medline: [32469894](https://pubmed.ncbi.nlm.nih.gov/32469894/)]
42. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013;1(1):e3 [FREE Full text] [doi: [10.2196/games.3139](https://doi.org/10.2196/games.3139)] [Medline: [25658754](https://pubmed.ncbi.nlm.nih.gov/25658754/)]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials

DOH: Department of Health

EBS: event-based surveillance

IBS: indicator-based surveillance

JSPS: Japan Society for the Promotion of Science

MARS: Mobile Application Rating Scale

mHealth: mobile health

uMARS: Mobile Application Rating Scale–User Version

WHO: World Health Organization

Edited by Y Khader; submitted 01.04.20; peer-reviewed by A Sudaryanto, D Leightley; comments to author 20.04.20; revised version received 18.05.20; accepted 03.02.21; published 01.03.21.

Please cite as:

Herbuela VRDM, Karita T, Carvajal TM, Ho HT, Lorena JMO, Regalado RA, Sobrepeña GD, Watanabe K
Early Detection of Dengue Fever Outbreaks Using a Surveillance App (Mozzify): Cross-sectional Mixed Methods Usability Study
JMIR Public Health Surveill 2021;7(3):e19034

URL: <https://publichealth.jmir.org/2021/3/e19034>

doi: [10.2196/19034](https://doi.org/10.2196/19034)

PMID: [33646128](https://pubmed.ncbi.nlm.nih.gov/33646128/)

©Von Ralph Dane Marquez Herbuela, Tomonori Karita, Thaddeus Marzo Carvajal, Howell Tsai Ho, John Michael Olea Lorena, Rachele Arce Regalado, Girly Dirilo Sobrepeña, Kozo Watanabe. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 01.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Mobile App and Dashboard for Early Detection of Infectious Disease Outbreaks: Development Study

Euijoon Ahn^{1,2}, PhD; Na Liu³, PhD; Tej Parekh^{1,4}, MInfoTech; Ronak Patel¹, MInfoTech; Tanya Baldacchino², BN; Tracy Mullavey², BN; Amanda Robinson⁵, BN; Jinman Kim^{1,2}, PhD

¹School of Computer Science, The University of Sydney, Darlington, Australia

²Telehealth Technology Centre, Nepean Hospital, Nepean Blue Mountains Local Health District, Kingswood, Australia

³The University Sydney Business School, Darlington, Australia

⁴Tej Consultancy, Sydney, Australia

⁵Public Health Unit, Nepean Hospital, Nepean Blue Mountains Local Health District, Kingswood, Australia

Corresponding Author:

Jinman Kim, PhD

School of Computer Science

The University of Sydney

Rm 340, Level 3, J12, 1 Cleveland St

Darlington, 2006

Australia

Phone: 61 290369804

Email: jinman.kim@sydney.edu.au

Abstract

Background: Outbreaks of infectious diseases pose great risks, including hospitalization and death, to public health. Therefore, improving the management of outbreaks is important for preventing widespread infection and mitigating associated risks. Mobile health technology provides new capabilities that can help better capture, monitor, and manage infectious diseases, including the ability to quickly identify potential outbreaks.

Objective: This study aims to develop a new infectious disease surveillance (IDS) system comprising a mobile app for accurate data capturing and dashboard for better health care planning and decision making.

Methods: We developed the IDS system using a 2-pronged approach: a literature review on available and similar disease surveillance systems to understand the fundamental requirements and face-to-face interviews to collect specific user requirements from the local public health unit team at the Nepean Hospital, Nepean Blue Mountains Local Health District, New South Wales, Australia.

Results: We identified 3 fundamental requirements when designing an electronic IDS system, which are the ability to capture and report outbreak data accurately, completely, and in a timely fashion. We then developed our IDS system based on the workflow, scope, and specific requirements of the public health unit team. We also produced detailed design and requirement guidelines. In our system, the outbreak data are captured and sent from anywhere using a mobile device or a desktop PC (web interface). The data are processed using a client-server architecture and, therefore, can be analyzed in real time. Our dashboard is designed to provide a daily, weekly, monthly, and historical summary of outbreak information, which can be potentially used to develop a future intervention plan. Specific information about certain outbreaks can also be visualized interactively to understand the unique characteristics of emerging infectious diseases.

Conclusions: We demonstrated the design and development of our IDS system. We suggest that the use of a mobile app and dashboard will simplify the overall data collection, reporting, and analysis processes, thereby improving the public health responses and providing accurate registration of outbreak information. Accurate data reporting and collection are a major step forward in creating a better intervention plan for future outbreaks of infectious diseases.

(JMIR Public Health Surveill 2021;7(3):e14837) doi:[10.2196/14837](https://doi.org/10.2196/14837)

KEYWORDS

public health; infectious disease reporting; mobile app; disease notification; mobile phone

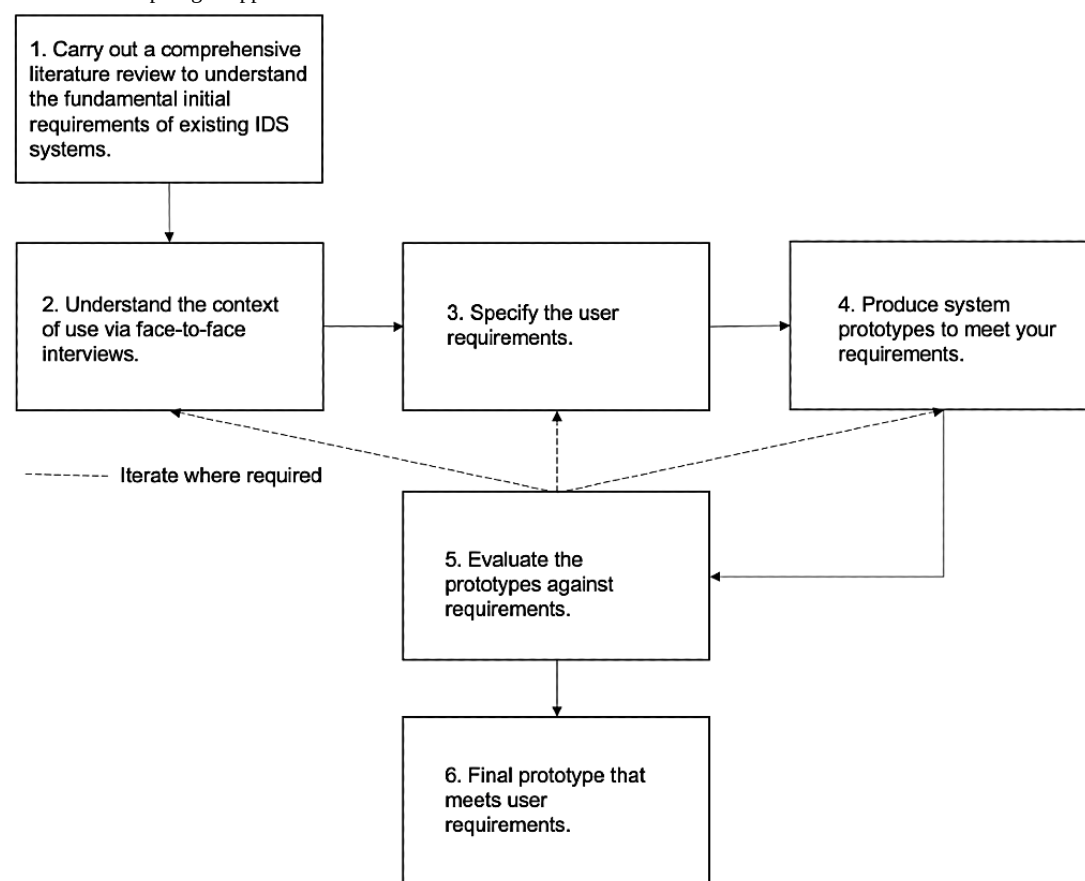
Introduction

Outbreaks of infectious diseases, such as gastroenteritis, measles, influenza, and COVID-19, pose great risks, including hospitalization and death, to public health. Outbreaks have been directly associated with increased utilization of health care resources and negative social impacts [1,2]. Therefore, improving the management of these outbreaks is important for preventing widespread infection and mitigating associated risks. A core strategy to minimize the risk posed by widespread infectious diseases is the development of a surveillance system to quickly identify potential outbreaks as early as possible accurately, such that an effective intervention plan can be implemented before the substantial costs associated with the spread of diseases have been incurred.

Many infectious disease surveillance (IDS) systems have been developed to quickly collect, interpret, and respond to outbreaks. Organizations such as the US Centers for Disease Control and Prevention [3], Germany's SurvNet@RKI [4], and Australia's Health Victoria [5] have established digital and web-based disease notification systems. These systems are beneficial for timely reporting of disease outbreaks between distributed sites (often dispersed geographical locations) such as hospitals or community facilities (eg, aged care and childcare centers) and local public health units (PHUs). However, each system has been developed to satisfy particular needs (eg, specific to certain diseases [6] or patients only) and is thus limited in its applicability to a wider domain. Meanwhile, local health authorities in many countries still rely on conventional manual paper forms and communicate them via fax or email (eg, New South Wales [NSW] Ministry of Health [7]). Reporting in such a way between geographically dispersed sites is susceptible to lost, incomplete, and incorrect information [8]. This is problematic because of the increased public health costs associated with infectious diseases.

Mobile health (mHealth) technology is an emerging concept that uses mobile devices coupled with wireless technology for health care purposes [9-11]. It has gained great popularity in recent years and has been used in various clinical environments. Communication between clinicians and hospitals has been improved greatly using *push alert* and automatic notification (eg, paging) features [12]. In addition, a number of innovative mobile apps have been developed, such as remote patient monitoring systems [13] and orthopedic decision support apps [14]. With continued innovation in mHealth technologies, dashboards have also been introduced to summarize and integrate key information into a visual display for better decision making [15-17]. Clinical dashboards are generally designed to improve the quality of care by measuring performance metrics such as the doctor-to-bed ratio, hospital infection rates, and mortality ratios [15]. These data are then shared among clinicians in real time to react to patient needs as well as to efficiently capture and understand underlying patterns and trends.

Limited research has been conducted to investigate the use of mHealth technology coupled with dashboards to improve the management of the outbreak of infectious diseases. In this paper, we propose a new IDS system that accurately captures outbreak data using a mobile app and analyzes them by visualizing interactively using a dashboard. We developed our system using a 2-pronged approach; we firstly carried out a comprehensive literature review to understand the initial fundamental requirements of existing IDS systems and then conducted iterative face-to-face interviews to understand the context of use and collect user requirements. A flowchart of our 2-pronged approach is shown in Figure 1. We suggest that the use of mHealth technology and dashboard can improve the management of outbreaks between distributed sites and identify new opportunities for better health outcomes and planning.

Figure 1. Flowchart of our 2-pronged approach. IDS: infectious disease surveillance.

Methods

Study Site and Setting

This study was conducted at Nepean Blue Mountains Local Health District (NBMLHD), a Nepean hospital in NSW, Australia. NBMLHD covers both urban and semirural areas, covering approximately 9179 km². The estimated resident population of NBMLHD in 2017 was 381,704 and is projected to increase by 30% by 2036. The PHU at NBMLHD identifies and prevents or minimizes public health risks to the communities by working with local general practitioners, community health workers, schools and childcare centers, and aged care facilities.

All health facilities (eg, childcare centers and aged care facilities) in NBMLHD (and also across the NSW state) currently use a paper-based form to report outbreaks to local PHU teams, which is error prone and time consuming. The increasing populations along with large covering areas of the district will introduce new and unique challenges in the management of infectious disease outbreaks.

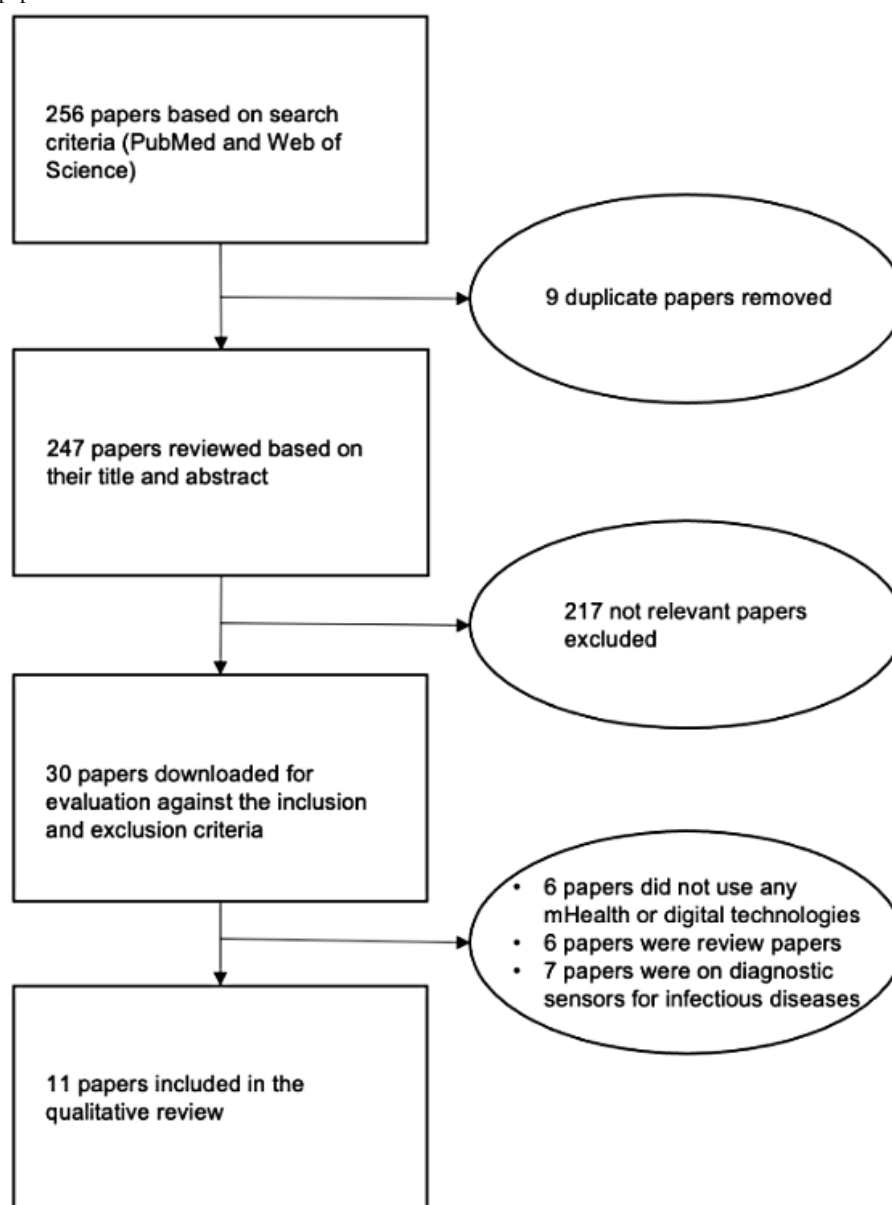
Identification of Fundamental Requirements of the IDS System Through a Literature Review

We conducted a literature review on existing IDS systems to analyze and define the overall scopes and fundamental

requirements of our system, following the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-analyses Protocols) guidelines [18]. We also analyzed the design trends and incorporated the relevant technologies into our system, with a focus on studies that used mobile devices and dashboard systems.

We conducted a systematic search of published papers from the following scientific databases: PubMed and Web of Science. We searched the papers published in the last 15 years from 2004, concerning infectious disease reporting and analysis using mHealth technologies. The search terms used were “infectious” AND “disease” AND “surveillance” AND “smartphone” OR “smartphone application” OR “mobile app” OR “mobile phones” OR “dashboard” OR “mHealth.” Figure 2 provides an overview of the study selection process.

The inclusion criteria were as follows: (1) infectious disease reporting systems must have been designed to use an electronic device (mobile phone, tablet, or personal computer) and (2) reporting was in the form of a specific app or website. We excluded papers that described the use of mobile-based systems for noninfectious diseases. Review papers were excluded. Only papers published in English were included.

Figure 2. Flowchart of paper selections. mHealth: mobile health.

Identification of Specific User Requirements of the IDS System Through Interviews

Face-to-face interviews were conducted to collect the user requirements of our IDS system. Multiple interviews were carried out iteratively until we had a comprehensive understanding of user requirements. The participants included public health nurses from the local PHU team and telehealth providers from the NBMLHD. The interviews were conducted by experienced researchers, including project coordinators and software developers. First, we identified different types of end users and their corresponding roles. We then investigated and defined specific workflows and use case scenarios of our system. Our study did not require any patient involvement, as ethical approval or informed consent was not needed.

User Interface Design and Agile Development

Agile development uses an interactive system framework that aims to provide a good user experience. The initial user interfaces of the system were designed based on the user

requirements and the understanding of the context of use. The software developers and health care team were in constant communication, discussing and enhancing the contents and user interfaces of the system. We developed a functional mobile app and dashboard prototype that was used to collect further feedback from users.

Results

Phase 1: Literature Review of the State of the Art

Our initial search identified 256 papers using the search criteria. We removed 9 duplicate papers and screened the remaining 247 papers based on their titles and abstracts. We then downloaded 30 relevant papers and examined them based on the inclusion and exclusion criteria. A total of 11 relevant papers were selected for qualitative review.

The papers described new digital technologies for better detection of, reporting on, and response to the outbreak of the diseases. A summary of our literature review of the existing

IDS systems is shown in [Table 1](#). Through analyzing the common and central aims of selected studies, we have identified 3 fundamental requirements for infectious disease reporting systems, which is the ability to capture and report outbreak data accurately, completely, and in a timely fashion [19-25]. Fundamental requirements were also used as the basis for identifying subsequent specific user requirements.

The ability to accurately capture and report data, such as symptoms and age, is important to create an optimal disease-specific intervention plan [21,26,27]. It is also crucial to identify and understand the unique characteristics of certain diseases; some diseases have a higher number of reported cases, but only a few patients have serious illnesses (eg, measles). Another electronic reporting system was used in Ireland to reduce the underreporting of notifiable diseases [28]. Underreporting has been associated with a number of factors, including clinicians' lack of time and motivation and the use of complicated and inconsistent paper forms. The electronic system, which was consistent and user friendly, created an environment where clinicians could better engage in reporting and understand their notification obligations [21,25,29]. An

interactive data visualization technique (eg, web dashboard) was also used to understand the unique characteristics of certain infectious diseases, thereby improving the accuracy and completeness of outbreak reporting [27]. Kamadjeu et al [30] developed an electronic dashboard for the poliovirus outbreak response in Somalia. The dashboard captured complete outbreak data and was able to improve the daily analysis and sharing of outbreak information between different stakeholders, such as outbreak response managers and immunization partners located in various areas.

The internet-based reporting system OSIRIS was introduced in the Netherlands to reduce delays in receiving outbreak data and improve the completeness of the data [20]. The system was able to reduce the delay from 10 days to 1 day and had a higher completeness of data with over 10% improvements compared with traditional paper-based approaches. Similarly, Angues et al [21] used a mobile-based data collection platform and real-time cloud-based storage technology to reduce delayed recognition and intervention. Fahnrich et al [23] also developed a mobile-based system that can efficiently and promptly control Ebola in West Africa.

Table 1. The descriptions of existing infectious disease surveillance systems found from our literature review.

Study	Infectious disease addressed	Description of main aims or features	Main outcomes reported	Fundamental requirements
Ward et al 2005 [20]	Meningitis, malaria, hepatitis A, B and C, etc	To improve timeliness and completeness of surveillance data on infectious diseases reported from regional to national level	The primary outcome measure was whether the internet-based reporting system improves timeliness and completeness in receiving the outbreak data in comparison with conventional paper-based forms. The internet-based system was able to reduce the delay and had a higher completeness of data.	Timeliness and completeness
Angues et al 2018 [21]	Nodding syndrome	To reduce delays in receiving outbreak data from rural regions in Uganda	The primary outcome measure was whether the real-time mobile-based system can improve the timeliness in reporting outbreak data. The system was able to accurately capture the data and reduce the delayed recognition and intervention. The system also created an environment where clinicians can better engage with the reporting and understand their notification obligations.	Accuracy, completeness, and timeliness
Severi et al 2014 [22]	Food poisoning, chickenpox, Q fever, etc	To improve sensitivity and timeliness of infectious disease data	An event-based surveillance system was developed. The reports from the system were timely and had complete information.	Timeliness and completeness
Fahnrich et al 2015 [23]	Ebola	To achieve efficient and timely control of communicable Ebola in West Africa	A standard digital Ebola outbreak management system was developed. The system can handle big data in combination with its mobile interface for bidirectional information exchange between different users.	Timeliness
Waarbeek et al 2011 [24]	Hepatitis, A, B, and C, etc	To design a real-time IDS ^a system for cross-border collaboration	The collaborative disease protocols and risk assessments were embedded into the IDS system that provides real-time surveillance and geographical information.	Timeliness
Diwan et al 2015 [25]	Symptoms such as fever, cough, body ache, headache, and runny nose	To achieve efficient and timely collection of data in a resource-limited setting in rural India	The study demonstrated that a mobile-based system allowed to reduce the underreporting of notifiable diseases.	Accuracy and timeliness
Blanchon [26]	Acute diarrhea and influenza-like illness	To identify gaps in the French Sentinelles IDS system	The main finding is that it is important to maintain a high level of motivation and participation among physicians to improve the completeness and accuracy of outbreak data.	Accuracy and completeness
Yan et al 2013 [27]	Symptoms such as cough, sore throat, and fever	To design and pilot implementation of a syndromic surveillance system in rural China	A new web-based syndromic surveillance system was developed, which allows accurate and immediate reporting of syndromic data and automated detection of abnormal disease clusters using interactive visualization (ie, dashboard).	Accuracy, completeness, and timeliness
Brabazon et al 2015 [28]	Viral meningitis, viral encephalitis, bacterial meningitis, malaria, etc	To identify if there was significant underreporting of hospitalized notifiable infectious diseases in Ireland	A new computerized IDS system was developed, which reduced underreporting of outbreak data.	Accuracy and completeness
Karimuribo et al 2017 [29]	Any symptoms	To design a participatory disease surveillance system for community-based health reporters in East and Southern Africa	A mobile-based disease surveillance system was developed, which improved the engagement of people in their own communities in the detection of infectious human and animal disease threats.	Completeness

Study	Infectious disease addressed	Description of main aims or features	Main outcomes reported	Fundamental requirements
Kamadjeu et al 2017 [30]	Poliovirus	To design and implement an electronic dashboard for polio outbreak response management	The Somalia Polio Room Dashboard was developed; it integrates various outbreak and surveillance response data sources into rich graphical interfaces. The dashboard was able to reduce the need for daily analysis and sharing of outbreak information between different stakeholders.	Accuracy

^aIDS: infectious disease surveillance.

Phase 2: User Research and User Requirements

The results of the literature reviews allowed us to create a list of appropriate questions to be asked for subsequent face-to-face interviews. Some of the questions were designed to derive answers that could potentially help improve the ability to capture and report the outbreak data accurately, completely, and in a

timely fashion. We conducted several interviews with 2 public health nurses and 1 clinical doctor from the PHU team to understand the overall workflow of how outbreak data were communicated between community facilities and the PHU team. The list of interview questions and their corresponding answers are presented in Table 2. Testimony from users and experts was used to help developers better understand user requirements.

Table 2. The list of interview questions and corresponding answers to understand the context of use.

Question ID number	Question themes	Interview questions	Answers (user requirements)	Sources
1	Data security and user access	Can anyone use the mobile app?	No; only registered clinicians should use the mobile app and the registration should be limited.	This was a common answer among all interviewees.
2	Data capturing	What are the required fields when entering outbreak data for a patient?	Patient name, date of birth, sex, vaccination details, symptoms and laboratory testing details, date of onset of symptoms, and date of last symptoms are required.	One of the public health nurses said some of the data such as laboratory testing details might not be available at the initial stage of the outbreak.
3	Data capturing (accuracy and completeness)	What are the most important fields (ie, mandatory fields) that need to be captured?	Patient names, case number, symptoms, symptoms onset date, sex, date of birth, disease type, and occupancy need to be at least captured.	This was a common answer among all interviewees.
4	Data update and review (timeliness)	How do you update and track the details of certain outbreak data?	Existing patient outbreak data must be reviewed and monitored to identify any changes in symptoms.	This was a common answer among all interviewees.
5	Data update and review (timeliness)	How often do you need to check the status of the outbreak or patients?	On the outbreak of infectious disease in a particular facility, the PHU ^a nurses need to contact the facility every day to check the status of the outbreak and conditions of patients.	This was a common answer among all interviewees.
6	Data capturing (accuracy and completeness)	What are the common errors when entering outbreak data?	The symptoms of patients and symptoms onset date are often inconsistent. This is because there are multiple nursing staff at the facilities recording data every day, and there can be errors during handover.	This was a common answer among all interviewees.
7	Data report	What are the key fields that allow you to search and identify right patient?	Facility name, age, and patient name.	This was a common answer among all interviewees.
8	Data report	What are the fields of patients that you would like to see on the web dashboard? Would those fields be able to help you identify right information?	We felt that the most important field to see would be patient name, facility name, hospitalized (yes or no), deceased (yes or no), treatment or prophylaxis type, and pathogen (influenza or gastro). We thought these would be the most important to see first as long as full patient details can be provided from expanded page view if needed.	This was a common answer among all interviewees.
9	Data report	What are the most important reports that you would like to see when you log in to the web dashboard?	It is important to view the weekly total number of patients reported under each infectious disease for the past 6 months.	This was specifically requested by the clinical doctor.

^aPHU: public health unit.

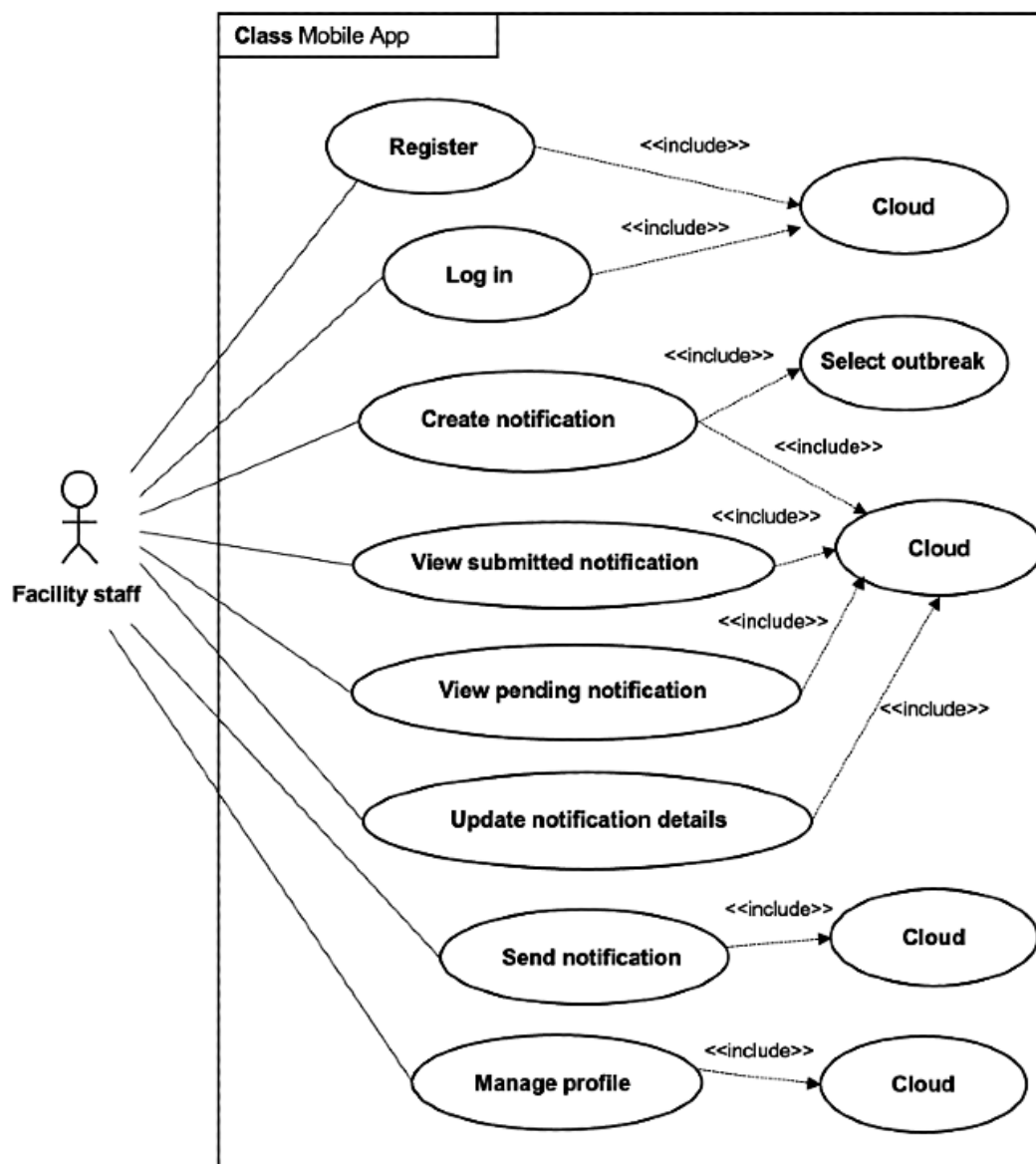
Phase 3: Extraction of Function Requirements

The results of the user interviews were then used to formulate functional requirements. A list of functional requirements is

presented in Table 3. The use case diagram of our mobile app is shown in Figure 3. The diagram represents the main functionalities of the mobile app: data capturing and notification.

Table 3. The list of functional requirements formulated based on the outcomes of interviews.

Question ID number (consistent with Table 2)	Function type	Functional requirements
1	User verification for new registration	Send a request to the web server to verify the verification passcode entered by the user. If the passcode is validated, then the user can be registered; otherwise, display an error message.
2	Data capturing	Patient name, date of birth, sex, vaccination details, symptoms and laboratory testing details, date of onset of symptoms, and date of last symptoms. Validate the fields against the required fields. Show an error message if the user has not entered the mandatory field.
3	Data mandatory fields	Create mandatory fields on patient names, case number, symptoms, symptoms onset date, sex, date of birth, disease type, and occupancy.
4	Data update	Existing patient information can be easily retrieved, and users should be able to easily submit a new report based on previous outbreak data.
5	User alert for regular data submission	Create a push alert to inform users to regularly submit outbreak data for individual patient.
6	User error reducing functions	Most recently reported data can be prefilled, and the symptom onset date should not be editable once submitted.
7	User convenience functions	Create search bars for each field so that users can easily retrieve relevant information.
8	User report	Create an overview web page of selected infectious diseases and list all the patients in the table with the columns showing patient name, facility name, hospitalized (yes or no), deceased (yes or no), and treatment or prophylaxis type. Create an individual page to show all the details of the case reported when clicked by users.
9	User report	Create a table that displays the number of cases reported for each infectious disease during the outbreak; create landing pages that contain a 2-dimensional graph if there has been increase or decrease in the number of cases reported.

Figure 3. Use case diagram for the system (mobile app).

Phase 4: Prototype Building and Iteration

On the basis of communication between clinicians and development teams, we developed an app and dashboard prototype. The app and dashboard were presented back to clinicians for further feedback, and we identified new user requirements that were not found in phase 1. A list of new

requirements identified is shown in [Table 4](#). We updated and improved the prototype system iteratively based on the feedback from 3 clinicians from the PHU team and 2 clinicians from the telehealth team over the period between January 2017 and December 2017. A description of the main functionalities of our prototype is presented in [Table 5](#).

Table 4. The list of changes to functional requirements based on the exposure of prototypes to clinicians.

Question ID number	New functional requirements
8	- Create a separate report designed for individual facility; summary of current active outbreaks within a facility. The data can also be visualized interactively, showing historical or specific details of an outbreak. - Create a summary of current active outbreaks across the whole local health district. Specific outbreak details of certain facilities can also be visualized.
9	- The data need to be better presented by highlighting important metrics such as the number of reported cases and patient details, including their symptoms and vaccinations. - An automatic email notification to the PHU^a team is required when there are any new reported outbreaks.

^aPHU: public health unit.

Table 5. The description of the main functionalities of our prototype.

Type	Main functionalities
Function 1: Data capturing	• Outbreak details need to be entered, edited, and submitted using a mobile app or desktop PC (web interface) without geographical restrictions (within registered facilities managed by local PHU^a). • Clinical information collected includes patient name, date of birth, sex, vaccination details, symptoms and laboratory testing details, date of onset of symptoms, and date of last symptoms. • Enforce correct and mandatory data entry (eg, facility care type and symptoms onset date).
Function 2: Outbreak notification	• The data need to be sent and displayed in real time, highlighting important metrics such as the number of reported cases and patient details including their symptoms and vaccinations. An automatic email notification will be sent to the local PHU team.
Function 3: Data report for facilities	• Summary of current active outbreaks within a facility. The data can also be visualized interactively, showing historical or specific details of an outbreak.
Function 4: Data report for PHU	• Summary of current active outbreaks across the whole local health district. Specific outbreak details of certain facilities can be also visualized.
Function 5: Data storage	• The data need to be processed and stored in a secure server.

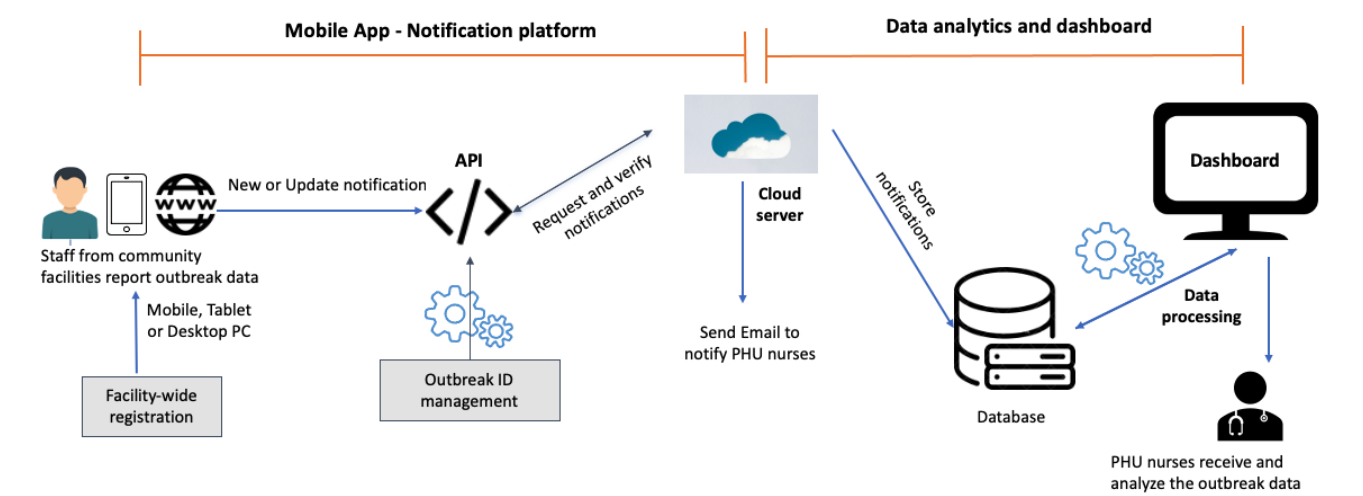
^aPHU: public health unit.

System Overview

Figure 4 shows an overview of our proposed system. Our system consists of a mobile app, a cloud server, and a dashboard. On the outbreak of an infectious disease in a facility, staff from the facility capture the patient data, including symptoms and

vaccination details, using a mobile device or desktop PC (web interface) and send them to the cloud server. The data are then processed and displayed on the dashboard in real time. The cloud server automatically notifies PHU nurses by sending an email. The PHU nurses then use the web dashboard to receive and analyze the outbreak data.

Figure 4. Overview of our proposed system. Data analysis and dashboard. API: application programming interface; PHU: public health unit.



Mobile App

The mobile app was developed for both Android and iOS devices using a cross-platform app development framework. We specifically used a development platform called cross-compilation [31], which automatically transforms source code into platform-specific apps. We also developed a mobile web app so that it could be run in a standard web browser (Figure 5).

Users can create a new outbreak detail and add or edit the details of existing information, as shown in Figure 6 (function 1). During an outbreak of infectious diseases, it is important to monitor the symptoms of patients daily; therefore, our app

allows us to review the history of previous patient details of outbreaks. Our app then sends them to a cloud server using an application programming interface gateway. To prevent unauthorized data submissions, the app is only accessible to users with a preregistered verification passcode, which is controlled by the PHU. All data are encrypted and sent using a JavaScript Object Notation (JSON) Web Token (JWT), which can only be validated at the cloud server (function 5). Automatic email notification is also sent to the PHU once a new outbreak is reported (function 2). The accessibility and user registration of the app are controlled by the cloud server. We used the local cloud server, and the data were encrypted using JWTs to enforce the security of communication.

Figure 5. A web version of our mobile app with a consistent user interface. Web technologies such as HTML 5, Cascading Style Sheet 3, and JavaScript were used.

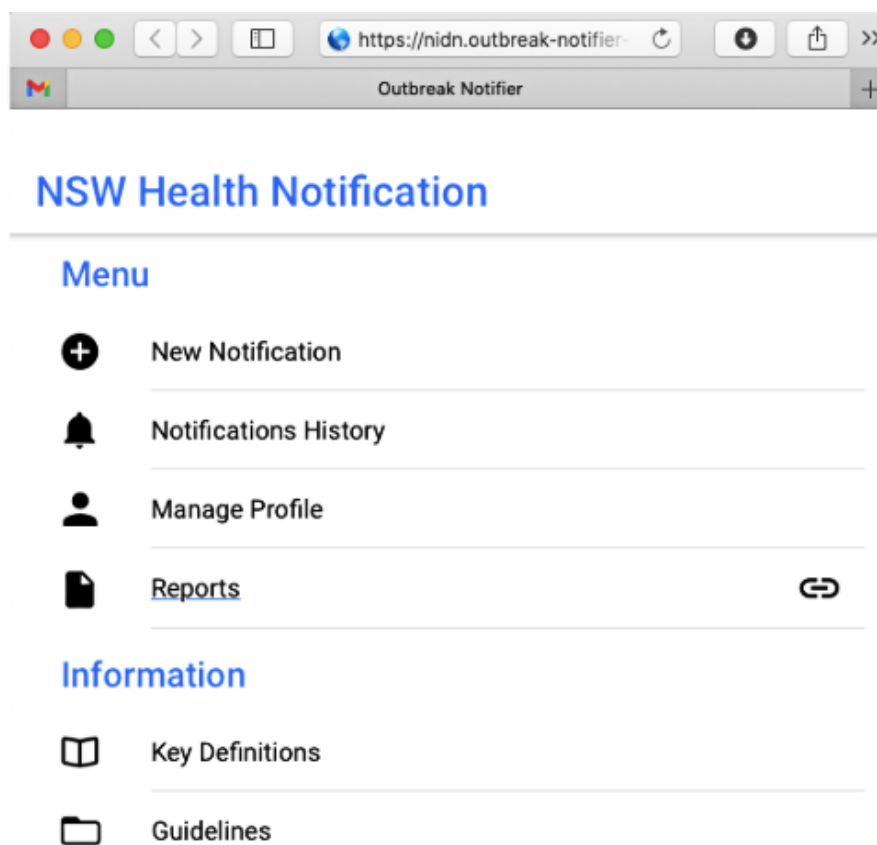
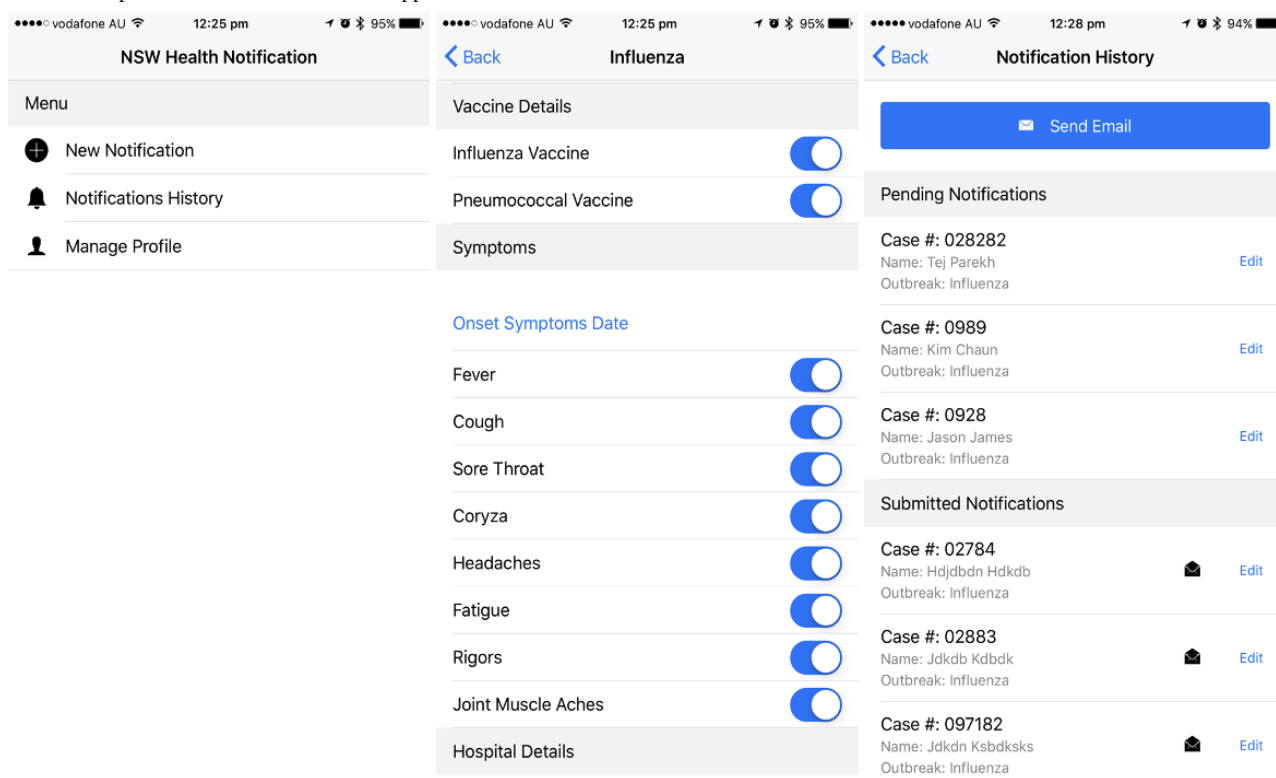


Figure 6. Example screenshots of our mobile app.



Dashboard

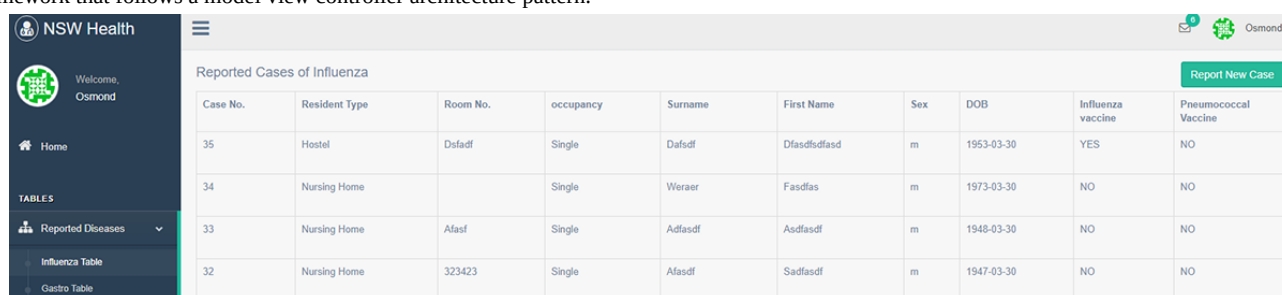
We developed a dashboard using an open-source Hypertext Preprocessor (PHP) Laravel framework. It provides rich built-in features, such as authentication, authorization, and a database management system. The use of the Laravel framework allowed us to iteratively develop our system and efficiently refine it based on the feedback received from the PHU team.

The dashboard displays outbreak data sent via our mobile app in real time (function 2). [Figure 7](#) displays a quick summary of the key information (eg, facility name, room number, patient details, and vaccination details). [Figure 8](#) shows a summary of

current active outbreaks, providing daily, weekly, and monthly summaries of reported cases (functions 3 and 4). To present data that are visually intuitive, key information is color coded and graphical charts are also used. More specific and historical data of an outbreak can also be retrieved. [Figure 9](#) shows the trend of the selected outbreak over multiple years.

To make our system more applicable to wider domains (eg, different states or countries), data attributes such as different types of infectious diseases, symptoms, and registered community facilities can be added, modified, and deleted through the dashboard.

Figure 7. Details of the reported cases of infectious diseases. The Hypertext Preprocessor (PHP) Laravel framework is a popular and standard development framework that follows a model view controller architecture pattern.



Case No.	Resident Type	Room No.	occupancy	Surname	First Name	Sex	DOB	Influenza vaccine	Pneumococcal Vaccine
35	Hostel	Dsfadf	Single	Dafsd	Dfadsdfdsd	m	1953-03-30	YES	NO
34	Nursing Home		Single	Weraer	Fasdfas	m	1973-03-30	NO	NO
33	Nursing Home	Alasf	Single	Adfasdf	Asdfasdf	m	1948-03-30	NO	NO
32	Nursing Home	323423	Single	Alasf	Sadfasdf	m	1947-03-30	NO	NO

Figure 8. Daily, weekly, and monthly summary of current active outbreaks.

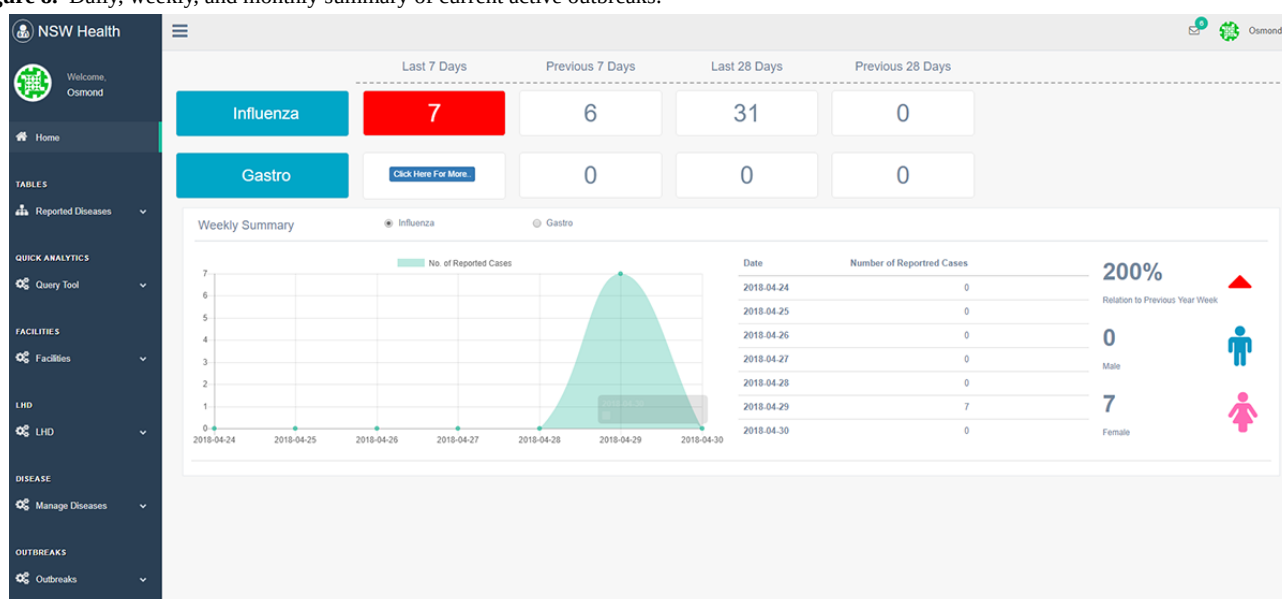
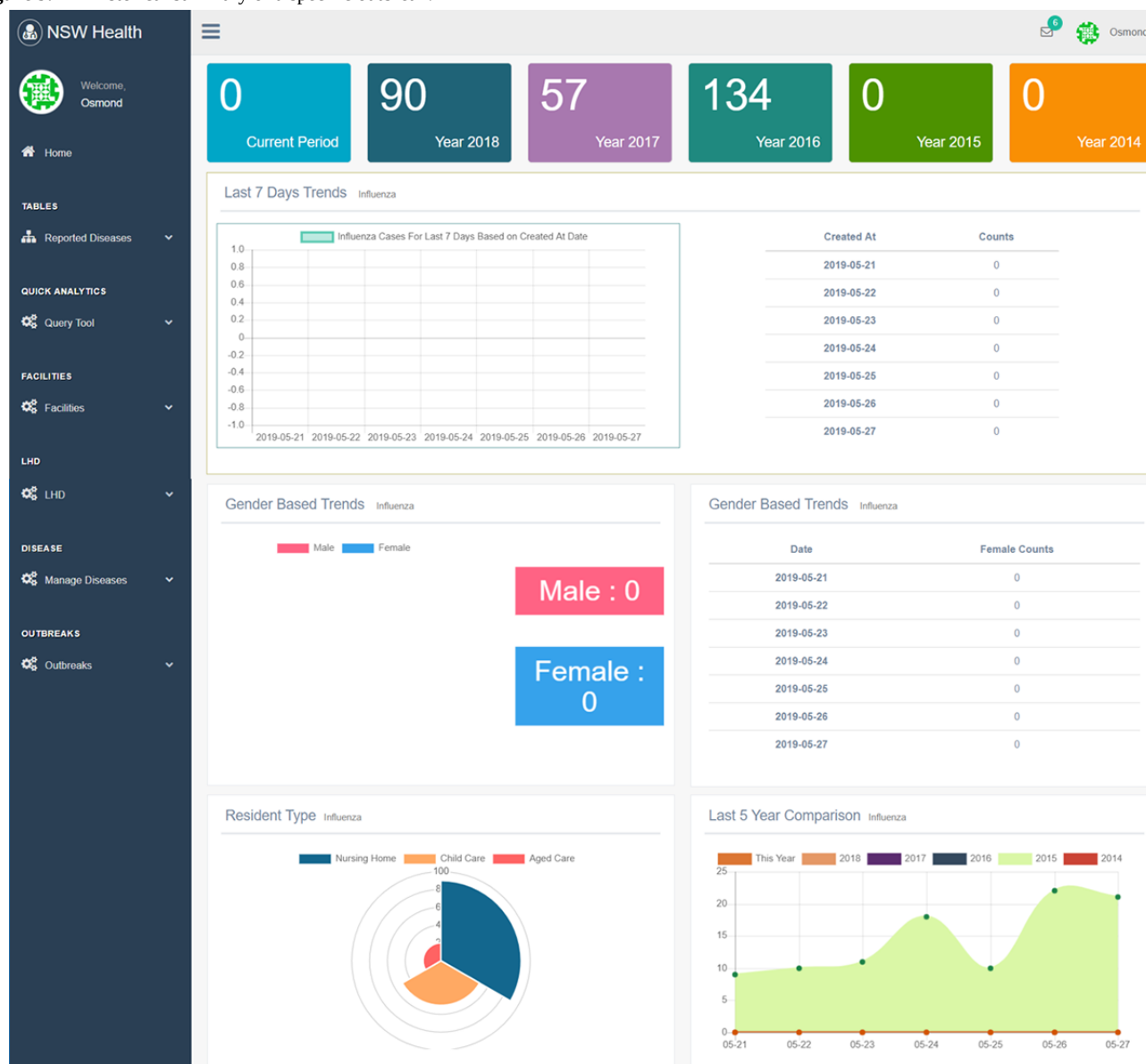


Figure 9. An historical summary of a specific outbreak.

Discussion

Principal Findings

This study collected user requirements using a 2-pronged approach: a literature review and face-to-face interviews. We identified 3 fundamental requirements when designing an electronic IDS system through the literature review and developed a new IDS system using iterative face-to-face interviews. We also presented the development and implementation details of a new IDS system using mHealth technologies. Our study focused on the development of a generic mobile app or dashboard framework for capturing and reporting outbreak data, which is different from conventional existing works that are limited to addressing specific infectious diseases such as influenza [6,32-34] or patient cohorts [35].

With the portability of using a mobile device, outbreak data can be captured and sent from anywhere, and the use of the cross-platform framework further allows users to record the data with different types of devices (eg, mobile, tablet, and PC).

The data are transferred using a client-server architecture and, therefore, can be analyzed in real time. Our dashboard is designed to provide a daily, weekly, monthly, and historical summary of outbreak information, highlighting important metrics such as the number of reported cases, male-to-female ratio, and number of reported specific symptoms. Specific information about certain outbreaks can also be visualized interactively, which can help understand the unique characteristics of emerging infectious diseases. Some of the key feedback we received from the PHU were as follows: (1) ability to enforce correct and mandatory data entry (eg, facility care type and symptom onset date) and (2) identification of key metrics that need to be visualized in the dashboard. The capability to customize data attributes is important for adaptively refining the system for general applications.

We suggest that the use of a mobile app and dashboard will simplify the overall data collection, reporting, and analysis processes, thereby improving the public health responses and providing accurate registration of outbreak information, which can potentially reduce associated morbidity and mortality [8].

This can also improve communication and engagement between clinicians and community facilities, providing more accessible information and education about infectious diseases.

We will evaluate the effectiveness of our proposed system by piloting the system at local community facilities, such as aged care and childcare centers. Preimplementation and postimplementation surveys will be conducted. Our system has a few limitations. Although our PHU team had extensive clinical experience, the system was developed and refined based on feedback from a single hospital; expanding across multiple hospitals will help generalize overall scopes.

Conclusions

We suggest that accurate data reporting and collection, enabled by our system, are a major step forward in the early detection

of a future outbreak of infectious diseases. With the availability of large-scale outbreak data, population-based analysis can be conducted to understand the overall characteristics of the outbreaks. This can potentially help prevent widespread infection and associated risks. Effective intervention plans can also be created based on the patterns identified and provided before the substantial avoidable spread of diseases has been incurred. Another key advantage of using large-scale data is the application of machine learning algorithms (eg, convolutional neural networks [36]) to predict infectious diseases. These predictive models can be derived and embedded in our notification system.

Acknowledgments

This research was supported by the Telehealth and Technology Center, Nepean Hospital, NBMLHD, NSW, Australia.

Authors' Contributions

EA, NL, TB, TM, AR, and JK were responsible for the conceptualization of this project. TP, RP, EA, and JK were responsible for developing the proposed system. All authors contributed to the preparation of this manuscript.

Conflicts of Interest

None declared.

References

1. Smith PW, Rusnak PG. Infection Prevention and Control in the Long-Term-Care Facility. *Infection Control and Hospital Epidemiology* 1997 Dec;18(12):831-849. [doi: [10.2307/30141342](https://doi.org/10.2307/30141342)]
2. Montoya A, Mody L. Common infections in nursing homes: a review of current issues and challenges. *Aging Health* 2011 Dec;7(6):889-899. [doi: [10.2217/ahe.11.80](https://doi.org/10.2217/ahe.11.80)]
3. Coates RJ, Jajosky RA, Stanbury M, Macdonald SC. Summary of Notifiable Noninfectious Conditions and Disease Outbreaks: Introduction to the Summary of Notifiable Noninfectious Conditions and Disease Outbreaks — United States. *MMWR Morb. Mortal. Wkly. Rep* 2015 Oct 23;62(54):1-4. [doi: [10.15585/mmwr.mm6254a1](https://doi.org/10.15585/mmwr.mm6254a1)]
4. Faensen D, Claus H, Benzler J, Ammon A, Pfoch T, Breuer T, et al. SurvNet@RKI--a multistate electronic reporting system for communicable diseases. *Euro Surveill* 2006;11(4):100-103. [Medline: [16645245](https://pubmed.ncbi.nlm.nih.gov/16645245/)]
5. Notifications procedures for infectious diseases. Victoria State Government. URL: <https://www2.health.vic.gov.au/public-health/infectious-diseases/notification-procedures> [accessed 2021-02-11]
6. Iribarren SJ, Schnall R, Stone PW, Carballo-Diéguez A. Smartphone Applications to Support Tuberculosis Prevention and Treatment: Review and Evaluation. *JMIR mHealth uHealth* 2016 May 13;4(2):e25. [doi: [10.2196/mhealth.5022](https://doi.org/10.2196/mhealth.5022)]
7. Disease notification - Infectious Diseases. NSW Ministry of Health. URL: <http://www.health.nsw.gov.au/Infectious/Pages/notification.aspx> [accessed 2021-02-11]
8. Revere D, Hills RH, Dixon BE, Gibson PJ, Grannis SJ. Notifiable condition reporting practices: implications for public health agency participation in a health information exchange. *BMC Public Health* 2017 Mar 11;17(1). [doi: [10.1186/s12889-017-4156-4](https://doi.org/10.1186/s12889-017-4156-4)]
9. Onodera R, Sengoku S. Innovation process of mHealth: An overview of FDA-approved mobile medical applications. *International Journal of Medical Informatics* 2018 Oct;118:65-71. [doi: [10.1016/j.ijmedinf.2018.07.004](https://doi.org/10.1016/j.ijmedinf.2018.07.004)]
10. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. *P T* 2014 May;39(5):356-364 [FREE Full text] [Medline: [24883008](https://pubmed.ncbi.nlm.nih.gov/24883008/)]
11. Al Ayubi SU, Pelletier A, Sunthara G, Gujral N, Mittal V, Bourgeois FC. A Mobile App Development Guideline for Hospital Settings: Maximizing the Use of and Minimizing the Security Risks of. *JMIR mHealth uHealth* 2016 May 11;4(2):e50. [doi: [10.2196/mhealth.4424](https://doi.org/10.2196/mhealth.4424)]
12. Wu RC, Morra D, Quan S, Lai S, Zanjani S, Abrams H, et al. The use of smartphones for clinical communication on internal medicine wards. *J. Hosp. Med* 2010 Aug 05;5(9):553-559. [doi: [10.1002/jhm.775](https://doi.org/10.1002/jhm.775)]

13. Liu N, Kim J, Jung Y, Arisy A, Nicdao MA, Mikaheal M, et al. Remote Monitoring Systems for Chronic Patients on Home Hemodialysis: Field Test of a Copresence-Enhanced Design. *JMIR Hum Factors* 2017 Aug 29;4(3):e21. [doi: [10.2196/humanfactors.7078](https://doi.org/10.2196/humanfactors.7078)]
14. Franko OI. Smartphone apps for orthopaedic surgeons. *Clin Orthop Relat Res* 2011 Jul;469(7):2042-2048 [FREE Full text] [doi: [10.1007/s11999-011-1904-0](https://doi.org/10.1007/s11999-011-1904-0)] [Medline: [21547414](https://pubmed.ncbi.nlm.nih.gov/21547414/)]
15. Dowding D, Randell R, Gardner P, Fitzpatrick G, Dykes P, Favela J, et al. Dashboards for improving patient care: Review of the literature. *International Journal of Medical Informatics* 2015 Feb;84(2):87-100. [doi: [10.1016/j.ijmedinf.2014.10.001](https://doi.org/10.1016/j.ijmedinf.2014.10.001)]
16. Wang Y, Kung L, Byrd TA. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change* 2018 Jan;126:3-13. [doi: [10.1016/j.techfore.2015.12.019](https://doi.org/10.1016/j.techfore.2015.12.019)]
17. Zhang X, Gallagher K, Goh S. Bi Application: Dashboards for healthcare. *AMCIS 2011 Proceedings - All Submissions* 2011:442 [FREE Full text]
18. Moher D. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Ann Intern Med* 2009 Aug 18;151(4):264. [doi: [10.7326/0003-4819-151-4-200908180-00135](https://doi.org/10.7326/0003-4819-151-4-200908180-00135)]
19. Milinovich GJ, Williams GM, Clements ACA, Hu W. Internet-based surveillance systems for monitoring emerging infectious diseases. *The Lancet Infectious Diseases* 2014 Feb;14(2):160-168. [doi: [10.1016/s1473-3099\(13\)70244-5](https://doi.org/10.1016/s1473-3099(13)70244-5)]
20. Ward M, Brandsema P, van Straten E, Bosman A. Electronic reporting improves timeliness and completeness of infectious disease notification, The Netherlands, 2003. *Euro Surveill* 2005 Jan;10(1):7-8. [doi: [10.2807/esm.10.01.00513-en](https://doi.org/10.2807/esm.10.01.00513-en)] [Medline: [29183539](https://pubmed.ncbi.nlm.nih.gov/29183539/)]
21. Valdes Angues R, Suits A, Palmer VS, Okot C, Okot RA, Atonywalo C, et al. A real-time medical cartography of epidemic disease (Nodding syndrome) using village-based lay mHealth reporters. *PLoS Negl Trop Dis* 2018 Jun 15;12(6):e0006588. [doi: [10.1371/journal.pntd.0006588](https://doi.org/10.1371/journal.pntd.0006588)]
22. Severi E, Kitching A, Crook P. Evaluation of the Health Protection Event-Based Surveillance for the London 2012 Olympic and Paralympic Games. *Euro Surveill* 2014 Jun 19;19(24):20832 [FREE Full text] [doi: [10.2807/1560-7917.es2014.19.24.20832](https://doi.org/10.2807/1560-7917.es2014.19.24.20832)] [Medline: [24970374](https://pubmed.ncbi.nlm.nih.gov/24970374/)]
23. Fähnrich C, Denecke K, Adeoye OO, Benzler J, Claus H, Kirchner G, et al. Surveillance and Outbreak Response Management System (SORMAS) to support the control of the Ebola virus disease outbreak in West Africa. *Euro Surveill* 2015 Mar 26;20(12):21071 [FREE Full text] [doi: [10.2807/1560-7917.es2015.20.12.21071](https://doi.org/10.2807/1560-7917.es2015.20.12.21071)] [Medline: [25846493](https://pubmed.ncbi.nlm.nih.gov/25846493/)]
24. ter Waarbeek H, Hoebe C, Freund H, Bochat V, Kara-Zaïr C. Strengthening infectious disease surveillance in a Dutch-German crossborder area using a real-time information exchange system. *J Bus Contin Emer Plan* 2011 Jun;5(2):173-184. [Medline: [21835755](https://pubmed.ncbi.nlm.nih.gov/21835755/)]
25. Diwan V, Agnihotri D, Hulth A. Collecting syndromic surveillance data by mobile phone in rural India: implementation and feasibility. *Global Health Action* 2015 Apr 02;8(1):26608. [doi: [10.3402/gha.v8.26608](https://doi.org/10.3402/gha.v8.26608)]
26. Blanchon T. Web - based sentinel provider surveillance network in France. *New Jersey: Wiley*; 2013:418-425.
27. Yan W, Palm L, Lu X, Nie S, Xu B, Zhao Q, et al. ISS-An Electronic Syndromic Surveillance System for Infectious Disease in Rural China. *PLoS ONE* 2013 Apr 23;8(4):e62749. [doi: [10.1371/journal.pone.0062749](https://doi.org/10.1371/journal.pone.0062749)]
28. Brabazon ED, Sheridan A, Finnegan P, Carton MW, Bedford D. Under-reporting of notifiable infectious disease hospitalizations: significant improvements in the Irish context. *Epidemiol. Infect* 2014 Jul 18;143(6):1166-1174. [doi: [10.1017/s0950268814001733](https://doi.org/10.1017/s0950268814001733)]
29. Karimuribo ED, Mutagahywa E, Sindato C, Mboera L, Mwabukusi M, Kariuki Njenga M, et al. A Smartphone App (AfyaData) for Innovative One Health Disease Surveillance from Community to National Levels in Africa: Intervention in Disease Surveillance. *JMIR Public Health Surveill* 2017 Dec 18;3(4):e94. [doi: [10.2196/publichealth.7373](https://doi.org/10.2196/publichealth.7373)]
30. Kamadjeu R, Gathenji C. Designing and implementing an electronic dashboard for disease outbreaks response - Case study of the 2013-2014 Somalia Polio outbreak response dashboard. *Pan Afr Med J* 2017;27. [doi: [10.11604/pamj.supp.2017.27.3.11062](https://doi.org/10.11604/pamj.supp.2017.27.3.11062)]
31. Hartmann G, Stead G, DeGani A. Cross-platform mobile development. *Mobile Learning Environment*, Cambridge 2011;16(9):158-171 [FREE Full text]
32. Chunara R, Goldstein E, Patterson-Lomba O, Brownstein JS. Estimating influenza attack rates in the United States using a participatory cohort. *Sci Rep* 2015 Apr 2;5(1). [doi: [10.1038/srep09540](https://doi.org/10.1038/srep09540)]
33. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining Search, Social Media, and Traditional Data Sources to Improve Influenza Surveillance. *PLoS Comput Biol* 2015 Oct 29;11(10):e1004513. [doi: [10.1371/journal.pcbi.1004513](https://doi.org/10.1371/journal.pcbi.1004513)]
34. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wójcik O, et al. Flu Near You: Crowdsourced Symptom Reporting Spanning 2 Influenza Seasons. *Am J Public Health* 2015 Oct;105(10):2124-2130. [doi: [10.2105/ajph.2015.302696](https://doi.org/10.2105/ajph.2015.302696)]
35. Lawpoolsri S, Khamsiriwatchara A, Liulark W, Taweeseeneepitch K, Sangvichean A, Thongprarong W, et al. Real-Time Monitoring of School Absenteeism to Enhance Disease Surveillance: A Pilot Study of a Mobile Electronic Reporting System. *JMIR mHealth uHealth* 2014 May 12;2(2):e22. [doi: [10.2196/mhealth.3114](https://doi.org/10.2196/mhealth.3114)]
36. Chen M, Hao Y, Hwang K, Wang L, Wang L. Disease Prediction by Machine Learning Over Big Data From Healthcare Communities. *IEEE Access* 2017;5:8869-8879. [doi: [10.1109/access.2017.2694446](https://doi.org/10.1109/access.2017.2694446)]

Abbreviations

CSS: Cascading Style Sheet
IDS: infectious disease surveillance
JSON: JavaScript Object Notation
JWT: JSON Web Token
mHealth: mobile health
NBMLHD: Nepean Blue Mountains Local Health District
NSW: New South Wales
PHU: public health unit

Edited by G Eysenbach; submitted 28.05.19; peer-reviewed by D Mogul, RB Bentvlsen, E Almeida, R Harte; comments to author 30.09.19; revised version received 23.11.19; accepted 24.01.21; published 09.03.21.

Please cite as:

Ahn E, Liu N, Parekh T, Patel R, Baldacchino T, Mullavey T, Robinson A, Kim J
A Mobile App and Dashboard for Early Detection of Infectious Disease Outbreaks: Development Study
JMIR Public Health Surveill 2021;7(3):e14837
URL: <https://publichealth.jmir.org/2021/3/e14837>
doi: [10.2196/14837](https://doi.org/10.2196/14837)
PMID: [33687334](https://pubmed.ncbi.nlm.nih.gov/33687334/)

©Euijoon Ahn, Na Liu, Tej Parekh, Ronak Patel, Tanya Baldacchino, Tracy Mullavey, Amanda Robinson, Jinman Kim. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 09.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Combining Wearable Devices and Mobile Surveys to Study Child and Youth Development in Malawi: Implementation Study of a Multimodal Approach

Onicio Leal Neto¹, MSc, PhD; Simon Haenni¹, PhD; John Phuka², PhD; Laura Ozella³, MSc, PhD; Daniela Paolotti³, PhD; Ciro Cattuto^{3,4}, PhD; Daniel Robles⁵, MA; Guilherme Lichand¹, MSc, PhD

¹Department of Economics, University of Zurich, Zurich, Switzerland

²College of Medicine, University of Malawi, Lilongwe, Malawi

³ISI Foundation, Turin, Italy

⁴Department of Computer Science, University of Torino, Turin, Italy

⁵Department of Psychology, University of Alberta, Edmonton, AB, Canada

Corresponding Author:

Onicio Leal Neto, MSc, PhD

Department of Economics

University of Zurich

Schönberggasse 1

Zurich, 8001

Switzerland

Phone: 41 44 634 55 80

Email: onicio@gmail.com

Abstract

Background: Multimodal approaches have been shown to be a promising way to collect data on child development at high frequency, combining different data inputs (from phone surveys to signals from noninvasive biomarkers) to understand children's health and development outcomes more integrally from multiple perspectives.

Objective: The aim of this work was to describe an implementation study using a multimodal approach combining noninvasive biomarkers, social contact patterns, mobile surveying, and face-to-face interviews in order to validate technologies that help us better understand child development in poor countries at a high frequency.

Methods: We carried out a mixed study based on a transversal descriptive analysis and a longitudinal prospective analysis in Malawi. In each village, children were sampled to participate in weekly sessions in which data signals were collected through wearable devices (electrocardiography [ECG] hand pads and electroencephalography [EEG] headbands). Additionally, wearable proximity sensors to elicit the social network were deployed among children and their caregivers. Mobile surveys using interactive voice response calls were also used as an additional layer of data collection. An end-line face-to-face survey was conducted at the end of the study.

Results: During the implementation, 82 EEG/ECG data entry points were collected across four villages. The sampled children for EEG/ECG were 0 to 5 years old. EEG/ECG data were collected once a week. In every session, children wore the EEG headband for 5 minutes and the ECG hand pad for 3 minutes. In total, 3531 calls were sent over 5 weeks, with 2291 participants picking up the calls and 984 of those answering the consent question. In total, 585 people completed the surveys over the course of 5 weeks.

Conclusions: This study achieved its objective of demonstrating the feasibility of generating data through the unprecedented use of a multimodal approach for tracking child development in Malawi, which is one of the poorest countries in the world. Above and beyond its multiple dimensions, the dynamics of child development are complex. It is the case not only that no data stream in isolation can accurately characterize it, but also that even if combined, infrequent data might miss critical inflection points and interactions between different conditions and behaviors. In turn, combining different modes at a sufficiently high frequency allows researchers to make progress by considering contact patterns, reported symptoms and behaviors, and critical biomarkers all at once. This application showcases that even in developing countries facing multiple constraints, complementary technologies can

leverage and accelerate the digitalization of health, bringing benefits to populations that lack new tools for understanding child well-being and development.

(*JMIR Public Health Surveill* 2021;7(3):e23154) doi:[10.2196/23154](https://doi.org/10.2196/23154)

KEYWORDS

child development; wearables; participatory surveillance; proximity sensors; mobile surveying

Introduction

Background

Research on many determinants of child development, including its underlying biological mechanisms, has accelerated at a fast pace over the last decade. It is now well known that early childhood development is a maturational and interactive process, resulting in an ordered progression of perceptual motor, cognitive, language, socioemotional, and self-regulation skills [1,2]. It is also clear that early childhood nutrition crucially affects educational outcomes [3] that are, in turn, important for adult productivity and wealth. Having said that, while a lot is known about potential critical or sensitive periods in child development, we know much less about how the effects of different interventions interact, either within or across critical periods [3-5]. For instance, we know that decreasing malnutrition from ages 2 to 4 years leads to a higher likelihood of future employment [6]. We also know that early stimulation from ages 0 to 3 years leads to a higher likelihood of future employment [7]. However, the following question is important: do the two interventions combined have effects that are larger than the sum of its parts, because of complementarities, or are they substitutes, partly crowding out the effects of each other?

In addition, knowledge on the impacts of disease outbreaks on child development is still scarce, particularly when it comes to infection susceptibility disrupting the environment in which children grow [8]. Specifically, on the impact of infectious diseases, close proximity between caregivers and children under 5 years of age could increase the exposure risk in case caregivers fall ill [9]. Moreover, the consequences of disease outbreaks might interact with other features of child development, such as nutritional status, since disease burden can affect nutrition requirements to sustain child development under adverse conditions [10,11].

Panel studies in developing countries typically employ face-to-face data collection (ie, traditional survey instruments), which happens at a low frequency [12,13]. The Demographic and Health Survey, for instance, is updated only every 3 to 5 years. However, studying phenomena, such as those cited above, requires reliable high-frequency data. To accommodate this need, panel designs need to be reimagined to create better data streams.

Today's technologies allow collecting a multiplicity of users' features at scale and at a higher frequency and lower cost than face-to-face data collection, even if typically restrained to developed country settings. If their application could be extended to poor countries to track child development throughout different stages of their lives, such data could give us more power to evaluate interventions and improve causal

inferences, and could help identify underlying mechanisms [14]. In particular, having access to high-frequency data on children's biomarkers could allow for timely responses in public health (such as early warning systems), with immediate potential to decrease child mortality and morbidity, and to increase children's cognitive development [15]. For instance, electroencephalography (EEG) is considered an optimal neuroimaging technique to record developmental changes during childhood [16]. Such markers of cognitive development include changes in oscillatory rhythm frequencies responsible for brain connectivity and cognitive functions [17]. Moreover, having access to such data would allow learning about interactions between critical periods, as in the examples of nutrition and early stimulation mentioned previously. Importantly, if the frequency of data is high enough, it has the potential to feed machine learning models that predict child-level impacts of different interventions, allowing policymakers to customize child development programs at a high frequency.

While some of those technologies, including wearables, have been applied to track children's conditions in developing country settings, this has typically been done either (1) in isolation, rather than combining multiple data streams, or (2) at a low frequency. The problem of the former is that each data stream in isolation has important limitations, for example, symptoms might be underreported in phone surveys, wearable readings might correspond to a multiplicity of states when considered in isolation, and contact patterns might be insufficient to characterize disease transmission dynamics. The problem of the latter is that above and beyond being multidimensional, the dynamics of child development are complex. It is the case not only that no data stream in isolation can accurately characterize it, but also that even if combined, infrequent data might miss critical inflection points and interactions between different conditions and behaviors.

Instead, multimodal approaches have been shown to be a promising alternative for high-frequency monitoring, combining different inputs of noninvasive biomarkers and creating multiple angles to understand health and clinical outcomes [18,19]. These data signals also include other types of data streams that can help understand more integrally the different aspects in each patient or subject. Despite the potential, there is a lack of studies showing the feasibility of high-frequency data collection in poor settings, which combines a multimodal approach to track child development at a high frequency.

Aim

The aim of this study was to describe a multimodal approach combining noninvasive biomarkers, social contact patterns, mobile surveying, and face-to-face interviews in order to validate technologies that help us better understand child

development, as well as epidemiological events at the community level.

Methods

Study Design, Location, and Timeline

We carried out a mixed study based on a transversal descriptive analysis and a longitudinal prospective analysis in Malawi. The field work took place in the following four villages in Dowa District: Mdoliro, Chidothi, Mitalanje, and Mkuwani. A pilot was run from October 1 to October 31, 2019. In this pilot, the team sampled villagers and users as described in Table 1. From November 1 to November 8, 2019, the field work team prepared the activities on the ground. Training sessions were conducted

with health surveillance assistants and community watchers from selected villages. In this period, local authorities and stakeholders, such as the District Health Office, the District Health Management Team, and the District Emergency Committee, were involved. Moreover, a demonstration for village chiefs was performed. From November 22 to December 23, 2019, mobile surveys using interactive voice response (IVR) calls and text messages were started. In parallel, the internal team of researchers and surveyors ran fieldwork activities from November 29 to January 10, 2020. A research assistant supervised this fieldwork, making weekly spot checks for over a month. Wearable sensors were used from December 6, 2019, to January 10, 2020. The end-line face-to-face survey was conducted from January 6 to January 12, 2020.

Table 1. Sampling for the implementation phase.

Mode of data collection	Wearable devices (children 0-5 years old), n	Wearable proximity sensors (children + teenagers + adults), n	Face-to-face interviews (caregivers), n
Village 1	27	99	61
Village 2	12	N/A ^a	58
Village 3	21	N/A	61
Village 4	22	N/A	55
Total	82	99	235

^aN/A: not applicable.

Sampling

In each village, children were sampled to use wearable devices. Additionally, wearable proximity sensors to elicit the social network were deployed in one village. Table 1 provides details about the sampling strategy.

Ethical Aspects

A key ethical concern is the protection of human subjects from physical or psychological harm. All participants in the study were fully informed about the nature and purpose of the research and their requested involvement. Only participants who provided written consent (documented) were included in the research. In the case of children (subjects under 10 years old), consent was obtained from their guardians. In the case of adolescents (subjects between 10 and 19 years old), consent was obtained from both themselves and their guardians. Specific safeguards were in place to protect the safety (both physical and psychological) of respondents and those collecting the data. Personal identification information, including the names of the children, the names of the caregivers, and the phone numbers of the caregivers, were only accessible to local field workers and to one research associate in Malawi. No researcher was able to see any personal identification information as data sets were anonymized. Before the study, every participant was informed about the content of the study, including the objective of the study and the different types of data collected. Thereafter, every participant read a consent form (in both Chichewa and English) and signed it to indicate that they agreed to participate or let their children participate in this study. For those who could not read and write, our field workers read the content of the consent form to the participant and the participant provided his/her

fingerprint to indicate consent. Enumerators were trained in collecting sensitive information, including information about children and health. Data collection tools, including wearable devices, were tested and designed in a way that was culturally appropriate (eg, no stigma) and did not create distress for respondents (eg, not invasive). Data collection visits were organized at the appropriate time (eg, did not disturb the regular life of study participants) and place to minimize the risk to respondents. Enumerators were able to provide information on how individuals in situations of risk can seek support. Respondents were always given the option to withdraw from the study at any time. All data from the study were kept without identifiers to protect subject privacy. No data generated by this study, even if deidentified, can be exploited for commercial purposes by any of the counterparts involved. For the inception phase, two institutional review board (IRB) approvals were obtained. One was obtained from the Ethics Committee at the University of Zurich (OEC IRB #2018-046), and the other was obtained from the College of Medicine Research and Ethics Committee at the University of Malawi (P.10/19/2825).

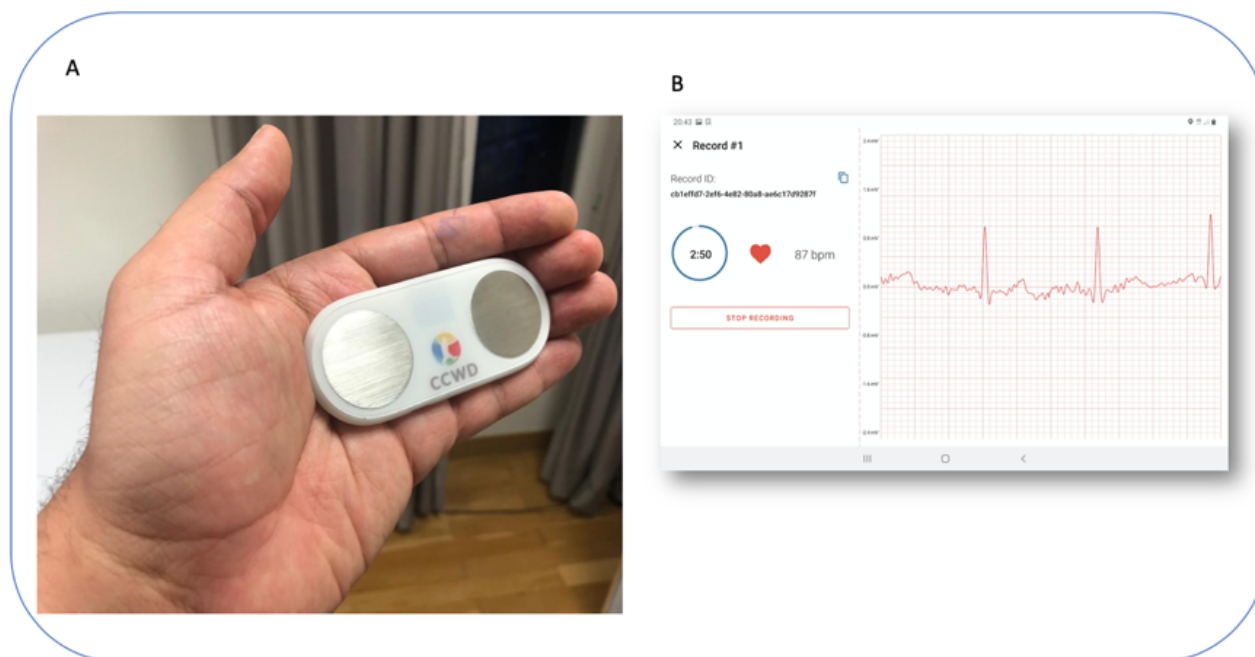
Multimodal Approach

Wearables: Electrocardiography Pads

Electrocardiography (ECG) pads (Figure 1) collect ECG data at a 500-Hz sample rate in a resting state. ECG is a recording of the electrical activity of the heart using electrodes placed on the skin. The signal is described in a graph of voltage versus time. The electrodes detect the small electrical changes that are a consequence of cardiac muscles followed by repolarization during each cardiac cycle (eg, heartbeat). The reason for including ECG data is to understand through heart rate and heart

rate variability if identified patterns could be related to infectious diseases, as an increase in heart rate is considered a proxy when the body temperature increases.

Figure 1. Electrocardiography pads with electrodes (A) and the tablet interface collecting the data (B).



During data collection, the study participants held the pad with their hands in a resting state and placed their thumbs on the metal electrodes for around 3 minutes. For small children who could not hold the pad, the health surveillance assistants and volunteers put two clips with electrodes on the child's wrists, connected the clips to the pad with a cable, and waited for 3 minutes. The ECG data appeared on an app on a tablet, and the data were stored on the tablet, which could be later uploaded to the cloud.

Wearables: EEG Headbands

Muse headbands (Figure 2) were used to collect EEG data (research resource identifier: SCR_014418; provided by Interaxon). It is a portable scalp EEG system that can be used to measure brain activity. It is battery powered and has four active electrodes located at 10 to 20 coordinates (TP9, AF7, AF8, and TP10), a common mode reference, and a driven right leg. Muse headsets initially oversample EEG data and subsequently down sample the data to yield a selectable output at a sampling rate of 220 Hz to 500 Hz, with 2 μ V (root mean square) noise. The input range of the AC-coupled signal (low cutoff at 1 Hz) is 2 mV [20-22].

Figure 2. Muse headband used to collect electroencephalography data from children.



The main research question that we were interested in is whether EEG data can show proxies that could be related to cerebral malaria, which has a high incidence in the country or whether those metrics could be related to cognitive development,

enabling systematic long-term monitoring in children at the community level.

During recording sessions, health surveillance assistants and volunteers placed the headband on the child's head, fitting the frontal electrodes above the eyebrow bridge and adjusting the posterior electrodes behind each mastoid. A block of 5 minutes of EEG data was recorded for each child while in a resting state with eyes open. The size of the headband could be adjusted to fit the size of the child's head. For very young children whose head circumferences were too small to fit the headband, an elastic gauze material was used to wrap up around the headband to maintain firm connectivity. The EEG data were shown on an app on a tablet, and the data were stored on the tablet and later synced to an encrypted server for retrieval and analysis.

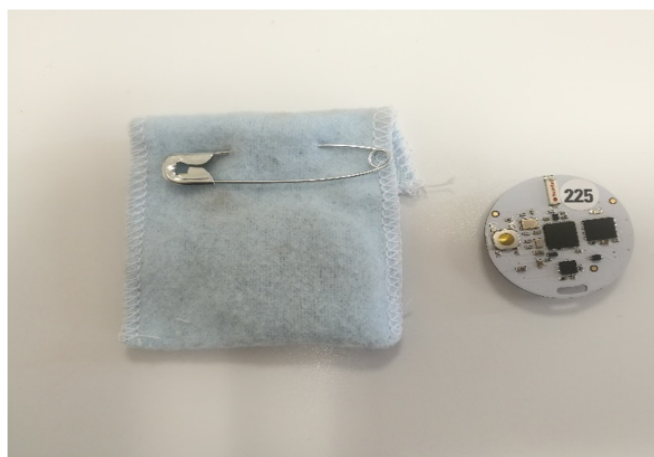
Both EEG and ECG signals were integrated in a mobile app used on a tablet by health surveillance assistants and community watchers. Additionally, portable solar power kits were distributed to make sure the devices would be recharged when necessary without relying on local electricity availability.

Wearables: Proximity Sensors

Proximity sensors (Figure 3) using low-powered radio frequencies have been used in different settings, such as

hospitals [23], households [24], and schools [25]. In Africa, proximity sensors have been applied in studies to characterize contact networks in households [26], and more recently, they have been used in households and schools [24,25]. Sensors provide contact data with high temporal resolution that can be used to investigate plausible characteristics of infection spread on network structures weighted by frequency and duration of contact [25,27]. Sensors exchange ultralow-power radio packets in a peer-to-peer fashion [28]. We considered that "contact" occurred between two individuals during a time slice duration of 20 seconds in a range of 1 to 1.5 m if the proximity sensors exchanged at least one radio packet during that interval. After contact is established, it is considered ongoing as long as the devices continue to exchange at least one packet for every subsequent 20-second interval [28]. All participants wore the device pinned to the front of a blouse/shirt in the chest area, and only face-to-face proximity relations were detected as the human body acts as a radiofrequency shield at the carrier frequency used for communication. Each device has a unique identification number that was used to link the information on the contacts established by the person carrying the device with his/her profile.

Figure 3. Proximity sensors and fabric pockets used to track social contact patterns.



As a research aspect, we were interested in social network structures at the community level to help understand when and how a disease would spread through the villages. Additionally, we wanted to assess if it is possible to rely on network sensing to evaluate the effect of nudges sent during the study.

Mobile Surveying

A mobile-based system of data collection, where individuals were asked to participate in a survey via IVR, was used as an additional data collection layer [29]. Respondents were asked to report their personal assessments of the condition of their children. There were the following five different survey modules over the course of 5 weeks: (1) medical symptoms, (2) nutrition, (3) wash and habits, (4) early stimulation, and (5) child development. For symptoms, a syndromic protocol was used [30,31], where the occurrences of the following clinical outcomes were collected: fever, headache, diarrhea, rash, joint pain, sore throat, cough, body aches, itching, shortness of breath, and red eyes. For nutrition, the system collected information on food intake, vegetable and fruit intake, protein-rich food

intake, food stocks, and breastfeeding. For wash, the reported information involved water availability and quality, available toilet facilities, and hand washing with soap. For early stimulation, we collected information on books and toys that might be available for a child, times when the child is left alone or in the care of another child, and activities that parents perform with the child, such as singing songs to the child. Finally, for child development, we collected information on the different types of activities that the child is currently able to perform. The survey questions and modules differed for children at different ages [32-34].

During the 1-month implementation, the selected participants received a call on their mobile phones every Friday for 5 weeks. One day before the phone call, they received an SMS text message that reminded them of the upcoming call [35-37]. If they picked up the call, they answered the survey questions by pressing different buttons on their phones. If they missed the call, they received another call later. They could call back the number to complete the survey at any time without any charge.

Besides mobile surveying by IVR, we sent nudges by SMS text message to the participants. By sending nudges on different topics, we wanted to test how participants react to the messages and whether there is any change in their behavior in response to the provided information (ie, treatment effect of nudges). It was decided to randomly split each of the four villages into two equal groups (control and treatment). The individuals in the treatment group received the nudges every week for 4 weeks. Conversely, the individuals in the control group did not receive the nudges. The weekly nudges asked caregivers to take their children to the health or nutrition clinic for check-ups, requested households to talk with other households about handwashing behavior, or asked caregivers to take their children on play dates with children from other households.

Face-to-Face Surveys

This study included the administration of (1) a light census and (2) a follow-up survey. These two data collections facilitated the interpretation of data coming from the devices.

First, in order to figure out who is eligible for the study and sample participants, the team performed a census survey in the four selected villages before the inception phase started. The survey only collected data about demographic information of the households and information about the medical situation of the children in the past 4 weeks, as well as their functioning mobile phone number and their willingness to receive SMS text messages/IVR calls.

After the fieldwork activities, the team deployed another face-to-face survey to collect information about child development, including health, wash, nutrition, education, child protection, and environmental data. Similarly, this follow-up survey collected reported information on device usability and acceptability, and similar ethical concerns that the population might raise because of the deployment of technological devices. We were interested to understand how the population receives a mobile message and if individuals share this information

within the household and with neighbors. For this, we measured the ratio of information received through phone messages during the face-to-face data collection among segments of the target population that were not sent phone messages.

The follow-up survey contained the following four sections: demographics of the households and measured children; questions about the Child Development Study, including the reactions of the participants toward the study and the ethical concerns of the study; questions about the user experiences of the different wearable devices; and questions about the medical symptoms and medical treatments of the targeted children.

Data Analysis

Descriptive analyses were undertaken showcasing the adherence and number of sessions involving wearable devices. For EEG data, we performed a grand average frequency spectra analysis evaluating the frequency range distribution. For proximity sensor data, we created a social contact matrix to evaluate relevant contacts during the study period. In addition, the contact patterns were compared with nudges sent by the IVR through a time series. For the end-line survey, we performed linear regression for medical symptoms. In-depth analyses were not the objective in this study and will be covered in future publications.

Results

Participation Characteristics

During the implementation, 82 EEG/ECG data entry points were collected in the four villages (Table 2). The sampled children for EEG/ECG were 0 to 5 years old. EEG/ECG data were collected weekly, meaning that health surveillance assistants/volunteers used the wearable devices to collect data on the children once a week. For every collection session, the children needed to wear the EEG headband for 5 minutes and needed to wear the ECG hand pad for 3 minutes. We report the frequency at which EEG and ECG data were collected in Tables 2 and 3, respectively.

Table 2. Frequency table for electroencephalography data (weeks).

Number of successful sessions per child	Frequency	Percentage	Cumulative
0	5	6.10	6.10
1	23	28.05	34.15
2	11	13.41	47.56
3	18	21.95	69.51
4	25	30.49	100
Total	82	100	N/A ^a

^aN/A: not applicable.

Table 3. Frequency table for electrocardiography data (weeks).

Number of successful sessions per child	Frequency	Percentage	Cumulative
0	2	2.44	2.44
1	22	26.83	29.27
2	10	12.20	41.46
3	21	25.61	67.07
4	27	32.93	100
Total	82	100	N/A ^a

^aN/A: not applicable.

Outcomes

Regarding EEG data collection, the goal of this study was to measure the feasibility of the Muse headband in recording the signal from the children. Figure 4 shows the grand average EEG spectra for the frontal and posterior Muse sensors, demonstrating that it is possible to collect reliable signals from our sample. Figure 5 presents a sample of raw EEG data for an entire

recording period. The raw signal is displayed across frontal and posterior channels, showing changes in EEG over time and movement/blink artifacts. Figure 6 presents the EEG spectrogram for frontal and posterior electrodes, showing an overall trend of higher oscillatory activity in low frequencies over time. Taken together, these figures suggest that reliable EEG recordings can be obtained using this portable Muse methodology.

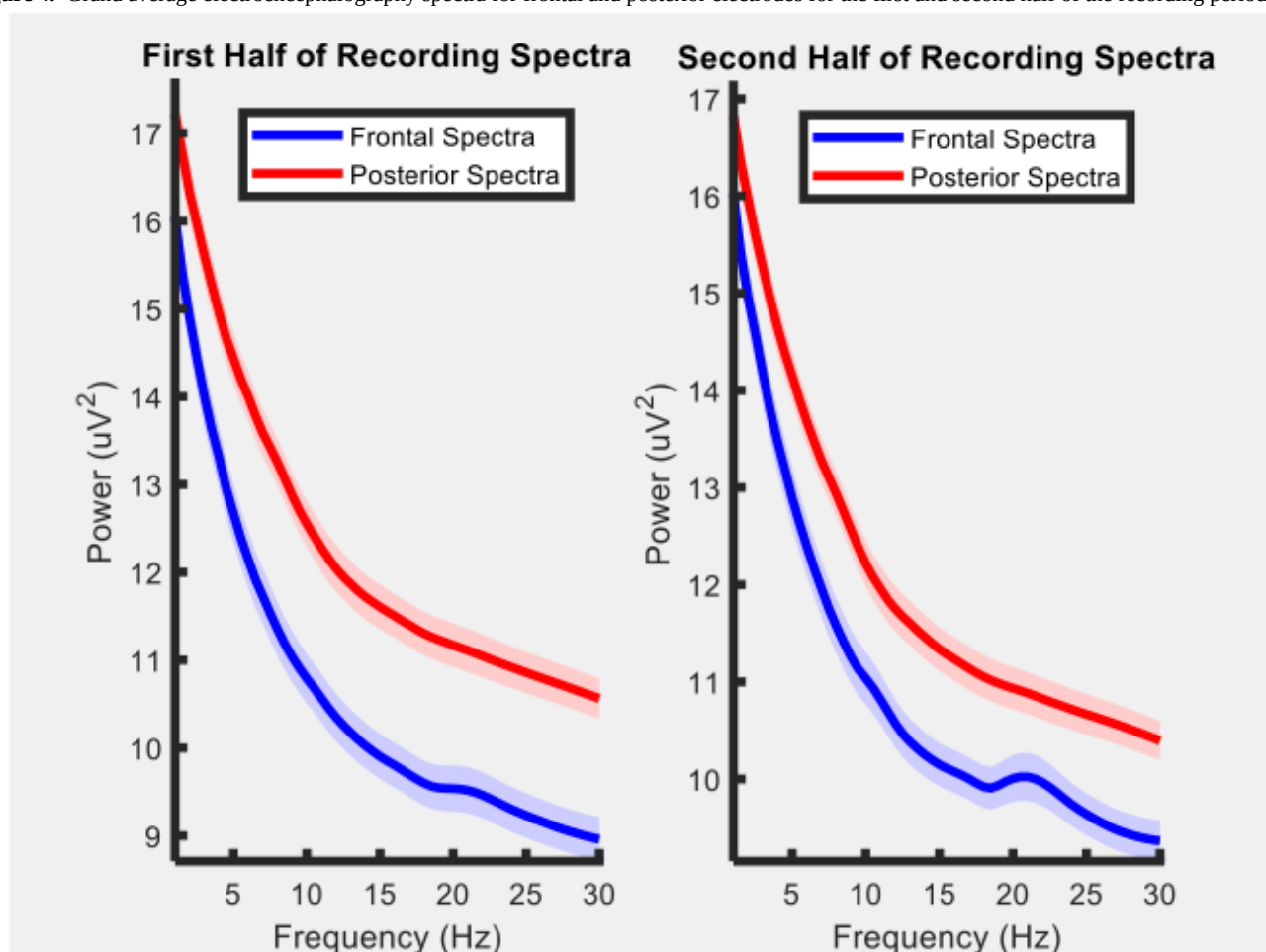
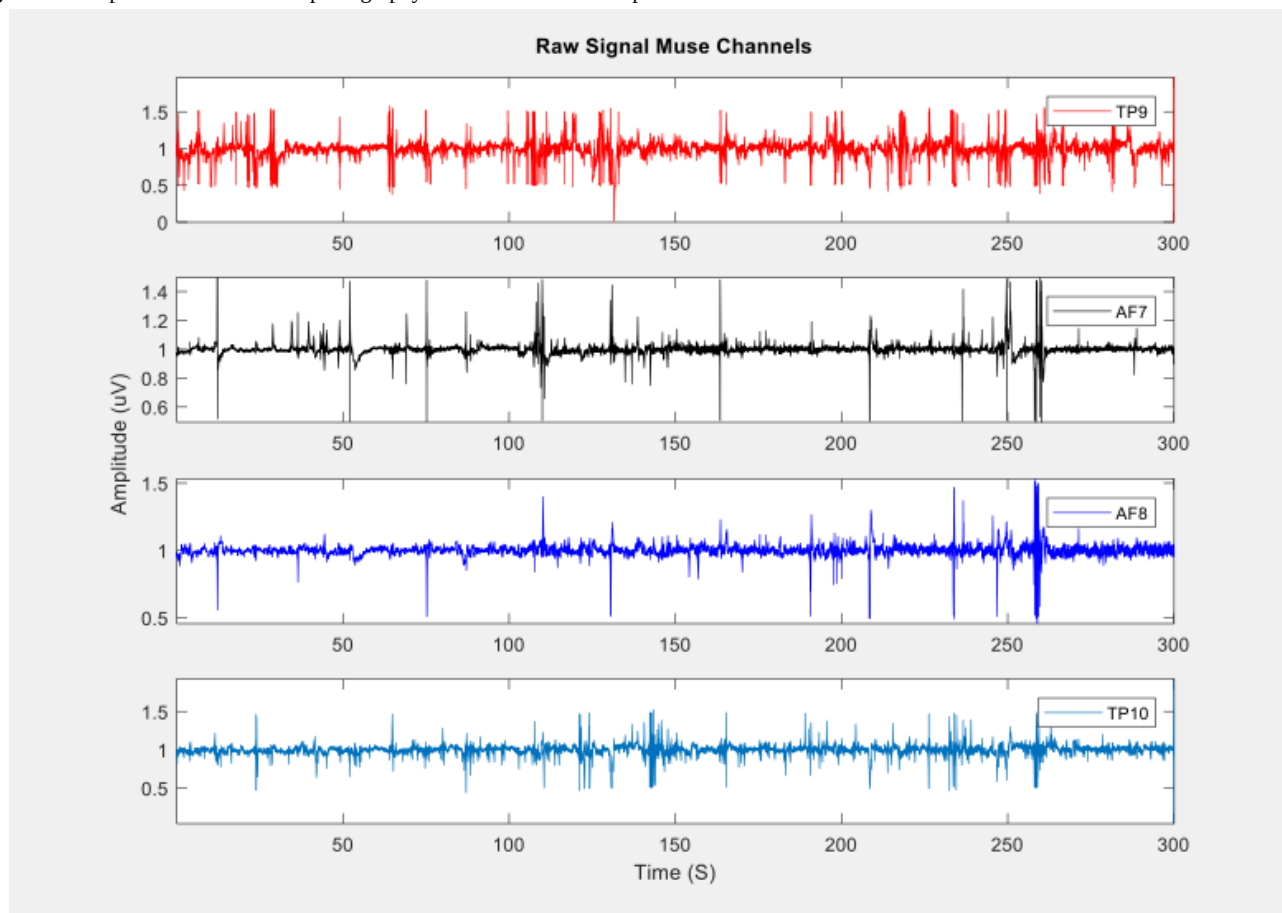
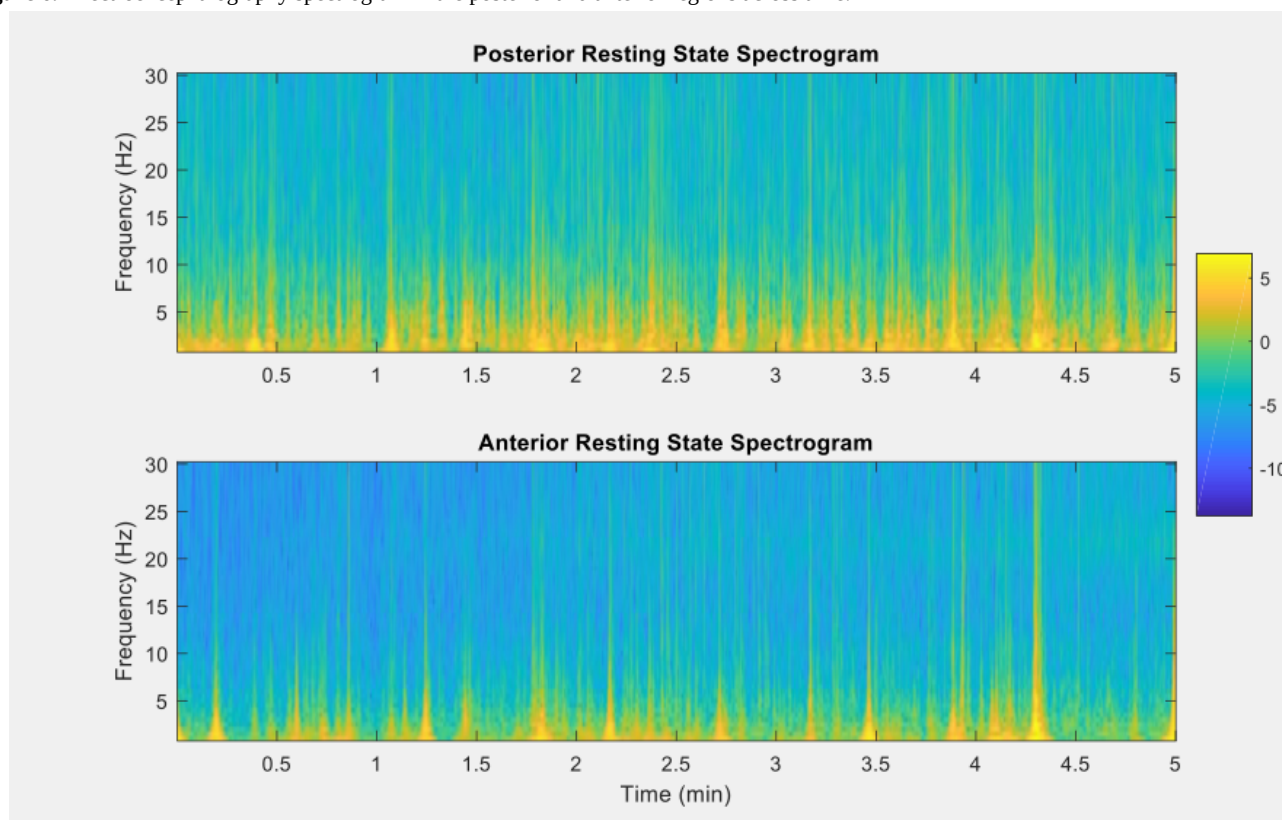
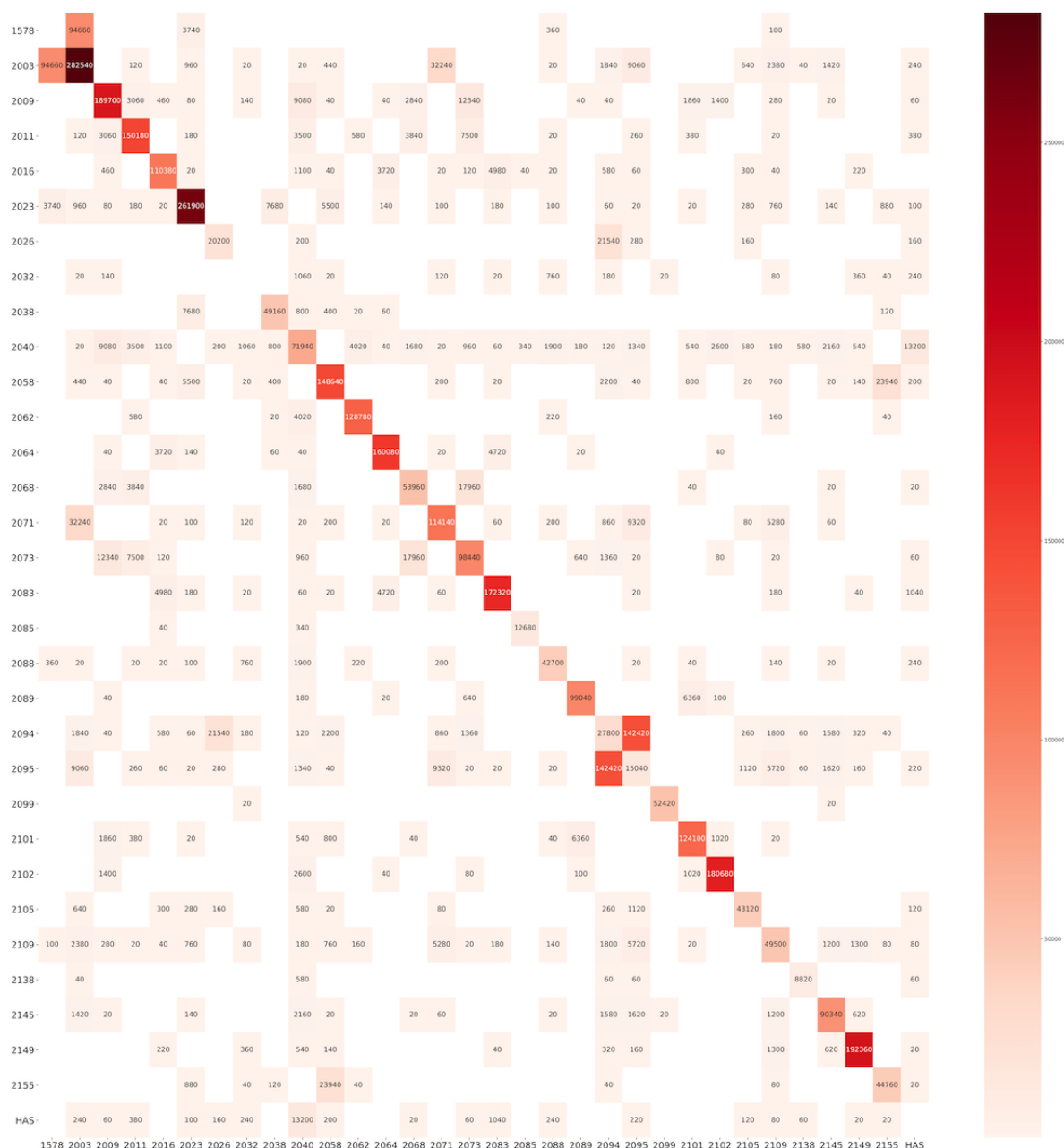
Figure 4. Grand average electroencephalography spectra for frontal and posterior electrodes for the first and second half of the recording period.

Figure 5. Sample of raw electroencephalography data at the frontal and posterior Muse electrodes.**Figure 6.** Electroencephalography spectrogram in the posterior and anterior regions across time.

For the proximity sensors, the matrix (Figure 7) shows the intrahousehold and interhousehold social contacts.

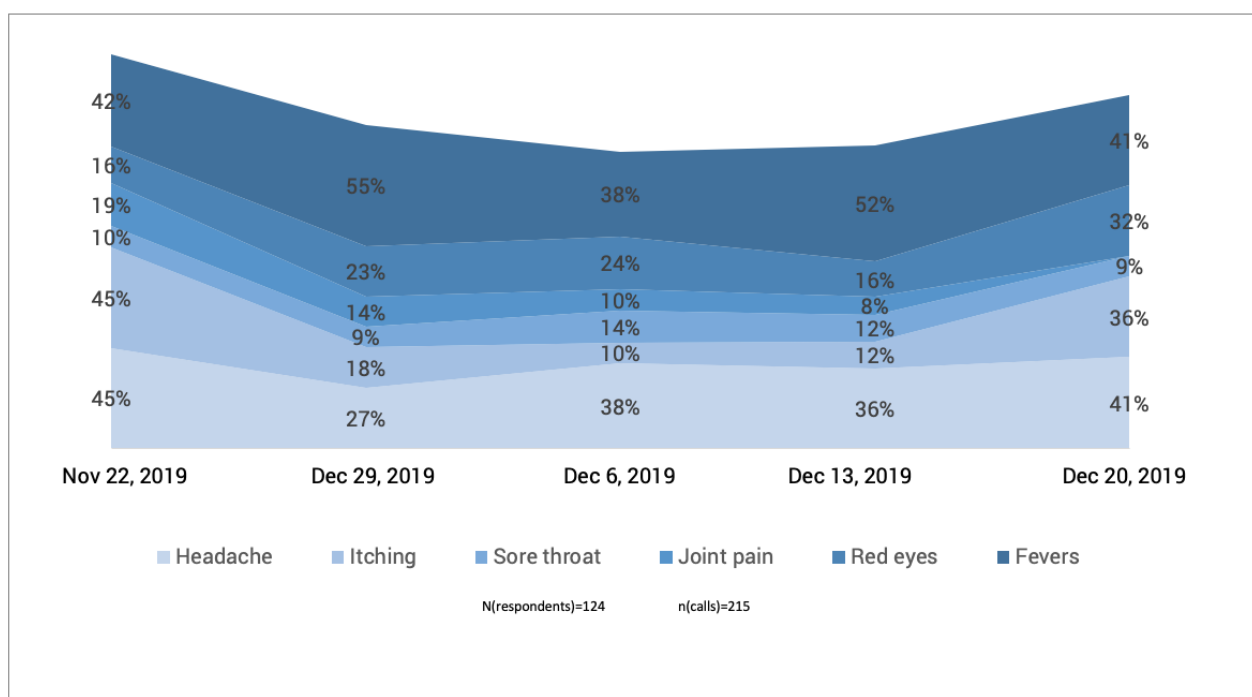
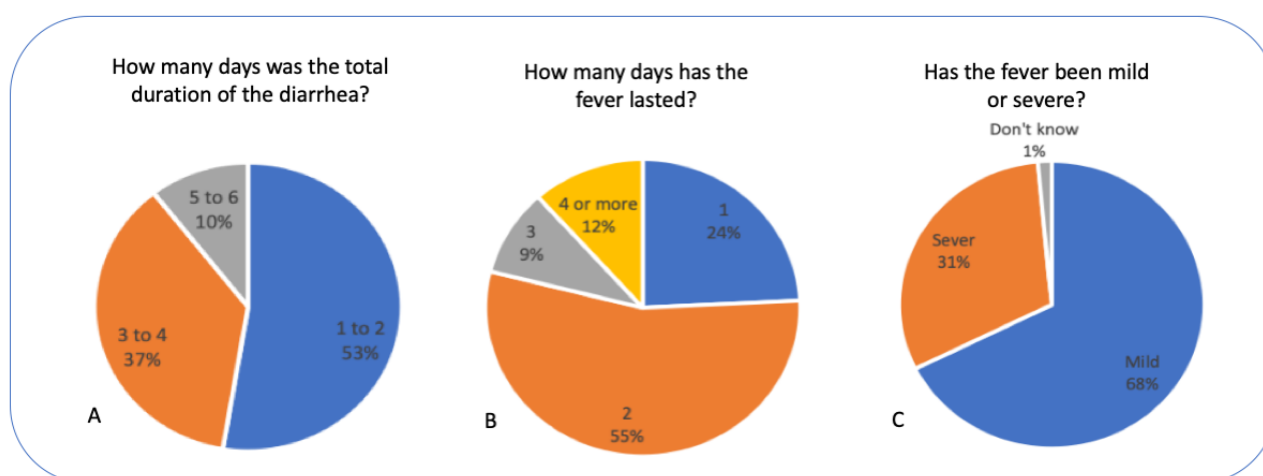
Figure 7. Intrahousehold and interhousehold total time in proximity (in seconds) registered by proximity sensors. The numbers on the Y axis and X axis represent the household ID. Each square represents the social contacts between the different households on the Y axis and the household on the X axis. The number in the square is the total time in proximity (in seconds). The numbers in the diagonal cells represent the total time in proximity among the members in the same household.



In total, 3531 calls were sent over the 5 weeks the project was live, with 2291 people picking up the calls and 984 answering the consent question. In total, 585 people completed the surveys over the 5 weeks. This implies a pickup rate of 64.88% (2291/3531), a completion rate after pickup of 25.53% (585/2291), and a completion rate after answering the consent question of 59.5% (585/984). Completion rates dropped for each village after the first week, regardless of the incentive provided both in that week and the following weeks.

First, in order to understand if IVR surveys are feasible, we studied phone ownership in the reference area. Phone ownership ranges from 56% in Mkuwani to 81% in Mdoliro. The probability of owning a phone might be related to other characteristics of a household.

Regarding the symptoms collected during the mobile surveying, Figure 8 shows the temporal proportion distribution along the study period. As shown in Figure 9, data on the severity of symptoms were collected, which can help distinguish between minor issues and major diseases.

Figure 8. Temporal proportion distribution of most reported symptoms in the studied villages.**Figure 9.** Severity of symptoms collected during mobile surveying.

Besides symptoms, we collected data on some dimensions of child development, including information on nutrition (Figure 10) and early stimulation (Figure 11). These data were collected

for three villages. High-frequency nutrition data can help interpret symptoms and support recommendations to caregivers.

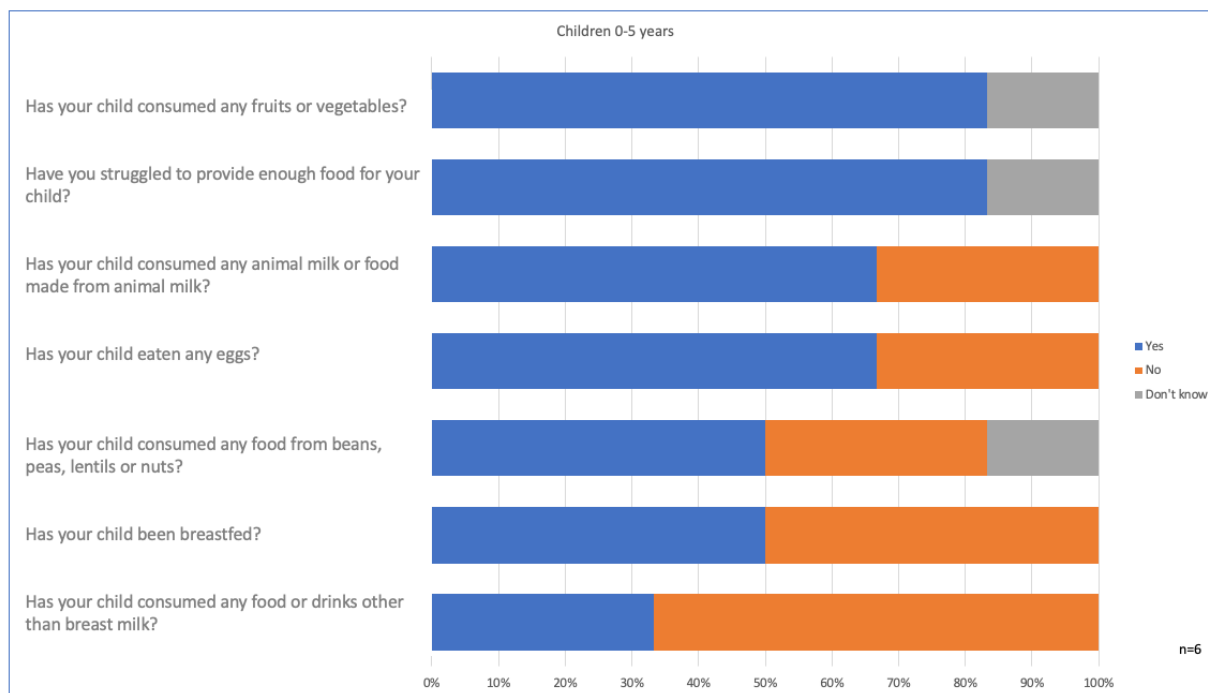
Figure 10. Nutrition aspects collected by mobile surveying.

Figure 11 shows the data collected on children in terms of early stimulation. Early stimulation was shown to be crucial for the intellectual development of children. Although not directly related to emergency action, these data show how mobile

surveys based on IVR can be used to explore a wide range of determinants of a child's long-term well-being.

Finally, Figure 12 shows the results from the end-line investigation related to symptoms and nudges for the village participants.

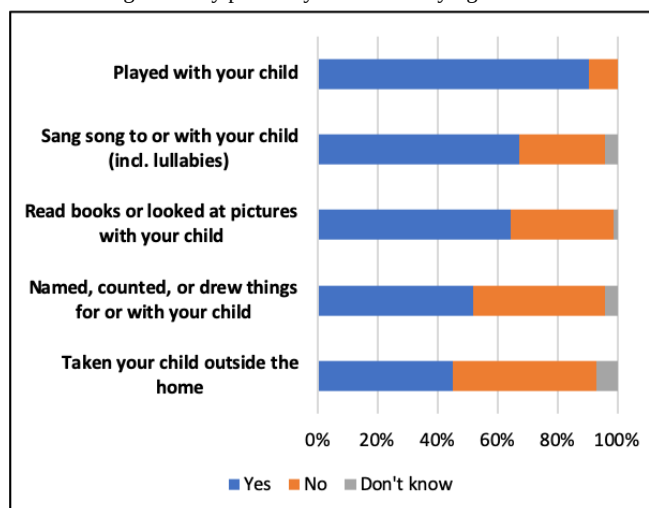
Figure 11. Early stimulation questions asked during the study period by mobile surveying.

Figure 12. Linear regression for nudge participants and their symptoms. (***p<0.01, ** p<0.05, * p<0.1).

Symptoms over the last 7 days											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Diarrhea	Headache	Itching	Sore throat	Joint pain	Red eyes	Fever	Skin rashes	Coughing	Body aches	Shortness of breath
Nudged	-0.0140 [0.0571]	0.0537 [0.0455]	0.0243 [0.0496]	-0.00990 [0.00990]	0.0210 [0.0244]	-0.0189 [0.0279]	-0.0345 [0.0559]	0.0358 [0.0554]	-0.0942 [0.0576]	-0.0681* [0.0371]	0.0108 [0.0224]
Constant	0.208*** [0.0406]	0.0891*** [0.0285]	0.129*** [0.0335]	0.00990 [0.00990]	0.0198 [0.0139]	0.0495** [0.0217]	0.208*** [0.0406]	0.168*** [0.0374]	0.257*** [0.0437]	0.109*** [0.0312]	0.0198 [0.0139]
Observations	199	199	199	199	199	199	199	199	199	199	199
R-squared	0.000	0.007	0.001	0.005	0.004	0.002	0.002	0.002	0.013	0.017	0.001
Symptoms over the last 30 days											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Diarrhea	Headache	Itching	Sore throat	Joint pain	Red eyes	Fever	Skin rashes	Coughing	Body aches	Shortness of breath
Nudged	-0.0674 [0.0704]	0.0490 [0.0647]	0.0685 [0.0633]	0.0108 [0.0224]	0.0321 [0.0353]	-0.00839 [0.0325]	-0.0167 [0.0708]	0.00788 [0.0626]	-0.0455 [0.0712]	0.0787 [0.0636]	0.0111 [0.0263]
Constant	0.465*** [0.0499]	0.267*** [0.0443]	0.238*** [0.0426]	0.0198 [0.0139]	0.0495** [0.0217]	0.0594** [0.0236]	0.455*** [0.0498]	0.257*** [0.0437]	0.515*** [0.0500]	0.238*** [0.0426]	0.0297* [0.0170]
Observations	199	199	199	199	199	199	199	199	199	199	199
R-squared	0.005	0.003	0.006	0.001	0.004	0.000	0.000	0.000	0.002	0.008	0.001

Discussion

Principal Results

This study demonstrated the feasibility of the unprecedented use of a multimodal approach for collecting data related to child development in Malawi settings. The preliminary data collected through our different data streams showcase a feasible methodology for tracking children's health and development in poor country settings at a high frequency.

Comparison With Prior Work

While several studies applied in developing countries have already proven the feasibility of using wearable devices to understand how diseases behave and affect certain populations [38-41], including children [42], much less has been shown about the feasibility of multimodal approaches with the multiple data streams that we combined at the high frequency that we implemented within the complex setting that we targeted, which are all necessary elements to understand the intricacies of child development in poor countries and to unlock the potential of interventions powered by these data, with the potential to change the life trajectories of those children. The complexity and innovation of the study can be perceived by the scarcity of literature that describes the methods adopted with precision and reproducibility.

The execution of studies like this is still a challenge owing to the logistical and infrastructure issues that need to be overcome. However, the portability of the devices used favored the measurement of noninvasive biomarkers, such as EEG and ECG data, in environments without any structure, which is an important finding that was demonstrated. The EEG and ECG data showed that in addition to the feasibility of collection, the quality achieved has great potential for scaling interventions that previously did not consider these types of biomarkers. As seen in Figures 4 to 6, the quality of measurements indicated a feasible way to collect EEG readings using a wearable device, although further tests are still needed to find out whether the premise of the relationship of brain wave patterns with the involvement in infectious diseases or psychological trauma can

be addressed by specific patterns. Toth et al [43] suggested the use of EEG combined with facial coding for improved evaluation of facial expressions. This type of approach could provide inputs for understanding stress levels through the assessment of facial patterns [44,45]. With regard to ECG data, studies [46-48] have demonstrated that this noninvasive biomarker can be used to measure malnutrition, in addition to the use of some specific standards for proxies related to temperature increase. The data from the proximity sensors indicated that nudges sent via SMS text messaging played a role in activating behavior change, and despite being an observation not yet reported in the literature, a similar approach used in Africa showed promising results for social network pattern studies [49]. Going further into this part, data from proximity sensors in combination with the syndromic data from mobile surveying could bring new insights into transmission routes for infectious diseases, supporting the understanding of how intrahousehold and interhousehold contacts play roles in disease spread. This sort of combination is important and opens up opportunities for future research involving extensive experiments on both modules. Synthetic models and simulations used to be reliable ways to create scenarios for understanding disease outbreaks, but periodically, reports from mobile surveys could really help in understanding which features of contact networks are the most important for disease spread. A doable app that can be mentioned involves the use of social networks for digital contact tracing of COVID-19 spreaders, measuring exposure and contagion, and complementing existing traditional disease surveillance systems.

Limitations

The syndromic approach seems to be an important tool for assessing the risk and existence of outbreaks of infectious diseases in the studied region owing to the self-report of symptoms present in the children informed by their caregivers. The alignment of syndromic approaches with participatory surveillance plays an important role in understanding epidemiological settings at the community level [50], and even though there is a limitation in the implementation of more sophisticated tools, such as mobile apps, they are still strategic instruments for health surveillance systems. Nevertheless, the

premise of the combination of not only information related to symptoms but also nutritional or developmental aspects has the possibility of deepening with the construction of syndemic models [51]. Even the observation of several factors that coexist in Malawi shows that they have been addressed as syndemics in similar contexts like Kenya or in similar social fabric related to food insecurity, maternal mental health, and domestic violence [52,53].

Conclusions

The development of research that uses different technologies for data collection in children in the context of underdeveloped countries is challenging, requiring a recurrent search for instruments and strategies that enable and balance the needs and potential of certain areas [54]. Wearable technologies have demonstrated wide application in society, integrating the internet

of things with telemetry systems to bring the context of remote management in the monitoring of health situations [55-57]. Even considering the observations mentioned by Piwek et al [58], where concerns related to security, reliability, and privacy in the use of these technologies are exposed, we conclude that the application of multimodal approaches like our approach is promising. In addition, they are able to achieve the immense challenge of collecting high-frequency data and feeding decision makers strategic information for planning policies and programs. Wearable technologies have already created a health revolution and are gradually reaching developed countries [59]. In addition, we understand that it is a good time for developing countries, even those that are in a critical scarcity scenario, to use this opportunity to leverage and accelerate the digitalization of health, bringing benefits to populations that lack new tools for understanding child well-being and development.

Acknowledgments

We thank the Center for Child Development team members Nicolás Tomaselli, Walter Dzerouniam, Jiajing Feng, Faith Milongo, and Julien Christen; the UNICEF Malawi Country Office; UNICEF Innovation; and ARM. We specially acknowledge Interaxon for donating 10 Muse devices for the study. Finally, Mawi and VIAMO are very important stakeholders in this study, providing the technology that allowed data collection. LO, CC, and DP acknowledge support from the Lagrange Project of the Institute for Scientific Interchange Foundation funded by the Cassa di Risparmio di Torino Foundation.

Conflicts of Interest

None declared.

References

1. Black MM, Walker SP, Fernald LCH, Andersen CT, DiGirolamo AM, Lu C, et al. Early childhood development coming of age: science through the life course. *The Lancet* 2017 Jan;389(10064):77-90. [doi: [10.1016/s0140-6736\(16\)31389-7](https://doi.org/10.1016/s0140-6736(16)31389-7)]
2. Engle PL, Fernald LC, Alderman H, Behrman J, O'Gara C, Yousafzai A, et al. Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *The Lancet* 2011 Oct;378(9799):1339-1353. [doi: [10.1016/s0140-6736\(11\)60889-1](https://doi.org/10.1016/s0140-6736(11)60889-1)]
3. Cunha F, Heckman J, Schennach S. Estimating the technology of cognitive and noncognitive skill formation. National Bureau of Economic Research. 2010. URL: https://www.nber.org/system/files/working_papers/w15664/w15664.pdf [accessed 2021-02-19]
4. Bridgman A. Child Development. In this issue. *Child Dev* 2015;86(2):337-341. [doi: [10.1111/cdev.12373](https://doi.org/10.1111/cdev.12373)] [Medline: [25809234](https://pubmed.ncbi.nlm.nih.gov/25809234/)]
5. Attanasio O, Meghir C, Nix E. Human capital development and parental investment in India. National Bureau of Economic Research. 2015. URL: https://www.nber.org/system/files/working_papers/w21740/w21740.pdf [accessed 2021-02-19]
6. Maluccio JA, Hoddinott J, Behrman JR, Martorell R, Quisumbing AR, Stein AD. The Impact of Improving Nutrition During Early Childhood on Education among Guatemalan Adults. *The Economic Journal* 2009;119(537):734-763. [doi: [10.1111/j.1468-0297.2009.02220.x](https://doi.org/10.1111/j.1468-0297.2009.02220.x)]
7. Gertler P, Heckman J, Pinto R, Zanolini A, Vermeersch C, Walker S, et al. Labor market returns to an early childhood stimulation intervention in Jamaica. *Science* 2014 May 30;344(6187):998-1001 [FREE Full text] [doi: [10.1126/science.1251178](https://doi.org/10.1126/science.1251178)] [Medline: [24876490](https://pubmed.ncbi.nlm.nih.gov/24876490/)]
8. Fischer H, Elliott L, Bertrand S. Guidance Note: Protection of Children During Infectious Outbreaks. New York, New York: The Alliance for Child Protection in Humanitarian Action; 2018.
9. Mupere E, Kaducu OF, Yoti Z. Ebola haemorrhagic fever among hospitalised children and adolescents in northern Uganda: epidemiologic and clinical observations. *Afr Health Sci* 2001 Dec;1(2):60-65 [FREE Full text] [Medline: [12789118](https://pubmed.ncbi.nlm.nih.gov/12789118/)]
10. Rice A, Sacco L, Hyder A, Black RE. Malnutrition as an underlying cause of childhood deaths associated with infectious diseases in developing countries. *Bull World Health Organ* 2000;78(10):1207-1221 [FREE Full text] [Medline: [11100616](https://pubmed.ncbi.nlm.nih.gov/11100616/)]
11. Page A, de Rekeneire N, Sayadi S, Aberrane S, Janssens AC, Rieux C, et al. Infections in children admitted with complicated severe acute malnutrition in Niger. *PLoS One* 2013;8(7):e68699 [FREE Full text] [doi: [10.1371/journal.pone.0068699](https://doi.org/10.1371/journal.pone.0068699)] [Medline: [23874731](https://pubmed.ncbi.nlm.nih.gov/23874731/)]

12. Rubio-Codina M, Grantham-McGregor S. Evolution of the wealth gap in child development and mediating pathways: Evidence from a longitudinal study in Bogota, Colombia. *Dev Sci* 2019 Sep;22(5):e12810. [doi: [10.1111/desc.12810](https://doi.org/10.1111/desc.12810)] [Medline: [30742349](https://pubmed.ncbi.nlm.nih.gov/30742349/)]
13. Jensen S, Obradović J, Nelson CA. Introduction to special issue on global child development studies. *Dev Sci* 2019 Sep;22(5):e12888. [doi: [10.1111/desc.12888](https://doi.org/10.1111/desc.12888)] [Medline: [31323172](https://pubmed.ncbi.nlm.nih.gov/31323172/)]
14. Müller J, Hoch AM, Zoller V, Oberhoffer R. Feasibility of Physical Activity Assessment with Wearable Devices in Children Aged 4-10 Years-A Pilot Study. *Front Pediatr* 2018;6:5 [FREE Full text] [doi: [10.3389/fped.2018.00005](https://doi.org/10.3389/fped.2018.00005)] [Medline: [29435438](https://pubmed.ncbi.nlm.nih.gov/29435438/)]
15. Bian L, Leslie SJ, Cimpian A. Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science* 2017 Jan 27;355(6323):389-391. [doi: [10.1126/science.aah6524](https://doi.org/10.1126/science.aah6524)] [Medline: [28126816](https://pubmed.ncbi.nlm.nih.gov/28126816/)]
16. Bell MA, Cuevas K. Using EEG to Study Cognitive Development: Issues and Practices. *J Cogn Dev* 2012 Jul 10;13(3):281-294 [FREE Full text] [doi: [10.1080/15248372.2012.691143](https://doi.org/10.1080/15248372.2012.691143)] [Medline: [23144592](https://pubmed.ncbi.nlm.nih.gov/23144592/)]
17. Stroganova TA, Orekhova EV, Posikera IN. EEG alpha rhythm in infants. *Clinical Neurophysiology* 1999 Jun;110(6):997-1012. [doi: [10.1016/s1388-2457\(98\)00009-1](https://doi.org/10.1016/s1388-2457(98)00009-1)] [Medline: [10402087](https://pubmed.ncbi.nlm.nih.gov/10402087/)]
18. Saby JN, Marshall PJ. The utility of EEG band power analysis in the study of infancy and early childhood. *Dev Neuropsychol* 2012;37(3):253-273 [FREE Full text] [doi: [10.1080/87565641.2011.614663](https://doi.org/10.1080/87565641.2011.614663)] [Medline: [22545661](https://pubmed.ncbi.nlm.nih.gov/22545661/)]
19. Kumari P, Mathew L, Syal P. Increasing trend of wearables and multimodal interface for human activity monitoring: A review. *Biosens Bioelectron* 2017 Apr 15;90:298-307. [doi: [10.1016/j.bios.2016.12.001](https://doi.org/10.1016/j.bios.2016.12.001)] [Medline: [27931004](https://pubmed.ncbi.nlm.nih.gov/27931004/)]
20. Huang Y, Yang J, Liao P, Pan J. Fusion of Facial Expressions and EEG for Multimodal Emotion Recognition. *Comput Intell Neurosci* 2017;2017:2107451 [FREE Full text] [doi: [10.1155/2017/2107451](https://doi.org/10.1155/2017/2107451)] [Medline: [29056963](https://pubmed.ncbi.nlm.nih.gov/29056963/)]
21. Kovacevic N, Ritter P, Tays W, Moreno S, McIntosh A. 'My Virtual Dream': Collective Neurofeedback in an Immersive Art Environment. *PLoS One* 2015;10(7):e0130129 [FREE Full text] [doi: [10.1371/journal.pone.0130129](https://doi.org/10.1371/journal.pone.0130129)] [Medline: [26154513](https://pubmed.ncbi.nlm.nih.gov/26154513/)]
22. Bhayee S, Tomaszewski P, Lee DH, Moffat G, Pino L, Moreno S, et al. Attentional and affective consequences of technology supported mindfulness training: a randomised, active control, efficacy trial. *BMC Psychol* 2016 Nov 29;4(1):60 [FREE Full text] [doi: [10.1186/s40359-016-0168-6](https://doi.org/10.1186/s40359-016-0168-6)] [Medline: [27894358](https://pubmed.ncbi.nlm.nih.gov/27894358/)]
23. Isella L, Romano M, Barrat A, Cattuto C, Colizza V, Van den Broeck W, et al. Close encounters in a pediatric ward: measuring face-to-face proximity and mixing patterns with wearable sensors. *PLoS One* 2011 Feb 28;6(2):e17144 [FREE Full text] [doi: [10.1371/journal.pone.0017144](https://doi.org/10.1371/journal.pone.0017144)] [Medline: [21386902](https://pubmed.ncbi.nlm.nih.gov/21386902/)]
24. Ozella L, Gesualdo F, Tizzoni M, Rizzo C, Pandolfi E, Campagna I, et al. Close encounters between infants and household members measured through wearable proximity sensors. *PLoS One* 2018;13(6):e0198733 [FREE Full text] [doi: [10.1371/journal.pone.0198733](https://doi.org/10.1371/journal.pone.0198733)] [Medline: [29879196](https://pubmed.ncbi.nlm.nih.gov/29879196/)]
25. Stehlé J, Voirin N, Barrat A, Cattuto C, Isella L, Pinton JF, et al. High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS One* 2011;6(8):e23176 [FREE Full text] [doi: [10.1371/journal.pone.0023176](https://doi.org/10.1371/journal.pone.0023176)] [Medline: [21858018](https://pubmed.ncbi.nlm.nih.gov/21858018/)]
26. Kiti M, Melegaro A, Cattuto C, Nokes DJ. Study design and protocol for investigating social network patterns in rural and urban schools and households in a coastal setting in Kenya using wearable proximity sensors. *Wellcome Open Res* 2019;4:84 [FREE Full text] [doi: [10.12688/wellcomeopenres.15268.2](https://doi.org/10.12688/wellcomeopenres.15268.2)] [Medline: [31489381](https://pubmed.ncbi.nlm.nih.gov/31489381/)]
27. Barrat A, Cattuto C, Tozzi AE, Vanhems P, Voirin N. Measuring contact patterns with wearable sensors: methods, data characteristics and applications to data-driven simulations of infectious diseases. *Clin Microbiol Infect* 2014 Jan;20(1):10-16 [FREE Full text] [doi: [10.1111/1469-0691.12472](https://doi.org/10.1111/1469-0691.12472)] [Medline: [24267942](https://pubmed.ncbi.nlm.nih.gov/24267942/)]
28. Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton JF, Vespignani A. Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS One* 2010 Jul 15;5(7):e11596 [FREE Full text] [doi: [10.1371/journal.pone.0011596](https://doi.org/10.1371/journal.pone.0011596)] [Medline: [20657651](https://pubmed.ncbi.nlm.nih.gov/20657651/)]
29. Ruikar V. Interactive Voice/Web Response System in clinical research. *Perspect Clin Res* 2016;7(1):15-20 [FREE Full text] [doi: [10.4103/2229-3485.173781](https://doi.org/10.4103/2229-3485.173781)] [Medline: [26952178](https://pubmed.ncbi.nlm.nih.gov/26952178/)]
30. Leal Neto O, Dimech GS, Libel M, de Souza WV, Cesse E, Smolinski M, et al. Saúde na Copa: The World's First Application of Participatory Surveillance for a Mass Gathering at FIFA World Cup 2014, Brazil. *JMIR Public Health Surveill* 2017 May 04;3(2):e26 [FREE Full text] [doi: [10.2196/publichealth.7313](https://doi.org/10.2196/publichealth.7313)] [Medline: [28473308](https://pubmed.ncbi.nlm.nih.gov/28473308/)]
31. Leal Neto O, Cruz O, Albuquerque J, Nacarato de Sousa M, Smolinski M, Pessoa Cesse E, et al. Participatory Surveillance Based on Crowdsourcing During the Rio 2016 Olympic Games Using the Guardians of Health Platform: Descriptive Study. *JMIR Public Health Surveill* 2020 Apr 07;6(2):e16119 [FREE Full text] [doi: [10.2196/16119](https://doi.org/10.2196/16119)] [Medline: [32254042](https://pubmed.ncbi.nlm.nih.gov/32254042/)]
32. Liu L, Leslie HH, Joshua M, Kruk ME. Exploring the association between sick child healthcare utilisation and health facility quality in Malawi: a cross-sectional study. *BMJ Open* 2019 Jul 27;9(7):e029631 [FREE Full text] [doi: [10.1136/bmjopen-2019-029631](https://doi.org/10.1136/bmjopen-2019-029631)] [Medline: [31352421](https://pubmed.ncbi.nlm.nih.gov/31352421/)]
33. Flores-Quispe MDP, Restrepo-Méndez MC, Maia MFS, Ferreira LZ, Wehrmeister FC. Trends in socioeconomic inequalities in stunting prevalence in Latin America and the Caribbean countries: differences between quintiles and deciles. *Int J Equity Health* 2019 Oct 15;18(1):156 [FREE Full text] [doi: [10.1186/s12939-019-1046-7](https://doi.org/10.1186/s12939-019-1046-7)] [Medline: [31615530](https://pubmed.ncbi.nlm.nih.gov/31615530/)]

34. Boerma JT, Sommerfelt AE. Demographic and health surveys (DHS): contributions and limitations. *World Health Stat Q* 1993;46(4):222-226. [Medline: [8017081](#)]
35. Kannisto K, Koivunen MH, Välimäki MA. Use of mobile phone text message reminders in health care services: a narrative literature review. *J Med Internet Res* 2014 Oct 17;16(10):e222 [FREE Full text] [doi: [10.2196/jmir.3442](#)] [Medline: [25326646](#)]
36. Marcano Belisario JS, Jamsek J, Huckvale K, O'Donoghue J, Morrison CP, Car J. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. *Cochrane Database Syst Rev* 2015 Jul 27(7):MR000042. [doi: [10.1002/14651858.MR000042.pub2](#)] [Medline: [26212714](#)]
37. Tyler KA, Schmitz RM. Using Cell Phones for Data Collection: Benefits, Outcomes, and Intervention Possibilities with Homeless Youth. *Child Youth Serv Rev* 2017 Mar;76:59-64 [FREE Full text] [doi: [10.1016/j.childyouth.2017.02.031](#)] [Medline: [28970644](#)]
38. Horby P, Pham QT, Hens N, Nguyen TTY, Le QM, Dang DT, et al. Social contact patterns in Vietnam and implications for the control of infectious diseases. *PLoS One* 2011 Feb 14;6(2):e16965 [FREE Full text] [doi: [10.1371/journal.pone.0016965](#)] [Medline: [21347264](#)]
39. Rhodes G, Neumann M, Ewing L, Bank S, Read A, Engfors LM, et al. Ensemble coding of faces occurs in children and develops dissociably from coding of individual faces. *Dev Sci* 2018 Mar;21(2). [doi: [10.1111/desc.12540](#)] [Medline: [28544105](#)]
40. Bian J, Guo Y, Xie M, Parish AE, Wardlaw I, Brown R, et al. Exploring the Association Between Self-Reported Asthma Impact and Fitbit-Derived Sleep Quality and Physical Activity Measures in Adolescents. *JMIR Mhealth Uhealth* 2017 Jul 25;5(7):e105 [FREE Full text] [doi: [10.2196/mhealth.7346](#)] [Medline: [28743679](#)]
41. Shull PB, Damian DD. Haptic wearables as sensory replacement, sensory augmentation and trainer - a review. *J Neuroeng Rehabil* 2015 Jul 20;12(1):59 [FREE Full text] [doi: [10.1186/s12984-015-0055-z](#)] [Medline: [26188929](#)]
42. Chandler J, Beets M, Saint-Maurice P, Weaver R, Cliff D, Drenowatz C, et al. Wrist-Based Accelerometer Cut-Points to Identify Sedentary Time in 5 11-Year-Old Children. *Children (Basel)* 2018 Sep 26;5(10):137 [FREE Full text] [doi: [10.3390/children5100137](#)] [Medline: [30261646](#)]
43. Toth J, Arvaneh M. Facial expression classification using EEG and gyroscope signals. 2017 Presented at: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 11-15, 2017; Jeju, South Korea p. 1018-1021. [doi: [10.1109/embc.2017.8036999](#)]
44. Kwon J, Kim DH, Park W, Kim L. A wearable device for emotional recognition using facial expression and physiological response. 2016 Presented at: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); August 16-20, 2016; Orlando, FL, USA p. 5765-5768. [doi: [10.1109/embc.2016.7592037](#)]
45. Eisermann M, Kaminska A, Moutard M, Soufflet C, Plouin P. Normal EEG in childhood: from neonates to adolescents. *Neurophysiol Clin* 2013 Jan;43(1):35-65. [doi: [10.1016/j.neucli.2012.09.091](#)] [Medline: [23290174](#)]
46. Barreto GSC, Vanderlei FM, Vanderlei LCM, Leite Á. Impact of malnutrition on cardiac autonomic modulation in children. *J Pediatr (Rio J)* 2016 Nov;92(6):638-644 [FREE Full text] [doi: [10.1016/j.jpeds.2016.03.005](#)] [Medline: [27234037](#)]
47. Mohshim SA, Ismail NS, Fadzir MF. Body temperature and ECG monitoring system for cardiac problem patient. In: AIP Conference Proceedings 2129. 2019 Presented at: 5th International Conference on Green Design and Manufacture; April 29-30, 2019; Jawa Barat, Indonesia. [doi: [10.1063/1.5118113](#)]
48. Sun X, Zhang Y. Design and Implementation of Portable ECG and Body Temperature Monitor. 2014 Presented at: International Symposium on Computer, Consumer and Control; June 10-12, 2014; Taichung, Taiwan p. 910-913. [doi: [10.1109/is3c.2014.239](#)]
49. 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 Oct 11;3(4):e62 [FREE Full text] [doi: [10.2196/publichealth.7540](#)] [Medline: [29021131](#)]
50. Singer M, Bulled N, Ostrach B, Mendenhall E. Syndemics and the biosocial conception of health. *The Lancet* 2017 Mar;389(10072):941-950. [doi: [10.1016/s0140-6736\(17\)30003-x](#)]
51. Kibicho J, Campbell JK. Community perspectives of second-generation alcohol misuse and HIV risk in rural Kenya: A gendered syndemic lens. *Glob Public Health* 2019 Dec 10;14(12):1733-1743. [doi: [10.1080/17441692.2019.1638958](#)] [Medline: [31291832](#)]
52. Laurenzi C, Field S, Honikman S. Food Insecurity, Maternal Mental Health, and Domestic Violence: A Call for a Syndemic Approach to Research and Interventions. *Matern Child Health J* 2020 Apr;24(4):401-404. [doi: [10.1007/s10995-019-02872-8](#)] [Medline: [32009230](#)]
53. Kohrt BA, Rai S, Vilakazi K, Thapa K, Bhardwaj A, van Heerden A. Procedures to Select Digital Sensing Technologies for Passive Data Collection With Children and Their Caregivers: Qualitative Cultural Assessment in South Africa and Nepal. *JMIR Pediatr Parent* 2019 Jan 16;2(1):e12366 [FREE Full text] [doi: [10.2196/12366](#)] [Medline: [31518316](#)]
54. Mertz L. Convergence Revolution Comes to Wearables: Multiple Advances are Taking Biosensor Networks to the Next Level in Health Care. *IEEE Pulse* 2016 Jan;7(1):13-17. [doi: [10.1109/mpul.2015.2498475](#)]

55. Metcalf D, Milliard ST, Gomez M, Schwartz M. Wearables and the Internet of Things for Health: Wearable, Interconnected Devices Promise More Efficient and Comprehensive Health Care. *IEEE Pulse* 2016 Sep;7(5):35-39. [doi: [10.1109/mpul.2016.2592260](https://doi.org/10.1109/mpul.2016.2592260)]
56. Marouf M, Vukomanovic G, Saranovac L, Bozic M. Multi-purpose ECG telemetry system. *Biomed Eng Online* 2017 Jun 19;16(1):80 [FREE Full text] [doi: [10.1186/s12938-017-0371-6](https://doi.org/10.1186/s12938-017-0371-6)] [Medline: [28629374](https://pubmed.ncbi.nlm.nih.gov/28629374/)]
57. Yetisen AK, Martinez-Hurtado JL, Ünal B, Khademhosseini A, Butt H. Wearables in Medicine. *Adv Mater* 2018 Jun 11;30(33):e1706910 [FREE Full text] [doi: [10.1002/adma.201706910](https://doi.org/10.1002/adma.201706910)] [Medline: [29893068](https://pubmed.ncbi.nlm.nih.gov/29893068/)]
58. Piwek L, Ellis DA, Andrews S, Joinson A. The Rise of Consumer Health Wearables: Promises and Barriers. *PLoS Med* 2016 Feb 2;13(2):e1001953 [FREE Full text] [doi: [10.1371/journal.pmed.1001953](https://doi.org/10.1371/journal.pmed.1001953)] [Medline: [26836780](https://pubmed.ncbi.nlm.nih.gov/26836780/)]
59. Dunn J, Runge R, Snyder M. Wearables and the medical revolution. *Per Med* 2018 Sep;15(5):429-448 [FREE Full text] [doi: [10.2217/pme-2018-0044](https://doi.org/10.2217/pme-2018-0044)] [Medline: [30259801](https://pubmed.ncbi.nlm.nih.gov/30259801/)]

Abbreviations

ECG: electrocardiography

EEG: electroencephalography

IRB: institutional review board

IVR: interactive voice response

Edited by H Bradley; submitted 03.08.20; peer-reviewed by M Woldegorgis, H Luo; comments to author 25.09.20; revised version received 05.11.20; accepted 02.02.21; published 05.03.21.

Please cite as:

Leal Neto O, Haenni S, Phuka J, Ozella L, Paolotti D, Cattuto C, Robles D, Lichand G

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

URL: <https://publichealth.jmir.org/2021/3/e23154>

doi: [10.2196/23154](https://doi.org/10.2196/23154)

PMID: [33536159](https://pubmed.ncbi.nlm.nih.gov/33536159/)

©Onicio Leal Neto, Simon Haenni, John Phuka, Laura Ozella, Daniela Paolotti, Ciro Cattuto, Daniel Robles, Guilherme Lichand. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 05.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Diagnostic Accuracy of Detecting Diabetic Retinopathy by Using Digital Fundus Photographs in the Peripheral Health Facilities of Bangladesh: Validation Study

Tahmina Begum¹, MD, MPH; Aminur Rahman², MD, MPH; Dilruba Nomani², MSc; Abdullah Mamun¹, PhD; Alayne Adams³, PhD; Shafiqul Islam⁴, MD; Zara Khair⁵, MSc; Zareen Khair⁵, PhD; Iqbal Anwar², MD, PhD

¹Institute for Social Science Research, The University of Queensland, Brisbane, Australia

²icddr,b, Dhaka, Bangladesh

³McGill University, Montreal, QC, Canada

⁴Barisal Medical College Hospital, Barisal, Bangladesh

⁵The Fred Hollow Foundation, Dhaka, Bangladesh

Corresponding Author:

Tahmina Begum, MD, MPH

Institute for Social Science Research

The University of Queensland

80 Meiers Rd

Brisbane, 4068

Australia

Phone: 61 469705278

Email: tahminashakera@gmail.com

Abstract

Background: Diabetic retinopathy can cause blindness even in the absence of symptoms. Although routine eye screening remains the mainstay of diabetic retinopathy treatment and it can prevent 95% of blindness, this screening is not available in many low- and middle-income countries even though these countries contribute to 75% of the global diabetic retinopathy burden.

Objective: The aim of this study was to assess the diagnostic accuracy of diabetic retinopathy screening done by non-ophthalmologists using 2 different digital fundus cameras and to assess the risk factors for the occurrence of diabetic retinopathy.

Methods: This validation study was conducted in 6 peripheral health facilities in Bangladesh from July 2017 to June 2018. A double-blinded diagnostic approach was used to test the accuracy of the diabetic retinopathy screening done by non-ophthalmologists against the gold standard diagnosis by ophthalmology-trained eye consultants. Retinal images were taken by using either a desk-based camera or a hand-held camera following pupil dilatation. Test accuracy was assessed using measures of sensitivity, specificity, and positive and negative predictive values. Overall agreement with the gold standard test was reported using the Cohen kappa statistic (κ) and area under the receiver operating curve (AUROC). Risk factors for diabetic retinopathy occurrence were assessed using binary logistic regression.

Results: In 1455 patients with diabetes, the overall sensitivity to detect any form of diabetic retinopathy by non-ophthalmologists was 86.6% (483/558, 95% CI 83.5%-89.3%) and the specificity was 78.6% (705/897, 95% CI 75.8%-81.2%). The accuracy of the correct classification was excellent with a desk-based camera (AUROC 0.901, 95% CI 0.88-0.92) and fair with a hand-held camera (AUROC 0.710, 95% CI 0.67-0.74). Out of the 3 non-ophthalmologist categories, registered nurses and paramedics had strong agreement with kappa values of 0.70 and 0.85 in the diabetic retinopathy assessment, respectively, whereas the nonclinical trained staff had weak agreement ($\kappa=0.35$). The odds of having retinopathy increased with the duration of diabetes measured in 5-year intervals ($P<.001$); the odds of having retinopathy in patients with diabetes for 5-10 years (odds ratio [OR] 1.81, 95% CI 1.37-2.41) and more than 10 years (OR 3.88, 95% CI 2.91-5.15) were greater than that in patients with diabetes for less than 5 years. Obesity was found to have a negative association ($P=.04$) with diabetic retinopathy.

Conclusions: Digital fundus photography is an effective screening tool with acceptable diagnostic accuracy. Our findings suggest that diabetic retinopathy screening can be accurately performed by health care personnel other than eye consultants. People with more than 5 years of diabetes should receive priority in any community-level retinopathy screening program. In a

country like Bangladesh where no diabetic retinopathy screening services exist, the use of hand-held cameras can be considered as a cost-effective option for potential system-wide implementation.

(*JMIR Public Health Surveill* 2021;7(3):e23538) doi:[10.2196/23538](https://doi.org/10.2196/23538)

KEYWORDS

diabetic retinopathy; diagnostic accuracy; digital fundus photograph; Bangladesh; diabetes; retinopathy; retina; ophthalmology

Introduction

Diabetic retinopathy, a progressive eye complication of diabetes mellitus, which affects 9.3% of the people globally, is considered the fifth leading cause of global blindness [1,2]. Results from a recent systematic review estimated that globally, the prevalence of retinopathy among patients with diabetes is 35% [3]. Nearly all patients with type 1 diabetes and more than 60% of the patients with type 2 diabetes develop retinopathy within 20 years of diabetes onset [4]. However, like many other diabetic complications, retinopathy remains asymptomatic until significant damage has occurred [4,5]. Periodic eye screening is essential for diagnosing the disease in early stages and to enable timely initiation of treatment [6]. Current recommendations are that retinal screening occur once a year for all patients with diabetes and that more frequent examinations take place if abnormal findings are identified [7]. However, periodic retinal screening is not available in many low- and middle-income countries although they account for 75% of the global burden of diabetic retinopathy [3,8]. The reported barriers are multifactorial. A recent systematic review reported that poor knowledge and attitudes to asymptomatic eye screening are prevalent both among health care providers and patients [9]. At the health care system level, the lack of equipment, insufficiently skilled professionals, nonfunctioning referral mechanisms, and inadequate data within national management information systems are the main barriers [10]. Moreover, routine diabetic retinopathy screening is not always feasible for eye consultants, given their availability versus the load of patients with diabetes.

As an alternative to eye consultants, different cadres of non-ophthalmologists such as general practitioners, opticians, and diabetologists have been successfully involved in diabetic retinopathy screening in many high-income countries [11]. Particular attention has been focused on developing simple algorithms and technologies suitable for non-ophthalmologists. Among these screening tools, digital fundus photography has been identified as one of the best and lowest cost options [12]. A digital camera allows repeated images to be taken until the best one is captured. Final retinal images can be stored and sent for expert opinion by using a web-based interface [13]. With this technology, high-income countries have shown increased diagnostic accuracy even without pupil dilatation (mydriasis) [12,14]. However, the nonmydriatic approach has shown low accuracy (12%-25%) in the Southeast Asian context [6]. Patients having a dark iris and reporting to a hospital at an advanced age with comorbid eye diseases such as cataracts are the commonly reported explanations for poor vision in the nonmydriatic approach [6].

Like many low- and middle-income countries, Bangladesh demonstrates a substantial diabetic retinopathy disease burden and an array of health system challenges that complicate the routine implementation of diabetic retinopathy screening services. According to the International Diabetic Federation statistics, around 8.4% of the total population in Bangladesh had diabetes in 2017 [15], which puts the country among the top 10 high diabetes burden countries in the world [15]. At the same time, Bangladesh has a critical shortage and “geographic maldistribution” of health workforce with an increased concentration in urban areas even though 70% of the population resides in the rural region [16]. Eye care services are provided predominantly by ophthalmologists, and diabetic retinopathy screening programs are not readily available under the current health system [17]. Given the growing burden of diabetes in Bangladesh, blindness prevention has become a strategic priority [17]. To support this effort, the “Integrated Model of Care for Diabetic Retinopathy within the Health System of Bangladesh” was initiated as a collaborative program between the Fred Hollows Foundation (FHF), a nongovernmental organization and the Government of Bangladesh [18]. This program seeks to establish a care pathway for diabetes and diabetic retinopathy by training non-ophthalmologist health cadres to conduct diabetic retinopathy screening by using digital fundus photography [18].

The aim of this study was to test the diagnostic accuracy of detecting any grade of diabetic retinopathy by non-ophthalmologists using digital fundus cameras against the gold standard diagnosis of ophthalmologists. We also explored the risk factors of diabetic retinopathy in the rural sites under investigation.

Methods

Study Setting and Design

This validation study was conducted at the project implementation sites of FHF in 4 districts under 2 administrative divisions of Bangladesh. Six health facilities were randomly chosen from these 4 districts for our study: one medical college hospital, one district hospital, and 4 health centers of the Diabetic Association of Bangladesh (DAB). The medical college hospital and the district hospital are government-funded general hospitals that provide eye care services for patients with or without diabetes. DAB centers are autonomous health care organizations focused on the treatment of patients with diabetes only. Retinopathy screening for asymptomatic cases is not a regular clinical care option in any of these health facilities. Eye consultants were available in all government hospitals and in 1 out of the 4 DAB centers. As part of project activities, the FHF established a memorandum of understanding with the concerned health facilities to establish a diabetic retinopathy screening

corner within the eye department to provide equipment supplies and to organize relevant local and national-level trainings on diabetic retinopathy diagnosis and treatment for hospital staff. A parallel referral mechanism linking DAB centers and the closest public health facility was also established.

Diabetic Retinopathy Screening Process

Patients with diabetes attending the outdoor eye clinic at the selected study health facilities between July 2017 and June 2018 were included as the study participants. Two different types of digital fundus photography instruments were used for diabetic retinopathy screening: a desk-based high-resolution fundus camera and a hand-held low-cost fundus camera. Initial screening was done by non-ophthalmologists such as nurses, paramedics, and nonclinical trained staff. For the gold standard diagnosis, 2 eye consultants, one from each study division, were assigned to evaluate the screening done by the non-ophthalmologists for the respective administrative divisions. The desk-based camera is used in medical colleges and district hospitals, and registered nurses and paramedics are the primary diabetic retinopathy screening providers. Hand-held cameras are used in DAB centers. Other than 1 DAB center, none could deploy their own nurses/paramedics due to the high turnover rate. Thus, new project staff were recruited from FHF, one in each of the 3 DAB centers. These were nonclinical personnel with graduate degrees in any discipline. All the nonclinical staff, including nurse and paramedics, received hands-on training for 3 days from the eye consultant in the respective study division. Nurses working in the medical college obtained an opportunity to attend a month-long training at the national level. Once the 2 eye consultants certified that images taken by the non-ophthalmologists were satisfactory and their ability to detect diabetic retinopathy from the retinal images was accurate, the formal data collection process started.

Initially, the non-ophthalmologist staff took the retinal images and performed diabetic retinopathy grading independently. A short-acting pupil dilator was used prior to taking the retinal image and then, a single-field macula-focused image was taken. Subsequently, the same study participant with a referral slip (indicating patient ID and date of diabetic retinopathy screening) was referred to the eye department of the respective district or medical college hospital to be examined by the eye consultant. Retinal images were also sent to the eye consultant through a web-based interface or a portable computer disk. The eye consultants checked the gradeability of the retinal images provided through the web-based interface and performed diabetic retinopathy grading independently. The entire screening process was double-blinded, that is, no one had access to the findings of the others. The project-appointed staff monitored the data collection and retrieved data from the hospital records with a diagnostic accuracy checklist.

Sampling Strategy

The inclusion criteria were patients with type 1 diabetes older than 18 years or with type 2 diabetes having no previous diagnosis of diabetic retinopathy and images taken from both eyes. Exclusion criteria were patients with significant physical or mental disabilities that could hamper photography and having mature cataract and corneal opacity in any eye. Considering the

current facility-based prevalence of diabetic retinopathy as 27% [19], a sample size calculation was performed for sensitivity and specificity. The final sample size was the maximum of these two [20]. For the anticipated sensitivity and specificity, we considered the Canadian and British Ophthalmology Society guidelines. Both guidelines recommend at least 80% sensitivity and 90%-95% specificity for any alternative approach of diabetic retinopathy grading [21,22]. Taking all these into account, our required sample size was 1138, and we distributed them proportionately across the 6 study health facilities, considering the patient turnover rate.

Outcome Measures

The outcome variable was the presence of any form of diabetic retinopathy in either eye of a patient, which was confirmed by the eye consultant. The Airlie house classification was used for diabetic retinopathy staging, which is a validated tool for the diabetic retinopathy screening program [22]. This classification divides the diagnosis of diabetic retinopathy into 5 stages: no diabetic retinopathy, mild nonproliferative diabetic retinopathy, moderate nonproliferative diabetic retinopathy, severe nonproliferative diabetic retinopathy, and proliferative diabetic retinopathy. A diabetic retinopathy positive case referred to a patient who had any kind of diabetic retinopathy (mild nonproliferative diabetic retinopathy/moderate nonproliferative diabetic retinopathy/severe nonproliferative diabetic retinopathy/proliferative diabetic retinopathy) in any of the two eyes.

Covariates

We considered the patient's sociodemographic and clinical characteristics, both of which have been identified as risk factors of diabetic retinopathy in the literature [3]. Among the sociodemographic features, patient age, gender, education, and occupation were considered. Clinical covariates were BMI, duration of diabetes, recent blood sugar test result, and presence of hypertension. BMI was calculated from height and weight measurements performed on the day of the clinic visit by using the following formula: weight in kilograms/height in meters squared. The World Health Organization criteria of BMI classification for Asian people was used to categorize our sample into 4 groups [23]. For diabetes test results, we considered any form of blood sugar test done within 3 months with written documentation provided by the patient during diabetic retinopathy screening. We categorized diabetes test results into normal limit and uncontrolled blood sugar by using reference values provided by the available blood glucose tests [24].

Statistical Analysis

We calculated the test accuracy of non-ophthalmologists against the gold standard value and reported measures of sensitivity (true positive diabetic retinopathy/[true positive + false negative]) and specificity (true negative diabetic retinopathy/[true negative + false positive cases]) at 95% CI. The positive predictive values and negative predictive values were also calculated from the true diabetic retinopathy positive and diabetic retinopathy negative results out of the total positive and negative test results, respectively [25]. All diagnostic accuracy results were compared by instrument types (desk-based

camera vs hand-held camera). Additionally, differences in diagnostic accuracy were measured across the different diabetic retinopathy grades. Overall agreement and disagreement were tested using the Cohen kappa statistic and area under the receiver operating curve (AUROC). The AUROC is an index of accuracy [26] presented as a plot of true positive rates against false positive rates for different possible cut-off points of a diagnostic test [26]. Descriptive analysis reports the distribution of the study sample by covariates. The statistical association of diabetic retinopathy positive status with all covariates were tested initially through a bivariate analysis using the chi-square test. Covariates that were significant at a P value $<.05$ in bivariate analysis were included in the multivariate analysis. Binary logistic regression findings were presented as odds ratio (OR) at 95% CI. All the analyses were performed using the STATA software (Release 16, College Station, StataCorp LLC) [27].

Ethics Approval

We obtained ethical clearance from the Institutional Review Board of icddr,b, Dhaka, Bangladesh under protocol number 17003. Verbal informed consent was obtained from respective hospital authorities and study participants to access medical records. Data collected from the study participants were reported anonymously to maintain privacy and confidentiality. Study

participants with poor-quality retinal images or confusing findings were referred to an eye consultant for further evaluation, and transport costs were remunerated. Any patient requiring laser treatment obtained this service from a tertiary-level public hospital free of cost.

Results

Characteristics of the Participants in This Study

In total, 1511 patients with diabetes were screened, which was slightly higher than our required sample size, and we included all of them. However, 3.7% (56/1511) of the images (including both eyes) were of poor quality and discarded from analysis. Thus, our final analytic sample consisted of 1455 patients. The mean (SD) age of the study participants was 53.23 (11.84) years, and a slightly larger proportion of the total study participants was females. More than half of the study participants had diabetes for more than 5 years. Blood sugar test results were missing for 7.0% (102/1455) of the patients. Blood sugar levels were reported 2 hours after breakfast in 52.5% (764/1455) of the patients and $<1\%$ (2/1455) of the patients by the hemoglobin A_{1C} test. The mean BMI of the patients was 25.38 (4.21). Nearly half of the patients with diabetes were screened by a nurse, and half were assessed using a desk-based camera (Table 1).

Table 1. Characteristics of the study participants who attended diabetic retinopathy screening from July 2017 to June 2018 in 6 selected peripheral hospitals in Bangladesh (N=1455).

Variable of interest	Values, n (%)
Patient age (years)	
<40 years	247 (16.98)
41-50 years	399 (27.42)
51-60 years	471 (32.37)
>60 years	338 (23.23)
Gender	
Female	814 (55.95)
Male	641 (44.05)
Education	
No schooling	256 (17.59)
Primary school completed	309 (21.24)
Higher secondary and above	890 (61.17)
Occupation	
Unemployed	1001 (68.80)
Service	266 (18.28)
Business	188 (12.92)
Body mass index	
Normal and underweight	429 (29.71)
Overweight	632 (43.77)
Obese	383 (26.52)
Duration of diabetes	
<5 years	513 (35.26)
5-10 years	459 (31.55)
>10 years	483 (33.19)
Blood sugar level	
Controlled	613 (45.31)
Not controlled	740 (54.69)
Hypertensive patient	
No	659 (45.29)
Yes	796 (54.71)
Type of non-ophthalmologist	
Nurse	766 (52.65)
Paramedics	276 (18.97)
Nonclinical trained staff	413 (28.38)
Instrument used	
Hand-held camera	576 (39.59)
Desk-based camera	879 (60.41)
Place of training for non-ophthalmologists	
Local	824 (56.63)
National	631 (43.37)

Diabetic Retinopathy Accuracy Test Results

The prevalence of diabetic retinopathy was 38.39% (558/1455). As shown in Table 2, the diagnostic accuracy of non-ophthalmologists was 86.6% (483/558, 95% CI 83.5%-89.3%) for diabetic retinopathy positive case detection (sensitivity) and 78.6% (705/897, 95% CI 75.8%-81.2%) for

diabetic retinopathy negative case detection (specificity). Further, non-ophthalmologists could correctly identify 71.6% (483/675) of the total diabetic retinopathy positive cases and correctly exclude 90.4% (705/780) of the total diabetic retinopathy negative cases. As the kappa value suggested, moderate agreement was observed with the gold standard ($\kappa=0.6$).

Table 2. Diagnostic accuracy of diabetes retinopathy by instrument type in 6 selected peripheral hospitals in Bangladesh from July 2017 to June 2018.

Indicators	Overall proportion (%) 95% CI	Hand-held camera proportion (%) 95% CI	Desk-based camera proportion (%) 95% CI
Sensitivity	86.56 (83.45-89.30)	85.60 (80.30-89.89)	87.19 (83.19-90.60)
Specificity	78.60 (75.78-81.2)	56.56 (51.19-61.70)	93.01 (90.50-95.01)
Positive predictive value	71.56 (68.78-74.10)	55.19 (51.89-58.40)	88.50 (84.89-91.30)
Negative predictive value	90.38 (88.39-92.10)	86.19 (81.70-89.67)	92.19 (89.89-93.89)
Accuracy	81.56 (79.56-83.60)	67.70 (63.70-71.65)	90.78 (88.70-92.56)
Kappa	0.63 (0.58-0.78)	0.38 (0.31-0.46)	0.80 (0.74-0.87)

Accuracy was further reported by AUROC findings. As shown in Figure 1, the ability of non-ophthalmologists to correctly classify diabetic retinopathy by using a hand-held camera was “Fair” (AUROC 0.710, 95% CI 0.67-0.74) (Figure 1, Panel A) and “Excellent” (AUROC 0.901, 95% CI 0.88-0.92) by using a desk-based camera (Figure 1, Panel B). We also assessed the agreement of different cadres of non-ophthalmologists against the gold standard value. Strong agreement was noted between the diabetic retinopathy classification of registered nurses and paramedics and that of the gold standard diagnosis by eye consultants, with kappa values of 0.70 and 0.85, respectively, whereas nonclinical trained staff had weak agreement ($\kappa=0.35$). As shown in Table 3, non-ophthalmologists were particularly good at detecting the presence or absence of diabetic

retinopathy. However, their accuracy differed by the diabetic retinopathy grading status particularly across the different grades of nonproliferative diabetic retinopathy. Accuracy also varied slightly depending on the type of instrument used. A desk-based camera was more likely to identify diabetic retinopathy correctly when diabetic retinopathy was present than when it was absent. In contrast, the probability of a correct diabetic retinopathy diagnosis was lower among those with a negative diabetic retinopathy finding versus those without. Although the hand-held camera was less successful in identifying diabetic retinopathy correctly in the presence of diabetic retinopathy, the probability that a person showed a negative finding with a hand-held camera for diabetic retinopathy was lower than that with a desk-based camera.

Figure 1. Diagnostic accuracy of digital fundus photography by area under receiver operating curve. Panel A: hand-held camera and Panel B: desk-based camera.

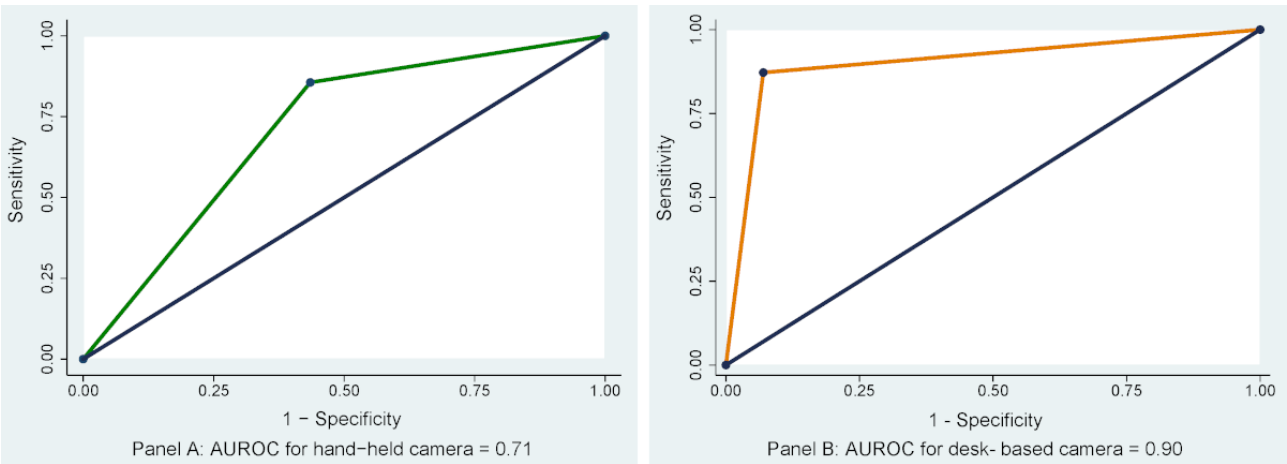


Table 3. Diagnostic accuracy of diabetic retinopathy screening by instrument types across different diabetic retinopathy stages in 6 selected peripheral hospitals in Bangladesh from July 2017 to June 2018.

Stage of diabetic retinopathy	Sensitivity/specificity (%) of hand-held camera	Sensitivity/specificity (%) of desk-based camera
No diabetic retinopathy	85.59 (56.50)	87.20 (93.00)
Mild NPDR ^a	69.35 (62.17)	59.26 (92.09)
Moderate NPDR	50.75 (92.53)	49.02 (96.91)
Severe NPDR	52.00 (96.37)	72.00 (93.03)
PDR ^b	33.33 (97.19)	49.02 (99.52)

^aNPDR: nonproliferative diabetic retinopathy.

^bPDR: proliferative diabetic retinopathy.

Determinants of Diabetic Retinopathy

Table 4 shows the determinants of diabetic retinopathy identified through bivariate and logistic regression analyses. Patient's age, education, BMI, diabetes duration, and controlled blood sugar levels were significant factors at 5% significance level while predicting diabetic retinopathy occurrence. After controlling for all significant covariates from bivariate results, the probability of diabetic retinopathy was found to increase with

increasing educational level, BMI, duration of diabetes for more than 5 years, and uncontrolled blood sugar levels. A graded response was observed between longer duration of diabetes and probability of diabetic retinopathy occurrence. However, BMI showed a negative association with diabetic retinopathy; the odds of having diabetic retinopathy decreased significantly with high BMI (OR 0.59, 95% CI 0.43-0.81) compared to the reference group consisting of normal weight and underweight patients.

Table 4. Factors associated with diabetic retinopathy occurrence in Bangladesh from July 2017 to June 2018.

Variable of interest	No diabetic retinopathy ^a (n=897), n (%)	Diabetic retinopathy present (n=558), n (%)	P value	Adjusted effect odds ratio (95% CI)	P value
Patient age (years)			.003 ^a		
<40 years	174 (70.5)	73 (29.6)		Ref ^b	
41-50 years	236 (59.1)	163 (40.9)		1.49 (1.03-2.16)	.03 ^a
51-60 years	270 (57.3)	201 (42.7)		1.32 (0.91-1.90)	.14
>60 years	217 (64.2)	121 (35.8)		0.99 (0.67-1.48)	.97
Gender			.09		
Male	380 (59.3)	261 (40.7)		N/A ^c	N/A
Female	517 (63.5)	297 (36.5)		N/A	N/A
Education			.001 ^a		
No schooling	185 (72.3)	71 (27.7)		Ref	
Primary completed	186 (60.2)	123 (39.8)		1.48 (1.00-2.19)	.05 ^a
≥Higher secondary	526 (59.1)	364 (40.9)		1.45 (1.03-2.03)	.03 ^a
Occupation			.43		
Unemployed	628 (62.7)	373 (37.3)		N/A	N/A
Service	159 (59.8)	107 (40.2)		N/A	N/A
Business	110 (58.5)	78 (41.5)		N/A	N/A
Body mass index			.04 ^a		
Normal and under-weight	254 (59.2)	175 (40.8)		Ref	
Overweight	378 (59.8)	254 (40.2)		0.92 (0.70-1.22)	.58
Obese	256 (66.8)	127 (33.2)		0.59 (0.43-0.81)	.001 ^a
Duration of diabetes			<.001 ^a		
<5 years	395 (77.0)	118 (23.0)		Ref	
5-10 years	292 (63.6)	167 (36.4)		1.98 (1.47-2.66)	<.001 ^a
>10 years	210 (43.5)	273 (56.5)		4.36 (3.23-5.89)	<.001 ^a
Diabetes controlled			<.001 ^a		
Yes	410 (66.9)	203 (33.1)		Ref	
No	418 (56.5)	322 (43.5)		1.45 (1.15-1.84)	.002 ^a
Hypertensive patient			.27		
No	501 (62.9)	295 (37.1)		N/A	N/A
Yes	396 (60.1)	263 (39.9)		N/A	N/A

^aResults are significant at $P < .05$.^bRef: reference category, that is, <40 years, no schooling, normal and underweight category of BMI, diabetes duration <5 years and diabetes controlled for the corresponding covariates. Diabetic retinopathy present was the reference value for the outcome variable.^cN/A: not applicable. Implies not included in the multivariate model.

Discussion

Conducted in 6 peripheral health facilities in Bangladesh, this validation study showed that non-ophthalmologists can detect the presence of diabetic retinopathy with reasonable diagnostic

accuracy. The sensitivity and specificity of using a single-field macular photograph after pupillary dilatation against the referral gold standard diagnosis by an eye consultant were 86.6% (483/558) and 78.6% (705/897), respectively. The degree of diagnostic agreement between a non-ophthalmologist and an

ophthalmologist was excellent with a desk-based camera and satisfactory with a hand-held camera.

To our knowledge, this is the first ever study in Bangladesh to engage health care personnel other than eye consultants in routine retinal examination. The reported true positive diabetic retinopathy case detection rate was within the international standard of more than 80% [21,22]. However, our overall specificity (79%) was lower than the international recommendation of 90%-95% [21,22]. This implies that 19.9% (173/869) of the total diabetic retinopathy negative cases were referred to an eye consultant when they did not have diabetic retinopathy [26]. In generic terms, lower specificity implies greater health system burden with a greater number of false positive cases referred to the next level [26]. Nevertheless, in a country like Bangladesh where no formal diabetic retinopathy screening services are available, a large proportion of people with diabetes remain undiagnosed until opportunistic diagnosis occurs at a very advanced stage [28]. Considering the increasing burden of diabetic retinopathy, this will result in increasing eye care-related costs and a greater risk of blindness [29]. By training and engaging non-ophthalmologist health staff in diabetic retinopathy screening, limited resources can be maximized and coverage increased [25]. However, evidence to support the incorporation of non-ophthalmologists into the diabetic retinopathy screening pathway remains scarce in low- and middle-income settings [30]. Some studies in India, Sri Lanka, and Pakistan report satisfactory test accuracy for non-ophthalmologist screening [6,9,31,32]. Using digital fundus photographs and a nonmydriatic approach, “physician graders” in Sri Lanka showed a sensitivity of 88.7% and specificity of 94.9% [10]. In Pakistan, optometrists showed 72% sensitivity and 86.3% specificity [32]. A decision must also be made about which instrument and imaging technique to use and whether to dilate or not dilate [33]. We validated the diagnostic agreement of hand-held versus desk-based cameras by calculating the AUROC and kappa values. The AUROC was “close to 1” for both instruments [26], suggesting a fair-to-excellent amount of agreement. However, in low- and middle-income country settings, instrument costs should also be considered. Here, the hand-held camera performs better—being comparatively less expensive than a desk-based camera and easier to carry and employ in the context of community-level screening [34].

The variation in test accuracy by instrument type further emphasizes the need to explore other factors influencing diagnostic test accuracy. Here, the gradeability of images plays an important part in successful diabetic retinopathy screening. In this study, a lower rate of ungradable images was reported (56/1511, 4.1%) than that reported in similar health facility-based studies in the South Asian context. A technical failure rate of 7.5% was noted in another study in Bangladesh [35], which rose to 12% in Sri Lanka [10]. The higher level of gradeability observed in our study is probably a function of the exclusion criteria or the image-taking technique that was employed. Considering the higher prevalence of cataracts in low-income country settings [29,36], we made the presence of cataract an exclusion criterion in our sample selection and chose to deploy a mydriatic approach for imaging. Although the gold standard for diabetic retinopathy screening is a 7-field

photograph using a nonmydriatic approach [37], the applicability of this technique in low-income country settings is highly controversial [38]. When using a nonmydriatic approach, the body’s autonomic nervous system becomes hyperactive with reflex pupillary constriction in the second eye after taking an image of the first eye [39,40]. Further, a Brazilian study showed that longer screening time and more referrals to ophthalmologists occurred in the absence of pupil dilation due to a larger number of ungradable images [41]. Further supporting our approach is the evidence of increased provider compliance when using a single-field photograph [42].

To identify other factors affecting diabetic retinopathy diagnosis, we compared test accuracy across different cadres of non-ophthalmologist personnel. We observed that nurses and paramedics showed higher accuracy in detecting any form of diabetic retinopathy than the nonclinical trained staff [43]. The relatively poorer performance of the nonclinical trained staff may be due to their lack of comfort with the hand-held camera. We expect that with more onsite supportive training, performance can be increased. Supporting the call for more hands-on training, a study suggested that the diagnostic accuracy of the fundus camera is highly dependent on the user’s technique in taking a correct image and their ability to do proper grading—both of which can be improved with more hands-on support [44]. Variations in the test accuracy by type of provider were similarly observed in a systematic review of 22 observational studies [45]. Noting that the sensitivity of detecting any form of retinopathy using a mydriatic approach ranged between 87% and 100% for general practitioners, >91% for optometrists, and 89% and 93% for ophthalmologists or their assistants, the authors of a study concluded that using appropriate technology and ensuring quality are more important than the type of provider in diabetic retinopathy screening programs [45].

We also explored factors predicting retinopathy occurrence among patients with diabetes. Increased duration of diabetes and obesity were identified as significant predictors of diabetic retinopathy. The odds of having diabetic retinopathy was found to increase 2 times among patients with diabetes for 5-10 years and 4 times among those with diabetes for more than 10 years compared to the odds of having diabetic retinopathy in those with diabetes for less than 5 years. While the increased duration of diabetes for the development of diabetic retinopathy is well-established [3,8,46,47], the observed negative association between obesity and diabetic retinopathy in this study was inconsistent with that reported in the literature. Only 1 study conducted in an urban slum in India reported a similar result; those in the overweight category had lower odds of diabetic retinopathy (OR 0.6, 95% CI 0.4-0.9) than people with normal BMI [31]. In general, however, obesity in the Asian context has been identified as a risk factor for developing diabetic retinopathy [1].

Although results from this pilot study are supportive of mainstreaming non-ophthalmologist health staff in diabetic retinopathy screening services, some limitations in the diagnostic processes need to be considered before generalizing this recommendation to other contexts. One important limitation was our decision to exclude patients with cataract and to employ

a mydriatic and single-field macular photographic approach due to the pilot nature of this research. In this regard, further clinical trials may be useful for determining the test accuracy by type of health professional, type of instrument used, use of pupil dilation, and number of fields chosen for screening. Finally, recommendations emerging from a study conducted in hospitals with specialized diabetes or eye departments may be less pertinent in nonspecialist hospitals where the prevalence of diabetic retinopathy is likely lower.

Considering the rising burden of diabetes in Bangladesh, routine retinal screening of patients with diabetes is not feasible by eye consultants alone. Our study results suggest that

non-ophthalmologist staff such as nurses and paramedics could function as frontline health workers in diabetic retinopathy screening programs. As the first step, the engagement of non-ophthalmologist cadres in diabetic retinopathy screening should be limited to categorization based on the presence or absence of diabetic retinopathy only. Their involvement in the more technical area of diabetic retinopathy grading requires further specific training and health system level support. Regarding instrument choice, although the desk-based camera shows better accuracy in detecting diabetic retinopathy, the choice of instrument type should be a function of the capacity of the health facility and the health care provider performing the diagnosis.

Acknowledgments

We received assistance from Md Rajibul Islam and Dr. Lal B Rawal, former Research Fellow and Senior Associate in Research of icddr,b, Dhaka, Bangladesh, respectively. We would like to acknowledge their contributions during the early phase of project implementation. The financial support to conduct this research was obtained from the Fred Hollow Foundation, Bangladesh through a competitive research grant proposal.

Authors' Contributions

TB, IA, AA, and AR developed the research proposal. TB, DN, and SI curated the data. TB, DN, and AR analyzed the data. All authors drafted, wrote, reviewed, and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Bourne RRA, Stevens GA, White RA, Smith JL, Flaxman SR, Price H, et al. Causes of vision loss worldwide, 1990–2010: a systematic analysis. *The Lancet Global Health* 2013 Dec;1(6):e339–e349. [doi: [10.1016/s2214-109x\(13\)70113-x](https://doi.org/10.1016/s2214-109x(13)70113-x)]
2. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, IDF Diabetes Atlas Committee. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9 edition. *Diabetes Res Clin Pract* 2019 Nov;157:107843. [doi: [10.1016/j.diabres.2019.107843](https://doi.org/10.1016/j.diabres.2019.107843)] [Medline: [31518657](https://pubmed.ncbi.nlm.nih.gov/31518657/)]
3. Yau JWY, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, Meta-Analysis for Eye Disease (META-EYE) Study Group. Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care* 2012 Mar;35(3):556–564 [FREE Full text] [doi: [10.2337/dc11-1909](https://doi.org/10.2337/dc11-1909)] [Medline: [22301125](https://pubmed.ncbi.nlm.nih.gov/22301125/)]
4. Klein BEK. Overview of epidemiologic studies of diabetic retinopathy. *Ophthalmic Epidemiol* 2007;14(4):179–183. [doi: [10.1080/09286580701396720](https://doi.org/10.1080/09286580701396720)] [Medline: [17896294](https://pubmed.ncbi.nlm.nih.gov/17896294/)]
5. Fong DS, Aiello L, Gardner TW, King GL, Blankenship G, Cavallerano JD, American Diabetes Association. Retinopathy in diabetes. *Diabetes Care* 2004 Jan;27 Suppl 1:S84–S87. [doi: [10.2337/diacare.27.2007.s84](https://doi.org/10.2337/diacare.27.2007.s84)] [Medline: [14693935](https://pubmed.ncbi.nlm.nih.gov/14693935/)]
6. Gupta V, Bansal R, Gupta A, Bhansali A. Sensitivity and specificity of nonmydriatic digital imaging in screening diabetic retinopathy in Indian eyes. *Indian J Ophthalmol* 2014 Aug;62(8):851–856 [FREE Full text] [doi: [10.4103/0301-4738.141039](https://doi.org/10.4103/0301-4738.141039)] [Medline: [25230960](https://pubmed.ncbi.nlm.nih.gov/25230960/)]
7. Raman R, Srinivasan S, Roy R. Screening practices for diabetic retinopathy. *Expert Review of Ophthalmology* 2015 Aug 03;10(6):519–521. [doi: [10.1586/17469899.2015.1073108](https://doi.org/10.1586/17469899.2015.1073108)]
8. Lin S, Ramulu P, Lamoureux EL, Sabanayagam C. Addressing risk factors, screening, and preventative treatment for diabetic retinopathy in developing countries: a review. *Clin Exp Ophthalmol* 2016 May;44(4):300–320. [doi: [10.1111/ceo.12745](https://doi.org/10.1111/ceo.12745)] [Medline: [26991970](https://pubmed.ncbi.nlm.nih.gov/26991970/)]
9. Piyasena MMPN, Murthy GVS, Yip JLY, Gilbert C, Zuurmond M, Peto T, et al. Systematic review on barriers and enablers for access to diabetic retinopathy screening services in different income settings. *PLoS One* 2019;14(4):e0198979 [FREE Full text] [doi: [10.1371/journal.pone.0198979](https://doi.org/10.1371/journal.pone.0198979)] [Medline: [31013274](https://pubmed.ncbi.nlm.nih.gov/31013274/)]
10. Piyasena MMPN, Yip JLY, MacLeod D, Kim M, Gudlavalleti VSM. Diagnostic test accuracy of diabetic retinopathy screening by physician graders using a hand-held non-mydratic retinal camera at a tertiary level medical clinic. *BMC Ophthalmol* 2019 Apr 08;19(1):89 [FREE Full text] [doi: [10.1186/s12886-019-1092-3](https://doi.org/10.1186/s12886-019-1092-3)] [Medline: [30961576](https://pubmed.ncbi.nlm.nih.gov/30961576/)]

11. Hutchinson A, McIntosh A, Peters J, O'Keeffe C, Khunti K, Baker R, et al. Effectiveness of screening and monitoring tests for diabetic retinopathy--a systematic review. *Diabet Med* 2000 Jul;17(7):495-506. [doi: [10.1046/j.1464-5491.2000.00250.x](https://doi.org/10.1046/j.1464-5491.2000.00250.x)] [Medline: [10972578](#)]
12. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016 Dec 13;316(22):2402-2410. [doi: [10.1001/jama.2016.17216](https://doi.org/10.1001/jama.2016.17216)] [Medline: [27898976](#)]
13. Bu D, Pan E, Walker J, Adler-Milstein J, Kendrick D, Hook JM, et al. Benefits of information technology-enabled diabetes management. *Diabetes Care* 2007 May;30(5):1137-1142. [doi: [10.2337/dc06-2101](https://doi.org/10.2337/dc06-2101)] [Medline: [17322483](#)]
14. Baeza M, Orozco-Beltrán D, Gil-Guillen VF, Pedrera V, Ribera MC, Pertusa S, et al. Screening for sight threatening diabetic retinopathy using non-mydratic retinal camera in a primary care setting: to dilate or not to dilate? *Int J Clin Pract* 2009 Mar;63(3):433-438. [doi: [10.1111/j.1742-1241.2008.01921.x](https://doi.org/10.1111/j.1742-1241.2008.01921.x)] [Medline: [19222628](#)]
15. International Diabetes Federation Diabetes Atlas eighth edition. 2017. URL: <https://www.diabetesatlas.org> [accessed 2019-11-07]
16. Performance Report (2009-2010) and HRD (2011)/HRH (2013,2014) Data Sheet. Ministry of Health and Family Welfare, Government of the People's Republic of Bangladesh. URL: http://www.mohfw.gov.bd/index.php?option=com_content&view=article&id=112 [accessed 2019-11-07]
17. Islam SMS, Tabassum R, Colet P, Cruz JP. Human resources for non-communicable diseases in Bangladesh. *International Journal of Perceptions in Public Health* 2017;1(2):98-101.
18. Sayeed MA, Rahman AS, Ali MH, Rhaman MM, Haq JA, Banu A. Diabetic retinopathy and visual impairment in disaster prone coastal population of Bangladesh. *IMC J Med Sci* 2017 Jan 12;10(1):10-17. [doi: [10.3329/imcjs.v10i1.31100](https://doi.org/10.3329/imcjs.v10i1.31100)]
19. Thapa R, Twyana SN, Paudyal G, Khanal S, van Nispen R, Tan HS, et al. Prevalence and risk factors of diabetic retinopathy among an elderly population with diabetes in Nepal: the Bhaktapur Retina Study. *Ophth* 2018 Mar; Volume 12:561-568. [doi: [10.2147/opth.s157560](https://doi.org/10.2147/opth.s157560)]
20. Malhotra R, Indrayan A. A simple nomogram for sample size for estimating sensitivity and specificity of medical tests. *Indian J Ophthalmol* 2010;58(6):519-522 [FREE Full text] [doi: [10.4103/0301-4738.71699](https://doi.org/10.4103/0301-4738.71699)] [Medline: [20952837](#)]
21. Harding S, Greenwood R, Aldington S, Gibson J, Owens D, Taylor R, Diabetic Retinopathy GradingDisease Management Working Party. Grading and disease management in national screening for diabetic retinopathy in England and Wales. *Diabet Med* 2003 Dec;20(12):965-971. [doi: [10.1111/j.1464-5491.2003.01077.x](https://doi.org/10.1111/j.1464-5491.2003.01077.x)] [Medline: [14632697](#)]
22. Hooper P, Boucher MC, Cruess A, Dawson KG, Delpero W, Greve M, et al. Canadian Ophthalmological Society evidence-based clinical practice guidelines for the management of diabetic retinopathy. *Can J Ophthalmol* 2012 Apr;47(2 Suppl):S1-30, S31. [doi: [10.1016/j.jcjo.2011.12.025](https://doi.org/10.1016/j.jcjo.2011.12.025)] [Medline: [22632804](#)]
23. Global strategy on diet, physical activity and health: a framework to monitor and evaluate implementation. World Health Organization. 2006. URL: <https://apps.who.int/iris/handle/10665/43524> [accessed 2017-05-06]
24. American Diabetes Association. Classification and Diagnosis of Diabetes. *Diabetes Care* 2020 Jan;43(Suppl 1):S14-S31. [doi: [10.2337/dc20-S002](https://doi.org/10.2337/dc20-S002)] [Medline: [31862745](#)]
25. Trevethan R. Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice. *Front Public Health* 2017;5:307 [FREE Full text] [doi: [10.3389/fpubh.2017.00307](https://doi.org/10.3389/fpubh.2017.00307)] [Medline: [29209603](#)]
26. Florkowski C. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *Clin Biochem Rev* 2008 Aug;29 Suppl 1:S83-S87. [Medline: [18852864](#)]
27. Citing Stata software, documentation, and FAQs. STATACorp Release 16. College Station, TX: StataCorp LLC; 2019. URL: <https://www.stata.com/support/faqs/resources/citing-software-documentation-faqs/> [accessed 2019-11-19]
28. Diabetes retinopathy prevalent in Bangladesh. 2014. URL: <https://www.diabetes.co.uk/news/2014/oct/diabetes-retinopathy-prevalent-in-bangladesh-99436687.html> [accessed 2019-09-16]
29. Stevens GA, White RA, Flaxman SR, Price H, Jonas JB, Keeffe J, Vision Loss Expert Group. Global prevalence of vision impairment and blindness: magnitude and temporal trends, 1990-2010. *Ophthalmology* 2013 Dec;120(12):2377-2384. [doi: [10.1016/j.ophtha.2013.05.025](https://doi.org/10.1016/j.ophtha.2013.05.025)] [Medline: [23850093](#)]
30. Piyasena MMPN, Murthy GVS, Yip JLY, Gilbert C, Peto T, Gordon I, et al. Systematic review and meta-analysis of diagnostic accuracy of detection of any level of diabetic retinopathy using digital retinal imaging. *Syst Rev* 2018 Nov 07;7(1):182 [FREE Full text] [doi: [10.1186/s13643-018-0846-y](https://doi.org/10.1186/s13643-018-0846-y)] [Medline: [30404665](#)]
31. Wadhvani M, Vashist P, Singh SS, Gupta N, Malhotra S, Gupta A, et al. Diabetic retinopathy screening programme utilising non-mydratic fundus imaging in slum populations of New Delhi, India. *Trop Med Int Health* 2018 Apr;23(4):405-414 [FREE Full text] [doi: [10.1111/tmi.13039](https://doi.org/10.1111/tmi.13039)] [Medline: [29430785](#)]
32. Fahadullah M, Memon N, Salim S, Ahsan S, Fahim MF, Mumtaz SN, et al. Diagnostic accuracy of non-mydratic fundus camera for screening of diabetic retinopathy: A hospital based observational study in Pakistan. *J Pak Med Assoc* 2019 Mar;69(3):378-382 [FREE Full text] [Medline: [30890831](#)]
33. Martínez-Vizcaino V, Cavero-Redondo I, Álvarez-Bueno C, Rodríguez-Artalejo F. The Accuracy of Diagnostic Methods for Diabetic Retinopathy: A Systematic Review and Meta-Analysis. *PLoS One* 2016;11(4):e0154411 [FREE Full text] [doi: [10.1371/journal.pone.0154411](https://doi.org/10.1371/journal.pone.0154411)] [Medline: [27123641](#)]

34. Scarpa G, Urban F, Vujosevic S, Tessarin M, Gallo G, Visentin A, et al. The Nonmydriatic Fundus Camera in Diabetic Retinopathy Screening: A Cost-Effective Study with Evaluation for Future Large-Scale Application. *J Ophthalmol* 2016;2016:4625096 [FREE Full text] [doi: [10.1155/2016/4625096](https://doi.org/10.1155/2016/4625096)] [Medline: [27885337](https://pubmed.ncbi.nlm.nih.gov/27885337/)]
35. Ahsan S, Basit A, Ahmed KR, Ali L, Shaheen F, Ulhaque MS, et al. Diagnostic accuracy of direct ophthalmoscopy for detection of diabetic retinopathy using fundus photographs as a reference standard. *Diabetes Metab Syndr* 2014;8(2):96-101. [doi: [10.1016/j.dsx.2014.04.015](https://doi.org/10.1016/j.dsx.2014.04.015)] [Medline: [24907174](https://pubmed.ncbi.nlm.nih.gov/24907174/)]
36. Khairallah M, Kahloun R, Bourne R, Limburg H, Flaxman SR, Jonas JB, Vision Loss Expert Group of the Global Burden of Disease Study. Number of People Blind or Visually Impaired by Cataract Worldwide and in World Regions, 1990 to 2010. *Invest Ophthalmol Vis Sci* 2015 Oct;56(11):6762-6769. [doi: [10.1167/iov.15-17201](https://doi.org/10.1167/iov.15-17201)] [Medline: [26567788](https://pubmed.ncbi.nlm.nih.gov/26567788/)]
37. Roser P, Kalscheuer H, Groener JB, Lehnhoff D, Klein R, Auffarth GU, et al. Diabetic Retinopathy Screening Ratio Is Improved When Using a Digital, Nonmydriatic Fundus Camera Onsite in a Diabetes Outpatient Clinic. *J Diabetes Res* 2016;2016:4101890 [FREE Full text] [doi: [10.1155/2016/4101890](https://doi.org/10.1155/2016/4101890)] [Medline: [26904690](https://pubmed.ncbi.nlm.nih.gov/26904690/)]
38. Li HK, Horton M, Bursell SE, Cavallerano J, Zimmer-Galler I, Tennant M, American Telemedicine Association Diabetic Retinopathy Telehealth Practice Recommendations Working Group, et al. Telehealth practice recommendations for diabetic retinopathy, second edition. *Telemed J E Health* 2011 Dec;17(10):814-837 [FREE Full text] [doi: [10.1089/tmj.2011.0075](https://doi.org/10.1089/tmj.2011.0075)] [Medline: [21970573](https://pubmed.ncbi.nlm.nih.gov/21970573/)]
39. Perrier M, Boucher MC, Angioi K, Gresset JA, Olivier S. Comparison of two, three and four 45° image fields obtained with the Topcon CRW6 nonmydriatic camera for screening for diabetic retinopathy. *Canadian Journal of Ophthalmology* 2003 Dec;38(7):569-574. [doi: [10.1016/s0008-4182\(03\)80110-2](https://doi.org/10.1016/s0008-4182(03)80110-2)]
40. Liesenfeld B, Kohner E, Piehlmeier W, Kluthe S, Aldington S, Porta M, et al. A telemedical approach to the screening of diabetic retinopathy: digital fundus photography. *Diabetes Care* 2000 Mar;23(3):345-348 [FREE Full text] [doi: [10.2337/diacare.23.3.345](https://doi.org/10.2337/diacare.23.3.345)] [Medline: [10868863](https://pubmed.ncbi.nlm.nih.gov/10868863/)]
41. Quellec G, Bazin L, Cazuguel G, Delafoy I, Cochener B, Lamard M. Suitability of a Low-Cost, Handheld, Nonmydriatic Retinograph for Diabetic Retinopathy Diagnosis. *Transl Vis Sci Technol* 2016 Apr;5(2):16 [FREE Full text] [doi: [10.1167/tvst.5.2.16](https://doi.org/10.1167/tvst.5.2.16)] [Medline: [27134775](https://pubmed.ncbi.nlm.nih.gov/27134775/)]
42. Williams GA, Scott IU, Haller JA, Maguire AM, Marcus D, McDonald H. Single-field fundus photography for diabetic retinopathy screening: a report by the American Academy of Ophthalmology. *Ophthalmology* 2004 May;111(5):1055-1062. [doi: [10.1016/j.ophtha.2004.02.004](https://doi.org/10.1016/j.ophtha.2004.02.004)] [Medline: [15121388](https://pubmed.ncbi.nlm.nih.gov/15121388/)]
43. Šimundić AM. Measures of Diagnostic Accuracy: Basic Definitions. *EJIFCC* 2009 Jan;19(4):203-211 [FREE Full text] [Medline: [27683318](https://pubmed.ncbi.nlm.nih.gov/27683318/)]
44. Ahmed J, Ward TP, Bursell S, Aiello LM, Cavallerano JD, Vigersky RA. The sensitivity and specificity of nonmydriatic digital stereoscopic retinal imaging in detecting diabetic retinopathy. *Diabetes Care* 2006 Oct;29(10):2205-2209. [doi: [10.2337/dc06-0295](https://doi.org/10.2337/dc06-0295)] [Medline: [17003294](https://pubmed.ncbi.nlm.nih.gov/17003294/)]
45. Montori VM. Mydriatic retinal photography is the most effective test for detecting diabetic retinopathy. *BMJ* 2001 Mar;6(2):56-56. [doi: [10.1136/ebm.6.2.56](https://doi.org/10.1136/ebm.6.2.56)]
46. Varma R, Choudhury F, Klein R, Chung J, Torres M, Azen SP, Los Angeles Latino Eye Study Group. Four-year incidence and progression of diabetic retinopathy and macular edema: the Los Angeles Latino Eye Study. *Am J Ophthalmol* 2010 May;149(5):752-61.e1 [FREE Full text] [doi: [10.1016/j.ajo.2009.11.014](https://doi.org/10.1016/j.ajo.2009.11.014)] [Medline: [20149342](https://pubmed.ncbi.nlm.nih.gov/20149342/)]
47. Ding J, Wong TY. Current epidemiology of diabetic retinopathy and diabetic macular edema. *Curr Diab Rep* 2012 Aug;12(4):346-354. [doi: [10.1007/s11892-012-0283-6](https://doi.org/10.1007/s11892-012-0283-6)] [Medline: [22585044](https://pubmed.ncbi.nlm.nih.gov/22585044/)]

Abbreviations

AUROC: area under the receiver operating curve
DAB: Diabetic Association of Bangladesh
FHF: Fred Hollows Foundation
OR: odds ratio

Edited by T Sanchez; submitted 15.08.20; peer-reviewed by S Rahman, G Lim; comments to author 11.09.20; revised version received 01.11.20; accepted 16.12.20; published 09.03.21.

Please cite as:

Begum T, Rahman A, Nomani D, Mamun A, Adams A, Islam S, Khair Z, Khair Z, Anwar I
 Diagnostic Accuracy of Detecting Diabetic Retinopathy by Using Digital Fundus Photographs in the Peripheral Health Facilities of Bangladesh: Validation Study
JMIR Public Health Surveill 2021;7(3):e23538
 URL: <https://publichealth.jmir.org/2021/3/e23538>
 doi:[10.2196/23538](https://doi.org/10.2196/23538)
 PMID:[33411671](https://pubmed.ncbi.nlm.nih.gov/33411671/)

©Tahmina Begum, Aminur Rahman, Dilruba Nomani, Abdullah Mamun, Alayne Adams, Shafiqul Islam, Zara Khair, Zareen Khair, Iqbal Anwar. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 09.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>