

JMIR Public Health and Surveillance

A multidisciplinary journal that focuses on public health and technology, public health informatics, mass media campaigns, surveillance, participatory epidemiology, and innovation in public health practice and research

Volume 5 (2019), Issue 1 ISSN: 2369-2960 Editor in Chief: Travis Sanchez, PhD, MPH

Contents

Original Papers

Analytics for Investigation of Disease Outbreaks: Web-Based Analytics Facilitating Situational Awareness in Unfolding Disease Outbreaks (e12032) Nileena Velappan, Ashlynn Daughton, Geoffrey Fairchild, William Rosenberger, Nicholas Generous, Maneesha Chitanvis, Forest Altherr, Lauren Castro, Reid Priedhorsky, Esteban Abeyta, Leslie Naranjo, Attelia Hollander, Grace Vuyisich, Antonietta Lillo, Emily Cloyd, Ashvini Vaidya, Alina Deshpande.	4
Differences in Emotional and Pain-Related Language in Tweets About Dentists and Medical Doctors: Text Analysis of Twitter Content (e10432) Jan-Are Johnsen, Trude Eggesvik, Thea Rørvik, Miriam Hanssen, Rolf Wynn, Per Kummervold.	22
Forecasting the West Nile Virus in the United States: An Extensive Novel Data Streams–Based Time Series Analysis and Structural Equation Modeling of Related Digital Searching Behavior (e9176) Abdulla Watad, Samaa Watad, Naim Mahroum, Kassem Sharif, Howard Amital, Nicola Bragazzi, Mohammad Adawi.	31
Google Trends Predicts Present and Future Plague Cases During the Plague Outbreak in Madagascar: Infodemiological Study (e13142) Nicola Bragazzi, Naim Mahroum.	42
Generating Engagement on the Make Healthy Normal Campaign Facebook Page: Analysis of Facebook Analytics (e11132) James Kite, Anne Grunseit, Vincy Li, John Vineburg, Nathan Berton, Adrian Bauman, Becky Freeman.	50
Spatial Access and Willingness to Use Pre-Exposure Prophylaxis Among Black/African American Individuals in the United States: Cross-Sectional Survey (e12405) Bisola Ojikutu, Laura Bogart, Kenneth Mayer, Thomas Stopka, Patrick Sullivan, Yusuf Ransome.	72
Fraud Detection Protocol for Web-Based Research Among Men Who Have Sex With Men: Development and Descriptive Evaluation (e12344) April Ballard, Trey Cardwell, April Young.	80
Barriers to Implementation of Perinatal Death Audit in Maternity and Pediatric Hospitals in Jordan: Cross-Sectional Study (e11653) Yousef Khader, Mohammad Alyahya, Anwar Batieha.	90

Applying National Estimates of Adults With Indications for Pre-Exposure Prophylaxis to Populations of Men Who Have Sex With Men and People Who Inject Drugs in Colorado: Modeling Study (e11113)	110
Jennifer Donnelly, Thomas Deem, Megan Duffy, Anita Watkins, Alia Al-Tayyib, Daniel Shodell, Mark Thrun, Sarah Rowan.	
Programmatic Mapping to Estimate Size, Distribution, and Dynamics of Key Populations in Kosovo (e11194)	120
Dafina Gexha Bunjaku, Edona Deva, Luljeta Gashi, Pranvera Kaçaniku-Gunga, Carly Comins, Faran Emmanuel.	
Population Size Estimation of Venue-Based Female Sex Workers in Ho Chi Minh City, Vietnam: Capture-Recapture Exercise (e10906)	
Giang Le, Nghia Khuu, Van Tieu, Phuc Nguyen, Hoa Luong, Quang Pham, Hau Tran, Thuong Nguyen, Meade Morgan, Abu Abdul-Quader.	
1 3 5	
Estimating the Population Size of Female Sex Workers in Namibia Using a Respondent-Driven Sampling Adjustment to the Reverse Tracking Method: A Novel Approach (e11737)	142
Paul Wesson, Rajatashuvra Adhikary, Anna Jonas, Krysta Gerndt, Ali Mirzazadeh, Frieda Katuta, Andrew Maher, Karen Banda, Nicholus Mutenda, Willi McFarland, David Lowrance, Dimitri Prybylski, Sadhna Patel.	
Population Size Estimations Among Hidden Populations Using Respondent-Driven Sampling Surveys: Case Studies From Armenia (e12034)	154
Katherine McLaughlin, Lisa Johnston, Laura Gamble, Trdat Grigoryan, Arshak Papoyan, Samvel Grigoryan.	
Novel Approaches for Estimating Female Sex Worker Population Size in Conflict-Affected South Sudan (e11576)	166
Alfred Okiria, Alex Bolo, Victoria Achut, Golda Arkangelo, Acaga Michael, Joel Katoro, Jennifer Wesson, Steve Gutreuter, Lee Hundley, Avi Hakim.	
Kuantim mi tu ("Count me too"): Using Multiple Methods to Estimate the Number of Female Sex Workers, Men Who Have Sex With Men, and Transgender Women in Papua New Guinea in 2016 and 2017 (e11285)	174
Damian Weikum, Angela Kelly-Hanku, Parker Hou, Martha Kupul, Angelyne Amos-Kuma, Steven Badman, Nick Dala, Kelsey Coy, John Kaldor, Andrew Vallely, Avi Hakim.	
Period of Measurement in Time-Series Predictions of Disease Counts from 2007 to 2017 in Northern Nevada: Analytics Experiment (e11357)	184
Amir Talaie-Khoei, James Wilson, Seyed-Farzan Kazemi.	
Relationship Between Media Coverage and Measles-Mumps-Rubella (MMR) Vaccination Uptake in Denmark: Retrospective Study (e9544)	210
Niels Hansen, Kåre Mølbak, Ingemar Cox, Christina Lioma.	
Characterizing Trends in Human Papillomavirus Vaccine Discourse on Reddit (2007-2015): An Observational Study (e12480)	222
Yuki Lama, Dian Hu, Amelia Jamison, Sandra Quinn, David Broniatowski.	
A Syndrome-Based Surveillance System for Infectious Diseases Among Asylum Seekers in Austrian Reception Centers, 2015-2018: Analysis of Reported Data (e11465)	235
Ziad El-Khatib, Karin Taus, Lukas Richter, Franz Allerberger, Daniela Schmid.	
Evaluating the Usefulness of Translation Technologies for Emergency Response Communication: A Scenario-Based Study (e11171)	247
Anne Turner, Yong Choi, Kristin Dew, Ming-Tse Tsai, Alyssa Bosold, Shuyang Wu, Donahue Smith, Hendrika Meischke.	

Review

- Considering the Impact of Social Media on Contemporary Improvement of Australian Aboriginal Health: Scoping Review ([e11573](#))
Troy Walker, Claire Palermo, Karen Klassen. 64

Rapid Surveillance Report

- The Annual American Men's Internet Survey of Behaviors of Men Who Have Sex With Men in the United States: 2016 Key Indicators Report ([e11313](#))
Maria Zlotorzynska, Patrick Sullivan, Travis Sanchez. 98

Viewpoints

- The Continuing Value of CD4 Cell Count Monitoring for Differential HIV Care and Surveillance ([e11136](#))
Brian Rice, Andrew Boule, Sandra Schwarcz, Amir Shroufi, George Rutherford, James Hargreaves. 129
- Innovative Approaches to Improve Public Health Practice in the Eastern Mediterranean Region: Findings From the Sixth Eastern Mediterranean Public Health Network Regional Conference ([e11382](#))
Bashiruddin Noormal, Elmuez Eltayeb, Mohannad Al Nsour, Ezzeddine Mohsni, Yousef Khader, Mark Salter, Scott McNabb, Dionisio Herrera Guibert, Salman Rawaf, Amrsh Baidjoe, Aamer Ikram, Christophe Longuet, Abdulwahed Al Serouri, Faris Lami, Asmae Khattabi, Sami AlMudarra, Ibrahim Iblan, Sahar Samy, Nissaf Bouafif ép Ben Alaya, Qahtan Al-Salihi. 199

Original Paper

Analytics for Investigation of Disease Outbreaks: Web-Based Analytics Facilitating Situational Awareness in Unfolding Disease Outbreaks

Nileena Velappan^{1*}, MS; Ashlynn Rae Daughton^{1*}, MPH; Geoffrey Fairchild¹, PhD; William Earl Rosenberger¹, BS; Nicholas Generous¹, MPH; Maneesha Elizabeth Chitanvis¹, MPH; Forest Michael Altherr¹, MPH; Lauren A Castro¹, BA; Reid Priedhorsky¹, PhD; Esteban Luis Abeyta¹; Leslie A Naranjo^{1,2}, MS; Attelia Dawn Hollander¹, BS; Grace Vuyisich¹, BS; Antonietta Maria Lillo¹, PhD; Emily Kathryn Cloyd^{1,3}, MPH; Ashvini Rajendra Vaidya¹, BS; Alina Deshpande¹, PhD

¹Los Alamos National Laboratory, Los Alamos, NM, United States

²Specifica Inc, New Mexico Consortium Biological Laboratory, Los Alamos, NM, United States

³University of Virginia, Charlottesville, VA, United States

*these authors contributed equally

Corresponding Author:

Alina Deshpande, PhD
Los Alamos National Laboratory
TA-43, bldg1, MS - M888
Los Alamos, NM, 87545
United States
Phone: 1 505 667 9938
Email: deshpande_a@lanl.gov

Abstract

Background: Information from historical infectious disease outbreaks provides real-world data about outbreaks and their impacts on affected populations. These data can be used to develop a picture of an unfolding outbreak in its early stages, when incoming information is sparse and isolated, to identify effective control measures and guide their implementation.

Objective: This study aimed to develop a publicly accessible Web-based visual analytic called Analytics for the Investigation of Disease Outbreaks (AIDO) that uses historical disease outbreak information for decision support and situational awareness of an unfolding outbreak.

Methods: We developed an algorithm to allow the matching of unfolding outbreak data to a representative library of historical outbreaks. This process provides epidemiological clues that facilitate a user's understanding of an unfolding outbreak and facilitates informed decisions about mitigation actions. Disease-specific properties to build a complete picture of the unfolding event were identified through a data-driven approach. A method of analogs approach was used to develop a short-term forecasting feature in the analytic. The 4 major steps involved in developing this tool were (1) collection of historic outbreak data and preparation of the representative library, (2) development of AIDO algorithms, (3) development of user interface and associated visuals, and (4) verification and validation.

Results: The tool currently includes representative historical outbreaks for 39 infectious diseases with over 600 diverse outbreaks. We identified 27 different properties categorized into 3 broad domains (population, location, and disease) that were used to evaluate outbreaks across all diseases for their effect on case count and duration of an outbreak. Statistical analyses revealed disease-specific properties from this set that were included in the disease-specific similarity algorithm. Although there were some similarities across diseases, we found that statistically important properties tend to vary, even between similar diseases. This may be because of our emphasis on including diverse *representative* outbreak presentations in our libraries. AIDO algorithm evaluations (similarity algorithm and short-term forecasting) were conducted using 4 case studies and we have shown details for the Q fever outbreak in Bilbao, Spain (2014), using data from the early stages of the outbreak. Using data from only the initial 2 weeks, AIDO identified historical outbreaks that were very similar in terms of their epidemiological picture (case count, duration, source of exposure, and urban setting). The short-term forecasting algorithm accurately predicted case count and duration for the unfolding outbreak.

Conclusions: AIDO is a decision support tool that facilitates increased situational awareness during an unfolding outbreak and enables informed decisions on mitigation strategies. AIDO analytics are available to epidemiologists across the globe with access to internet, at no cost. In this study, we presented a new approach to applying historical outbreak data to provide actionable information during the early stages of an unfolding infectious disease outbreak.

(*JMIR Public Health Surveill* 2019;5(1):e12032) doi:[10.2196/12032](https://doi.org/10.2196/12032)

KEYWORDS

epidemiology; infectious diseases; algorithm; public health informatics; web browser

Introduction

Challenges in Outbreak Investigation

Infectious diseases continue to be a leading cause of mortality worldwide despite substantial advances in public health [1]. Disease outbreaks are not bound by national borders and can have far-reaching economic and social impacts. Therefore, early detection and monitoring are key to curtailing unfolding outbreaks. Tools and analytics that improve situational awareness can aid communication in the initial stages of an outbreak and enable effective decisions for outbreak control [2].

Traditionally, historical outbreak data have been used to enhance and develop disease forecasting models. For example, when modeling influenza, Viboud et al utilized the method of analogs, which uses weighted vectors of historical time series data that match current activity to build forecasts 1 to 10 weeks ahead [3]. Relatedly, Sugihara and May utilized a library of historical measles and chickenpox outbreaks to understand historical patterns in variation and develop short-term forecasting models [4]. In addition, extrapolation of retrospective data has proven useful in resource-poor areas, for example, to establish levels of antimicrobial resistance in areas with minimal surveillance [5]. More broadly, historical disease data are often used for model parameter estimation [6-8]. Data used are often not confined to epidemiological data. For example, climate data have been used to develop outbreak models that are affected by environmental factors, such as malaria, dengue, and cholera [9]. However, these data are typically difficult to use owing to poor organization and integration. Extracting relevant information from official reports is time consuming.

Epidemiological data are rarely easily accessible or available in an easily analyzed format. There are efforts to circumvent these issues, such as Project Tycho [10,11] and Gideon [12], but these Web applications have limitations. Project Tycho is currently limited to data from the National Notifiable Disease Surveillance System in the United States and Gideon offers its data for a fee, which may prove to be prohibitive for some public health users. Furthermore, neither provide information to contextualize or describe historical outbreaks nor the tools to meaningfully relate a present situation to a past one.

Rapid, Facile Decision Support Using Historical Data

To the authors' knowledge, there have been no previous formal attempts to build a decision-support application based on historical data. However, historical analogies have been utilized for disease projection and forecasting. In 2010, Haiti experienced its first cholera outbreak in over 100 years, months

after being struck by an earthquake that damaged its infrastructure and gave rise to poor sanitary conditions [13]. In the days following the first case notifications, the US Centers for Disease Control and Prevention (CDC) contextualized the Haiti outbreak by comparing it with a similar outbreak in Peru [14]. As more surveillance data became available and forecasting models were developed, more complex projections were established. However, quick comparisons, such as the historical analogy model that provided a notification to the Ministere de la Sante Publique de la Population of Haiti of the need to prepare for a large epidemic [15].

Currently, internet access is widely available around the globe. More than 40% of the world's population has access to the World Wide Web [16]. Hence, Web-based analytics that facilitate outbreak investigation have the potential to improve outbreak control around the world. In this study, we presented a Web-based visual analytic called Analytics for the Investigation of Disease Outbreaks (AIDO) available under the domain name bsvgateway.org [17], which has been developed to facilitate situational awareness in the early stages of an infectious disease outbreak. AIDO allows matching of unfolding outbreak data to a representative library of historical outbreaks and provides epidemiological clues that facilitate a user's understanding of an unfolding outbreak and enables informed decisions on mitigation actions. This tool currently has analytics for 28 infectious diseases and contains a *browse library* for an additional 11 diseases. We described the methodology used to develop this tool and illustrated the utility of the tool using 4 case studies, one of which is described in detail. We offer AIDO as a rapid, easy-to-use no cost analytic for outbreak investigation and short-term forecasting.

Methods

To develop AIDO, we used the following iterative process: (1) collect historical outbreak data; (2) develop AIDO algorithms; (3) develop the user interface, additional visuals, and functionalities; and (4) perform verification and validation. In this section, we have described the process for developing a disease-specific representative historical outbreak library, the associated algorithms, and the AIDO interface.

Collect Historical Outbreak Data

AIDO contains representative outbreak data for 39 diseases ([Multimedia Appendix 1](#)). We defined a representative library as one that includes outbreaks with a broad range of cumulative case counts, outbreak durations, diverse circumstances, and which occur in a variety of locations. Outbreak data were identified using official academic and government data, as well

as retrospective studies. Data sources included ProMED [18,19], CDC [20], World Health Organization [21,22], Eurosurveillance [23], European Centre for Disease Prevention and Control Disease Reports [24], and government Ministry of Health databases (available from Biosurveillance Resource Directory under the domain name bsvgateway.org) [25], as well as other scholarly journals available on PubMed and Google Scholar. If data were only available in the graphical form (eg, graphs in a PDF report or peer-reviewed publication), plots were digitized using PlotDigitizer [26].

To be considered for inclusion in a disease library, outbreak data must contain (1) a time series of case counts, (2) enough associated data to perform property analysis (described below), and (3) enough metadata to annotate the outbreak (described below). To apply analytics to a disease library, there must be (1) a minimum of 10 outbreaks included per disease library and (2) sufficient data to complete the property identification protocol (described below). Our analytics rely on a robust library; therefore, it is necessary to have a minimum threshold. A total of 10 outbreaks were considered to be a reasonable lower limit and produced reasonable results. In general, outbreaks that occurred during or after the year 2000 were prioritized for inclusion to represent current natural and built environments. However, in cases of rare diseases, outbreaks from previous years were included to achieve the minimum threshold for analysis (eg, both Ebola and Marburg libraries include outbreaks that occurred before 2000).

In addition to the outbreak time series, detailed information on factors that influenced the outbreak was also collected and used to describe each outbreak. Information collected included index case, important dates, the vector (if applicable), transmission routes, pathogen classification, case definition, geographic and

historical information, and identified risk factors and control measures that were implemented.

Algorithm Development

Similarity Algorithm

The similarity algorithm identifies outbreaks similar to the user's unfolding situation through a similarity score that is computed as a sum of values assigned to weighted properties specific to a disease. The algorithm has 3 components: (1) disease-specific properties; (2) weights calculated for each property; and (3) computation of the weighted sum by the AIDO algorithm. Here, we have described the statistical process used to identify properties, the procedure used to weight properties for relative importance, and the final equation used in this algorithm. Diseases with less than 10 outbreaks do not have an associated similarity algorithm and are represented in AIDO as *browse-only*.

Property selection: In AIDO, properties are characteristics that influence the case count or duration of outbreaks. In essence, these are the population-level signatures that help match a user's situation to outbreaks in our library. There were 3 types of properties: (1) categorical (eg, vaccination status: 90% to 100%, 80% to 89%, 50% to 79%, and less than 50%), (2) continuous (eg, physician density: range of values 0.1 to 10), and (3) binary (eg, population movement: yes or no). Properties were discretized if extant literature supported categorization of continuous variables. We identified 27 different properties that were used to evaluate outbreaks across all diseases for their effect on case count and duration of an outbreak. These properties and their definitions are included in [Table 1](#). Properties were further categorized into 3 domains.

Table 1. Properties collected for statistical analysis.

Name (variable type)	Description
Domain: population	
Population (continuous)	Population of affected location as a continuous variable
Population (categorical)	Population of affected location. Discretized into groups based on orders of magnitude
Disease status (binary)	Endemicity status (ie, endemic or nonendemic to the region) during the time of the outbreak
Rural versus urban (binary)	Binary categorization of the relative population density of the outbreak's location
Age stratification (categorical)	Relevant age categories or median age of reported cases. (Varies by disease; groupings are identified using published literature)
Special population group (binary)	Binary (yes or no) indicator describing if the outbreak occurred in the general population or a particular group with a specific common exposure or risk factor
Vaccination status (categorical)	Vaccine coverage (%) of the country and/or region of interest
Population movement (binary)	Indication of whether or not large-scale population movement (eg, mass migration and influx of a refugee population) was an influential component of the outbreak
Sex (continuous)	Fraction of cases in males (identified using the literature)
Domain: location	
Climate (categorical)	Climate type corresponding to the location of interest, represented as the first letter of the Köppen-Geiger climate classification key (A, B, C, D, and E) [27,28]
Season (categorical)	Time of year (Spring, Summer, Autumn, and Winter) during which the majority of the outbreak occurred
Precipitation (categorical)	Precipitation category corresponding to the location of interest, represented as the second letter in the Köppen-Geiger climate classification (f, m, w, s, W, S, T, and F)
Rainy versus dry (binary)	Binary indicator describing the typical weather patterns (ie, rainy or dry) in the location at the start of the outbreak
Natural disaster (binary)	Binary indicator describing if a natural disaster appeared to be associated with the onset of the outbreak
Human Development Index (HDI; categorical or continuous)	HDI in the location and year of interest [29]. Both categorical and continuous values were tested. In the event that both properties were significantly related to outcomes, categorical values were preferred because of they are more usable within the user interface.
Physician density (continuous)	Physician density per 1000 persons in the year of interest, or the most recent year reported [30]
Domain: disease	
Pathogen source (categorical)	Main source of exposure to the pathogen of interest
Outbreak curve (categorical)	Type of outbreak as reflected in the outbreak curve shape (point source, common source, and propagated outbreak)
Vector type (categorical)	The most relevant genus/species/classification of the disease vector
Case fatality rate (CFR; continuous)	Percent of fatal cases
Attack rate (continuous)	Number of new cases per 1000 persons
Case definition (categorical)	Classification of reported cases (suspected, probable, confirmed, or any combination thereof)
Disease presentation classification (categorical)	Description of clinical disease presentation (eg, bubonic plague and pneumonic plague)
Animal contact (binary or categorical)	Reported contact with potentially infectious animal (used for zoonotic diseases). Can be binary (yes or no) or categorical (ie, contact with particular animal), depending on the level of data available in literature
Contamination source (categorical)	Product or site epidemiologically linked to the outbreak (used for foodborne illnesses)
Transmission mode (categorical)	Mode of transmission that best characterizes the majority of disease spread during the outbreak (eg, airborne and direct contact)
Outbreak source proximity (categorical)	Geographic proximity of cases to a known or likely source of contamination
Outbreak pathogen (categorical)	Etiological agent

Available data for all properties were collected for each outbreak in AIDO. Statistical analyses were used to identify which properties separated outbreaks on the basis of outbreak size or duration for use in the similarity algorithm described in the following equation.



Outbreak similarity scores are generated using this simple weighted sum such that $0 \leq w_i, m_i \leq 1$ and the sum of all $w_i = 1$, which ensure that $0 \leq s \leq 1$. Here, s is the outbreak's similarity score, K is the number of parameters considered, w_i is the weight of parameter i , and m_i is the outbreak's match score of parameter i (ie, how well the outbreak's value for parameter i matches the user's value provided in the query form). Note that although this equation returns a score, s , between 0 and 1, AIDO displays scores as percentages (ie, $s \cdot 100$).

Statistical analyses were conducted to determine which properties to include in the disease specific similarity algorithm based on relationships with outbreak magnitude and duration. A conclusion that there is a meaningful association (ie, statistical significance or a strong enough correlation) indicates that this property helps to distinguish outbreaks of different magnitude or length from one another, and thus can be used to find outbreaks most similar to a user's data. The results from the property analysis do not reflect or intend to imply any causative relationship between a property and outbreaks of a given disease. A statistical association is merely reflective of a property's relative ability to discriminate between outbreaks of different sizes and lengths.

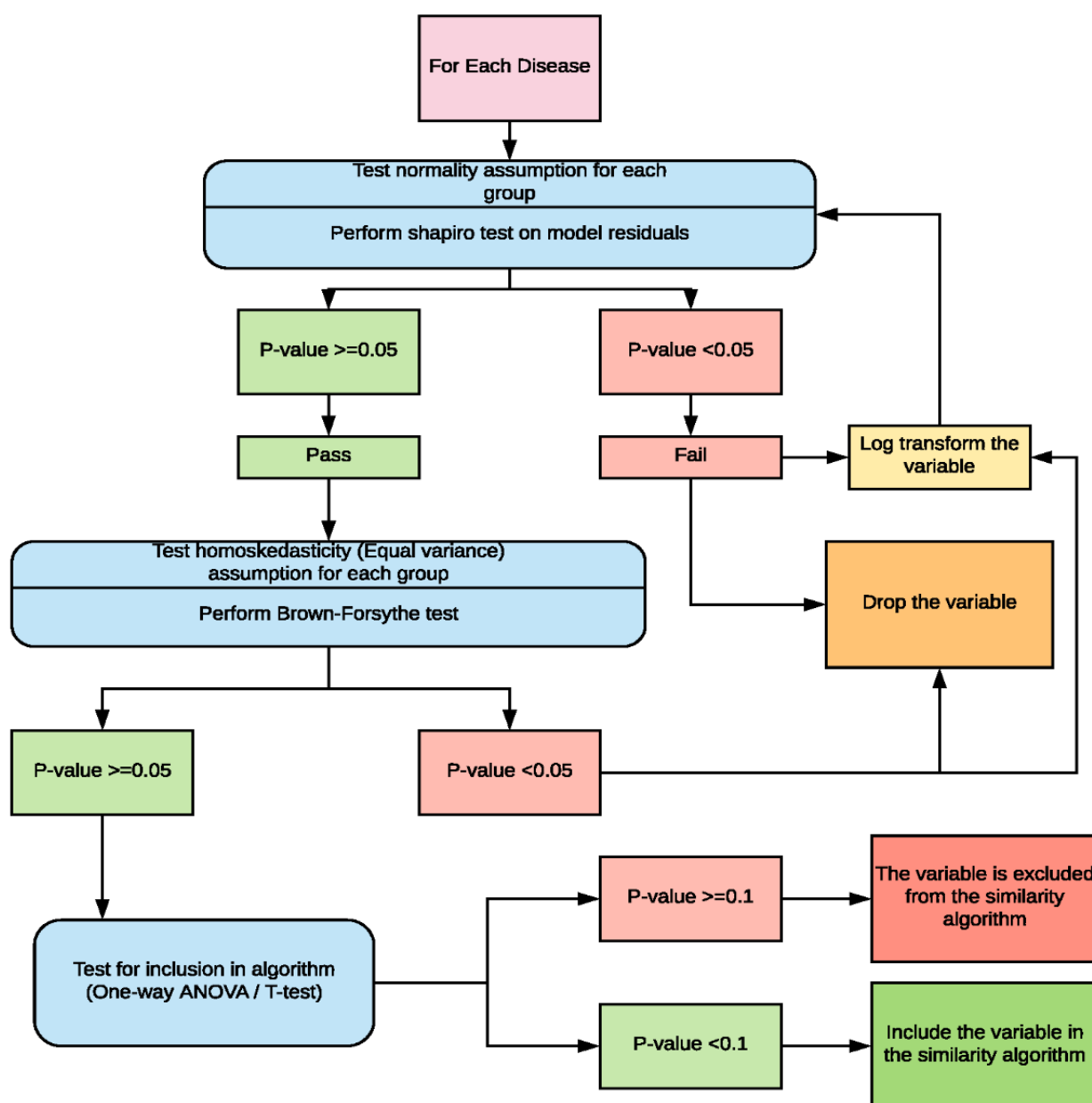
To perform the statistical analysis, a series of statistical tests were automated using R (R-foundation) [31]. Properties were segregated into their variable type: binary, categorical (multilevel), or continuous. Figure 1 illustrates the process by which these statistical analyses were completed for categorical variables. For a given disease, after identifying all properties listed in Table 1, statistical assessments first measured whether or not the data met the assumptions of normality for the distribution of residuals as well as equality of variance (homoscedasticity) by performing the Shapiro-Wilk and Brown-Forsythe tests, respectively. For continuous variables, only the distribution of the residuals was assessed, and for instances where the distribution was normal (Shapiro-Wilk test), a Pearson correlation was run. A Spearman correlation was performed on properties with non-normal distributions. We assume that all values for dependent property variables (both case count and duration) are independent, as we did not have any prior knowledge that these values are dependent on one

another in any way. As our data are curated from available scientific literature and published official reports from different locations across different years, we assume all individual values to be independent from other values and outbreaks in our library.

Following assessments of normality and homoscedasticity, statistical tests of association, depending on variable type, were performed (either T-test, one-way analysis of variance (ANOVA), or Pearson or Spearman correlation) to assess each property's relationship with both case count and duration, independently. t test and ANOVA F test results were assessed at a significance level of 0.05, whereas a correlation above 0.30 was considered a meaningful association [32]. Properties that did not meet these criteria were excluded from the similarity algorithm, whereas those that did meet these criteria were included.

Weight calculation for selected properties: In the previous sections, analyses of outbreak data have been dedicated to determining which properties correlate most strongly with the duration and number of cases in the outbreaks of a given disease. The ultimate goal of these properties is to allow AIDO to sort the historical outbreaks based on the user's input. The exact process used to create this ranking is described in later sections, but in short, the similarity algorithm is a function that maps the case count, duration, and disease-specific properties of the user's input and the historical outbreak to a number between 0 and 1. The set of values that are used by the similarity score are referred to as *parameters*. These include the user's input for case count and duration.

After determining which parameters would be included in the algorithm, a modified sensitivity analysis was used to determine the weights for each parameter. The sensitivity analysis determines the relative importance of selected parameters to size and duration of the outbreak. For all algorithms, the first 2 parameters are case count and duration, which were given the greatest weight and were not considered in the processing done for the weighting algorithm. For the additional disease-specific properties, weights were calculated using an automatic algorithm that compares each outbreak to all other outbreaks for a given disease. The effect of each property on the ranking of outbreaks in the library was evaluated. If removing a property had a large effect on the order of the outbreaks, it was inferred that the excluded property was important and should be given a greater weight. This evaluation was conducted for all properties for a given disease, and relative ranking or weights were determined. The sum of all weights was set to 1. A more detailed description of the process is given in Multimedia Appendix 2.

Figure 1. Schematic depicting the flow of statistical operations performed on categorical properties during property analysis.

Additional Analytics

In addition to the similarity algorithm, AIDO includes a number of other visual analytics designed to enhance situational awareness. Use of these analytics is illustrated in the Results section, Case Study, and associated figures.

Short-term forecasting: The AIDO disease library can also be used to perform short-term forecasting using the method of analogs (similar to, but simplified from, the study of Viboud et al described above [3]) approach. As a first step in this method, the cumulative case count curves for all outbreaks for the disease of interest were aligned in time. The distance criterion used in AIDO is simpler than that used in Viboud et al; in AIDO, because outbreaks are first deliberately aligned in time and because AIDO stores *representative* outbreaks and not *all* outbreaks, all available outbreaks are used in the analysis (ie, without using all outbreaks, there would not be enough data to

attempt the analysis). The mean and SD at each time point were calculated and fit to a normal distribution. This distribution was then used to find the median, 50% prediction interval (PI) and 90% PI at each time unit. AIDO institutes a lower bound of zero (case counts cannot be below zero). To customize the forecast to the user, each case count is weighted in proportion to its outbreak similarity score; thus, case counts in outbreaks that are scored higher weigh more than case counts in outbreaks that are scored lower. To achieve this, AIDO computed weighted mean and SD values, which were then used as the normal distribution's parameters.

AIDO currently requires at least 10 data points at each time point. Once there are fewer than 10 data points, the forecast stops. If necessary, AIDO will use cubic B-spline interpolation to handle time series interval granularity issues. For example, if there are both monthly and weekly data, cubic B-spline interpolation will be used to fill in the gaps in the monthly data

so that they can be used alongside the weekly data. AIDO interpolates at the finest resolution present in each outbreak library; although interpolating epidemiological data may introduce errors into the forecasting algorithm, the authors felt that because the method of analogs is already a rough forecasting measure that in practice yields reasonably large PIs, the additional error provided by interpolation would be unsubstantial.

The method of analogs approach has been applied in fields such as meteorology, climatology, and epidemiology [3,33,34]. When a user sorts results by similarity score, a short-term forecast is displayed at the top of the sorted outbreak results. This graph presents a simple custom forecast of cumulative disease incidence based on user input and our library of historical outbreak curves.

Anomaly detection: AIDO features an anomaly detection component that enables a user to compare their values to the values in AIDO's historical outbreak database. Rather than employing a specific anomaly detection algorithm, AIDO offers a simple visual approach to guide the user's analysis and identify anomalies. We display the user's value among all outbreak values so that the user can visually see where their values lie. For example, for discrete variables, we draw a pie chart and highlight the section that the user falls into; if the user's slice represents only 5% of all outbreaks, for example, then this may be indicative of an anomaly. For continuous variables, we plot all outbreak points and show the corresponding box plot and highlight the user's value; here, an anomaly may be if the user's point lies within the outliers. An anomaly for the user's data is easily seen through such visualizations and directs the user's attention to key features of their unfolding event that may warrant further investigation.

Owing to data sparsity issues and significant differences between disease presentations, AIDO focuses on a qualitative, rather than quantitative, anomaly detection approach that requires user engagement and interpretation. Future work, however, could include automated quantitative anomaly detection results. There are a number of unsupervised anomaly detection algorithms that could be explored [35]. For example, classifier-adjusted density estimation (CADE) is a promising nonparametric unsupervised anomaly detection algorithm [36,37]; CADE and many other such algorithms, however, often require a significant volume of data for analysis, which can prove to be difficult when epidemiological data are used.

Developing the User Interface, Additional Visuals, and Functionalities

AIDO functionalities are written in Python [38] using the Django [39] Web framework and PostgreSQL [40] for the backend. Bootstrap [41], jQuery [42], and Plotly [43] are used on the front end for overall user interface design or functionality and graphs. The AIDO homepage and various features of the user interface are described in the Results section.

Outbreak comparisons, browsing and related outbreaks: An outbreak comparison graph is displayed on each result page that

shows the point estimate for user's values in comparison to the 5 outbreaks listed on that page. In addition to the analytics provided, the interface allows users to explore the outbreaks available in each library without making use of the analytic components. We refer to this as *browsing* the library. Utility of these features were evaluated in the case study detailed in the Results section.

We noted in our analysis of representative libraries that some outbreaks in our libraries have meaningful relationships between each other. For example, there are some instances of outbreaks that were related because of common exposure to contaminated food that is spread to multiple locations in a country. Other times, outbreaks might be related by virtue of an individual seeding a new outbreak of the same illness in a new location. AIDO provides such information to the user if this option is selected in the analytic.

Performing Verification and Validation

Verification of AIDO was performed using more than 200 automated tests. These tests are run automatically every time the code base is changed, allowing an alarm to be raised before deployment if an error in the codebase is detected. In addition, we performed two manual tests after implementation of disease-specific algorithms in AIDO. The first uses data for outbreaks already in the library to verify a 100% similarity match between identical outbreaks. The second uses outbreaks that were not included in the library as test case scenarios. Here, we qualitatively observe if the highest matching outbreaks are similar to the situation from the test outbreak.

AIDO also underwent extensive user interface evaluation and user experience testing by several external entities such as Massachusetts Institute of Technology Lincoln Labs, a user interface design class from the University of Washington, the Fusion Division within the Office of the Assistant Secretary for Preparedness and Response in the US Department of Health Human Services, Science and Technology Directorate in the US Department of Homeland Security, community health epidemiologists in the state of New Mexico, US CDC, and the National Bio-surveillance Integration Center.

Results

Disease Libraries and Algorithm Properties

Currently, AIDO contains 673 outbreaks across 39 different diseases. Figure 2 illustrates the geographic breadth and multicontinent coverage of outbreaks for 4 diseases (measles, Q fever, dengue, and chikungunya). Multimedia Appendix 1 provides data on the total number of outbreaks, geographical distribution, and algorithm properties for all diseases in AIDO. The properties identified for the similarity algorithms for the 4 diseases are shown in Table 2. Unsurprisingly, precipitation and climate were identified as relevant properties for mosquito-borne diseases such as chikungunya and dengue. Vaccination coverage was found to be important for measles outbreaks and animal-specific properties were considered important for Q fever (a zoonotic disease).

Figure 2. Geographic spread of historical libraries for four diseases. Points are proportional in size to the number of outbreaks in that country within our library.

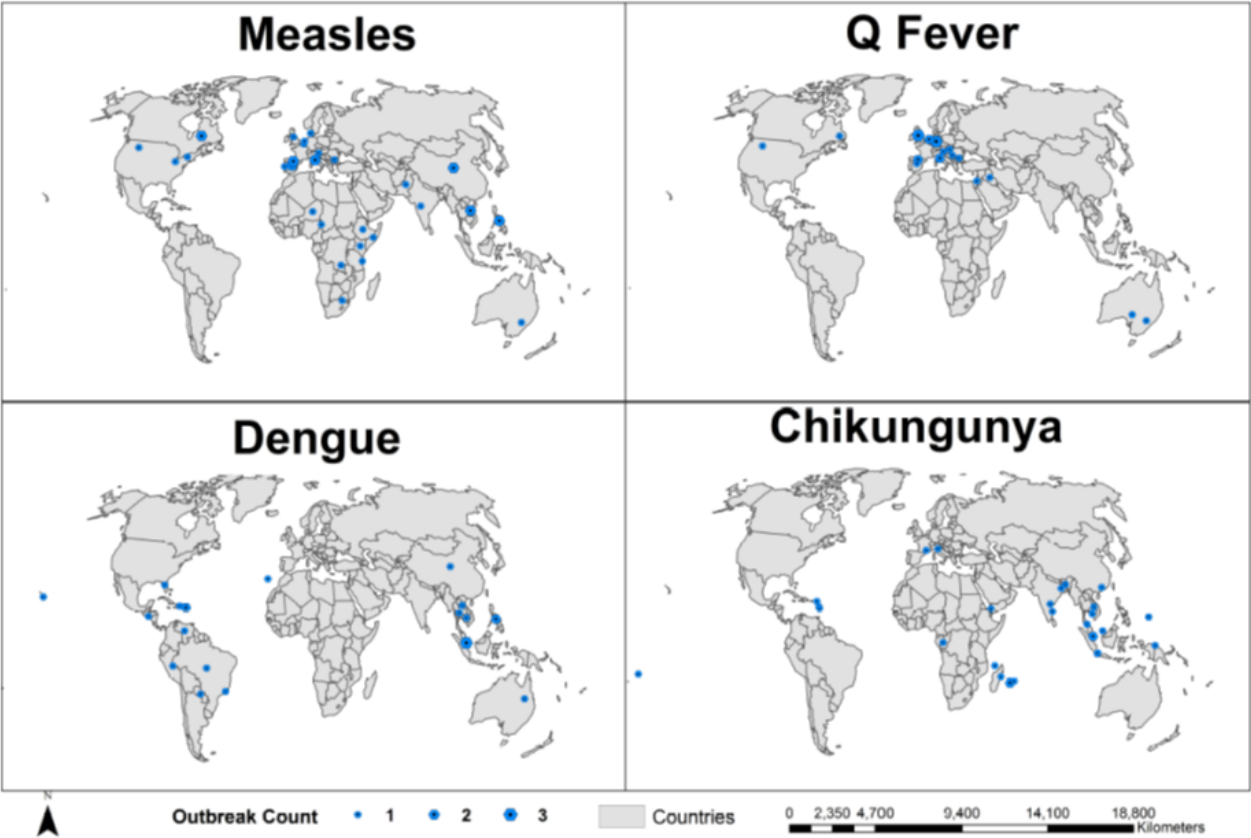


Table 2. Measles, Q fever, dengue, and chikungunya algorithm properties.

Disease	Algorithm Properties
Measles	Vaccination status (country)
	Vaccination status (region)
	Physician density
	Climate
Q fever	Animal contact
	HDI ^a
	Affected animal
	Outbreak source proximity
Dengue	Physician density
	Climate
	Population (discrete)
	Precipitation
Chikungunya	HDI

^aHDI: Human Development Index.

Statistically significant algorithm properties for the 4 diseases highlighted in the study are given.

An analysis of properties identified for diseases with similar characteristics was conducted to identify trends in properties across similar diseases [Multimedia Appendix 3](#) show the comparison of properties for mosquito-borne diseases and

vaccine preventable diseases that are part of the AIDO library. Although there are some similarities across diseases, we find that statistically important properties tend to vary, even between similar diseases, which one may not expect given common modes of transmission such as mosquitoes as vectors or vaccine-preventable diseases. This may be due to our emphasis

on including diverse *representative* outbreak presentations in our libraries.

Case Study: Q Fever Outbreak in Bilbao, Spain 2014

To describe a plausible use case in AIDO, we present a case study on a Q fever outbreak in Bilbao, Spain, in 2014. This outbreak was described in depth by Alonso et al [44]. This disease and outbreak combination was selected because available data were detailed enough to illustrate multiple features of AIDO.

AIDO user input data: Q fever was selected as the disease of interest from the AIDO homepage and the following information was used to populate the user input form (Figure 3). This outbreak occurred among workers at a waste-sorting plant in Bilbao, Spain, between February and April 2014. Approximately 10 cases were reported in the first 2 weeks of the outbreak [44]. The plant employed about 100 employees, was located in a municipality, and the patients did not report regular contact with animal hosts [44]. The human development index (HDI) for Spain in 2014 was 0.834 [28]. Information about animal hosts and proximity to the farm was not available during the early stages of this outbreak and was left blank. Figure 3 also shows the ability to sort the AIDO library for Q fever based on date and distance from location and illustrates the expanded *filter results* that can be used to narrow down outbreaks with specific properties.

Analysis of AIDO output: After completion of the similarity algorithm computation, the AIDO output showed the top matching outbreaks in the outbreak comparison graph and details of the five most similar outbreaks on page 1 (Figure 4). An example of information provided for each outbreak in AIDO is given here. In this case study, the outbreak comparison graph showed that the 5 most similar outbreaks had cumulative cases ranging from 10 to 100 cases and a total duration of 2 weeks to 6 months. The Q fever outbreak reported from Italy in 1993 and United Kingdom (2000) were the most similar outbreaks (80% and 79% similarity, respectively; Figure 4 shows the Italian outbreak). A radar plot is used to illustrate computation of the similarity score. This plot can be accessed through the *view this outbreak's detailed score* hyperlink under the epidemic curve graph. Further examination of the metadata showed that both outbreaks were caused by environmental exposure (sheep migration and contaminated strawboard in a factory office, respectively). This is similar to the probable cause of the outbreak in Bilbao, animal remains that had contaminated the waste-sorting environment at a factory [44].


The other 3 outbreaks with the highest similarity score included outbreaks from Hungary in 2013 (77%), Iraq in 2005 (75%), and Canada in 1987 (69%). The outbreak factors revealed that these epidemics also occurred among small groups of people associated by work and that the case count and duration were similar. Analysis using the anomaly detection features (Figure 5 top panel) showed that the case study outbreak parameters fell within the normal range for natural outbreaks and revealed high likelihood of the waste materials of sheep and goats as the cause of this outbreak. The short-term forecasting algorithm predicted a mean cumulative case count of 50 to 100 cases (Figure 5 bottom panel). Figure 6 shows AIDO outbreaks sorted by date and distance from the location (Spain).

Using AIDO's analysis, it could be hypothesized that the case study outbreak in Spain was likely to have 30 to 50 cases during the initial one to two months of the outbreak. This was confirmed by Alonso et al, who showed that the Spanish outbreak reported 45 cases from February 17 to April 27, 2014 [44]. The short-term forecast analysis in AIDO accurately predicted the expected case count and duration for the unfolding outbreak using only data from the initial 2 weeks. This case study shows that AIDO can be used during initial stages of an outbreak with minimum input data and information on expected case count and duration, and possible causes can be gleaned from the analysis. AIDO analysis can be performed iteratively as the outbreak progresses.

The AIDO Q fever library did not contain any related outbreaks. Figure 7 describes the related outbreak feature by showing the various outbreaks connected to the France 2008-2011 measles outbreak. These graphs provide information on the timeline associated with the start of related outbreaks and alerts the user on the possibility of the unfolding outbreak being part of a larger phenomenon. The *view related outbreaks* hyperlink under relevant outbreaks provides access to the related outbreak information.

All disease libraries in AIDO were evaluated using multiple case studies similar to the Q fever case study presented above. Details on 3 additional case studies for chikungunya [45], measles [46], and dengue [47-49] are given in [Multimedia Appendix 4](#). These case studies also demonstrate the utility of AIDO analysis during the unfolding stage (3 to 4 weeks) of an outbreak to identify possible case count, duration, and distinctive features during the epidemic.

A
ADO
Get Data
Get Data



Analytics for Investigation of Disease Outbreaks

ADO is a decision support tool designed to enhance situational awareness during unfolding disease outbreaks by providing the user with detailed background information on historical outbreaks.

Composed of over 100 outbreaks across a wide variety of diseases, the goal of ADO is to give the user a representative sample of how outbreaks of different diseases progress under various circumstances.


This tool was previously known as the DMAP (Disease Map) Application

[View](#) [Download](#) [Feedback](#)

Outbreak comparison

ADO offers a simple interface for viewing and interacting with a substantial outbreak library spanning many diseases. Users are able to search across this library to find outbreaks similar to what they are facing, allowing them to quickly identify historical outbreaks that may be significant to the ongoing event.


[Click here for an overview the data our outbreaks include.](#)



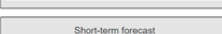
Anomalous outbreak detection

Given the breadth and depth of ADO's outbreak library, a user might wonder if their outbreak is anomalous compared to historical outbreaks. The "Anomaly detection" section is aimed at allowing the user to answer this question. Bar plots and pie charts are shown to provide the distribution of continuous and discrete properties, respectively.

[Click here for more information.](#)



Short-term forecast



Analytics for Investigation of Disease Outbreaks

ADO is a decision support tool designed to enhance situational awareness during unfolding disease outbreaks by providing the user with detailed background information on historical outbreaks.

Composed of over 100 outbreaks across a wide variety of diseases, the goal of ADO is to give the user a representative sample of how outbreaks of different diseases progress under various circumstances.

This tool was previously known as the DMAP (Disease Map) Application

[View](#) [Download](#) [Feedback](#)

Outbreak time series

Each outbreak in ADO provides detailed information on how the outbreak progressed under these circumstances. All outbreak records display an interactive time series and are annotated with contextual information.

[Click here for more information.](#)


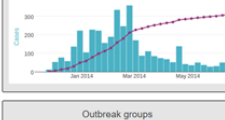


Chart export

ADO provides visualizations for much of the data it uses. All plots displayed in ADO can be quickly downloaded in various formats, allowing them to be easily included in reports.


[Download plot as PNG](#) [Download plot as PDF](#) [Download plot as CSV](#) [Download plot as JSON](#)



Outbreak groups

ADO includes several major outbreaks that spread across multiple regions. The reports from each of these regions have been collected and added to ADO, making it easy for users to gather data from the different governments and organizations that monitored the separate locations involved in the outbreak.

[Outbreak group time series](#)



Filter results

*** Disease**

*** Sort by** [Help](#)

*** Cases**

*** First case report**

*** Last case report**

Location

Reported contact with animal hosts

Human development index [Data sources](#)

Affected animal

Proximity to cattle, sheep, or goat farms or herds

Type of epidemic curve

Q Fever

Similarity score

10

2018-04-01

2018-04-15

Spain (Country)

No

0.834

[Similarity score](#)

Date

Distance from a location of concern

[Filter results](#)

Location

Rural or urban

Reported contact with animal hosts

Human development index [Data sources](#)

Population group

Affected animal

Proximity to cattle, sheep, or goat farms or herds

Type of epidemic curve

[Filter results](#)

A Q Fever: Outbreak comparison

B Q Fever: Italy (1993)
80% similar to your situation

C Score breakdown

Q Fever: Italy (1993)

Property	Outbreak value	Score	Weighted
Case count	---	92.0%	0.262 24.1%
Time	---	100%	0.262 26.2%
Reported contact	no	100%	0.100 19.0%
Affected animal	sheep	---	0.119 11.1%
Human development index	0.779	93.4%	0.119 11.1%
Proximity to cattle/sheep/goat farms or herds	<5km	---	0.040 7.6%
Total			86.6%
Max possible score			93.3%

Event features

- Important dates: In the summer and fall of 1993, a greater than expected number of cases of Q-Fever were reported following the transmigration of several flocks of sheep through heavily populated areas of Vicenza on their way to pre-alpine pastures. An outbreak of Q fever was subsequently investigated.
- Index case: No index case was reported.
- Affected animal: Sheep are the suspected affected animal.

Contextual information

- Case definition:
 - A case was defined as a person who presented with fever between June 4 and November 11 1993 and who had a positive serology for *C. burnetii*.
 - The graph shows cases identified during the study.
- Geographic/historic information:
 - There were only 3 Q fever cases officially reported in the province of Vicenza between 1983 and 1992.
 - The road followed by the flocks of sheep went from Vicenza to the Alps, crossing the western peripheral area of the town, and then passing through rural and industrial villages. Indirect exposure to flocks of sheep being herded to alpine pastures was a determinant of this outbreak (point source outbreak).
 - This outbreak occurred among general population that did not have regular contact with animals. Human development index (HDI) was 0.779.
- Risk factors: Indirect exposure to the flocks of sheep being herded in pastures was a determinant in this outbreak.

Actions or control measures

- No control measures were reported.

Figure 5. Additional analytics-anomaly detection and short-term forecast. Panel A illustrates two types of graphs used in the anomaly detection tab. Continuous variables (e.g., average cases per day, population at risk, or total cases) are shown as box plots. Discrete or categorical variables (e.g., season or affected animal) are shown as pie charts. The example presented shows that the case study outbreak is similar to other outbreaks included in our library. Panel B shows short term forecasting using the method of analogs. The data shown here can be used to estimate cumulative case count. This figure also demonstrates how data points are aligned for the short-term forecast.

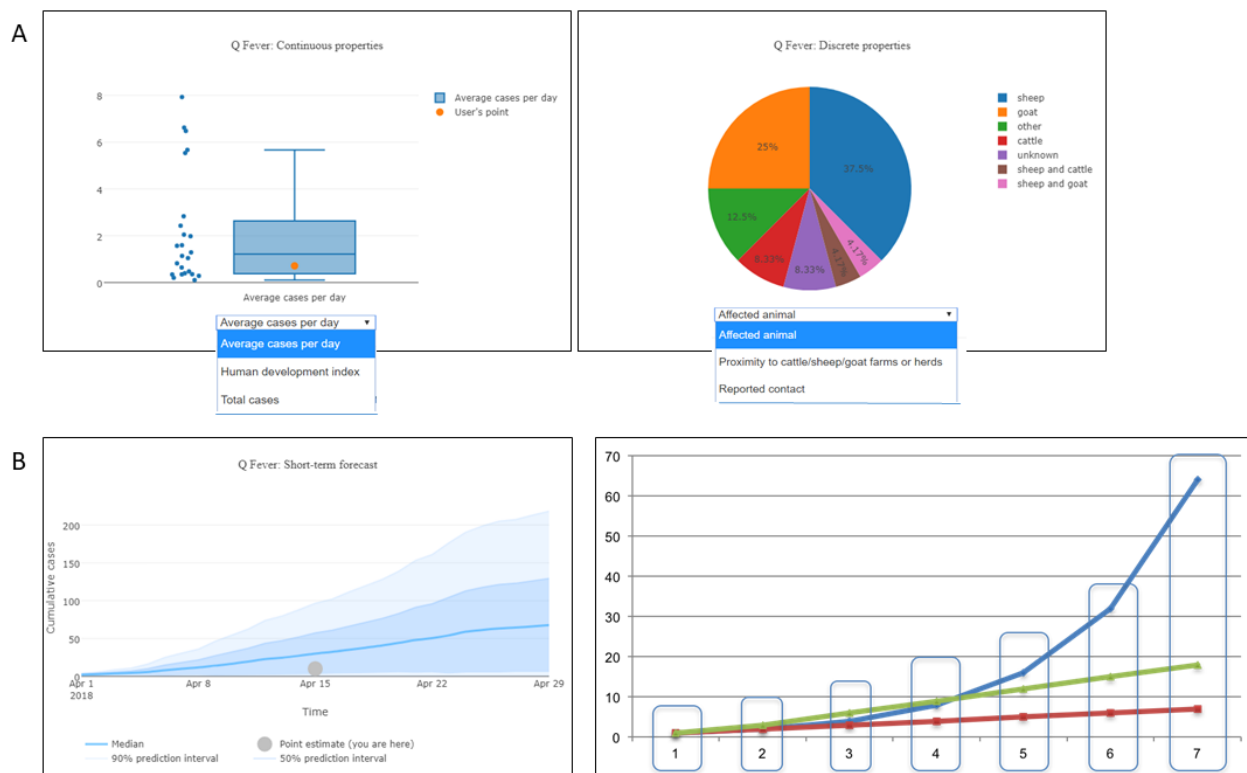


Figure 6. Browse functionality. This figure demonstrates browse functionality by date and by location available on AIDO.

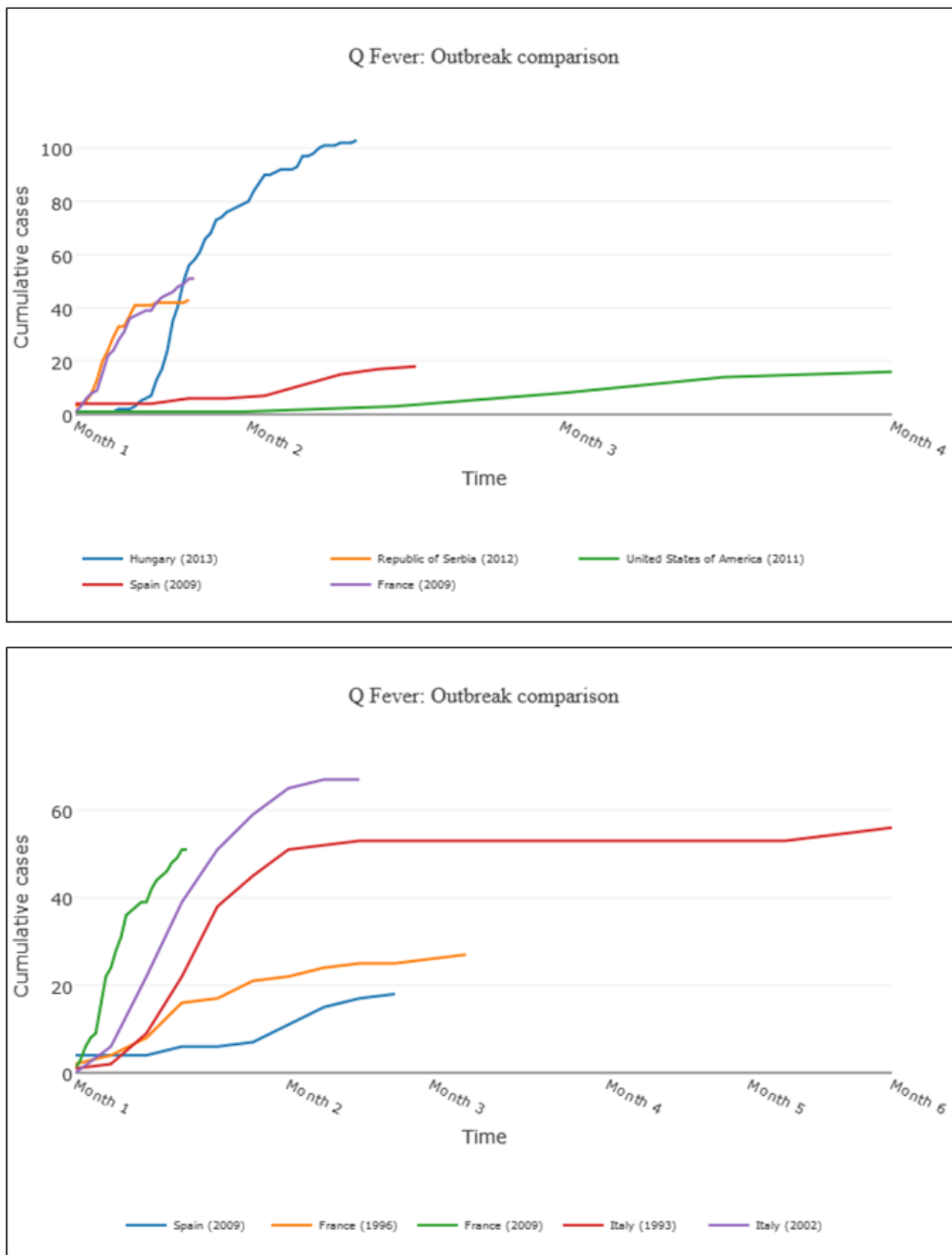
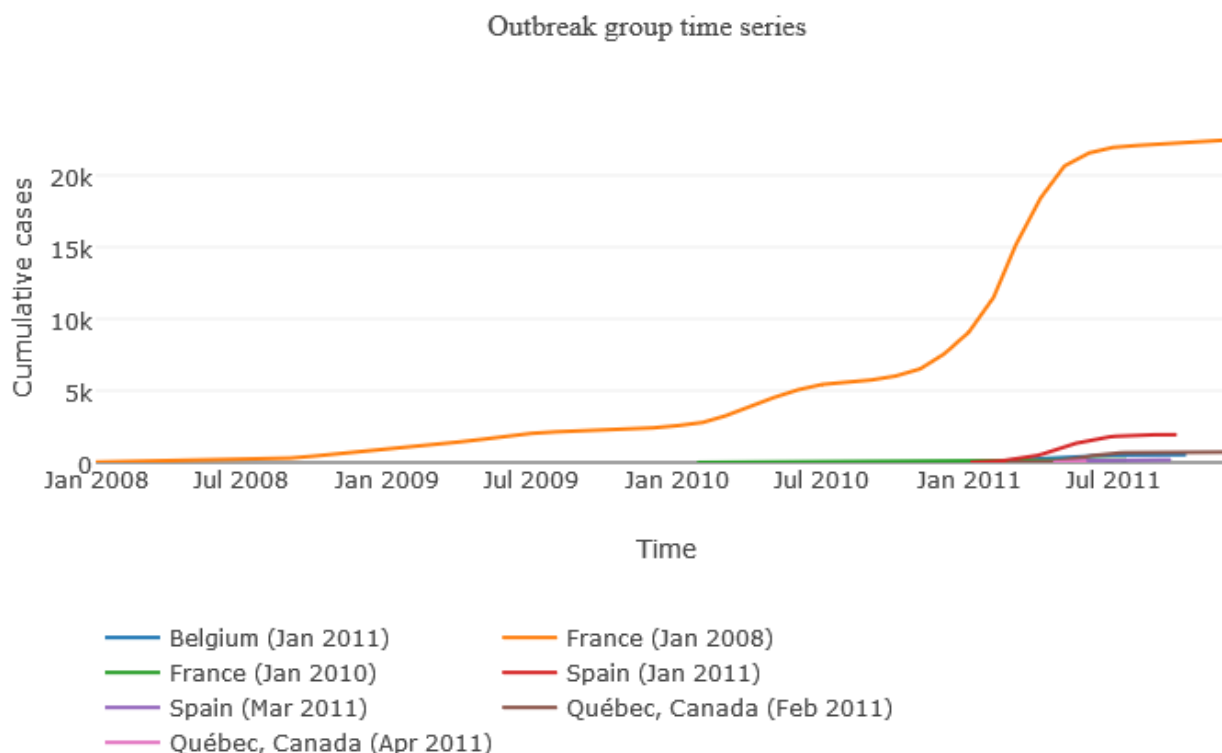


Figure 7. Example of the related outbreak interface. Here, we show outbreaks related to the France 2008-2011 measles epidemic. All graphs presented in AIDO are interactive and available for download in multiple formats.



Discussion

AIDO presents relevant outbreak data in a visually concise manner with graphs and point estimates of the user's input scenario in the context of the historical outbreak. Its analysis can provide an estimated case count and duration (short-term forecasting) and outbreak control measures that were effective in the past and AIDO facilitates delivery of outbreak information in an easy-to-interpret format.

AIDO is intended for members of the infectious disease surveillance community, both at the global level as well as the local level. A nongovernmental organization studying an ongoing outbreak can use this tool to analyze the scope of a current outbreak in any part of the world and devise control strategies that have proven effective in similar historical outbreaks. An individual physician may find this tool useful in understanding their case counts in a wider context and help facilitate their decision-making process on reporting relevant data to authorities. Individual analysts or local epidemiologists can use this tool as an aid in accessing the ongoing outbreak with increased situational awareness on what happens in their region and in similar regions around the world. In effect, this feature provides a projection of how the outbreak may unfold and could be considered a form of forecasting.

The algorithms and visuals available in AIDO inform users about the historical and geographical context of outbreaks for a given disease. The analysis also increases a user's knowledge on diverse outbreak scenarios associated with a given disease. This information may enhance understanding on possible routes

for outbreak progression (eg, transportation-associated global transmission). The related outbreak feature of AIDO can be utilized to generate hypotheses on next *hot spots* for a given outbreak and improve surveillance efforts in those locations. The anomaly detection algorithm on AIDO was specifically designed to detect anomalous presentations of an unfolding outbreak, perhaps biothreat scenarios. External evaluators of AIDO described it as *research at fingertips* for analysts or users. AIDO can also be used in education or training of epidemiologists.

Historical data have been used to develop epidemiological models. However, most models have challenges becoming operationalized owing to a variety of reasons [50]. AIDO is a tool that combines extensive epidemiological data with novel but simple analytics to contextualize and generate hypotheses about the trajectory of unfolding outbreaks. Rather than focusing on complex epidemiological models (which we recognize have substantial use and utility in the field), we take a historical approach and aim to identify relevant events that have already happened. We believe that this approach is novel and can provide complementary information to that derived from traditional modeling approaches.

Limitations

It is important to note that our approach relies on a few recognized limitations. First, we know from prior research that historical data are subject to change and are not fully complete [51]. To minimize this known bias, we used the most complete data available. However, it is likely that the data presented in our libraries include some known reporting bias.

Relatedly, the outbreak-matching algorithm depends on the diversity of historical outbreaks that exist in the library. This may particularly affect AIDO when investigating outbreaks with no precedence. For example, the 2014 Ebola outbreak in West Africa had no historical counterparts. Using AIDO in these situations may prove to be less reliable. However, we also believe that the anomaly detection functions provide use in these unprecedented situations.

As our libraries are created using publicly available data, it is also likely that our libraries are slightly skewed toward large or highly publicized outbreaks. Furthermore, because of this bias, data are much more prevalent in nations with robust surveillance systems. For example, Chase-Topping et al describe this trend with respect to worldwide incidence of *Escherichia coli* O157. Although food contamination of *E. coli* is present worldwide, outbreak investigations are conducted almost exclusively by developed nations [52]. It is because of this known bias that we aim for representative outbreak libraries that showcase the known breadth of disease outbreak presentation. However, it is unlikely that our effort to create representative libraries completely eliminates this bias.

AIDO data are largely limited to diseases affecting humans. Although we have libraries for Porcine Epidemic Diarrhea Virus and Foot and Mouth Disease, attempts to expand to other animal diseases have been challenging. This is simply due to a lack of data; time series data are difficult to obtain for animal diseases and almost nonexistent in plant outbreaks. These types of data are economically sensitive and, as a result, are not often reported.

Additionally, our short-term forecasting component relies on a method of analogs approach that assumes normally distributed data, which may not always be a good assumption. As a result, in some cases, the short-term forecast may not be reliable.

Finally, we wish to draw attention to some important considerations of our statistical analysis for property identification. As our procedure specifically identifies properties that are statistically related to outbreak size or duration, any updates to the disease library can, by design, change the related properties. Therefore, the property analysis must be performed any time there is a change in data. Automation of these processes is planned as a future project.

Future Work

Data in AIDO are constantly being updated. Much of this work is performed manually by a team of biologists and public health experts. However, we have also created the infrastructure to crowdsource data. AIDO includes a feedback form that allows the user to send us information about an outbreak currently not included in the library. The popularization of these types of decision support tools and inclusion of outbreak data from developing nations will facilitate enhanced disease surveillance and outbreak control in developing nations. In addition, we hope to promote AIDO as a training tool for epidemiologists. Owing to the breadth of information contained, and the wide array of analytics available, subject matter expert reviewers have mentioned that this is a logical next step in AIDO's development. We also envision application of AIDO for investigation of syndromic outbreaks through the development of a feature that would allow the identification of a causative agent for syndromic input into AIDO. These analyses would focus on identifying a pathogen within disease families (eg, gastroenteritis mosquito-borne disease family).

Conclusions

History often repeats itself. This is the simple underlying premise of AIDO. Rather than using limited elements of historical outbreak data to merely inform mathematical models, AIDO takes advantage of the entire story of a historical outbreak to offer examples of similar outbreaks for a current unfolding event. Importantly, it facilitates the use of limited and isolated information in the early stages of an outbreak to make decisions. Information obtained from AIDO can also be used to improve outbreak modeling parameters. AIDO aids in the investigation of disease outbreaks by contextualizing an unfolding outbreak with closely matching historical outbreaks. We posit that by providing diverse layers of information, visuals, and analytics, AIDO furnishes a comprehensive picture that may allow the user to make informed decisions about outbreak control. The tool allows *no cost* epidemiological evaluations, as it is freely available on the internet and can be used iteratively during the early stages of an outbreak. We offer this analytic to the global infectious disease surveillance community as a rapid and facile decision support tool that can be easily accessed—a simple yet useful resource that is the first of its kind.

Acknowledgments

AIDO was funded by the Defense Threat Reduction Agency, CBD ST, CB 10027, and the Department of Homeland Security Science and Technology Directorate, Homeland Security Advanced Research Projects Agency, Chemical Biological Division under agreement HSHQPM-16-X-00226. Dr Dave Osthus provided invaluable advice on the appropriate statistical tests to use in property analysis.

All data for this project were collected from previously published manuscripts or reports and was determined not to be human subjects research by the LANL Institutional Review Board.

NV, ARD, NG, MC, FA, LAC, EA, AH, GV, LN, MAL, EC, AV, and AD collected and analyzed outbreak data and developed the algorithms. NV prepared the case studies. GF, WR, ARD, and RP created the database, user interface, and relevant analytics. NV, ARD, FA, AD, MC, LAC, GF, and WR prepared the manuscript. All authors provided critical revisions.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Disease library details. This table shows all diseases currently included in AIDO and the number of outbreaks included in each disease. The number of countries represented in each library helps to describe the geographic variation represented therein. Properties included in the similarity algorithm for each disease (per the statistical procedure described in the main text) are also given. Some diseases had fewer than 10 outbreaks, which was determined to be too little data to run the statistical analysis. In addition, some other diseases did not have any statistically significant properties. These outbreaks are available to interact with via “browsing”, but do not include the similarity matching algorithm analytics.

[DOCX File, 17KB - [publichealth_v5i1e12032_app1.docx](#)]

Multimedia Appendix 2

Automatic weight calculation algorithm.

[DOCX File, 16KB - [publichealth_v5i1e12032_app2.docx](#)]

Multimedia Appendix 3

Supplementary Table 2A and 2B: Mosquito borne and vaccine preventable disease property comparison. Table 2A compares properties for eight mosquito-borne diseases. Table 2B shows a similar comparison for eight vaccine-preventable diseases. In general, we find that similar diseases may have some properties in common, but properties tend to be distinct across diseases. Note that no statistically significant properties were identified within our mumps library.

[DOCX File, 15KB - [publichealth_v5i1e12032_app3.docx](#)]

Multimedia Appendix 4

Additional case study examples; chikungunya, measles and dengue outbreaks in 2017-2018. AIDO was used to evaluate three recent outbreaks. AIDO input data and comparison of AIDO results and the actual outbreak data is given. Results showed that AIDO was able to provide estimates of case count, duration for the outbreak as well as identify distinctive features with only early stage data used as input.

[DOCX File, 13KB - [publichealth_v5i1e12032_app4.docx](#)]

References

1. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006 May 27;367(9524):1747-1757. [doi: [10.1016/S0140-6736\(06\)68770-9](#)] [Medline: [16731270](#)]
2. World Economic Forum. 2016. Global Disease Outbreaks URL: <http://reports.weforum.org/global-risks-2016/global-disease-outbreaks/> [accessed 2018-07-25] [WebCite Cache ID 71BUK7yDP]
3. Viboud C, Boëlle PY, Carrat F, Valleron AJ, Flahault A. Prediction of the spread of influenza epidemics by the method of analogues. *Am J Epidemiol* 2003 Nov 15;158(10):996-1006. [doi: [10.1093/aje/kwg239](#)] [Medline: [14607808](#)]
4. Sugihara G, May RM. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* 1990 Apr 19;344(6268):734-741. [doi: [10.1038/344734a0](#)] [Medline: [2330029](#)]
5. Alabi AS, Frielinghaus L, Kaba H, Kösters K, Huson MA, Kahl BC, et al. Retrospective analysis of antimicrobial resistance and bacterial spectrum of infection in Gabon, Central Africa. *BMC Infect Dis* 2013 Oct 02;13:455 [FREE Full text] [doi: [10.1186/1471-2334-13-455](#)] [Medline: [24083375](#)]
6. Mantilla G, Oliveros H, Barnston AG. The role of ENSO in understanding changes in Colombia's annual malaria burden by region, 1960-2006. *Malar J* 2009 Jan 08;8:6 [FREE Full text] [doi: [10.1186/1475-2875-8-6](#)] [Medline: [19133152](#)]
7. Adjemian JZ, Foley P, Gage KL, Foley JE. Initiation and spread of traveling waves of plague, *Yersinia pestis*, in the western United States. *Am J Trop Med Hyg* 2007 Feb;76(2):365-375. [doi: [10.4269/ajtmh.2007.76.365](#)] [Medline: [17297050](#)]
8. Campos NG, Burger EA, Sy S, Sharma M, Schiffman M, Rodriguez AC, et al. An updated natural history model of cervical cancer: derivation of model parameters. *Am J Epidemiol* 2014 Sep 01;180(5):545-555 [FREE Full text] [doi: [10.1093/aje/kwu159](#)] [Medline: [25081182](#)]
9. Fung IC. Cholera transmission dynamic models for public health practitioners. *Emerg Themes Epidemiol* 2014 Feb 12;11(1):1 [FREE Full text] [doi: [10.1186/1742-7622-11-1](#)] [Medline: [24520853](#)]

10. van Panhuis WG, Grefenstette J, Jung SY, Chok NS, Cross A, Eng H, et al. Contagious diseases in the United States from 1888 to the present. *N Engl J Med* 2013 Nov 28;369(22):2152-2158 [FREE Full text] [doi: [10.1056/NEJMms1215400](https://doi.org/10.1056/NEJMms1215400)] [Medline: [24283231](#)]
11. Tycho. 2018. Project Tycho URL: <https://www.tycho.pitt.edu/> [accessed 2018-07-26] [WebCite Cache ID 71EE2ERil]
12. GIDEON;. GIDEON Informatics. 2016. GIDEON URL: <https://www.gideononline.com/> [accessed 2018-07-23] [WebCite Cache ID 71BUgpG3O]
13. Orata FD, Keim PS, Boucher Y. The 2010 cholera outbreak in Haiti: how science solved a controversy. *PLoS Pathog* 2014 Apr;10(4):e1003967 [FREE Full text] [doi: [10.1371/journal.ppat.1003967](https://doi.org/10.1371/journal.ppat.1003967)] [Medline: [24699938](#)]
14. Centers for Disease ControlPrevention (CDC). Update: cholera outbreak---Haiti, 2010. *MMWR Morb Mortal Wkly Rep* 2010 Nov 19;59(45):1473-1479 [FREE Full text] [Medline: [21085088](#)]
15. Abrams JY, Copeland JR, Tauxe RV, Date KA, Belay ED, Mody RK, et al. Real-time modelling used for outbreak management during a cholera epidemic, Haiti, 2010-2011. *Epidemiol Infect* 2013 Jun;141(6):1276-1285. [doi: [10.1017/S0950268812001793](https://doi.org/10.1017/S0950268812001793)] [Medline: [22935487](#)]
16. Internet Live Stats. 2018. Internet Users URL: <http://www.internetlivestats.com/internet-users/> [accessed 2018-07-27] [WebCite Cache ID 71EEIK6tW]
17. Analytics for Investigation of Disease Outbreaks (AIDO). 2018. URL: <https://aido.bsvgateway.org/> [accessed 2018-11-19] [WebCite Cache ID 743vfBjix]
18. Woodall JP. Global surveillance of emerging diseases: the ProMED-mail perspective. *Cad Saude Publica* 2001;17(Suppl):147-154 [FREE Full text] [doi: [10.1590/S0102-311X2001000700024](https://doi.org/10.1590/S0102-311X2001000700024)] [Medline: [11426276](#)]
19. International Society for Infectious Diseases. 2017. ProMED-mail URL: <http://www.promedmail.org/> [accessed 2018-07-27] [WebCite Cache ID 71EFZ6Bnk]
20. Centers for Disease Control and Prevention. 2017. Morbidity and Mortality Weekly Report (MMWR) URL: <https://www.cdc.gov/mmwr/index.html> [accessed 2018-07-27] [WebCite Cache ID 71EFiMNy0]
21. World Health Organization. 2017. Disease Outbreak News (DONs) URL: <http://www.who.int/csr/don/en/> [accessed 2018-07-27] [WebCite Cache ID 71EFsT4Tk]
22. World Health Organization. 2017. Weekly Epidemiological Record (WER) URL: <http://www.who.int/wer/en/> [accessed 2018-07-27] [WebCite Cache ID 71EG1KLfQ]
23. Eurosurveillance. 2017. URL: <https://www.eurosurveillance.org/> [accessed 2018-07-27] [WebCite Cache ID 71EG5p35Z]
24. European Centre for Disease Prevention and Control. 2017. Annual Epidemiological Reports URL: <https://ecdc.europa.eu/en/annual-epidemiological-reports> [accessed 2018-07-27] [WebCite Cache ID 71EGZc3xE]
25. Biosurveillance Resource Directory (BRD). URL: <https://brd.bsvgateway.org/> [accessed 2018-07-25] [WebCite Cache ID 743eSiZru]
26. Automeris. WebPlotDigitizer URL: <https://automeris.io/WebPlotDigitizer/> [accessed 2018-07-27] [WebCite Cache ID 71EGfmAJs]
27. World Maps of Köppen-Geiger. Climate Classification URL: <http://koeppen-geiger.vu-wien.ac.at/present.htm> [accessed 2018-07-27] [WebCite Cache ID 71EHThHLW]
28. Climate Data. Climate data for cities worldwide URL: <https://en.climate-data.org/> [accessed 2018-07-26] [WebCite Cache ID 71EHISAQO]
29. Human Development Reports. Human Development Data (1990-2015) URL: <http://hdr.undp.org/en/data> [accessed 2018-07-27] [WebCite Cache ID 71ENfxSVs]
30. World Health Organization. Density of physicians (total number per 1000 population): Latest available year URL: http://gamapserver.who.int/gho/interactive_charts/health_workforce/PhysiciansDensity_Total/atlas.html [accessed 2019-01-31] [WebCite Cache ID 75qTXismJ]
31. The R Project for Statistical Computing. 2015. URL: <http://www.R-project.org/> [accessed 2018-07-25] [WebCite Cache ID 71EGq7iGB]
32. Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012 Sep;24(3):69-71 [FREE Full text] [Medline: [23638278](#)]
33. Mandel AS. Method of analogs in prediction of short time series: an expert-statistical approach. *Automation Remote Control* 2004 Apr;65(4):634-641 [FREE Full text] [doi: [10.1023/B:AURC.0000023540.97466.33](https://doi.org/10.1023/B:AURC.0000023540.97466.33)]
34. Vasechkina EF, Yarin VD. Prediction of time series by the method of analogs. *Phys Oceanogr* 2007 Jul;17(4):242-251 [FREE Full text] [doi: [10.1007/s11110-007-0019-3](https://doi.org/10.1007/s11110-007-0019-3)]
35. Goldstein M, Uchida S. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PLoS One* 2016;11(4):e0152173 [FREE Full text] [doi: [10.1371/journal.pone.0152173](https://doi.org/10.1371/journal.pone.0152173)] [Medline: [27093601](#)]
36. Friedland L, Gentzel A, Jensen D. Classifier-adjusted density estimation for anomaly detection and one-class classification. 2014 Presented at: Proceedings of the SIAM International Conference on Data Mining; 2014; Philadelphia, Pennsylvania, USA. [doi: [10.1137/1.9781611973440.67](https://doi.org/10.1137/1.9781611973440.67)]
37. Baseman E, Blanchard S, DeBardeleben N, Bonnie A, Morrow A. Interpretable anomaly detection for monitoring of high performance computing systems. 2016 Presented at: 22nd ACM SIGKDD Conference on Knowledge Discovery and Data

- Mining - Outlier Definition, Detection, and Description On-Demand Workshop; August 14, 2016; San Francisco, California, USA URL: http://lissalytics.com/static/kdd_2016_anomaly_workshop.pdf
38. Python. URL: <https://www.python.org/> [accessed 2018-07-27] [WebCite Cache ID 71EP4iL1h]
 39. Django. Django. URL: <https://www.djangoproject.com/> [accessed 2018-07-27] [WebCite Cache ID 71EPB0ARZ]
 40. PostgreSQL Global Development Group. PostgreSQL: Documentation URL: <https://www.postgresql.org/docs/> [accessed 2018-07-27] [WebCite Cache ID 71EPGCKso]
 41. Bootstrap. URL: <https://getbootstrap.com/docs/4.1/getting-started/introduction/> [accessed 2018-07-27] [WebCite Cache ID 71EPKmobm]
 42. jQuery: D. jQuery. URL: <https://jquery.com/> [accessed 2018-07-27] [WebCite Cache ID 71EPP70tT]
 43. Plotly. Plotly. Modern Analytics Apps for the Enterprise URL: <https://plot.ly/> [accessed 2018-07-25] [WebCite Cache ID 75qUR0uej]
 44. Alonso E, Lopez-Etxaniz I, Hurtado A, Liendo P, Urbaneja F, Aspririxaga I, et al. Q fever outbreak among workers at a waste-sorting plant. PLoS One 2015;10(9):e0138817 [FREE Full text] [doi: [10.1371/journal.pone.0138817](https://doi.org/10.1371/journal.pone.0138817)] [Medline: [26398249](https://pubmed.ncbi.nlm.nih.gov/26398249/)]
 45. Kabir I, Dhimal M, Müller R, Banik S, Haque U. The 2017 Dhaka chikungunya outbreak. Lancet Infect Dis 2017 Nov;17(11):1118. [doi: [10.1016/S1473-3099\(17\)30564-9](https://doi.org/10.1016/S1473-3099(17)30564-9)]
 46. Strait Times. 2018. Measles outbreak in Japan, Taiwan puts thousands in quarantine, delays travel plans URL: <https://www.straitstimes.com/asia/east-asia/measles-outbreak-in-japan-taiwan-puts-thousands-in-quarantine-delays-travel-plans> [accessed 2018-07-25] [WebCite Cache ID 71umPH7n1]
 47. Firstpost. Kerala dengue outbreak: Death toll in 2017 rises to 115, Congress blames state govt URL: <https://www.firstpost.com/india/kerala-dengue-outbreak-death-toll-in-2017-rises-to-115-congress-blames-state-govt-3709129.html> [accessed 2018-09-02] [WebCite Cache ID 71unZXjIF]
 48. Banerjee I. Dengue: the break-bone fever outbreak in Kerala, India. Nepal J Epidemiol 2017 Jun;7(2):666-669 [FREE Full text] [doi: [10.3126/nje.v7i2.17972](https://doi.org/10.3126/nje.v7i2.17972)] [Medline: [29181227](https://pubmed.ncbi.nlm.nih.gov/29181227/)]
 49. The Indian Express. 2018 Aug 24. Advertising Fever panic in Kerala: numbers and the reasons URL: <https://indianexpress.com/article/explained/fever-panic-in-kerala-numbers-and-the-reasons-4715870/> [accessed 2018-09-03] [WebCite Cache ID 71umix4wO]
 50. Margevicius KJ, Generous N, Abeyta E, Althouse B, Burkom H, Castro L, et al. The Biosurveillance Analytics Resource Directory (BARD): facilitating the use of epidemiological models for infectious disease surveillance. PLoS One 2016;11(1):e0146600 [FREE Full text] [doi: [10.1371/journal.pone.0146600](https://doi.org/10.1371/journal.pone.0146600)] [Medline: [26820405](https://pubmed.ncbi.nlm.nih.gov/26820405/)]
 51. Bahk CY, Scales DA, Mekaru SR, Brownstein JS, Freifeld CC. Comparing timeliness, content, and disease severity of formal and informal source outbreak reporting. BMC Infect Dis 2015 Mar 20;15:135 [FREE Full text] [doi: [10.1186/s12879-015-0885-0](https://doi.org/10.1186/s12879-015-0885-0)] [Medline: [25887692](https://pubmed.ncbi.nlm.nih.gov/25887692/)]
 52. Chase-Topping M, Gally D, Low C, Matthews L, Woolhouse M. Super-shedding and the link between human infection and livestock carriage of Escherichia coli O157. Nat Rev Microbiol 2008 Dec;6(12):904-912 [FREE Full text] [doi: [10.1038/nrmicro2029](https://doi.org/10.1038/nrmicro2029)] [Medline: [19008890](https://pubmed.ncbi.nlm.nih.gov/19008890/)]

Abbreviations

AIDO: Analytics for the Investigation of Disease Outbreaks
ANOVA: analysis of variance
CADE: classifier-adjusted density estimation
CDC: Centers for Disease Control and Prevention
HDI: Human Development Index
PI: prediction interval

Edited by T Sanchez; submitted 07.09.18; peer-reviewed by S Ghosh, S Ng; comments to author 21.10.18; revised version received 26.11.18; accepted 25.01.19; published 25.02.19.

Please cite as:

Velappan N, Daughton AR, Fairchild G, Rosenberger WE, Generous N, Chitanvis ME, Altherr FM, Castro LA, Priedhorsky R, Abeyta EL, Naranjo LA, Hollander AD, Vuyisich G, Lillo AM, Cloyd EK, Vaidya AR, Deshpande A
 Analytics for Investigation of Disease Outbreaks: Web-Based Analytics Facilitating Situational Awareness in Unfolding Disease Outbreaks

JMIR Public Health Surveill 2019;5(1):e12032

URL: <http://publichealth.jmir.org/2019/1/e12032/>

doi: [10.2196/12032](https://doi.org/10.2196/12032)

PMID: [30801254](https://pubmed.ncbi.nlm.nih.gov/30801254/)

©Nileena Velappan, Ashlynn Rae Daughton, Geoffrey Fairchild, William Earl Rosenberger, Nicholas Generous, Maneesha Elizabeth Chitanvis, Forest Michael Altherr, Lauren A Castro, Reid Priedhorsky, Esteban Luis Abeyta, Leslie A Naranjo, Attelia Dawn Hollander, Grace Vuyisich, Antonietta Maria Lillo, Emily Kathryn Cloyd, Ashvini Rajendra Vaidya, Alina Deshpande. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 25.02.2019. 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

Differences in Emotional and Pain-Related Language in Tweets About Dentists and Medical Doctors: Text Analysis of Twitter Content

Jan-Are K Johnsen¹, PhD; Trude B Eggesvik¹, MDent; Thea H Rørvik¹, MDent; Miriam W Hanssen¹, MDent; Rolf Wynn², MD, PhD; Per Egil Kummervold³, PhD

¹Department of Clinical Dentistry, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway

²Department of Clinical Medicine, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway

³Norut Northern Research Institute, Tromsø, Norway

Corresponding Author:

Jan-Are K Johnsen, PhD

Department of Clinical Dentistry

Faculty of Health Sciences

UiT The Arctic University of Norway

TANN Building

Hansine Hanssensvei 86

Tromsø, 9037

Norway

Phone: 47 776 49 131

Email: jan.a.johnsen@uit.no

Abstract

Background: Social media provides people with easy ways to communicate their attitudes and feelings to a wide audience. Many people, unfortunately, have negative associations and feelings about dental treatment due to former painful experiences. Previous research indicates that there might be a pervasive and negative occupational stereotype related to dentists and that this stereotype is expressed in many different venues, including movies and literature.

Objective: This study investigates the language used in relation to dentists and medical doctors on the social media platform Twitter. The purpose is to compare the professions in terms of the use of emotional and pain-related words, which might underlie and reflect the pervasive negative stereotype identified in relation to dentists. We hypothesized that (A) tweets about dentists will have more negative emotion-related words than those about medical doctors and (B) pain-related words occur more frequently in tweets about dentists than in those about medical doctors.

Methods: Twitter content (“tweets”) about dentists and medical doctors was collected using the Twitter application program interface 140Dev over a 4-week period in 2015, scanning the search terms “dentist” and “doctor”. Word content of the selected tweets was analyzed using Linguistic Inquiry and Word Count software. The research hypotheses were investigated using nonparametric Wilcoxon-Mann-Whitney tests.

Results: Over 2.3 million tweets were collected in total, of which about one-third contained the word “dentist” and about two-thirds contained the word “doctor.” Hypothesis A was supported since a higher proportion of negative words was used in tweets about dentists than in those about medical doctors ($z=-10.47$; $P<.001$). Similarly, tests showed a difference in the proportions of anger words ($z=-12.54$; $P<.001$), anxiety words ($z=-6.96$; $P<.001$), and sadness words ($z=-9.58$; $P<.001$), with higher proportions of these words in tweets about dentists than in those about doctors. Also, Hypothesis B was supported since a higher proportion of pain-related words was used in tweets about dentists than in those about doctors ($z=-8.02$; $P<.001$).

Conclusions: The results from this study suggest that stereotypes regarding dentists and dental treatment are spread through social media such as Twitter and that social media also might represent an avenue for improving messaging and disseminating more positive attitudes toward dentists and dental treatment.

(JMIR Public Health Surveill 2019;5(1):e10432) doi:[10.2196/10432](https://doi.org/10.2196/10432)

KEYWORDS

dental anxiety; dentistry; psychology; social media; internet; dental public health; Twitter; professional role; occupational stereotype

Introduction

An increasing number of people use social media, such as Facebook and Twitter; these are becoming central as news outlets and are even creating headline news themselves, for example, when the US President tweets and stirs controversy [1]. The ubiquity of social media has created an opportunity for researchers to use these tools as sources of data on a range of topics, including on the spread of illnesses and attitudes to health-related topics [2-4]. A few studies have also used social media to examine public opinion on dental health, especially fluoridation [5,6]. However, prior research has suggested that social media may be used to spread distorted or false information and that such information may have important negative consequences, for instance, when dangerous information is spread about how to contain epidemics or when mental disorders are associated with negative emotions and unsupportive tweets [7,8].

With new technologies and social media, there are new venues for health communication, as well as new venues for expressing stereotypes and social categories. In order to make sense of the world, people have a tendency to think about others in terms of stereotypes or categories [9,10], and the existence of occupational stereotypes is relatively well established [11,12]. This tendency has benefits for saving cognitive resources [13] but can pose problems concerning the accuracy of these impressions. Regarding health professions, strong stereotypes exist about medical doctors and nurses. For instance, nurses are often seen as good communicators, nurturing, feminine, and caring [14,15], while doctors are described as confident and decisive [16]. Regarding dentists, it has been documented that stereotypes related to gender, such as the belief that females are more emotionally and relationally competent than males, impact the expectations and impressions of male and female dentists alike. For instance, a study reported that female dentists were expected to spend more time talking to their patients, while male dentists were expected to value patients' tolerance of pain without complaints [17].

Although logic dictates that stereotypes could be either good or bad, evidence suggests that stereotypes are most often negative. In their review, Baumeister et al [18] argue that illusory correlation appears to form more easily between a social group and negative or bad (distinctive) behaviors compared with positive or neutral behavior, and bad information about a person has more impact on impression formation than good information. It appears to be easier to acquire bad reputations than good reputations because fewer instances of bad behavior are needed to confirm this to be indicating a bad trait or disposition compared with good behavior [18]. Stereotypes or social categories are quite easily learned by social learning processes, but interestingly, stereotypes appear to become more extreme and less variable through social learning processes [19]. Such distortions related to the social learning of stereotypes could then negatively influence people's thoughts about certain social groups or categories, including professions.

Negative or bad behaviors could be a major influence on professional stereotypes when such behaviors are perceived as distinctive to the profession. In case of dentists and dentistry, bad distinctive behaviors could include instances of painful treatment. For instance, in a study of Norwegian adults, 20%-30% rated their last dental visit as moderately painful or worse and 60% reported having at least one very painful experience at the dentist's office [20]. Also, a study of Canadian adults found that 42.5% reported having moderate to severe pain during their last dental treatment [21]. In light of these findings, it would be reasonable to assume that painful experiences might serve as a foundation for creating negative stereotypes in relation to dentists. This notion appears to be supported by the findings of Thibodeau and Mentasti [22], who reviewed 100 movies portraying dentists in Western culture. In this study, it was shown that visits to the dentist in movies are often portrayed as a negative and painful experience, where the dentist is being depicted as "...incompetent, menacing, sadistic, immoral, unethical, or corrupt, and one might assume that all dentists behave in this manner." This association between negative experiences and the public image of dentists has been found in large population studies as well [23], which would not only hamper the image of dentists [23] but also be regarded as a factor in both the maintenance and establishment of dental anxiety [24]. The impact of the negative occupational stereotype related to dentistry could be that people exposed to it are reluctant to seek dental care, and some authors have argued that the dental community should consider promotional campaigns or marketing strategies to dispel the negative images associated with dentistry and to influence reluctant patients [25,26].

Based on the findings that indicate the existence of negative emotions related to dentistry and dentists, we would expect that these associations and stereotypes influence how the profession is talked about in social media. The current study seeks to investigate the language used in Twitter posts about dentists and compare these posts with those about another well-known health profession (medical doctors). We hypothesize that (A) tweets about dentists will have more negative emotion-related words than those about medical doctors and (B) pain-related words are used more frequently in tweets about dentists than in those about medical doctors.

Methods**Data Source**

Text data were collected from Twitter over a 4-week period starting in the last week of May 2015. For data collection, we used 140Dev server software [27], which ran at a server of the Northern Research Institute in Tromsø, Norway. The server monitored and stored all tweets containing the search terms "dentist" and "doctor". During the study period, the server downloaded and stored 524,958 tweets containing the word "dentist" and 1,821,914 tweets containing the word "doctor." To preserve the tweeters' privacy, none of the supplemental user information available from Twitter was downloaded. The

study can, therefore, be said to build only upon nonidentifiable information.

The tweets were in English only, which enabled us to perform the analysis on content written in a single language. Using single, common English words as search terms resulted in over 2.3 million tweets collected over 4 weeks. This design, thus, had the advantage of achieving a large sample size over a relatively short time period.

Data Selection and Preparation

Because the research questions of the study were related to how most people use Twitter to communicate about health professionals, we made an informal analysis of the suitability of the collected material by browsing through a random selection of tweets from each database. Based on the screening process, we decided that selection criteria would have to be imposed on the material because the text data contained many entries that were outside of the scope of this study (eg, commercial content). In order to remove irrelevant content and to increase the likelihood that selected tweets were personal and relevant to the tweets' authors, we used personal and possessive pronouns to filter the data. Thus, we excluded all tweets without at least one of the following words present: "I," "me," "my," "mine," "we," "us," "our," and "ours." The process was automated using a simple custom-made algorithm ensuring that this selection was case insensitive. A selection of tweets with personal or possessive pronouns included are shown in [Textbox 1](#). In addition, only original tweets were chosen for analysis, that is, tweets tagged as being retweets were excluded.

To obtain approximately the same amount of text data for each target group, we saved a random selection of 10,000 rows of the databases as text files for each target group (eg, one text file containing 10,000 lines for dentists and a similar text file for doctors).

Data Analysis

To investigate the research hypotheses, we ran these files through Linguistic Inquiry and Word Count (LIWC) [28]. LIWC is a computer application that analyzes text files according to a predefined dictionary and gives information about the percentage of words in the text files that matches the dictionary. The current study used the built-in English language dictionary. Other authors have found LIWC to be a valid approach for measuring emotion in verbal expression [29]. LIWC differs from sentiment programs (such as SentiStrength), which typically give an overall (positive and negative) sentiment [30]. LIWC gives detailed information about the use of different categories of words (ie, "anger," "anxiety," and "sadness"), which allows for a lexicologically framed analysis. In order to investigate Hypothesis A, we looked specifically at the word categories in the LIWC dictionary related to negative emotions (Negative emotions and 3 categories of specific negative emotions: Anger, Anxiety, and Sadness) and an overall emotional category (Affective processes) and Positive emotions.

In order to investigate Hypothesis B, we selected synonyms of pain and pain-related words from a popular Web-based English dictionary [31], which were then added to the dictionary of the LIWC analysis software to provide a separate pain category ([Textbox 2](#)).

Descriptive analyses were performed with a single file for each target group, while specific hypotheses testing required segmentation of the text files to simulate individual tweets. For the purpose of this study, we used 1000 segments per text file, which was done automatically by choosing this option in LIWC (see [Figure 1](#) for a visualization of the segmentation process). Because the data is not normally distributed, nonparametric Wilcoxon-Mann-Whitney tests were used to test the equality of the distributions. Data were analyzed using Jeffreys's Amazing Statistics Program (version 0.8.6; JASP Team) [32] and SPSS (version 24; IBM Corp) [33].

Textbox 1. A selection of relevant tweets with pronoun filtering enabled; pronouns in italics.

Dentists as the target group:

"I hate the dentist man. Leave *my* wisdom teeth aloneeeee. They not bothering *me*"

"just went to the dentist and *my* mouth feels like someone has punched it"

"Someone come to the dentist with *me*. I'm scared"

Medical doctors as the target group:

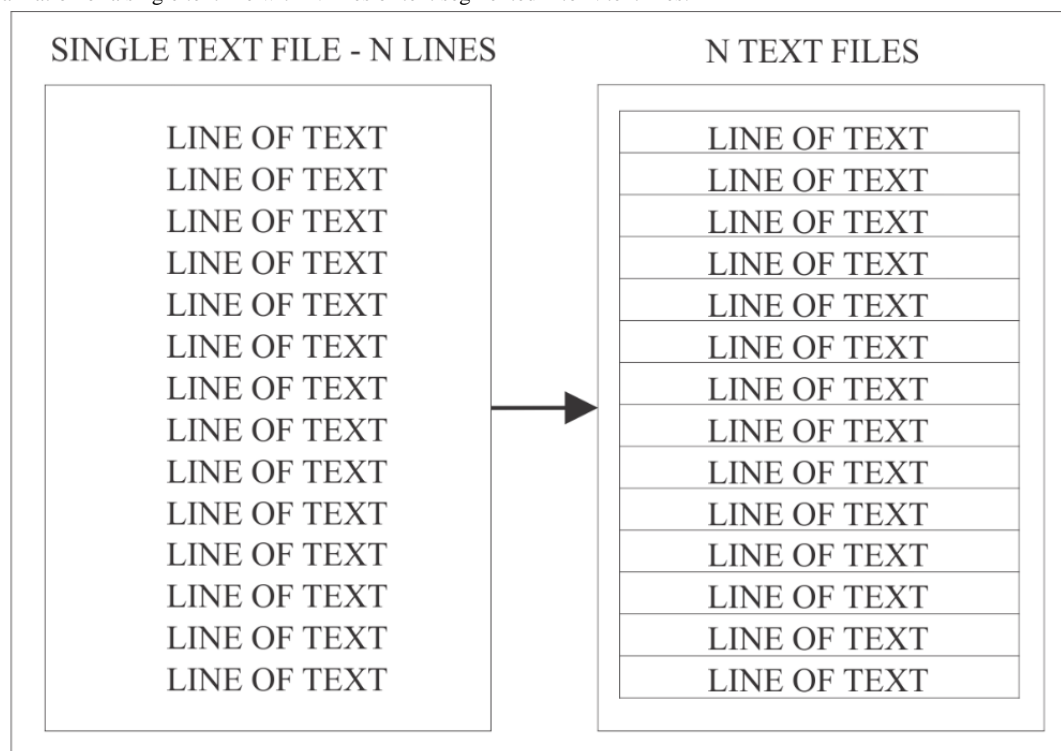
"I tweet this on a daily basis. But I truly dislike *my* doctor office."

"*My* ear is still ringing... Time to go back to the doctor."

"Hearing the doctor say I'm out for 6 weeks is probably the worst thing that has happened to *me*."

Textbox 2. Synonyms of pain used as a category in the Linguistic Inquiry and Word Count English language dictionary. Superscripted "a" indicates that a word stem was used and all words containing this word stem were counted.

ache, ached, aches, aching^a, affliction^a, agony, burn, burned, burns, burnin^a, burnt, cramp, cramped, cramps, discomfort^a, hurt, hurts, hurtful^a, illness^a, injur^a, irritation^a, maladies, malady, misery, pain, pain^a, pains, sickness^a, sore, soreness^a, sores^a, sting, stings, stingy, stitch, stitches, strain^a, suffer^a, tenderness^a, throbb^a, throe^a, tingle^a, torment^a, torture^a, trouble, troubles, troubled, twinge^a, wound^a

Figure 1. Visualization of a single text file with N lines of text segmented into N text files.

Results

The data selected for analyses contained a total of 166,266 words for “dentist” (704,397 characters without spaces) and 182,311 words for “doctor” (776,152 characters without spaces). Table 1 shows the mean proportions of emotional words for each professional category (group). In addition, grand mean baseline values for word categories across different writing tasks [28] have been included for comparison (Table 1).

The analysis (Table 1) showed that there was a difference between tweets about dentists and tweets about doctors, with more affective words used in tweets about dentists ($z=-6.80$;

$P<.001$), but a significant difference was not found for positive emotion-related words. Regarding Hypothesis A, the Wilcoxon-Mann-Whitney test showed that more negative words were used in tweets about dentists than in those about medical doctors ($z=-10.47$; $P<.001$). Similarly, tests of equality of distributions were performed for the specific emotion categories. These tests showed a difference in the proportions of words related to anger ($z=-12.54$; $P<.001$), anxiety ($z=-6.96$; $P<.001$), and sadness ($z=-9.58$; $P<.001$), with more words from these categories in tweets about dentists than in those about doctors. Also, Hypothesis B was supported since more pain-related words were used in tweets about dentists than in those about doctors ($z=-8.02$; $P<.001$).

Table 1. Mean proportions and SDs of emotional and pain-related word categories, grand mean baseline values from the LIWC documentation, and comparisons of the equality of distributions for the emotional word categories over professional categories.

Word category	Dentist (n=1000), mean (SD)	Doctor (n=1000), mean (SD)	Baseline, mean	P value ^a
Affective processes	6.29 (1.87)	5.63 (1.72)	4.41	<.001
Positive emotions	3.22 (1.46)	3.14 (1.43)	2.74	.15
Negative emotions	3.05 (1.41)	2.45 (1.22)	1.63	<.001 ^b
Anxiety	0.42 (0.53)	0.34 (0.47)	0.33	<.001 ^b
Anger	1.25 (0.92)	0.93 (0.76)	0.47	<.001 ^b
Sadness	0.53 (0.56)	0.41 (0.49)	0.37	<.001 ^b
Pain	0.34 (0.45)	0.25 (0.40)	N/A ^c	<.001 ^b

^aWilcoxon-Mann-Whitney test.

^bThe alternative hypothesis specifies that group dentist is greater than group doctor (1-tailed); other tests are 2-tailed.

^cN/A: not applicable; pain-related words are not included in the original LIWC dictionary.

Discussion

Principal Findings

This study demonstrated that more negative emotion words were used in tweets related to dentists than in those related to medical doctors. Thus, compared with medical doctors, dentists seemed to be associated with more negative emotions in tweets. Tweets about dentists did contain more affective words than tweets about medical doctors, but not more positive emotions. This study could be seen as supporting the idea that there is a negative stereotype related to dentists on Twitter. Twitter may, therefore, be one of several channels where the negative stereotype is transmitted, spread, and learned, as other research works indicate that emotions can be spread through both Web-based [34] and real-life social networks [35].

Dental anxiety is a widespread problem [36,37]. The results of the present study can perhaps be seen in unison with the idea that there are several pathways to developing dental anxiety. In a recent qualitative study of Web-based videos related to dental anxiety, 3 main pathways were outlined: direct experiences with aversive dental treatment, vicarious learning through parents and peers, and exposure to negative information [38]. In light of this, negative occupational stereotyping might be an important factor in the development of dental anxiety, as it creates negative associations and expectations irrespective of the individuals' own experiences. Thus, ambiguous stimuli or information in the dental situation might be interpreted negatively based on the negative emotions and expectations related to the stereotype. For instance, nonverbal communication such as the tone of voice used by the dentist when providing information might be considered as condescending or authoritative by some patients due to negative expectations, while patients without negative expectations are less prone to drawing similar conclusions.

This study supported the hypothesis that negative emotions would be more frequently used in relation to dentists than to doctors. For anxiety words, this was expected given that a potential occupational stereotype related to dentistry can be linked to the relatively widespread phenomena of subclinical dental anxiety or low-grade or moderate worry about dental treatment, which is believed to be quite prevalent in most societies [39,40]. More surprising, perhaps, was the differences observed for both anger and sadness, which can be more difficult to understand. However, it is a quite common finding that people are willing to share anger or anger-related materials on the Web [41] and that feelings of anger might be related to the idea that dental treatment is somehow unethical in the sense that it is expensive, painful, or administered without proper consent [42,43]. Also, the motivational direction of anger is argued to be different from some other feelings in that anger is an approach-oriented emotion concerned with removal of an obstacle rather than withdrawal or avoidance from the obstacle [44,45]. Thus, people might be motivated to write about (ie, approach) their angry feelings about dentists and dental treatment. This is, in part, supported by the fact that anger words are used more than other specific negative emotional words in this study's data. It is also noteworthy that positive words were more frequent than negative words for both doctors and dentists;

however, the existence of positive information or information that disconfirms stereotypes is often not effective in hindering the spread of stereotypes [46].

In addition, we found an expected difference between tweets concerning dentists and medical doctors for pain-related words, with more pain-related words used in relation to dentists. This might be a testament to the significance of pain in relation to dentistry [20,21], but it poses the question why pain is not as significant in relation to medical doctors. For instance, a visit to the doctor might very well be associated with pain and discomfort. A possible reason for the current results might be how pain is perceived in these different contexts. Pain related to a health problem is most often alleviated through interaction with a medical professional (ie, a doctor) either through a medical procedure or prescribed painkillers (eg, a visit to our general practitioner for help with acute back pain or a swollen knee after a fall). This also applies to acute dental problems. However, in the case of nonacute dentistry, it might be argued that pain is caused by the visit to the dentist rather than be a byproduct of necessary examinations or treatment. This might happen because we are expected to get frequent dental check-ups to prevent dental problems even though we are symptom free [47], while the notion of preventive medical check-ups appears to be related strongly to the concept of explicit risk factors [48]. Thus, we might end up receiving painful dental treatment and suffering both physical and financial discomforts for which there is no apparent reason (to the layperson), except for the professional opinion of the dentist. Such differences in the perception of pain related to these professions might provide us with some explanation for the differences observed in the current study, and it is important to consider pain experiences in the dental setting as a key factor in determining patient satisfaction [49,50].

Limitations

As is often the case with studies of language in natural settings, the results of this study will have to be viewed in light of the inherent challenges in interpreting language and language elements (eg, manifest content) in relation to social or psychological processes (eg, latent content). Specifically, we propose that more negative words in tweets about dentists are related to the existence of a negative occupational stereotype or negative expectations related to dentists. These findings might influence, or be a reflection of, people's behaviors, beliefs, or attitudes related to oral health. How differences in word categories influence real-life learning processes or reasoning, as suggested here, is not clear. However, the relevance of investigating linguistic data and word counts in relation to thinking and behavior has been demonstrated elsewhere for a wide range of issues [51-54]. While our study results support the existence of a negative occupational stereotype and negative expectations related to dentists, as others have argued previously [22,26,55], the actual impact of the stereotypes and expectations are outside the scope of this study. Also, the specificity of the search terms and single language content will impact the generalizability of the current results. In future studies, longer study periods, inclusion of more search terms, and a deeper look into ancillary data (ie, retweets and likes of tweets) could give larger sample sizes and additional insights.

Conclusions

In conclusion, our study suggests that dentists are tweeted about in more negative terms than medical doctors. More research is needed concerning the potential impact of this on dental patients' health-related behavior and beliefs. It is unclear what can be done to reduce the proportion of dentist-related tweets with negative emotion-related words or the potential impact of a negative occupational stereotype about dentists expressed in

social media. Potential interventions, however, could include informational campaigns on social media that could underline positive aspects of dental health and dentistry [56,57], interventions highlighting preventive dental care [25], interventions aimed at reducing both actual dental costs and uncertainty about dental cost [23,58], and increasing focus on the importance of provider-patient interaction in dental education [59].

Acknowledgments

The publication charges for this article have been funded by a grant from the publication fund of UiT The Arctic University of Norway.

Authors' Contributions

JAKJ, TBE, THR, MWH, and PEK designed the study, sampled the data, analyzed the data, drafted and revised the manuscript, and approved the final manuscript. RW analyzed the data, drafted and revised the manuscript, and approved the final manuscript.

Conflicts of Interest

None declared.

References

- Ott BL. The age of Twitter: Donald J. Trump and the politics of debasement. *Critical Studies in Media Communication* 2016 Dec 23;34(1):59-68. [doi: [10.1080/15295036.2016.1266686](https://doi.org/10.1080/15295036.2016.1266686)]
- Kendra RL, Karki S, Eickholt JL, Gandy L. Characterizing the Discussion of Antibiotics in the Twittersphere: What is the Bigger Picture? *J Med Internet Res* 2015 Jun 19;17(6):e154 [FREE Full text] [doi: [10.2196/jmir.4220](https://doi.org/10.2196/jmir.4220)] [Medline: [26091775](https://pubmed.ncbi.nlm.nih.gov/26091775/)]
- Prochaska JJ, Pechmann C, Kim R, Leonhardt JM. Twitter=quitter? An analysis of Twitter quit smoking social networks. *Tob Control* 2012 Jul;21(4):447-449 [FREE Full text] [doi: [10.1136/tc.2010.042507](https://doi.org/10.1136/tc.2010.042507)] [Medline: [21730101](https://pubmed.ncbi.nlm.nih.gov/21730101/)]
- Sinnenberg L, Bittenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a Tool for Health Research: A Systematic Review. *Am J Public Health* 2017 Dec;107(1):e1-e8. [doi: [10.2105/AJPH.2016.303512](https://doi.org/10.2105/AJPH.2016.303512)] [Medline: [27854532](https://pubmed.ncbi.nlm.nih.gov/27854532/)]
- Helmi M, Spinella MK, Seymour B. Community water fluoridation online: an analysis of the digital media ecosystem. *J Public Health Dent* 2018 Mar 30. [doi: [10.1111/jphd.12268](https://doi.org/10.1111/jphd.12268)] [Medline: [29603251](https://pubmed.ncbi.nlm.nih.gov/29603251/)]
- Curiel JA, Sanders AE, Christian TL, Lafferty-Hess S, Carsey TM, Lampiris LN, et al. Fluoridation advocacy in referenda where media coverage is balanced yet biased. *J Am Dent Assoc* 2018 Apr;149(4):273-280.e3. [doi: [10.1016/j.adaj.2017.10.016](https://doi.org/10.1016/j.adaj.2017.10.016)] [Medline: [29454461](https://pubmed.ncbi.nlm.nih.gov/29454461/)]
- Oyeyemi SO, Gabarron E, Wynn R. Ebola, Twitter, and misinformation: a dangerous combination? *BMJ* 2014 Oct 14;349:g6178. [doi: [10.1136/bmj.g6178](https://doi.org/10.1136/bmj.g6178)] [Medline: [25315514](https://pubmed.ncbi.nlm.nih.gov/25315514/)]
- Wynn R, Oyeyemi SO, Johnsen JK, Gabarron E. Tweets are not always supportive of patients with mental disorders. *Int J Integr Care* 2017 Jul 11;17(3):149. [doi: [10.5334/ijic.3261](https://doi.org/10.5334/ijic.3261)]
- Macrae CN, Bodenhausen GV. Social cognition: thinking categorically about others. *Annu Rev Psychol* 2000;51:93-120. [doi: [10.1146/annurev.psych.51.1.93](https://doi.org/10.1146/annurev.psych.51.1.93)] [Medline: [10751966](https://pubmed.ncbi.nlm.nih.gov/10751966/)]
- Bodenhausen GV, Todd AR, Becker AP. Categorizing the social world: affect, motivation, and self-regulation. In: Markman AB, Ross BH, editors. *Psychology of Learning and Motivation: Advances in Research and Theory*. Amsterdam: Elsevier; 2007:123-155.
- Glick P, Wilk K, Perreault M. Images of occupations: Components of gender and status in occupational stereotypes. *Sex Roles* 1995 May;32(9-10):565-582. [doi: [10.1007/BF01544212](https://doi.org/10.1007/BF01544212)]
- Hareli S, David S, Hess U. Competent and Warm but Unemotional: The Influence of Occupational Stereotypes on the Attribution of Emotions. *J Nonverbal Behav* 2013 Jul 10;37(4):307-317. [doi: [10.1007/s10919-013-0157-x](https://doi.org/10.1007/s10919-013-0157-x)]
- Macrae CN, Milne AB, Bodenhausen GV. Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology* 1994;66(1):37-47. [doi: [10.1037/0022-3514.66.1.37](https://doi.org/10.1037/0022-3514.66.1.37)]
- Bridges JM. Literature review on the images of the nurse and nursing in the media. *J Adv Nurs* 1990 Jul;15(7):850-854. [doi: [10.1111/j.1365-2648.1990.tb01917.x](https://doi.org/10.1111/j.1365-2648.1990.tb01917.x)]
- Cleary M, Dean S, Sayers JM, Jackson D. Nursing and Stereotypes. *Issues Ment Health Nurs* 2018 Feb;39(2):192-194. [doi: [10.1080/01612840.2017.1402626](https://doi.org/10.1080/01612840.2017.1402626)] [Medline: [29405774](https://pubmed.ncbi.nlm.nih.gov/29405774/)]
- Sollami A, Caricati L, Mancini T. Ambivalent stereotypes of nurses and physicians: impact on students' attitude toward interprofessional education. *Acta Biomed* 2015 Mar 13;86 Suppl 1:19-28. [Medline: [25835762](https://pubmed.ncbi.nlm.nih.gov/25835762/)]

17. Smith MK, Dundes L. The implications of gender stereotypes for the dentist-patient relationship. *J Dent Educ* 2008 May;72(5):562-570 [FREE Full text] [Medline: 18451079]
18. Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD. Bad is stronger than good. *Review of General Psychology* 2001;5(4):323-370. [doi: 10.1037/1089-2680.5.4.323]
19. Thompson MS, Judd CM, Park B. The Consequences of Communicating Social Stereotypes. *Journal of Experimental Social Psychology* 2000 Nov;36(6):567-599. [doi: 10.1006/jesp.1999.1419]
20. Vassend O. Anxiety, pain and discomfort associated with dental treatment. *Behaviour Research and Therapy* 1993 Sep;31(7):659-666. [doi: 10.1016/0005-7967(93)90119-F]
21. Maggiri J, Locker D. Psychological factors and perceptions of pain associated with dental treatment. *Commun Dent Oral Epidemiol* 2002 Apr;30(2):151-159. [doi: 10.1034/j.1600-0528.2002.300209.x]
22. Thibodeau E, Mentasti L. Who stole Nemo? *The Journal of the American Dental Association* 2007 May;138(5):656-660. [doi: 10.14219/jada.archive.2007.0238]
23. Hill KB, White DA, Morris AJ, Hall AC, Goodwin N, Burke FJT. National evaluation of personal dental services: a qualitative investigation into patients' perceptions of dental services. *Br Dent J* 2003 Dec 06;195(11):654-656. [doi: 10.1038/sj.bdj.4810784] [Medline: 14719009]
24. Willumsen T, Haukebo K, Raadal M. Aetiology of dental phobia. In: Öst LG, Skaret E, editors. *Cognitive behaviour therapy for dental phobia and anxiety*. Chichester: John Wiley & Sons; 2013:45-61.
25. Ashford R, Blinkhorn AS. Patient care: Marketing dental care to the reluctant patient. *Br Dent J* 1999 May 8;186(9):436-441. [doi: 10.1038/sj.bdj.4800135]
26. Land T. Market research: What patients think of dental services. *Br Dent J* 2000 Jul 8;189(1):21-24. [doi: 10.1038/sj.bdj.4800588]
27. 140DEV. 2013. Twitter API Programming Tips, Tutorials, Source Code Libraries and Consulting URL: <http://140dev.com/> [accessed 2018-03-16] [WebCite Cache ID 6xxZZ4m7f]
28. Pennebaker J, Chung C, Ireland M, Gonzales A, Booth R. The Development and Psychometric Properties of LIWC2007. Austin, TX: LIWC; 2007. URL: <http://www.liwc.net/LIWC2007LanguageManual.pdf> [accessed 2018-03-15] [WebCite Cache ID 6xwCrwQY9]
29. Kahn JH, Tobin RM, Massey AE, Anderson JA. Measuring emotional expression with the Linguistic Inquiry and Word Count. *Am J Psychol* 2007;120(2):263-286. [Medline: 17650921]
30. Gohil S, Vuik S, Darzi A. Sentiment Analysis of Health Care Tweets: Review of the Methods Used. *JMIR Public Health Surveill* 2018 Apr 23;4(2):e43 [FREE Full text] [doi: 10.2196/publichealth.5789] [Medline: 29685871]
31. Thesaurus.com / Philip Lief Group. Pain URL: <http://www.thesaurus.com/browse/pain> [accessed 2018-03-15] [WebCite Cache ID 6xwDMewzm]
32. JASP Team. 2018 Feb 28. JASP (Version 0.8.6) [Computer software] URL: <https://jasp-stats.org/> [accessed 2018-12-03] [WebCite Cache ID 74O9YYsSx]
33. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp; 2016. URL: https://www.ibm.com/products/spss-statistics?mhq=IBM%20SPSS%20Statistics&mhsr=ibmsearch_p [accessed 2018-12-03]
34. Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proc Natl Acad Sci USA* 2014 Dec 17;111(24):8788-8790 [FREE Full text] [doi: 10.1073/pnas.1320040111] [Medline: 24889601]
35. Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ* 2008 Dec 04;337:a2338 [FREE Full text] [doi: 10.1136/bmj.a2338] [Medline: 19056788]
36. Humphris G, King K. The prevalence of dental anxiety across previous distressing experiences. *J Anxiety Disord* 2011 Mar;25(2):232-236. [doi: 10.1016/j.janxdis.2010.09.007] [Medline: 20952156]
37. Edmunds R, Buchanan H. Cognitive vulnerability and the aetiology and maintenance of dental anxiety. *Community Dent Oral Epidemiol* 2012 Feb;40(1):17-25. [doi: 10.1111/j.1600-0528.2011.00636.x] [Medline: 21895734]
38. Gao X, Hamzah SH, Yiu CKY, McGrath C, King NM. Dental fear and anxiety in children and adolescents: qualitative study using YouTube. *J Med Internet Res* 2013 Feb 22;15(2):e29 [FREE Full text] [doi: 10.2196/jmir.2290] [Medline: 23435094]
39. Gatchel R. The prevalence of dental fear and avoidance: expanded adult and recent adolescent surveys. *The Journal of the American Dental Association* 1989 May;118(5):591-593. [doi: 10.14219/jada.archive.1989.0068]
40. Nicolas E, Collado V, Faulks D, Bullier B, Hennequin M. A national cross-sectional survey of dental anxiety in the French adult population. *BMC Oral Health* 2007 Oct 10;7:12 [FREE Full text] [doi: 10.1186/1472-6831-7-12] [Medline: 17927808]
41. Martin RC, Coyier KR, VanSistine LM, Schroeder KL. Anger on the internet: the perceived value of rant-sites. *Cyberpsychol Behav Soc Netw* 2013 Feb;16(2):119-122. [doi: 10.1089/cyber.2012.0130] [Medline: 23249241]
42. King J. Ethics & law: Consent: the patients' view – a summary of findings from a study of patients' perceptions of their consent to dental care. *Br Dent J* 2001 Jul 14;191(1):36-40. [doi: 10.1038/sj.bdj.4801086]
43. Newsome PR, Wright GH. A review of patient satisfaction: 2. Dental patient satisfaction: an appraisal of recent literature. *Br Dent J* 1999 Feb 27;186(4 Spec No):166-170. [Medline: 10205952]

44. Stemmler G, Aue T, Wacker J. Anger and fear: Separable effects of emotion and motivational direction on somatovisceral responses. *Int J Psychophysiol* 2007 Nov;66(2):141-153. [doi: [10.1016/j.ijpsycho.2007.03.019](https://doi.org/10.1016/j.ijpsycho.2007.03.019)] [Medline: [17544534](#)]
45. Carver CS, Harmon-Jones E. Anger is an approach-related affect: evidence and implications. *Psychol Bull* 2009 Mar;135(2):183-204. [doi: [10.1037/a0013965](https://doi.org/10.1037/a0013965)] [Medline: [19254075](#)]
46. Sherman J, Allen T, Sacchi D. Stereotype confirmation and disconfirmation. In: Gawronski B, Strack F, editors. *Cognitive consistency: A fundamental principle in social cognition*. New York, NY: Guilford Press; 2012:390-423.
47. Riley P, Worthington HV, Clarkson JE, Beirne PV. Recall intervals for oral health in primary care patients. *Cochrane Database Syst Rev* 2013 Dec 19(12):CD004346. [doi: [10.1002/14651858.CD004346.pub4](https://doi.org/10.1002/14651858.CD004346.pub4)] [Medline: [24353242](#)]
48. Culica D, Rohrer J, Ward M, Hilsenrath P, Pomrehn P. Medical checkups: who does not get them? *Am J Public Health* 2002 Jan;92(1):88-91. [Medline: [11772768](#)]
49. Luo JYN, Liu PP, Wong MCM. Patients' satisfaction with dental care: a qualitative study to develop a satisfaction instrument. *BMC Oral Health* 2018 Dec 30;18(1):15 [FREE Full text] [doi: [10.1186/s12903-018-0477-7](https://doi.org/10.1186/s12903-018-0477-7)] [Medline: [29382318](#)]
50. Skaret E, Berg E, Raadal M, Kvale G. Factors related to satisfaction with dental care among 23-year olds in Norway. *Community Dent Oral Epidemiol* 2005 Apr;33(2):150-157. [doi: [10.1111/j.1600-0528.2004.00208.x](https://doi.org/10.1111/j.1600-0528.2004.00208.x)] [Medline: [15725178](#)]
51. Tausczik YR, Pennebaker JW. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology* 2009 Dec 08;29(1):24-54. [doi: [10.1177/0261927X09351676](https://doi.org/10.1177/0261927X09351676)]
52. Sylwester K, Purver M. Twitter Language Use Reflects Psychological Differences between Democrats and Republicans. *PLoS One* 2015;10(9):e0137422 [FREE Full text] [doi: [10.1371/journal.pone.0137422](https://doi.org/10.1371/journal.pone.0137422)] [Medline: [26375581](#)]
53. Bazarova NN, Taft JG, Choi YH, Cosley D. Managing Impressions and Relationships on Facebook. *Journal of Language and Social Psychology* 2012 Aug 30;32(2):121-141. [doi: [10.1177/0261927X12456384](https://doi.org/10.1177/0261927X12456384)]
54. Han JY, Shaw BR, Hawkins RP, Pingree S, McTavish F, Gustafson DH. Expressing positive emotions within online support groups by women with breast cancer. *J Health Psychol* 2008 Nov;13(8):1002-1007 [FREE Full text] [doi: [10.1177/1359105308097963](https://doi.org/10.1177/1359105308097963)] [Medline: [18987072](#)]
55. Mandel ID. The image of dentistry in contemporary culture. *The Journal of the American Dental Association* 1998 May;129(5):607-613. [doi: [10.14219/jada.archive.1998.0279](https://doi.org/10.14219/jada.archive.1998.0279)]
56. Dyson MP, Newton AS, Shave K, Featherstone RM, Thomson D, Wingert A, et al. Social Media for the Dissemination of Cochrane Child Health Evidence: Evaluation Study. *J Med Internet Res* 2017 Dec 01;19(9):e308 [FREE Full text] [doi: [10.2196/jmir.7819](https://doi.org/10.2196/jmir.7819)] [Medline: [28864427](#)]
57. Gough A, Hunter RF, Ajao O, Jurek A, McKeown G, Hong J, et al. Tweet for Behavior Change: Using Social Media for the Dissemination of Public Health Messages. *JMIR Public Health Surveill* 2017 Mar 23;3(1):e14 [FREE Full text] [doi: [10.2196/publichealth.6313](https://doi.org/10.2196/publichealth.6313)] [Medline: [28336503](#)]
58. Elstad JI. Dental care coverage and income-related inequalities in foregone dental care in Europe during the great recession. *Community Dent Oral Epidemiol* 2017 Dec;45(4):296-302. [doi: [10.1111/cdoe.12288](https://doi.org/10.1111/cdoe.12288)] [Medline: [28220588](#)]
59. Ayn C, Robinson L, Nason A, Lovas J. Determining Recommendations for Improvement of Communication Skills Training in Dental Education: A Scoping Review. *J Dent Educ* 2017 Apr;81(4):479-488 [FREE Full text] [doi: [10.21815/JDE.016.003](https://doi.org/10.21815/JDE.016.003)] [Medline: [28365612](#)]

Abbreviations

LIWC: Linguistic Inquiry and Word Count

Edited by G Eysenbach; submitted 16.03.18; peer-reviewed by R Boyd, L Sinnenberg, M El Tantawi; comments to author 17.08.18; revised version received 12.10.18; accepted 30.10.18; published 06.02.19.

Please cite as:

Johnsen JAK, Eggesvik TB, Rørvik TH, Hanssen MW, Wynn R, Kummervold PE

Differences in Emotional and Pain-Related Language in Tweets About Dentists and Medical Doctors: Text Analysis of Twitter Content
JMIR Public Health Surveill 2019;5(1):e10432

URL: <http://publichealth.jmir.org/2019/1/e10432/>

doi:[10.2196/10432](https://doi.org/10.2196/10432)

PMID:[30724738](https://pubmed.ncbi.nlm.nih.gov/30724738/)

©Jan-Are K Johnsen, Trude B Eggesvik, Thea H Rørvik, Miriam W Hanssen, Rolf Wynn, Per Egil Kummervold. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 06.02.2019. 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

Forecasting the West Nile Virus in the United States: An Extensive Novel Data Streams–Based Time Series Analysis and Structural Equation Modeling of Related Digital Searching Behavior

Abdulla Watad^{1,2,3,4}, MD; Samaa Watad⁵, BSc; Naim Mahroum^{1,2,3}, MD; Kassem Sharif^{1,2,3}, MD; Howard Amital^{1,2,3}, MHA, MD; Nicola Luigi Bragazzi^{6*}, MPH, PhD, MD; Mohammad Adawi^{7,8,9*}, MHA, MD

¹Department of Medicine B, Sheba Medical Center, Tel-Hashomer, Ramat-Gan, Israel

²Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel

³Zabludowicz Center for Autoimmune Diseases, Sheba Medical Center, Tel-Hashomer, Ramat-Gan, Israel

⁴NIHR Leeds Musculoskeletal Biomedical Research Unit, Section of Musculoskeletal Disease, Leeds Institute of Molecular Medicine, Chapel Allerton Hospital, University of Leeds, Leeds, United Kingdom

⁵Department of Statistics and Operations Research, Tel-Aviv University, Tel-Aviv, Israel

⁶Department of Health Sciences, Postgraduate School of Public Health, University of Genoa, Genoa, Italy

⁷The Baruch Padeh Medical Center, Zefat, Israel

⁸Azrieli Faculty of Medicine, Bar-Ilan University, Zefat, Israel

⁹Ziv Medical Center, Zefat, Israel

*these authors contributed equally

Corresponding Author:

Nicola Luigi Bragazzi, MPH, PhD, MD

Department of Health Sciences

Postgraduate School of Public Health

University of Genoa

Via Antonio Pastore 1

Genoa, 16132

Italy

Phone: 39 0103538508

Fax: 39 0103538541

Email: robertobragazzi@gmail.com

Abstract

Background: West Nile virus is an arbovirus responsible for an infection that tends to peak during the late summer and early fall. Tools monitoring Web searches are emerging as powerful sources of data, especially concerning infectious diseases such as West Nile virus.

Objective: This study aimed at exploring the potential predictive power of West Nile virus–related Web searches.

Methods: Different novel data streams, including Google Trends, WikiTrends, YouTube, and Google News, were used to extract search trends. Data regarding West Nile virus cases were obtained from the Centers for Disease Control and Prevention. Data were analyzed using regression, times series analysis, structural equation modeling, and clustering analysis.

Results: In the regression analysis, an association between Web searches and “real-world” epidemiological figures was found. The best seasonal autoregressive integrated moving average model with explicative variable (SARIMAX) was found to be (0,1,1)x(0,1,1)₄. Using data from 2004 to 2015, we were able to predict data for 2016. From the structural equation modeling, the consumption of West Nile virus–related news fully mediated the relation between Google Trends and the consumption of YouTube videos, as well as the relation between the latter variable and the number of West Nile virus cases. Web searches fully mediated the relation between epidemiological figures and the consumption of YouTube videos, as well as the relation between epidemiological data and the number of accesses to the West Nile virus–related Wikipedia page. In the clustering analysis, the consumption of news was most similar to the Web searches pattern, which was less close to the consumption of YouTube videos and least similar to the behavior of accessing West Nile virus–related Wikipedia pages.

Conclusions: Our study demonstrated an association between epidemiological data and search patterns related to the West Nile virus. Based on this correlation, further studies are needed to examine the practicality of these findings.

(*JMIR Public Health Surveill* 2019;5(1):e9176) doi:[10.2196/publichealth.9176](https://doi.org/10.2196/publichealth.9176)

KEYWORDS

forecasting model; West Nile virus; Google Trends; infodemiology; infoveillance; seasonal autoregressive integrated moving average model with explicative variable (SARIMAX)

Introduction

West Nile virus, first isolated in Uganda in 1937, is a widely distributed arbovirus belonging to the *Flavivirus* genus and to the *Flaviviridae* family that can cause West Nile fever. This mosquito-borne infection has a seasonal trend with peaks during summer and autumn. In 70% to 80% of West Nile fever cases, no or few symptoms are reported [1]. Symptomatic infections generally consist of self-limited influenza-like illness with high-degree fever, chills, myalgia, and arthralgia, which last for approximately 5 days. Although rare at a rate of approximately 1%, neurological diseases, such as meningitis, encephalitis, meningoencephalitis, or poliomyelitis, can occur and are of great concern because they are characterized by many sequelae, especially among elderly hospitalized patients [2].

West Nile virus was detected in North America for the first time in August 1999 during an outbreak that occurred in College Point, Queens, in New York City. A cluster of human encephalitis cases, all residing in the same 16-square-mile area, was identified by Drs Deborah Asnis (a local physician based in Queens), Marcelle Layton, and Annie Fine (of the New York City Department of Health) [2]. This cluster was preceded by anecdotal reports of dead animals and birds at the Bronx Zoo, including American crows (*Corvus brachyrhynchos*), Chilean flamingos (*Phoenicopterus chilensis*), and a snowy owl (*Bubo scandiacus*). New human cases were subsequently diagnosed in Brooklyn, the Bronx, and Manhattan. Since then, West Nile virus has spread to the contiguous states, from the Mississippi River to the Pacific Coast, causing further outbreaks such as those that occurred during the summer of 2002. Cases decreased from 2008 to 2011, but in 2012 a major outbreak took place with its epicenter in the Dallas-Fort Worth, Texas metroplex [3,4]. The Texas Public Health authority issued a public health emergency, which attracted a lot of media coverage and public opinion reaction.

The widespread resurgence of human West Nile virus disease in 2012 following several years of relatively low incidence has highlighted the continued public health hazard posed by West Nile virus, and has emphasized the need for more accurate predictive models of when and where new West Nile virus outbreaks will occur.

Web-based tools are emerging as remarkable sources of data, especially for infectious diseases, by enabling Web search monitoring in real time and potentially capturing epidemiologically relevant information [5-7]. *Infodemiology* (a portmanteau of information and epidemiology) and *infoveillance* (a portmanteau of information and surveillance) indicate the emerging “science of distribution and determinants of

information in an electronic medium, specifically the internet, or in a population, with the ultimate aim to inform and improve public health and public policy” [8]. Systematically tracking and monitoring, collecting, and analyzing health-related demand data generated by novel data streams could have the potential to predict events relevant for public health purposes, such as epidemic outbreaks, as well as to investigate the effect of media coverage in terms of potential distortions, misinformation, and biases—the so-called “epidemics of fear” [9].

Little is known about West Nile virus-related digital behavior. To the best of our knowledge, only a few authors have investigated this topic. Carneiro and Mylonakis [10] reported a preliminary qualitative observation about a positive correlation between Google searches and West Nile virus epidemiological cases from 2004 to 2008. They found that the search volume exhibited a cyclical pattern, with regular peaks in August each year, reproducing the epidemiological figures. Also, Web searches related to West Nile fever symptoms (fever, headache, fatigue, rash, and eye pain) were characterized by seasonal patterns. The authors noticed increases in search volume for rash starting in May, just a month before the increases in cases. Furthermore, the top-ranked US cities in terms of West Nile virus-related search volumes were in states characterized by the highest epidemiological burden.

Bragazzi and collaborators [11] assessed the association between Web searches and cases in Italy from a quantitative standpoint from 2004 to 2015. They found a correlation of $r=.76$ ($P<.001$) and $r=.80$ ($P<.001$) between Google searches and “real-world” epidemiological cases in the same study period on a monthly basis and a yearly basis, respectively. The presence of a regular pattern in West Nile virus-related Web queries was confirmed by the partial autocorrelation function analysis and by spectral analysis. From a geospatial point of view, correlation between digital behavior and epidemiological figures yielded $r=.54$ ($P<.05$).

However, the potential predictive power of West Nile virus-related Web searches has not yet been explored. To fill this gap in knowledge, we conducted this study.

Methods

West Nile virus-related data were retrieved, downloaded, and analyzed from several novel data streams, including Google Trends, WikiTrends, YouTube, and Google News, as well as from epidemiological repositories.

Novel Data Streams

Google Trends (an open source tool) was mined from inception (2004) to 2015, by searching for West Nile virus in the United

States and using the “search topic” option. This strategy enables one to systematically collect all the searches related to a given keyword or list of keywords (in this case, West Nile virus), including synonyms and related terms, not just the precise string of characters typed by users [12]. WikiTrends is a freely available tool that could be used to investigate information seeking behavior concerning West Nile virus. It was mined from inception (2008) to 2015. The viewing of YouTube videos was investigated from 2008 to 2015, using Google Trends and selecting “YouTube” option. Finally, Google News is an open source news aggregator that can be used to explore the media coverage of a given topic. The consumption of West Nile virus-related news was explored from 2008 to 2015 using Google Trends and selecting “Google News” option. For further details concerning novel data streams, the reader is referred to Bragazzi et al [12].

Epidemiological Repositories

Epidemiological data related to West Nile virus cases in the United States were obtained from the Centers for Disease Control and Prevention (CDC) website and the bulletins of the *Morbidity and Mortality Weekly Report*, a publication of the CDC (data available on a trimester basis).

Statistical Analysis

Novel data streams-generated data were retrieved and downloaded from 2004 for Google Trends and 2008 for the other open source tools to 2015. All data were analyzed on a trimester basis. To detect a potential association with “real-world” epidemiological figures, regression analyses (with time as the confounding variable) were carried out. Furthermore, novel data streams-generated data were modeled as a time series

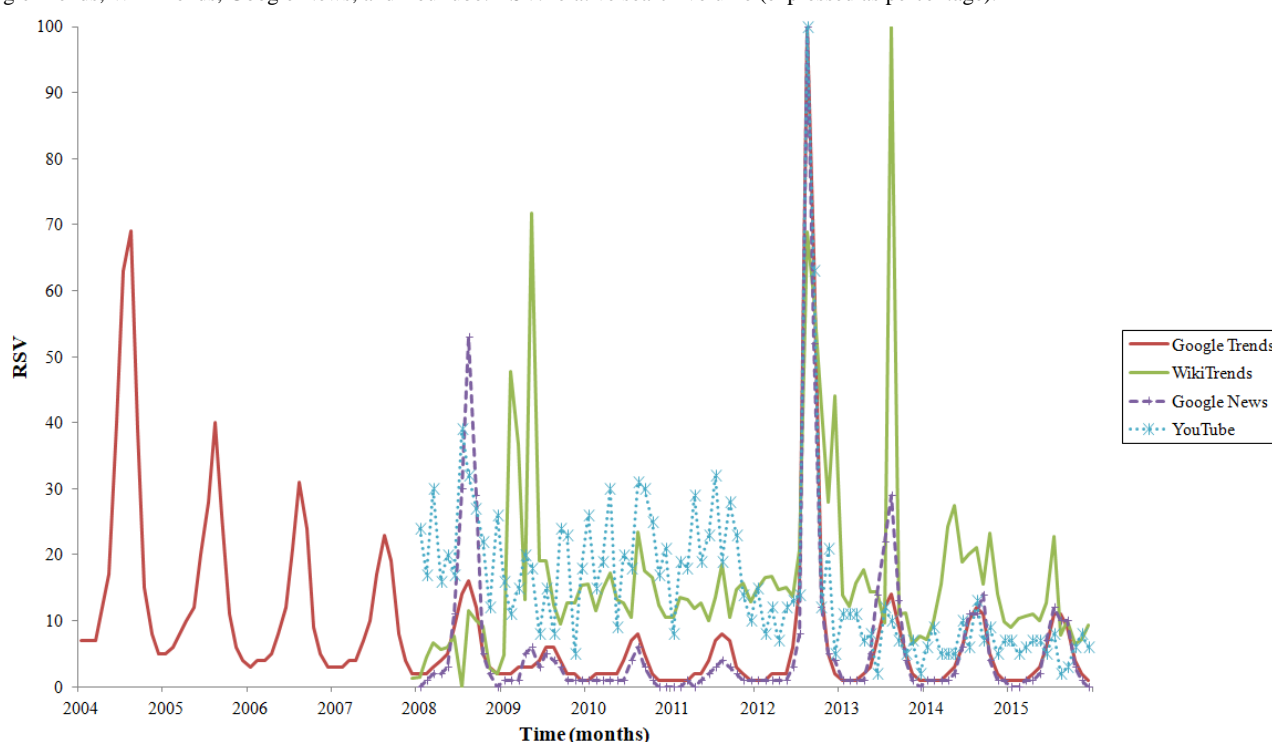
and analyzed using time series analyses. In particular, a seasonal autoregressive integrated moving average model with explicative variable (SARIMAX) was used. By visually inspecting the autocorrelogram and the partial autocorrelogram based on the autocorrelation and partial autocorrelation function, respectively, p (the order of the model or, in other words, the number of time lags), d (the degree of differencing of the model), q (the order of the moving average model), P (the order of the seasonal part of the model), D (the degree of differencing of the seasonal part of the model), and Q (the order of the moving average model for the seasonal part) coefficients and s (lag parameter) were determined. The explicative variable was the number of West Nile virus cases. Different models were run, and the best one was chosen based on the Akaike information criterion (AIC), corrected AIC, and Schwartz Bayesian information criterion values. The best model was used to forecast Google Trends-based relative search volumes for 2016. Furthermore, structural equation modeling and clustering analysis were used to capture the complex interplay between the different novel data streams.

Regression and clustering analyses were performed using SPSS version 24.0 (IBM Corp, Armonk, NY, USA), whereas the SARIMAX models and the structural equation modeling were carried out with XLSTAT (Addinsoft, Paris, France). A $P < .05$ was considered statistically significant.

Results

Visual inspection of novel data streams-based data showed that each tool captured a specific digital behavior, generating specific curves which were not perfectly superimposable (Figure 1).

Figure 1. Temporal pattern of searching behavior related to the West Nile virus in the United States, as captured by four different novel data streams: Google Trends, WikiTrends, Google News, and YouTube. RSV: relative search volume (expressed as percentage).



Concerning temporal trends, only West Nile virus-related Web searches pattern well-reproduced the epidemiological trend, with most Google queries concentrated in August. For instance, the number of accesses to the West Nile virus-related Wikipedia page (as captured by WikiTrends) and the consumption of YouTube videos exhibited high search volumes also during winter months compared to Google Trends (Figure 2).

Regression analyses showed a significant correlation between real-world epidemiological data and novel data streams-generated figures only for Google Trends data (Table 1 and Figure 3), with the effect of year ($P=.001$) and of West Nile virus cases ($P<.001$) reaching statistical significance.

Figure 2. Seasonal pattern of searching behavior related to the West Nile virus in the United States, as captured by four different novel data streams: Google Trends, WikiTrends, Google News, and YouTube. RSV: relative search volume (expressed as percentage).

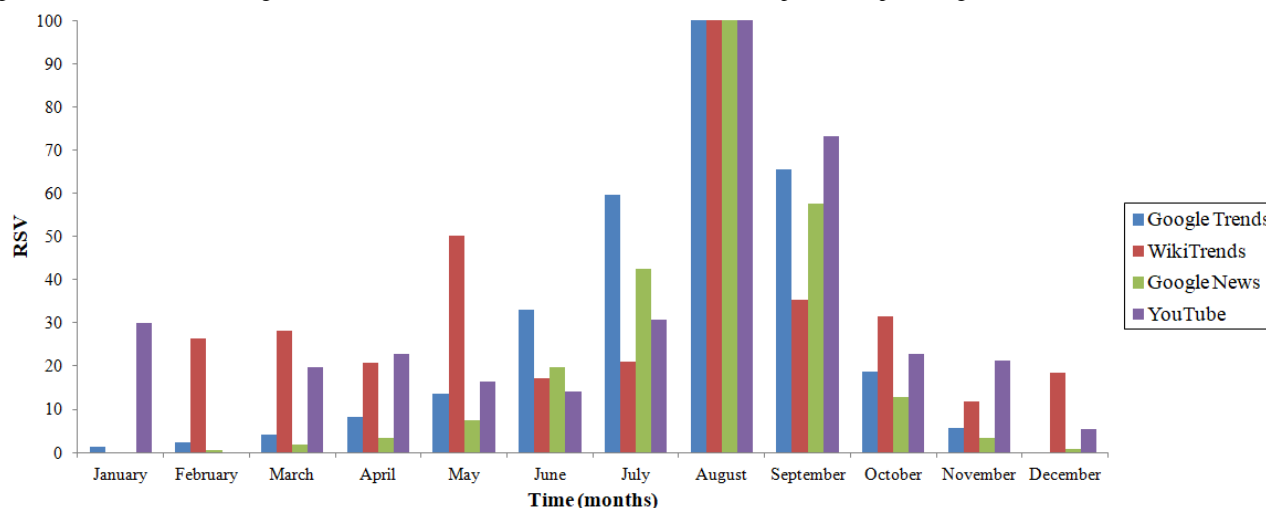
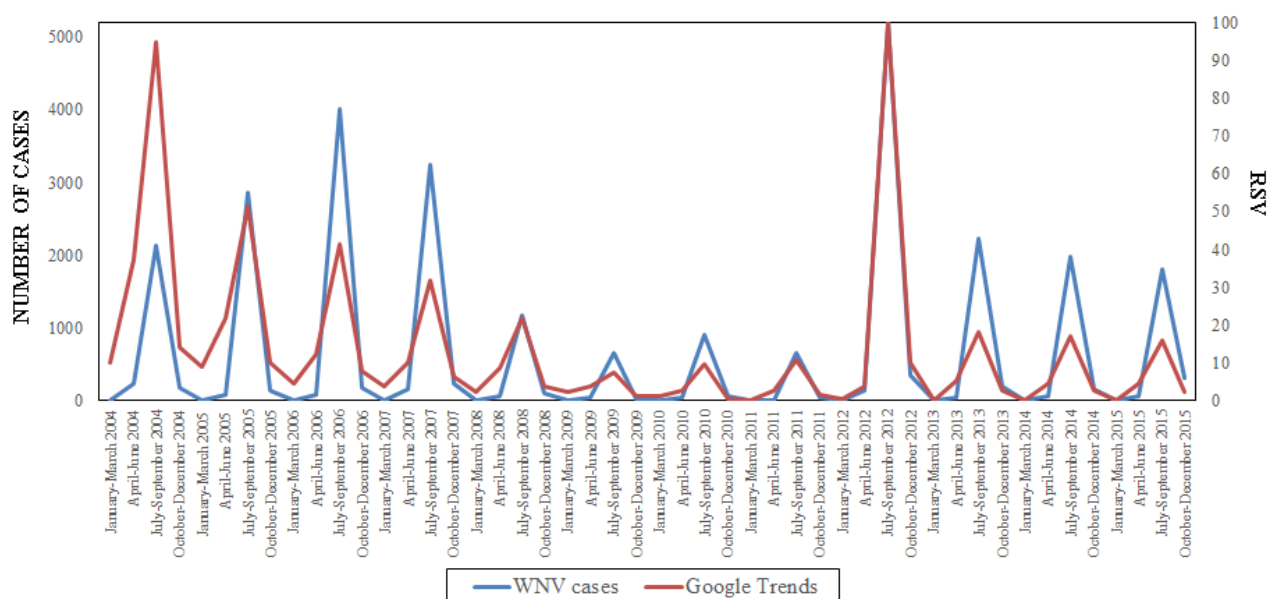


Table 1. Regression analyses to detect potential association between novel data streams (Google Trends, WikiTrends, Google News, and YouTube) and real-world epidemiological figures.

Source	Regression coefficient	SE	95% CI	t_{31}	P value
Google Trends					
Intercept	3327.876	966.213	1380.603, 5275.150	3.444	.001
Trimester	-0.531	1.535	-3.624, 2.561	-0.346	.73
Year	-1.653	0.481	-2.622, -0.684	-3.438	.001
West Nile virus cases	0.014	0.001	0.011, 0.017	9.629	<.001
WikiTrends					
Intercept	-7402.427	4130.337	-15863.039, 1058.184	-1.792	.08
Trimester	1.715	4.288	-7.068, 10.498	0.400	.69
Year	3.694	2.053	-0.512, 7.900	1.799	.08
West Nile virus cases	-0.003	0.005	-0.012, 0.007	-0.566	.58
Google News					
Intercept	8875.080	3807.169	1076.447, 16673.712	2.331	.03
Trimester	1.478	3.952	-6.618, 9.574	0.374	.71
Year	-4.407	1.893	-8.284, -0.530	-2.328	.03
West Nile virus cases	-0.001	0.004	-0.010, 0.008	-0.237	.81
YouTube					
Intercept	3297.355	3606.758	-4090.754, 10685.464	0.914	.37
Trimester	3.454	3.744	-4.216, 11.124	0.923	.36
Year	-1.622	1.793	-5.295, 2.051	-0.904	.37
West Nile virus cases	-0.002	0.004	-0.010, 0.007	-0.453	.65

Figure 3. Correlation between real-world epidemiological figures of West Nile virus (WNV) cases and digital searches. RSV: relative search volume (expressed as percentage).



A Google Trends-based autocorrelation and partial autocorrelation are reported in Figure 4. These show statistically significant positive spikes for lags 0, 4, and 8 and lags 0 and 4, respectively. Descriptive statistics for Google Trends-generated data modeled as a time series is shown in Table 2. The best SARIMAX model was found to be $(0,1,1) \times (0,1,1)_4$ (Multimedia Appendix 1 and Figure 5), or a “seasonal exponential smoothing” model, being $MA(1) \times SMA(1)$. This kind of model represents a variation of the seasonal random trend, with a fine tuning obtained adding the $MA(1)$ and the $SMA(1)$ components. Its parameters are reported in Table 3.

Concerning structural equation modeling, the consumption of West Nile virus-related news fully mediated the relationship between Google Trends and the consumption of YouTube videos, as well as the relation between this latter variable and the number of West Nile virus cases. Web searches as captured by Google Trends fully mediated the relation between West

Nile virus cases and the consumption of YouTube videos, as well as the relation between epidemiological data and the number of accesses to the West Nile virus-related Wikipedia page as captured by WikiTrends (Figure 6a). When adjusting for time as a potential confounding factor (Figure 6b), the consumption of YouTube videos mediated by the consumption of news was found to increase throughout time in a statistically significant way, although when mediated by the number of accesses to the West Nile virus-related Wikipedia page as captured by WikiTrends tended to decrease. Interestingly, the West Nile virus-related Web search behavior decreased over time (as captured by Google Trends and mediated by the number of epidemiological cases).

Clustering analysis showed that the consumption of news was most similar to the Web searches pattern (as captured by Google Trends), which was less close to the consumption of YouTube videos and least similar to accessing the West Nile virus-related Wikipedia page as captured by WikiTrends (as can be seen by the dendrogram in Figure 7).

Figure 4. Autocorrelation and partial autocorrelation of West Nile virus-related search volumes generated on Google Trends.

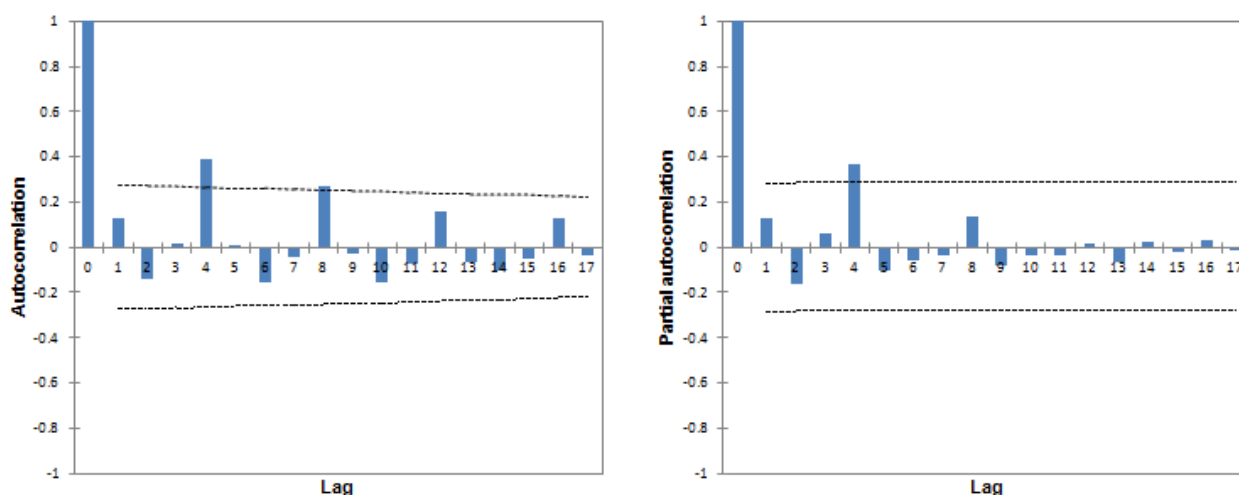


Table 2. Descriptive statistics of the Google Trends-generated data concerning Web queries related to the West Nile virus.

Lag	Autocovariance	Autocorrelation	SE	95% CI	Partial autocorrelation	SE	95% CI
0	430.22	1.00	0.00	Ref ^a	1.00	0.00	Ref
1	53.83	0.13	0.14	-0.27, 0.274	0.13	0.14	-0.28, 0.28
2	-60.95	-0.14	0.14	-0.27, 0.271	-0.16	0.14	-0.28, 0.28
3	7.85	0.02	0.14	-0.27, 0.268	0.06	0.14	-0.28, 0.28
4	166.97	0.38	0.14	-0.27, 0.265	0.37	0.14	-0.28, 0.28
5	3.73	0.01	0.13	-0.26, 0.262	-0.11	0.14	-0.28, 0.28
6	-66.73	-0.16	0.13	-0.26, 0.259	-0.06	0.14	-0.28, 0.28
7	-19.78	-0.05	0.13	-0.26, 0.256	-0.04	0.14	-0.28, 0.28
8	114.81	0.27	0.13	-0.25, 0.253	0.14	0.14	-0.28, 0.28
9	-13.47	-0.03	0.13	-0.25, 0.250	-0.08	0.14	-0.28, 0.28
10	-66.81	-0.16	0.13	-0.25, 0.247	-0.04	0.14	-0.28, 0.28
11	-30.33	-0.07	0.12	-0.24, 0.243	-0.04	0.14	-0.28, 0.28
12	67.99	0.16	0.12	-0.24, 0.240	0.01	0.14	-0.28, 0.28
13	-27.40	-0.06	0.12	-0.24, 0.237	-0.07	0.14	-0.28, 0.28
14	-45.30	-0.11	0.12	-0.23, 0.233	0.02	0.14	-0.28, 0.28
15	-22.67	-0.05	0.12	-0.23, 0.230	-0.02	0.14	-0.28, 0.28
16	56.78	0.13	0.12	-0.23, 0.226	0.03	0.14	-0.28, 0.28
17	-15.62	-0.04	0.11	-0.22, 0.223	-0.01	0.14	-0.28, 0.28

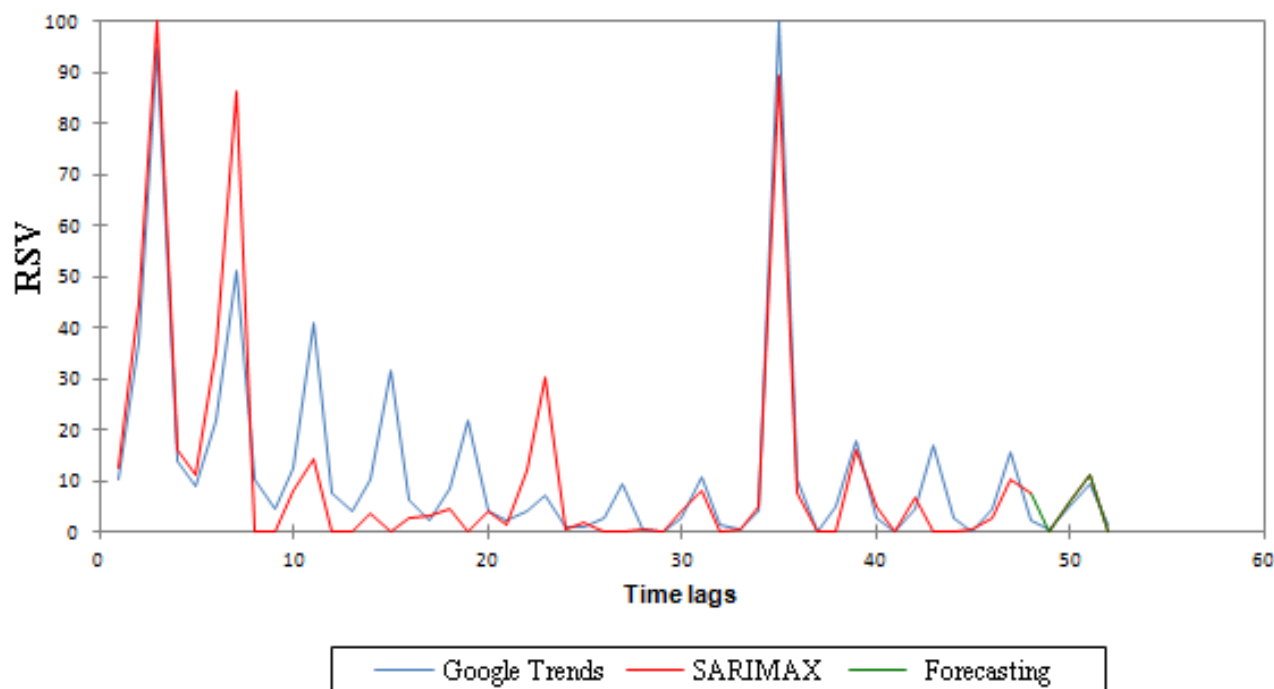
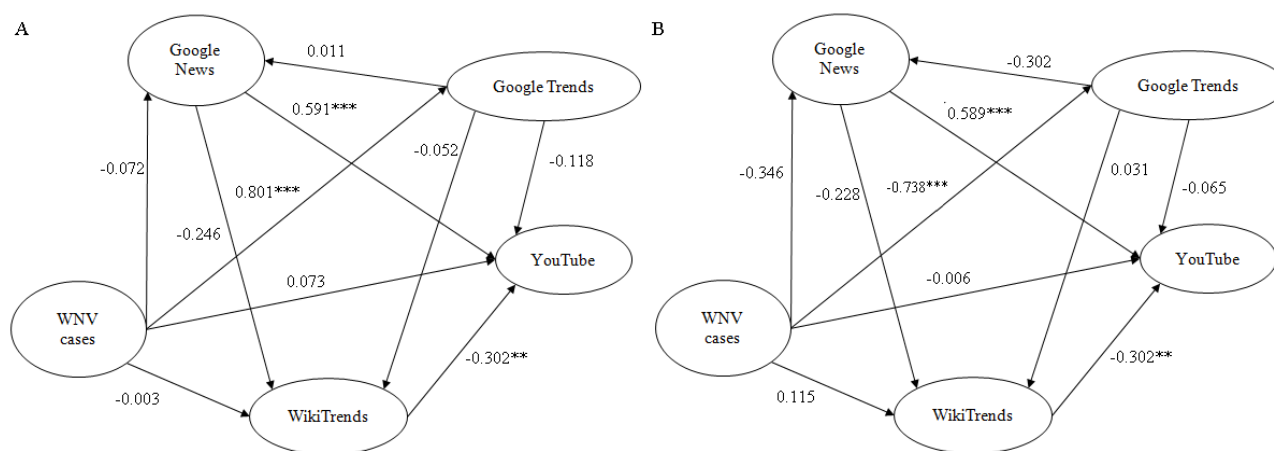
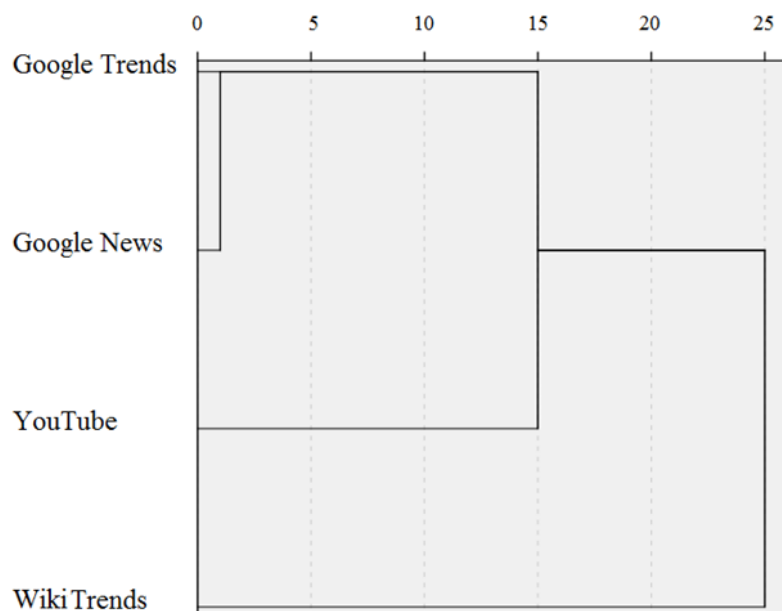
^aRef: reference.**Figure 5.** The outcome of the best seasonal autoregressive integrated moving average model with explicative variable (SARIMAX) forecasting the West Nile virus in the United States using Google Trends-generated data. RSV: relative search volume (expressed as percentage).

Table 3. Parameters of the best seasonal autoregressive integrated average model with explicative variable (SARIMAX) for forecasting West Nile virus in the United States using Google Trends-generated data.

Parameter	Value	Hessian SD	95% CI	Asymptotic SD	95% CI
Constant	4.261	Ref ^a	Ref	Ref	Ref
West Nile virus cases	0.022	0.055	-0.086, 0.130	Ref	Ref
MA ^b (1)	-0.867	0.101	-1.065, -0.670	0.124	-1.110, -0.624
SMA ^c (1)	0.672	0.120	0.436, 0.907	0.150	0.379, 0.965

^aRef: reference.^bMA: nonseasonal component.^cSMA: seasonal component.**Figure 6.** Structural equation model showing the interplay between the different novel data streams concerning West Nile virus-related searching behavior: (a) not adjusted and (b) adjusted for time as confounding variable.**Figure 7.** Dendrogram analysis of the four novel data streams (Google Trends, Google News, YouTube, and WikiTrends). Units are arbitrary.

Discussion

Principal Findings

Currently, arboviruses are re-emerging infectious agents. This is not a new phenomenon—it has been happening for

centuries—but today arboviral re-emergence and dispersion are more rapid and geographically extensive mainly due to globalization and to arthropod adaptation to its effects [13].

In the existing scholarly literature, different predictive models of West Nile virus have been reported. Kala and colleagues [14]

described a geographically weighted regression modeling of West Nile virus risk based on environmental parameters (ie, stream density, road density, and land surface temperature) being statistically superior to classical approaches relying on ordinary least squares regression analyses in terms of predictive power. Shand and coworkers [15] improved the performance of the mosquito-based surveillance approach by incorporating rainfall, temperature, and interaction terms between precipitation and temperature as predictors. Rochlin and coauthors [16] also exploited socioeconomic factors and found a higher proportion of the population with college education, increased habitat fragmentation, and proximity to West Nile virus-positive mosquito pools correlated with higher West Nile virus human risk, whereas Trawinski and Mackay [17] relied on meteorological parameters (including average, minimum, and maximum temperature values; precipitation; relative humidity; and evapotranspiration). Day and Shaman [18] used water table depth as a measure of drought and as a proxy of arboviral transmission in two peninsular Florida regions. In another investigation, Shaman and coworkers [19] exploited hydrological variables and found that wetter spring conditions and drier summer conditions predicted increased West Nile virus human risk in Colorado's eastern plains. Ghosh and Guha [20] performed a computational neural network analysis for capturing the nonlinear complex relationship between West Nile virus epidemiology and several different variables, including temperature, precipitation, wetlands, housing age, presence of catch basins and ditches, and mosquito abatement policies. Establishing an accurate predictive model is of crucial importance in arboviral infection control.

In our work, we exploited a variable (West Nile virus-related digital behavior), which has so far not been used in predicting West Nile virus epidemiology. We explored different novel data streams (Google Trends, WikiTrends, Google News, and YouTube) concerning seeking behavior and we were able to find a statistically significant association between epidemiological figures and digital behavior only in the case of Web searches as captured by Google Trends. Furthermore, we computed the best SARIMAX model for the period of 2004 to 2015, and we were able to forecast data related to 2016. Moreover, structural equation modeling and clustering analysis have enabled us to capture the complex interplay between the different novel data streams and the West Nile virus-related digital seeking behavior.

Even if our experience suggests the usefulness of using Google Trends for predicting West Nile virus, this should be considered as a pilot study, calling for the need for making our model more accurate and reliable, and maybe incorporating other variables (eg, environmental, socioeconomic, and ecological ones). This is of fundamental importance when designing and implementing a digital system for West Nile virus surveillance, which could complement the classical one or those actually under experimentation [21,22]. The combination of Google Trends and other predictors could reach an adequate temporal concordance with the real-world epidemiological figures and, therefore, could enable nowcasting or forecasting of new West Nile virus cases.

Our study has some limitations that should be recognized. Some of the novel data streams used provide users with relative, normalized figures, and not with raw, absolute data, thus hindering further mathematical processing and statistical analysis. Another drawback is given by the fact that Google Trends captures only a portion of the entire population, namely the percentage of people using Google as their preferred search engine (although Google is the most commonly used search engine worldwide). Furthermore, we did not perform a content analysis of the West Nile virus-related material; from the existing literature, it is known to be of rather poor quality and to exhibit some degrees of inconsistencies [23,24]. For instance, Birnbrauer and colleagues [23] explored how West Nile virus risk information was portrayed from its 1999 arrival in the United States through the year 2012, analyzing 428 articles obtained through Google News. Authors identified the following themes and topics: action, conflict, consequence, new evidence, reassurance, and uncertainty, with the action frame recurring most frequently. Moreover, West Nile virus risk was found to be improperly communicated, with statistical figures generally inaccurately reported. Dubey and coworkers [24] analyzed a total of 106 West Nile virus-related YouTube videos, 79.24% of which were found to contain useful information about the disease (60.71% related to disease prevention, and 34.52% concerning news and research updates). Videos were typically uploaded by individuals (54.6%) or news agencies (41.8%), but rarely by health care agencies (3.4%). Despite the usefulness of most West Nile virus-related videos, nonuseful videos received more views, both overall and on a daily basis. Moreover, West Nile virus-related digital behavior could have been influenced and, eventually, also distorted by extrinsic variables, such as the media coverage in terms of dissemination of imbalanced and biased information [25]. Some articles have shown that Google Trends does not always match with epidemiological data [25,26], such as in the case of Google Flu Trends [27], even though it is feasible to exploit some statistical techniques to externally revise novel data streams-generated figures, recalibrate them, and improve their accuracy and predictive power [28,29]. As such, the field of "behavioral medicine" remains largely unexplored [30], and because traditional surveillance is plagued by intrinsic limitations, enhanced methods for identification of real-time new cases and assessment of disease patterns and trends are urgently needed [31].

Conclusions

Statistically significant temporal correlations between West Nile virus epidemiological data and Google Trends suggest the feasibility of exploiting Google Trends as an internet-based monitoring tool. This is timely and of crucial importance given the recent re-emergence of arboviral infections. Workers in the field of public health and health authorities should be aware of the public interest and reaction to West Nile virus outbreaks in terms of Web searches. They could exploit the new information and communication technologies both for performing real-time monitoring of new population-based epidemic events and for carrying out a content analysis of the available online material, promptly replying to public concerns and correcting prejudices and inaccurate and misleading reports by disseminating

high-quality information. However, based on the previously mentioned limitations of this paper, further studies are warranted to make our model more useful and practical.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Different tested seasonal autoregressive integrated average (SARIMAX) models for forecasting the West Nile virus in the United States using Google Trends-generated data.

[PDF File (Adobe PDF File), 311KB - [publichealth_v5i1e9176_app1.pdf](#)]

References

1. Saxena V, Bolling BG, Wang T. West Nile virus. Clin Lab Med 2017 Jun;37(2):243-252. [doi: [10.1016/j.cll.2017.01.001](#)] [Medline: [28457348](#)]
2. Kleinschmidt-DeMasters BK, Beckham JD. West Nile virus encephalitis 16 years later. Brain Pathol 2015 Sep;25(5):625-633. [doi: [10.1111/bpa.12280](#)] [Medline: [26276026](#)]
3. Calisher CH. West Nile virus in the New World: appearance, persistence, and adaptation to a new econiche-an opportunity taken. Viral Immunol 2000;13(4):411-444. [doi: [10.1089/vim.2000.13.411](#)] [Medline: [11192287](#)]
4. Williamson PC, Custer B, Biggerstaff BJ, Lanciotti RS, Sayers MH, Eason SJ, et al. Incidence of West Nile virus infection in the Dallas-Fort Worth metropolitan area during the 2012 epidemic. Epidemiol Infect 2017 Sep;145(12):2536-2544. [Medline: [26829991](#)]
5. Pelat C, Turbelin C, Bar-Hen A, Flahault A, Valleron AJ. More diseases tracked by using Google Trends. Emerg Infect Dis 2009 Aug;15(8):1327-1328 [FREE Full text] [doi: [10.3201/eid1508.090299](#)] [Medline: [19751610](#)]
6. Alicino C, Bragazzi NL, Faccio V, Amicizia D, Panatto D, Gasparini R, et al. Assessing Ebola-related web search behaviour: insights and implications from an analytical study of Google Trends-based query volumes. Infect Dis Poverty 2015 Dec 10;4:54 [FREE Full text] [doi: [10.1186/s40249-015-0090-9](#)] [Medline: [26654247](#)]
7. Bragazzi NL, Barberis I, Rosselli R, Gianfredi V, Nucci D, Moretti M, et al. How often people google for vaccination: qualitative and quantitative insights from a systematic search of the web-based activities using Google Trends. Hum Vaccin Immunother 2017 Dec;13(2):464-469 [FREE Full text] [doi: [10.1080/21645515.2017.1264742](#)] [Medline: [27983896](#)]
8. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009 Mar 27;11(1):e11 [FREE Full text] [doi: [10.2196/jmir.1157](#)] [Medline: [19329408](#)]
9. Eysenbach G. Infodemiology: the epidemiology of (mis)information. Am J Med 2002 Dec 15;113(9):763-765. [Medline: [12517369](#)]
10. Carneiro HA, Mylonakis E. Google trends: a web-based tool for real-time surveillance of disease outbreaks. Clin Infect Dis 2009 Nov 15;49(10):1557-1564. [doi: [10.1086/630200](#)] [Medline: [19845471](#)]
11. Bragazzi NL, Bacigaluppi S, Robba C, Siri A, Canepa G, Brigo F. Infodemiological data of West-Nile virus disease in Italy in the study period 2004-2015. Data Brief 2016 Dec;9:839-845 [FREE Full text] [doi: [10.1016/j.dib.2016.10.022](#)] [Medline: [27872881](#)]
12. Bragazzi NL, Dini G, Toletone A, Brigo F, Durando P. Leveraging big data for exploring occupational diseases-related interest at the level of scientific community, media coverage and novel data streams: the example of silicosis as a pilot study. PLoS One 2016;11(11):e0166051 [FREE Full text] [doi: [10.1371/journal.pone.0166051](#)] [Medline: [27806115](#)]
13. Gould E, Pettersson J, Higgs S, Charrel R, de Lamballerie X. Emerging arboviruses: why today? One Health 2017 Dec;4:1-13 [FREE Full text] [doi: [10.1016/j.onehlt.2017.06.001](#)] [Medline: [28785601](#)]
14. Kala AK, Tiwari C, Mikler AR, Atkinson SF. A comparison of least squares regression and geographically weighted regression modeling of West Nile virus risk based on environmental parameters. PeerJ 2017 Mar 28;5:e3070 [FREE Full text] [doi: [10.7717/peerj.3070](#)] [Medline: [28367364](#)]
15. Shand L, Brown WM, Chaves LF, Goldberg TL, Hamer GL, Haramis L, et al. Predicting West Nile virus infection risk from the synergistic effects of rainfall and temperature. J Med Entomol 2016 Jul;53(4):935-944. [doi: [10.1093/jme/tjw042](#)] [Medline: [27113111](#)]
16. Rochlin I, Turbow D, Gomez F, Ninivaggi DV, Campbell SR. Predictive mapping of human risk for West Nile virus (WNV) based on environmental and socioeconomic factors. PLoS One 2011;6(8):e23280 [FREE Full text] [doi: [10.1371/journal.pone.0023280](#)] [Medline: [21853103](#)]
17. Trawinski PR, Mackay DS. Meteorologically conditioned time-series predictions of West Nile virus vector mosquitoes. Vector Borne Zoonotic Dis 2008 Aug;8(4):505-521. [doi: [10.1089/vbz.2007.0202](#)] [Medline: [18279008](#)]

18. Day JF, Shaman J. Using hydrologic conditions to forecast the risk of focal and epidemic arboviral transmission in peninsular Florida. *J Med Entomol* 2008 May;45(3):458-465. [Medline: [18533440](#)]
19. Shaman J, Day JF, Komar N. Hydrologic conditions describe West Nile virus risk in Colorado. *Int J Environ Res Public Health* 2010 Dec;7(2):494-508 [FREE Full text] [doi: [10.3390/ijerph7020494](#)] [Medline: [20616987](#)]
20. Ghosh D, Guha R. Using a neural network for mining interpretable relationships of West Nile risk factors. *Soc Sci Med* 2011 Feb;72(3):418-429. [doi: [10.1016/j.socscimed.2010.09.014](#)] [Medline: [20950908](#)]
21. Konrad SK, Zou L, Miller SN. A geographical information system-based web model of arbovirus transmission risk in the continental United States of America. *Geospat Health* 2012 Nov;7(1):157-159. [doi: [10.4081/gh.2012.114](#)] [Medline: [23242689](#)]
22. Carney R, Ahearn SC, McConchie A, Glasner C, Jean C, Barker C, et al. Early warning system for West Nile virus risk areas, California, USA. *Emerg Infect Dis* 2011 Aug;17(8):1445-1454 [FREE Full text] [doi: [10.3201/eid1708.100411](#)] [Medline: [21801622](#)]
23. Birnbrauer K, Frohlich DO, Treise D. Inconsistencies in reporting risk information: a pilot analysis of online news coverage of West Nile Virus. *Glob Health Promot* 2017 Sep;24(3):14-22. [doi: [10.1177/1757975915594603](#)] [Medline: [26265595](#)]
24. Dubey D, Amritphale A, Sawhney A, Dubey D, Srivastav N. Analysis of YouTube as a source of information for West Nile Virus infection. *Clin Med Res* 2014 Dec;12(3-4):129-132 [FREE Full text] [doi: [10.3121/cmr.2013.1194](#)] [Medline: [24573700](#)]
25. Bragazzi NL, Alicino C, Trucchi C, Paganino C, Barberis I, Martini M, et al. Global reaction to the recent outbreaks of Zika virus: insights from a Big Data analysis. *PLoS One* 2017 Sep 21;12(9):e0185263 [FREE Full text] [doi: [10.1371/journal.pone.0185263](#)] [Medline: [28934352](#)]
26. Adawi M, Bragazzi NL, Watad A, Sharif K, Amital H, Mahroum N. Discrepancies between classic and digital epidemiology in searching for the Mayaro virus: preliminary qualitative and quantitative analysis of Google Trends. *JMIR Public Health Surveill* 2017 Dec 01;3(4):e93 [FREE Full text] [doi: [10.2196/publichealth.9136](#)] [Medline: [29196278](#)]
27. Lazer D, Kennedy R, King G, Vespignani A. Big data. The parable of Google Flu: traps in big data analysis. *Science* 2014 Mar 14;343(6176):1203-1205. [doi: [10.1126/science.1248506](#)] [Medline: [24626916](#)]
28. Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLoS Comput Biol* 2013;9(10):e1003256 [FREE Full text] [doi: [10.1371/journal.pcbi.1003256](#)] [Medline: [24146603](#)]
29. Santillana M, Zhang DW, Althouse BM, Ayers JW. What can digital disease detection learn from (an external revision to) Google Flu Trends? *Am J Prev Med* 2014 Sep;47(3):341-347. [doi: [10.1016/j.amepre.2014.05.020](#)] [Medline: [24997572](#)]
30. Ayers JW, Althouse BM, Dredze M. Could behavioral medicine lead the web data revolution? *JAMA* 2014 Apr 09;311(14):1399-1400 [FREE Full text] [doi: [10.1001/jama.2014.1505](#)] [Medline: [24577162](#)]
31. Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, et al. The use of google trends in health care research: a systematic review. *PLoS One* 2014;9(10):e109583 [FREE Full text] [doi: [10.1371/journal.pone.0109583](#)] [Medline: [25337815](#)]

Abbreviations

AIC: Akaike information criterion

SARIMAX: seasonal autoregressive integrated moving average model with explicative variable

Edited by G Eysenbach; submitted 13.10.17; peer-reviewed by V Gianfredi, D Nucci, G Dini, T Cruvinel; comments to author 02.11.17; revised version received 09.11.17; accepted 19.11.17; published 28.02.19.

Please cite as:

Watad A, Watad S, Mahroum N, Sharif K, Amital H, Bragazzi NL, Adawi M

Forecasting the West Nile Virus in the United States: An Extensive Novel Data Streams-Based Time Series Analysis and Structural Equation Modeling of Related Digital Searching Behavior

JMIR Public Health Surveill 2019;5(1):e9176

URL: <http://publichealth.jmir.org/2019/1/e9176/>

doi: [10.2196/publichealth.9176](#)

PMID: [30601755](#)

©Abdulla Watad, Samaa Watad, Naim Mahroum, Kassem Sharif, Howard Amital, Nicola Luigi Bragazzi, Mohammad Adawi. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 28.02.2019. 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

Google Trends Predicts Present and Future Plague Cases During the Plague Outbreak in Madagascar: Infodemiological Study

Nicola Luigi Bragazzi¹, MD, MPH, PhD; Naim Mahroum², MD

¹Department of Health Sciences, Postgraduate School of Public Health, University of Genoa, Genoa, Italy

²Sackler Faculty of Medicine, Tel-Aviv University, Tel Aviv, Israel

Corresponding Author:

Nicola Luigi Bragazzi, MD, MPH, PhD

Department of Health Sciences

Postgraduate School of Public Health

University of Genoa

Via Antonio Pastore 1

Genoa, 16132

Italy

Phone: 39 0103538508

Fax: 39 0103538541

Email: robertobragazzi@gmail.com

Abstract

Background: Plague is a highly infectious zoonotic disease caused by the bacillus *Yersinia pestis*. Three major forms of the disease are known: bubonic, septicemic, and pneumonic plague. Though highly related to the past, plague still represents a global public health concern. Cases of plague continue to be reported worldwide. In recent months, pneumonic plague cases have been reported in Madagascar. However, despite such a long-standing and rich history, it is rather difficult to get a comprehensive overview of the general situation. Within the framework of electronic health (eHealth), in which people increasingly search the internet looking for health-related material, new information and communication technologies could enable researchers to get a wealth of data, which could complement traditional surveillance of infectious diseases.

Objective: In this study, we aimed to assess public reaction regarding the recent plague outbreak in Madagascar by quantitatively characterizing the public's interest.

Methods: We captured public interest using Google Trends (GT) and correlated it to epidemiological real-world data in terms of incidence rate and spread pattern.

Results: Statistically significant positive correlations were found between GT search data and confirmed ($R^2=0.549$), suspected ($R^2=0.265$), and probable ($R^2=0.518$) cases. From a geospatial standpoint, plague-related GT queries were concentrated in Toamasina (100%), Toliara (68%), and Antananarivo (65%). Concerning the forecasting models, the 1-day lag model was selected as the best regression model.

Conclusions: An earlier digital Web search reaction could potentially contribute to better management of outbreaks, for example, by designing ad hoc interventions that could contain the infection both locally and at the international level, reducing its spread.

(JMIR Public Health Surveill 2019;5(1):e13142) doi:[10.2196/13142](https://doi.org/10.2196/13142)

KEYWORDS

plague; infodemiology; infoveillance; infectious outbreaks; Google Trends; nowcasting and forecasting models; digital surveillance

Introduction

Plague and history have always been strongly interrelated since the earliest description of plague pandemic [1]. Excluding the so-called *plague of Athens*, which could have been caused by typhus or other microorganisms, the first well-authenticated mention of plague dates back to the 6th century: the *Justinian*

Plague, named after the Byzantine emperor. This outbreak took place in Egypt in 542 AD and spread across the Mediterranean basin regions, killing more than 25 million people [2]. Three large pandemics occurred afterward, including a major outbreak during the decline of the Eastern Roman Empire, the rapid weakening of the Persian Empire, and the subsequent tumultuous rise of the Islamic Empire [3]. Another major episode was the

Great Plague Pandemic, which began in China in 1333-1334 during a period of famine and spread to Europe by trade causing millions of deaths [4]. The third major outbreak was the Modern Plague Pandemic, which originally occurred in China in 1860, spreading worldwide by rats on trade ships leading to over 10 million deaths across the world [5,6].

This highly infectious zoonotic disease is caused by the bacillus *Yersinia pestis*, a member of the *Enterobacteriaceae* family [7]. Transmission of plague occurs when infected rodents' fleas bite humans. Human-to-human transmission is possible when transmitted by infected air droplets. Three major forms of the disease are known: (1) bubonic plague, the most common form of plague, which is characterized by acute febrile illness accompanied by enlarged and tender lymph nodes; (2) septicemic plague, characterized by sepsis manifested by fever and systemic illness, generally without preceding symptoms; and (3) pneumonic plague, which may be primary when acquired directly by infected air droplets and secondary when it spreads to the lung from other infected sites of the body—both of these forms lead to a highly contagious and lethal disease [8].

Diagnosis of plague is made based on the above-mentioned clinical findings accompanied with suspected history of exposure. Identification by culture as well as increased paired serum titers are suitable diagnostic methods. Plague can be fatal if left untreated. Besides supportive therapy for severely ill patients, treatment is composed of systemic antibiotics of aminoglycoside-based therapy.

Though highly associated with the past, plague still represents a global public health concern. Cases of plague continue to be reported worldwide, especially in Africa but also in Asia, South America, and even in the United States. According to the World Health Organization (WHO), 3248 cases of plague were reported from 2010 until 2015, with 584 deaths [9]. Recently, foci have been described in Libya and Algeria [8]. The highest incidence of plague cases in recent years is reported from Madagascar. In fact, cases of bubonic plague are reported annually in Madagascar since the first case was introduced in 1898; however, recent reports show a large outbreak of pneumonic plague occurring in major urban cities, which is different from what was previously reported (ie, cases concentrated mainly in rural areas). In recent months, we have witnessed the spread of plague to the Seychelles islands, reflecting a further escalation of the current outbreak [10-12].

However, despite such a long-standing and rich history, it is rather difficult to get a comprehensive overview of the general situation. Within the framework of electronic health (eHealth), in which people surf the internet more and more looking for health-related material, new information and communication technologies, such as Web 2.0, portable computers, mobile phones and devices, as well as social media and social networks, could enable researchers to get a wealth of data, which could complement traditional surveillance of infectious diseases.

In this study, we aimed to assess public reaction regarding the recent plague outbreak in Madagascar by quantitatively characterizing this interest and correlating it to epidemiological *real-world* data, in terms of incidence rate and spread pattern.

Methods

Google Trends (GT) is a free open-source tool used to track and observe internet search activity [13]. GT was used to assess recent search activity with regard to the recent plague outbreak in Madagascar. To that end, GT was mined from August 1 to November 17, 2017 [13]. This particular time frame was chosen in order to better capture the temporal dynamics of the plague outbreak, monitoring the internet-related activity before (ie, digital behavior at the baseline) and during the epidemic. During the drafting and production of this study, the latest available WHO situation report, released on November 17, 2017 [14], was utilized.

The GT search tool has two options for searching keywords: searches can be performed by *search term* or by *search topic*. While the former enables the user to search for exact keywords, the latter option uses a broader search that finds all Web searches containing the inserted keyword(s) or related terms.

The results given by GT are output as normalized values (ie, relative search volumes [RSVs]) rather than absolute, raw values. Every query is divided by the total searches performed in a given geographic region and time range and normalized to a scale between 0 and 100 based on the topic's popularity in comparison to all searches carried out in that region and time frame. For further details concerning GT, the reader is referred to Nuti et al's review of GT and its potential applications in the medical field [15]; the reader is also referred to Mavragani et al's recent systematic review concerning methods, tools, and statistical approaches and techniques in the field of GT research [16].

In this study, the second search option (ie, searching by topic with related terms) was used. Specifically, we searched for "Plague (Topic)." Searches were geographically limited to Madagascar. In Madagascar, the language spoken is Malagasy, while the second official language is French. English is spoken by less than 20% of the population. It is widely known that the sample when using online queries cannot be representative; however, in our case, searching for "Plague" and selecting the *search topic* option enabled us to overcome any linguistic issue related to the diffusion of the language. This approach ensured the robustness of our results.

Correlational analysis and multivariate regression models for nowcasting and forecasting, with lags up to 7 days, were performed based on the GT results with the number of confirmed, suspected, and probable cases of plague as reported by the WHO situation report. Different regression models were run, computing the different fitting parameters, including R^2 and adjusted R^2 , and the best model was chosen according to the Akaike Information Criterion (AIC) values.

Statistical analyses were performed with the commercial software XLSTAT 2017 (Addinsoft). All values with P values less than .05 were considered statistically significant.

Results

Average plague-related search activities, expressed as RSVs, are shown superimposed on the trends of new suspected, probable, and confirmed cases of plague in [Figure 1](#). The searches showed a small burst of activity on September 14, 2017, immediately after the official notification (ie, September 13, 2017) sent to the WHO by the Madagascar Ministry of Public Health of an outbreak of pneumonic plague in Madagascar. This notification followed the death of a young man some days before on September 11, 2017, who suffered from severe respiratory disease confirmed to be caused by plague. A very large spike was noticed during the first week of October 2017. Afterward, RSVs tended to decrease over time to slightly above baseline levels. Similarly, the incidence of suspected, probable, and confirmed cases of plague in Madagascar also exhibited a small spike in the third week of September 2017, and many more cases were confirmed during the recent outbreak starting in the first week of October.

The best nowcasting model in terms of AIC values (see [Multimedia Appendix 1](#)) showed that new confirmed cases of plague had a statistically significant association with GT-based RSVs ($P<.001$, beta coefficient 1.158), as shown in the multivariate regression analyses in [Multimedia Appendix 2](#). Scatterplots of incident cases showed similar and statistically significant positive correlations with GT search data ($R^2=0.549$, $P=.001$ for the confirmed cases; $R^2=0.265$, $P=.005$ for the suspected cases; and $R^2=0.518$, $P=.001$ for the probable cases; see [Figure 2](#)). From a geospatial standpoint, plague-related GT queries were concentrated in certain regions of Madagascar, most notably in Toamasina (100%), Toliara (68%), and Antananarivo (65%). A heat map of search density in different regions of Madagascar is shown in [Figure 3](#). Concerning the forecasting models, the 1-day lag model was selected for regression analysis due to optimal AIC values (see [Multimedia Appendix 3](#)). This forecasting model shows that we can predict new probable cases up to 1 day in advance with statistically significant certainty ($P<.001$; see [Multimedia Appendix 4](#)).

Figure 1. Time trends of plague cases (confirmed, probable, and suspected) and plague-related Google Trends (GT)-generated data. All data are normalized for comparison purposes. RSV: relative search volume.

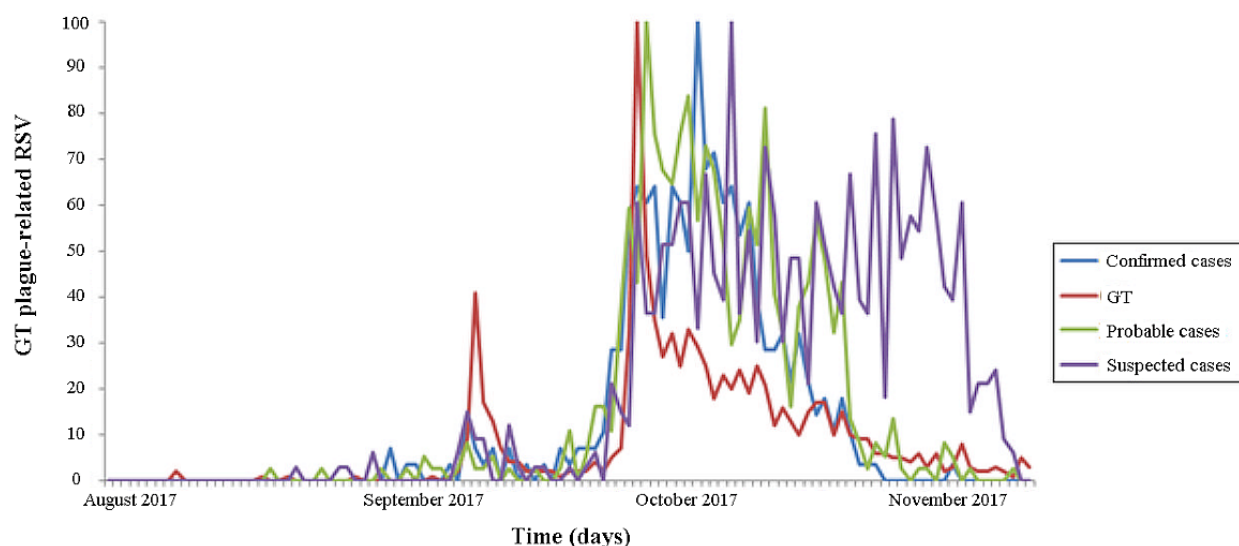


Figure 2. Scatterplots of the correlations between epidemiological values and the Google Trends (GT)–generated data related to the recent plague outbreak in Madagascar. RSV: relative search volume.

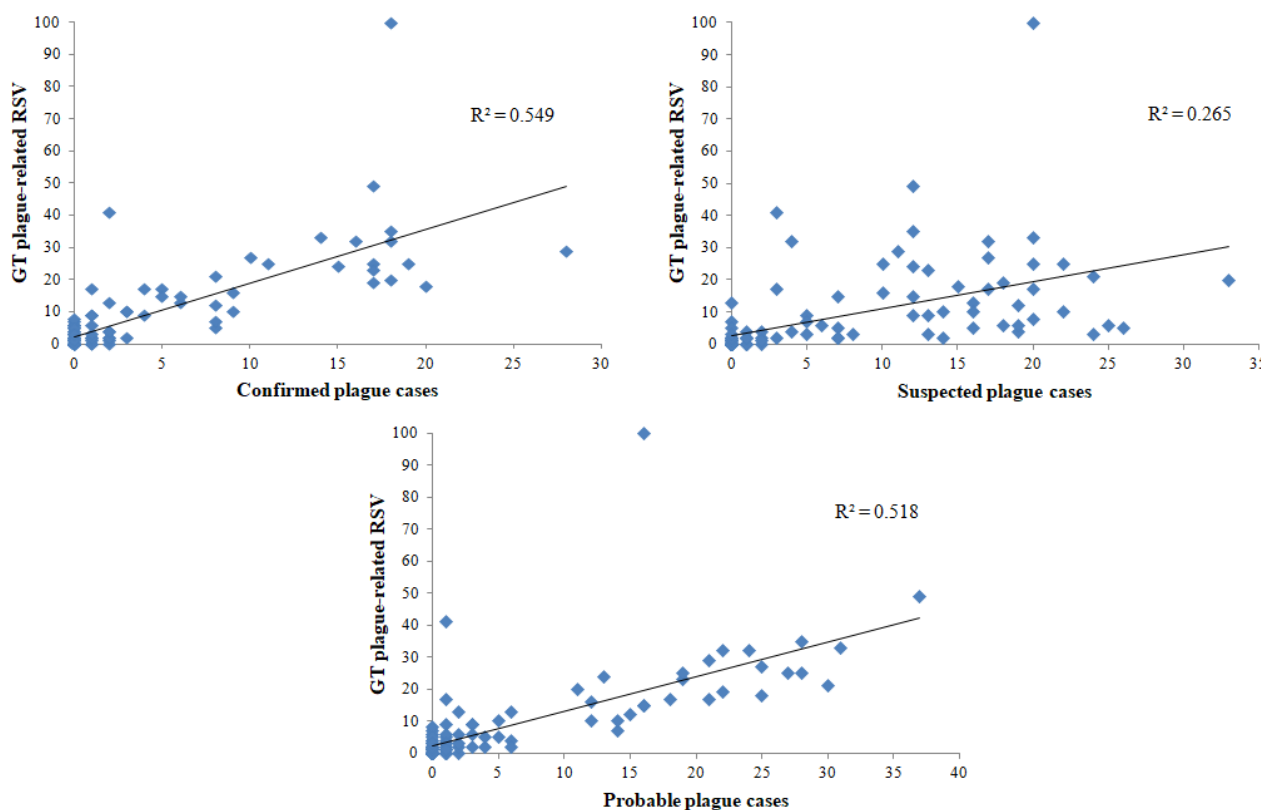
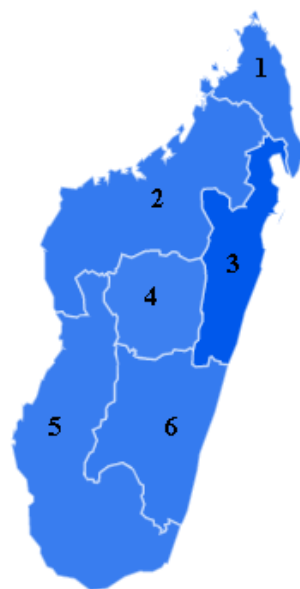


Figure 3. Spatial heat map of Google Trends–based plague-related data. The figure is based on the map provided by Google Trends. 1: Antsiranana Province; 2: Mahajanga Province; 3: Toamasina Province; 4: Antananarivo Province; 5: Toliara Province; and 6: Fianarantsoa Province. Color gradient correlates with the volume of plague-related Web searches.



Discussion

Principal Findings

Plague outbreak in Madagascar has drawn wide public attention shown here by our findings based on large Web search activity data analysis. Madagascar has been known as an endemic area of plague in its bubonic form with annually reported cases from

April to September, generally across rural areas. However, the recent outbreak is characterized by pneumonic plague occurring in larger and more crowded cities. The current outbreak, known to be a highly contagious form of plague and in combination with the recent spread to the Seychelles islands, has its own distinctive features.

Monitoring and analyzing Web search activity manifested by novel data streams (NDS), especially during outbreaks, is of

great importance in terms of surveillance as shown by O'Shea [17] in a recent systematic review. Big data or *vast digital data* analysis is, indeed, an opportunity to improve surveillance and epidemic intelligence, being inexpensive, transparent, and flexible. As such, *event-based internet biosurveillance* can act as an extension of traditional surveillance and monitoring systems and can be utilized as an additional data source, contributing, therefore, to a more comprehensive estimate of infectious diseases.

GT, based on Google search, is a freely accessed website tool, which provides data on how often a specific search item (ie, plague) is searched relative to total search volume worldwide or in specific areas and in different languages. For instance, in 2009, during the peanut butter-associated outbreak of *Salmonella enterica* subtype *Typhimurium*, GT provided preliminary evidence of an emerging infectious outbreak, enabling early disease detection [18]. Other research studies based on GT have shown the possibility of monitoring and tracking flu epidemics [19-21], as well as other infections [22-35].

In our study, bursts of searches of plague-related topics corresponded both spatially and temporally with the outbreak's spatiotemporal trends across the region studied (ie, Madagascar). The role of Web-based NDS for outbreak surveillance is crucial for workers in the field of public health and safety. Plague-related digital behavior as captured by GT analysis reflected rapid public response to the pneumonic plague outbreak in Madagascar, with some minor search peaks occurring even before the formal declaration by the WHO. Moreover, this reaction seemed to decline rapidly afterward, whereas the WHO continued to release the report of additional confirmed plague cases.

In our study, it is interesting to note that the potential influence of prior awareness of a clinical case of plague, which occurred on September 13, 2017, on search behaviors of a population was reflected by the rapid increase of searches found on September 14, 2017. From September 30, 2017, people were probably more able to recognize specific signs and symptoms related to plague due to news or public campaigns. In this case, the suspicion of disease may lead people to seek confirmatory Web information, contributing to the increase of the activity of internet users. These arguments could be used to explain the

highest value of R^2 when including confirmed, probable, and suspected cases in multivariate regression models (ie, searches were probably motivated or driven by personal impressions and knowledge of disease).

Findings from the regression analyses showed the feasibility of exploiting NDS for predicting (ie, nowcasting and forecasting) plague cases. Extant predictive models of plague are usually built within the *ecological-niche modeling framework*, in which geographic, environmental, and ecological parameters, such as landscape-scale environmental features, are utilized [36,37]. To the best of our knowledge, this is the first model incorporating plague-related information-seeking behavior in terms of Web-based NDS, such as GT. Even though a correlation between epidemiological values and Web searches could appear trivial, this is surprising, especially considering the poor internet penetration in Madagascar (ie, only 4%-5% of the population have access to the internet).

Despite its novelties, which are among the major strengths of the current investigation, our study suffers from some limitations, which should be properly recognized. The shortcomings include the fact that GT provides relative and not absolute values, thus hindering the possibility of further refining and processing them. Moreover, GT captures only Web searches carried out with the Google search engine, which is, on the other hand, the most utilized search tool. Another drawback was the relatively low values of R^2 . The limited internet penetration (ie, approximately 4%-5% of the entire population) as well as the short time frame chosen for the study could be among the factors explaining such values.

Conclusions

Our study has shown an increase in digital Web searches with a unique pattern induced by the recent outbreak of plague in Madagascar. GT plays a highly important role in outbreak tracking and monitoring, in that it can capture public reaction and interest toward infectious disorders in real time before cases are formally communicated by the WHO. This earlier digital Web search reaction could potentially contribute to better management of outbreaks, for example, by designing ad hoc interventions that could contain the infection both locally and at the international level, reducing its spread.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Regression analyses for the nowcasting models.

[PDF File (Adobe PDF File), 97KB - [publichealth_v5i1e13142_app1.pdf](#)]

Multimedia Appendix 2

Fitting parameters of the nowcasting models.

[PDF File (Adobe PDF File), 90KB - [publichealth_v5i1e13142_app2.pdf](#)]

Multimedia Appendix 3

Regression analysis of the best forecasting model.

[PDF File (Adobe PDF File), 124KB - [publichealth_v5i1e13142_app3.pdf](#)]

Multimedia Appendix 4

Fitting parameters of the forecasting models.

[PDF File (Adobe PDF File), 87KB - [publichealth_v5i1e13142_app4.pdf](#)]

References

1. Miller MJ. Plague: History and epidemiology. *Can J Comp Med Vet Sci* 1940 Jul;4(7):183-193 [FREE Full text] [Medline: 17647665]
2. Sabbatani S, Manfredi R, Fiorino S. The Justinian plague (part one) [Article in Italian]. *Infez Med* 2012 Jun;20(2):125-139 [FREE Full text] [Medline: 22767313]
3. Sabbatani S, Manfredi R, Fiorino S. The Justinian plague (part two). Influence of the epidemic on the rise of the Islamic Empire [Article in Italian]. *Infez Med* 2012 Sep;20(3):217-232 [FREE Full text] [Medline: 22992565]
4. Drancourt M, Raoult D. Molecular history of plague. *Clin Microbiol Infect* 2016 Nov;22(11):911-915 [FREE Full text] [doi: 10.1016/j.cmi.2016.08.031] [Medline: 27615720]
5. Peiffer-Smadja N, Thomas M. The plague: A disease that is still haunting our collective memory [Article in French]. *Rev Med Interne* 2017 Jun;38(6):402-406. [doi: 10.1016/j.revmed.2017.03.008] [Medline: 28457681]
6. Rechnitzer PA. A history of the plague. *Univ West Ont Med J* 1947 Jan;17(1):23-32. [Medline: 20296999]
7. Yang R, Cui Y, Bi Y. Perspectives on *Yersinia pestis*: A model for studying zoonotic pathogens. *Adv Exp Med Biol* 2016;918:377-391. [doi: 10.1007/978-94-024-0890-4_14] [Medline: 27722871]
8. Grácio AJDS, Grácio MAA. Plague: A millenary infectious disease reemerging in the XXI century. *Biomed Res Int* 2017;2017:5696542 [FREE Full text] [doi: 10.1155/2017/5696542] [Medline: 28904964]
9. Bertherat E. Plague around the world, 2010-2015. *Wkly Epidemiol Rec* 2016 Feb 26;91(8):89-93 [FREE Full text] [Medline: 26922822]
10. Nikiforov VV, Gao H, Zhou L, Anisimov A. Plague: Clinics, diagnosis and treatment. *Adv Exp Med Biol* 2016;918:293-312. [doi: 10.1007/978-94-024-0890-4_11] [Medline: 27722868]
11. Mead PS. Plague in Madagascar: A tragic opportunity for improving public health. *N Engl J Med* 2018 Jan 11;378(2):106-108. [doi: 10.1056/NEJMp1713881] [Medline: 29261403]
12. Drancourt M, Raoult D. Investigation of pneumonic plague, Madagascar. *Emerg Infect Dis* 2018 Dec;24(1):183 [FREE Full text] [doi: 10.3201/eid2401.170760] [Medline: 29260670]
13. Google Trends. Plague, Madagascar URL: <https://trends.google.com/trends/explore?date=2017-08-01%202017-11-17&geo=MG&q=%2Fm%2F01195hyq> [accessed 2019-03-04] [WebCite Cache ID 76dWwhKqB]
14. Plague Outbreak Madagascar: External Situation Report 11. Geneva, Switzerland: World Health Organization; 2017 Nov 17. URL: <http://apps.who.int/iris/bitstream/handle/10665/259479/Ex-PlagueMadagascar17112017.pdf?sequence=1> [accessed 2019-03-04] [WebCite Cache ID 76dXlvydi]
15. Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, et al. The use of Google Trends in health care research: A systematic review. *PLoS One* 2014;9(10):e109583 [FREE Full text] [doi: 10.1371/journal.pone.0109583] [Medline: 25337815]
16. Mavragani A, Ochoa G, Tsagarakis KP. Assessing the methods, tools, and statistical approaches in Google Trends research: Systematic review. *J Med Internet Res* 2018 Nov 06;20(11):e270 [FREE Full text] [doi: 10.2196/jmir.9366] [Medline: 30401664]
17. O'Shea J. Digital disease detection: A systematic review of event-based Internet biosurveillance systems. *Int J Med Inform* 2017 Dec;101:15-22. [doi: 10.1016/j.ijmedinf.2017.01.019] [Medline: 28347443]
18. 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):2153-2155, 2157 [FREE Full text] [doi: 10.1056/NEJMp0900702] [Medline: 19423867]
19. Eysenbach G. Infodemiology: Tracking flu-related searches on the Web for syndromic surveillance. *AMIA Annu Symp Proc* 2006;244-248 [FREE Full text] [Medline: 17238340]
20. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature* 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
21. Polgreen PM, Chen Y, Pennock DM, Nelson FD. Using Internet searches for influenza surveillance. *Clin Infect Dis* 2008 Dec 01;47(11):1443-1448. [doi: 10.1086/593098] [Medline: 18954267]
22. Kapitány-Fövény M, Ferenci T, Sulyok Z, Kegele J, Richter H, Vályi-Nagy I, et al. Can Google Trends data improve forecasting of Lyme disease incidence? *Zoonoses Public Health* 2019 Dec;66(1):101-107. [doi: 10.1111/zph.12539] [Medline: 30447056]

23. Verma M, Kishore K, Kumar M, Sondh AR, Aggarwal G, Kathirvel S. Google search trends predicting disease outbreaks: An analysis from India. *Healthc Inform Res* 2018 Oct;24(4):300-308 [FREE Full text] [doi: [10.4258/hir.2018.24.4.300](https://doi.org/10.4258/hir.2018.24.4.300)] [Medline: [30443418](https://pubmed.ncbi.nlm.nih.gov/30443418/)]
24. Ho HT, Carvajal TM, Bautista JR, Capistrano JDR, Viacrusis KM, Hernandez LFT, et al. Using Google Trends to examine the spatio-temporal incidence and behavioral patterns of dengue disease: A case study in Metropolitan Manila, Philippines. *Trop Med Infect Dis* 2018 Nov 11;3(4):118 [FREE Full text] [doi: [10.3390/tropicalmed3040118](https://doi.org/10.3390/tropicalmed3040118)] [Medline: [30423898](https://pubmed.ncbi.nlm.nih.gov/30423898/)]
25. Wu C, Kao SC, Shih CH, Kan MH. Open data mining for Taiwan's dengue epidemic. *Acta Trop* 2018 Jul;183:1-7. [doi: [10.1016/j.actatropica.2018.03.017](https://doi.org/10.1016/j.actatropica.2018.03.017)] [Medline: [29549012](https://pubmed.ncbi.nlm.nih.gov/29549012/)]
26. Gianfredi V, Bragazzi NL, Mahamid M, Bisharat B, Mahroum N, Amital H, et al. Monitoring public interest toward pertussis outbreaks: An extensive Google Trends-based analysis. *Public Health* 2018 Dec;165:9-15. [doi: [10.1016/j.puhe.2018.09.001](https://doi.org/10.1016/j.puhe.2018.09.001)] [Medline: [30342281](https://pubmed.ncbi.nlm.nih.gov/30342281/)]
27. Morsy S, Dang TN, Kamel MG, Zayan AH, Makram OM, Elhady M, et al. Prediction of Zika-confirmed cases in Brazil and Colombia using Google Trends. *Epidemiol Infect* 2018 Oct;146(13):1625-1627. [doi: [10.1017/S0950268818002078](https://doi.org/10.1017/S0950268818002078)] [Medline: [30056812](https://pubmed.ncbi.nlm.nih.gov/30056812/)]
28. Young SD, Zhang Q. Using search engine big data for predicting new HIV diagnoses. *PLoS One* 2018;13(7):e0199527 [FREE Full text] [doi: [10.1371/journal.pone.0199527](https://doi.org/10.1371/journal.pone.0199527)] [Medline: [30001360](https://pubmed.ncbi.nlm.nih.gov/30001360/)]
29. Mahroum N, Bragazzi NL, Brigo F, Waknin R, Sharif K, Mahagna H, et al. Capturing public interest toward new tools for controlling human immunodeficiency virus (HIV) infection exploiting data from Google Trends. *Health Informatics J* 2018 Apr 01;1460458218766573. [doi: [10.1177/1460458218766573](https://doi.org/10.1177/1460458218766573)] [Medline: [29638172](https://pubmed.ncbi.nlm.nih.gov/29638172/)]
30. van de Belt TH, van Stockum PT, Engelen LJLP, Lancee J, Schrijver R, Rodríguez-Baño J, et al. Social media posts and online search behaviour as early-warning system for MRSA outbreaks. *Antimicrob Resist Infect Control* 2018;7:69 [FREE Full text] [doi: [10.1186/s13756-018-0359-4](https://doi.org/10.1186/s13756-018-0359-4)] [Medline: [29876100](https://pubmed.ncbi.nlm.nih.gov/29876100/)]
31. Mahroum N, Adawi M, Sharif K, Waknin R, Mahagna H, Bisharat B, et al. Public reaction to Chikungunya outbreaks in Italy: Insights from an extensive novel data streams-based structural equation modeling analysis. *PLoS One* 2018;13(5):e0197337 [FREE Full text] [doi: [10.1371/journal.pone.0197337](https://doi.org/10.1371/journal.pone.0197337)] [Medline: [29795578](https://pubmed.ncbi.nlm.nih.gov/29795578/)]
32. Young SD, Torrone EA, Urata J, Aral SO. Using search engine data as a tool to predict syphilis. *Epidemiology* 2018 Dec;29(4):574-578. [doi: [10.1097/EDE.0000000000000836](https://doi.org/10.1097/EDE.0000000000000836)] [Medline: [29864105](https://pubmed.ncbi.nlm.nih.gov/29864105/)]
33. Oren E, Frere J, Yom-Tov E, Yom-Tov E. Respiratory syncytial virus tracking using Internet search engine data. *BMC Public Health* 2018 Dec 03;18(1):445 [FREE Full text] [doi: [10.1186/s12889-018-5367-z](https://doi.org/10.1186/s12889-018-5367-z)] [Medline: [29615018](https://pubmed.ncbi.nlm.nih.gov/29615018/)]
34. Wang J, Zhang T, Lu Y, Zhou G, Chen Q, Niu B. Vesicular stomatitis forecasting based on Google Trends. *PLoS One* 2018;13(1):e0192141 [FREE Full text] [doi: [10.1371/journal.pone.0192141](https://doi.org/10.1371/journal.pone.0192141)] [Medline: [29385198](https://pubmed.ncbi.nlm.nih.gov/29385198/)]
35. Adawi M, Bragazzi NL, Watad A, Sharif K, Amital H, Mahroum N. Discrepancies between classic and digital epidemiology in searching for the Mayaro virus: Preliminary qualitative and quantitative analysis of Google Trends. *JMIR Public Health Surveill* 2017 Dec 01;3(4):e93 [FREE Full text] [doi: [10.2196/publichealth.9136](https://doi.org/10.2196/publichealth.9136)] [Medline: [29196278](https://pubmed.ncbi.nlm.nih.gov/29196278/)]
36. Neerinx S, Peterson AT, Gulnick H, Deckers J, Kimaro D, Leirs H. Predicting potential risk areas of human plague for the Western Usambara Mountains, Lushoto District, Tanzania. *Am J Trop Med Hyg* 2010 Mar;82(3):492-500 [FREE Full text] [doi: [10.4269/ajtmh.2010.09-0426](https://doi.org/10.4269/ajtmh.2010.09-0426)] [Medline: [20207880](https://pubmed.ncbi.nlm.nih.gov/20207880/)]
37. Neerinx S, Bertherat E, Leirs H. Human plague occurrences in Africa: An overview from 1877 to 2008. *Trans R Soc Trop Med Hyg* 2010 Feb;104(2):97-103. [doi: [10.1016/j.trstmh.2009.07.028](https://doi.org/10.1016/j.trstmh.2009.07.028)] [Medline: [19716148](https://pubmed.ncbi.nlm.nih.gov/19716148/)]

Abbreviations

AIC: Akaike Information Criterion
eHealth: electronic health
GT: Google Trends
NDS: novel data streams
RSV: relative search volume
WHO: World Health Organization

Edited by G Eysenbach; submitted 14.12.18; peer-reviewed by T Cruvinel, A Mavragani; comments to author 17.01.19; revised version received 17.01.19; accepted 18.01.19; published 08.03.19.

Please cite as:

Bragazzi NL, Mahroum N

Google Trends Predicts Present and Future Plague Cases During the Plague Outbreak in Madagascar: Infodemiological Study

JMIR Public Health Surveill 2019;5(1):e13142

URL: <http://publichealth.jmir.org/2019/1/e13142/>

doi: [10.2196/13142](https://doi.org/10.2196/13142)

PMID: [30763255](https://pubmed.ncbi.nlm.nih.gov/30763255/)

©Nicola Luigi Bragazzi, Naim Mahroum. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 08.03.2019. 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

Generating Engagement on the Make Healthy Normal Campaign Facebook Page: Analysis of Facebook Analytics

James Kite¹, BA, MPH; Anne Grunseit^{1,2}, PhD; Vincy Li³, MPhil; John Vineburg⁴, MA; Nathan Berton⁵, DipM; Adrian Bauman¹, PhD; Becky Freeman¹, PhD

¹Prevention Research Collaboration, Sydney School of Public Health and Charles Perkins Centre, The University of Sydney, The University of Sydney, Australia

²The Australian Prevention Partnership Centre, The University of Sydney, Australia

³New South Wales Office of Preventive Health, Liverpool, Australia

⁴Centre for Population Health, New South Wales Ministry of Health, North Sydney, Australia

⁵Strategic Communications and Engagement, New South Wales Ministry of Health, North Sydney, Australia

Corresponding Author:

James Kite, BA, MPH

Prevention Research Collaboration

Sydney School of Public Health and Charles Perkins Centre

The University of Sydney

Australia

Phone: 61 0286270838

Email: james.kite@sydney.edu.au

Abstract

Background: Facebook is increasingly being used as part of mass media campaigns in public health, including the Make Healthy Normal (MHN) campaign in New South Wales, Australia. Therefore, it is important to understand what role Facebook can play in mass media campaigns and how best to use it to augment or amplify campaign effects. However, few studies have explored this.

Objective: This study aimed to investigate usage of and engagement with the MHN Facebook page and to identify influential factors in driving engagement with the page.

Methods: We examined both post-level and page-level analytic data from Facebook from the campaign's launch in June 2015 to September 2017. For post-level data, we conducted a series of negative binomial regressions with four different outcome measures (likes, shares, comments, post consumers), including some characteristics of Facebook posts as predictors. We also conducted time series analyses to examine associations between page-level outcomes (new page likes or "fans" and number of engaged users) and different measures of exposure to the page (number of unique users reached and total count of impressions) and to television advertising.

Results: Of the 392 posts reviewed, 20.7% (n=81) received a paid boost and 58.9% (n=231) were photo posts. We found that posts that received a paid boost reached significantly more users and subsequently received significantly more engagement than organic (unpaid) posts ($P<.001$). After adjusting for reach, we found the effect of being paid was incremental for all outcome measures for photos and links, but not videos. There were also associations between day of the week and time of post and engagement, with Mondays generally receiving less engagement and posts on a Friday and those made between 8 AM and 5 PM receiving more. At the page level, our time series analyses found that organic impressions predicted a higher number of new fans and engaged users, compared to paid impressions, especially for women. We also found no association between television advertising and engagement with the Facebook page.

Conclusions: Our study shows that paying for posts is important for increasing their reach, but that page administrators should look to maximize organic reach because it is associated with significantly higher engagement. Once reach is accounted for, video posts do not benefit from being paid, unlike the other post types. This suggests that page administrators should carefully consider how they use videos as part of a Facebook campaign. Additionally, the lack of association between television advertising and engagement suggests that future campaigns consider how best to link different channels to amplify effects. These results highlight the need for ongoing evaluation of Facebook pages if administrators are to maximize engagement.

KEYWORDS

social media; Facebook; overweight and obesity; mass media campaign; evaluation

Introduction

Background

Facebook is the largest social media platform in the world, with more than 1.4 billion daily users on average in December 2017 [1]. In Australia, nearly two-thirds of adults have a Facebook profile, making it the most popular social media platform in the country [2]. It is also the most intensely used social media platform; around 40% of Australian Facebook users log in 20 times or more per week. Further, Facebook is one of the most commonly used social media platforms for engaging with health issues [3]. It is no surprise then that public health organizations are using Facebook to communicate their messages, either as stand-alone campaigns or as an additional channel in a broader mass media campaign [4,5]. In both cases, organizations are seeking to capitalize on the wide reach of Facebook, the ability to engage directly with their target audiences, and the potential for generating marketing directly between consumers (word-of-mouth marketing), which lends credibility to a brand and is known to be one of the most trusted forms of marketing [6-8]. Within mass media campaigns specifically, the intention is that Facebook posts will augment or amplify campaign messages and, in so doing, increase the impact of the campaign [9].

The theory behind Facebook use for public health communication places “engagement” as a critical first step in achieving change. Creating engagement, defined as users “liking,” sharing, commenting, or clicking on any content, is important for two main reasons: it demonstrates that the content is attention grabbing and it directly influences the reach of the content and of future content through the Facebook algorithm [10]. The algorithm determines the amount of exposure a post receives and to whom it is shown, although it should be noted that Facebook has revealed little on the specific parameters it uses to prioritize posts. However, what is clear is that the characteristics of the post and the engagement it receives are factors in the algorithm’s calculations [4], making it essential to investigate what drives engagement in order to maximize Facebook’s marketing potential for public health campaigns. Facebook also allows page administrators to pay to increase the reach of a post, making it important to investigate the interaction between paying for posts and other post characteristics.

Despite the potential of Facebook and other social media for public health and health communication being well recognized [11-14], there is limited evidence available to guide practice. The evidence we do have is often either descriptive or based on small-scale trials [5,15-18], with suggestive but modest evidence that social media can be effective in changing health outcomes [19,20]. How to build engagement with health content on Facebook has been recognized as one area in particular need of more evidence given the role it plays in the theory of health communication on social media [21]. Currently, there is some evidence that testimonials, positive emotional appeals, and

informative posts are associated with higher engagement, whereas posts that evoke negative emotions, use conventional marketing techniques (eg, sponsorships), or are posted during or after work hours are associated with lower engagement [4,22-24]. Similarly, posts that use photos and videos appear to generate higher engagement, although this is most likely due to the Facebook algorithm preferencing such content over other post types. In addition, one study that examined 20 public health Facebook pages covering a range of health issues speculated that particular health issues may be more suitable to Facebook [4]. However, they lamented that they were unable to test this, highlighting it as an area worthy of further research.

In addition, the available evidence has limited relevance to mass-reach campaigns, creating the risk that social marketers will use Facebook without considering what strategy they should employ to best use the platform in a broader campaign [25,26]. It is therefore important to investigate associations between Facebook engagement and traditional communication channels such as television. To our knowledge, no study has examined these associations. The evaluation of the Tips From Former Smokers (Tips) antismoking campaign in the United States did provide some insights into the relationship between online and traditional television marketing for public health purposes, although how relevant this is to Facebook is uncertain. Tips showed an association between television advertising and online behaviors, including increased visits to the campaign website and other cessation-related websites and searches for cessation information [27,28]. The evaluation also found that digital video was more cost-efficient at generating awareness compared to television, although the authors note that television advertising is still important because it reaches more people [29]. Another study compared the cost-effectiveness of three media formats (television, online video, and online display advertising) for delivering an antismoking campaign [30]. This study found that online display advertising was the most cost-effective way of achieving Web page views, calls to the Quitline, online registrations for a cessation support service, and requests for the smoking cessation information pack. This was followed by a combination of online video and online display, with television alone the least cost-effective. Collectively, these studies suggest that online media present a potentially useful contribution to the reach and effectiveness of antismoking campaigns, but its role in other campaigns is yet to be explored.

To our knowledge, no population-level mass media campaign has reported specifically on their use of Facebook for public health purposes. Such information is only going to become more valuable as media consumption habits are changing rapidly [31], creating questions about the accuracy of conventional wisdom on “what works” in mass media campaigns. It will also help to understand how to optimize the use of Facebook as part of a wider mass media campaign. Here we report an evaluation of the Facebook page component of an obesity prevention lifestyle campaign, Make Healthy Normal (MHN).

The Make Healthy Normal Campaign

The MHN campaign was launched in New South Wales (NSW), Australia, in 2015, with the aim of challenging the normalization of being unhealthy and promoting physical activity, healthy eating, and healthy weight. The campaign initially targeted all adults but focused on parents with children aged 5 to 12 years and men aged 35 to 54 years from May 2017. The bulk of the advertising expenditure was directed toward television, but the campaign also made use of other channels, including Facebook. More details on the campaign are available elsewhere [32]. Briefly, the campaign was centered on two television commercials that juxtaposed unhealthy and healthy choices relating to nutrition and physical activity, while also making use of a number of other support channels, of which Facebook was one. The television commercials and most other campaign materials included the MHN website address but did not mention the Facebook page.

The MHN Facebook page had, at the time of writing, posted more than 400 times, generating over 100,000 likes, comments, and shares, and had over 32,000 page “likes” (hereafter “fans”). The page style is intended to be conversational and supportive, highlighting easy ways to eat healthier and increase physical activity, and promoting relevant NSW Government programs. The page uses both paid and organic posts (ie, content that is and is not paid advertising). The Ministry employed a strategy of paying for boosts on all posts during a specific period, as opposed to selectively boosting some posts and not others. This decision was based largely on practical considerations, especially the availability of funding.

This study aimed to investigate usage of and engagement with the MHN Facebook page as part of a broader multichannel campaign since its inception in 2015. Our research questions were: (1) What post characteristics influence the level of engagement a post receives and to what extent? (2) What page-level factors influence the number of fans, the characteristics of fans, and the engagement of fans with the MHN page over time? and (3) Is there a relationship between television advertising for the broader campaign and page-level engagement?

Methods

Study Overview

Facebook provides analytics (called “Insights”) to page administrators to help them monitor and understand usage of

their page. In this study, we analyzed the Insights data for the MHN page since June 2015 (when the campaign launched) through to September 2017. This study was approved by the University of Sydney’s Human Research Ethics Committee (protocol number: 2017/145).

Measures

Post-Level Data

We explored the characteristics of posts and their associations with engagement metrics (Table 1). Characteristics of posts included the post type, the date and time of the post, whether the post included a paid boost (paid posts tend to have a much greater increase in their reach), and the targeted behaviors. We also coded the content of the post using a modified version of the communication technique code frame developed in an earlier study [4]. The code frame was modified by collapsing some categories due to the relatively small number of posts compared to the original study. Engagement metrics were operationalized through the number of likes, shares, comments, and post consumers. Although likes technically include other Facebook “reactions” (eg, “love” and “haha”), we refer to this metric as “likes” because reactions were only introduced by Facebook a year into the campaign and the number of other reactions per post after that time was very low, typically zero.

Communication technique and target behavior were coded manually. Two coders independently coded each post, with interrater agreement for the communication techniques and target behaviors of 70% and 91%, respectively. Differences were resolved by discussion or referral to a third coder.

Page-Level Data

We used page-level data to examine the associations between the number of fans, the characteristics of fans, and the engagement of fans with campaign activity using the measures described in Table 2. Campaign activity was operationalized through weekly page impressions, separated by whether they were paid or organic, and weekly Target Audience Rating Points (TARPs). TARPs are an estimate of reach and frequency of exposure to television advertising, which is calculated by an external television ratings agency.

Table 1. Post-level measures and descriptions.

Variable	Description
Day of post	Day of the week the post first appeared
Time of post	Time post first appeared
Type	Whether the post is a photo, video, link, or text only
Paid/organic	Whether the post received a paid boost to its reach (“paid”) or not (“organic”)
Reach	The total number of unique users to whom the post was shown. Available in aggregate, as well as broken down by paid and organic reach
Consumers	The total number of unique users who clicked anywhere on the post
Likes	The number of “likes” and other “reactions” on a post. These are simple methods for users to indicate their response to a post, including to “like” the post, as well as other emotional reactions, including “love,” “haha,” “wow,” “sad,” and “angry”
Comment	The number of user comments (excluding replies) on the post
Share	The number of shares a post receives. The “share” button allows users to share the content with their Facebook friends
Communication technique	
Informative	Provides information on a health issue, its associated behaviors, and/or associated consequences or benefits
Call-to-action/instructive	Either provides instruction on how to do a behavior or encourages users to undertake a specific action (eg, call a helpline, make an appointment, register for a program or event). These were given coding precedence over informative messages
Emotional	Aims to elicit positive (eg, hope, excitement) or negative (eg, fear) emotions in users. Also includes posts that aim to generate a positive feeling about the brand. Emotional appeals took coding precedence over informative and call-to-action/instructive, reflecting evidence that emotive content is more powerful than nonemotive content [33]
Target behavior	
Eat	Information and encouragement to eat healthy food portions
Drink	Information and encouragement to make water the drink of choice and decrease sugar-sweetened beverage consumption
Act	Information and encouragement to be active daily and increase movement
Other	Posts that did not relate explicitly to one of the above categories, including changes to the profile picture and page banner image and posts that shared stories about fans and stakeholders

Table 2. Page-level measures and descriptions.

Variable	Description
Weekly new fans	The number of new page likes per week, overall and by gender
Weekly engaged users	The number of unique users who have engaged with the page per week, overall and by gender. This includes any click on the page or one of its post or any story ^a created by users
Weekly viral reach	The number of unique users who saw MHN or one of its posts from a story shared by a Facebook friend
Weekly paid impressions	Number of times a sponsored story or ad pointing to the page appeared in users’ News Feeds ^b . These impressions can be for fans and nonfans
Weekly organic impressions	Number of times MHN posts were displayed in News Feeds or on visits to the page. These impressions can be for fans and nonfans
Target Audience Rating Points (TARPs)	An estimate of the reach (how many people were exposed) and frequency (how often they were exposed) of the MHN television commercials per week, provided by an external ratings agency. This was used as an indicator of campaign advertising outside of Facebook

^aA user creates a “story” by liking the page, posting to the page’s timeline, liking, commenting on, or sharing one of the page’s posts, answering a question posted by the page, responding to an event, mentioning the page, or tagging the page in a photo.

^bNews Feed refers to the constantly updating list of stories in the middle of a user’s home page, including status updates, photos, videos, links, app activity, and likes from friends, pages, and groups that they follow.

Statistical Analysis

Post-Level Data

We conducted independent samples *t* tests to compare the means of engagement metrics (reach, likes, shares, comments, and post consumers) between paid and organic posts. In addition, we conducted a series of (separate) negative binomial regressions (the data were overdispersed), generating incidence rate ratios (IRRs) with the count of likes, comments, shares, and post consumers as the outcome variables, and post type, communication technique, and target behavior as categorical independent variables. The reference category for post type was photos (as this was the most populous category) and for communication technique was call-to-action/instructive because it represented a concrete action for users to take, as opposed to the other categories, which aimed to either inform or evoke emotion. For post day, each day was compared to the grand mean of all days, and for time of post during the day (8 am–5 pm; the most populous category) was used as the reference category. “Other” was used as the reference category for target behavior because these posts did not relate to specific behaviors. To examine whether the post being organic or paid interacted with other characteristics of the post, we entered two-way interaction terms for all covariates with paid/organic. Only significant two-way interactions were retained to generate the most parsimonious model. All models controlled for users’ exposure to the post by including an exposure or “offset” variable to estimate engagement with a post (ie, likes, comments) while accounting for the number of people each post was delivered to [34]. The relationship between post characteristics and engagement therefore becomes a rate per person reached.

Page-Level Data

To examine engagement with the MHN page over time as opposed to individual posts, we conducted time series analyses with page analytics. Time series analysis was used to account for the likely autocorrelation between observations (weekly counts) as Facebook users can view and react to content over an extended time. Further, prior engagement with content is a factor in the Facebook algorithm. Separate models were conducted for (1) new likes of the MHN page and (2) the number of unique users who “engaged” with the page for all users and for female and male users separately. In this context, “engagement” included any click on the MHN page or one of its posts or any “story” created, which would include actions such as liking the page; posting to the page’s timeline; liking, commenting on, or sharing a post; mentioning the page in one of their own posts; or tagging the page in a photo.

In addition to lag terms, each model initially included paid impressions, organic impressions, viral reach, a term for trend,

and the number of TARPs as predictors. Paid impressions, organic impressions, and viral reach were rescaled to the change in the outcome variable per 10,000 because the mean weekly counts were 167,857, 29,750, and 16,781, respectively. We used backward elimination (threshold of variable retention of $P=.10$). Modeling was preceded by tests for stationarity (Dickey-Fuller and Phillips-Perron) to ensure time series modeling was appropriate [35]. We examined autocorrelation with *q* tests and correlograms for each model [36].

To capture the impact of changing the post content in May 2017 to target men aged 35 to 54 years and families with children aged 5 to 12 years (operationalized as women aged 25–54 years), we conducted two interrupted time series (ITS) analyses with these subpopulations only, with weekly engaged users as the outcome. The same procedure as previously described was followed for the ITS analyses, only two terms were added to the models; namely, level change and change in trend [37]. These terms and the overall trend term were retained in the final models to examine whether there were significant effects of the change in campaign approach adjusted for other significant covariates.

Post- and page-level analyses were conducted using SPSS version 22.0 (*t* tests) and Stata version 15.0 (negative binomial regression, time series, and ITS analyses).

Results

Post-Level Data

In total, MHN posted 392 times during our analysis period, with 20.7% ($n=81$) of those posts receiving a paid boost (Table 3). The majority of posts (58.9%, $n=231$) were photos, whereas none were text only.

Posts that received a paid boost reached significantly more users and received significantly more likes, shares, comments, and post consumers than organic posts (Table 4). Across all measures, paid posts received at least 18 times the engagement compared to organic posts.

The significant interaction ($P<.001$) between organic/paid and post type indicated that the effect of paying was not the same across the three different types of posts (Table 5). Specifically, there was an incremental effect on likes, shares, comments, and post consumers for photos and links, but not for videos once adjusted for reach. For example, after adjusting for reach, both photo and link posts were predicted to receive more likes when paid (563 compared to 325 and 445 compared to 172, respectively), whereas paid video posts were predicted to receive only 53 likes compared to 211 for organic videos (Figure 1). A similar pattern was evident for all other engagement outcomes.

Table 3. Frequencies of post characteristics (N=392).

Post characteristic	Frequency, n (%)
Paid or organic	
Paid	81 (20.7)
Organic	311 (79.3)
Communication technique	
Instructive/call-to-action	204 (52.0)
Emotional	133 (33.9)
Informative	55 (14.0)
Post day	
Sunday	36 (9.2)
Monday	51 (13.0)
Tuesday	69 (17.6)
Wednesday	56 (14.3)
Thursday	69 (17.6)
Friday	60 (15.3)
Saturday	51 (13.0)
Post type	
Photo	231 (58.9)
Link	69 (17.6)
Video	92 (23.5)
Target behavior	
Act	118 (30.1)
Drink	67(17.1)
Eat	139 (35.5)
Other	68 (17.3)
Post time	
6 am to 8 am	111 (28.3)
8 am to 5 pm	202 (51.5)
After 5 pm	79 (20.2)

Table 4. Comparison of mean engagement for paid and organic posts using independent sample *t* tests.

Engagement metric	Paid mean (SD)	Organic mean (SD)	Mean difference (95% CI)	<i>P</i> value
Reach	107,764 (176,267)	3115 (2448)	104,649 (85,062-124,235)	<.001
Likes	886 (1175)	32 (33)	854 (723-985)	<.001
Shares	109 (205)	6 (8)	103 (80-126)	<.001
Comments	88 (137)	4 (6)	84 (68-99)	<.001
Consumers	1891 (3257)	86 (104)	1805 (1442-2167)	<.001

Table 5. Associations between post characteristics and engagement metrics per person reached calculated using negative binomial regressions adjusted for post reach.

Post characteristic	Likes, IRR ^a (95% CI)	Shares, IRR (95% CI)	Comments, IRR (95% CI)	Post consumers, IRR (95% CI)
Paid or organic				
Organic	Ref ^b	Ref	Ref	Ref
Paid	1.51 (1.17, 1.97)	0.84 (0.64, 1.09)	1.46 (1.05, 2.03)	1.02 (0.74, 1.39)
Post type				
Photo	Ref	Ref	Ref	Ref
Link	0.53 (0.44, 0.64)	0.67 (0.52, 0.86)	0.61 (0.44, 0.84)	0.72 (0.59, 0.88)
Video	0.65 (0.52, 0.81)	0.84 (0.63, 1.11)	0.85 (0.60, 1.21)	1.14 (0.91, 1.43)
Post day^c				
Sunday	0.93 (0.75, 1.15)	0.90 (0.69, 1.18)	0.92 (0.65, 1.29)	0.93 (0.74, 1.16)
Monday	0.73 (0.61, 0.88)	0.64 (0.51, 0.81)	0.81 (0.61, 1.09)	0.72 (0.60, 0.87)
Tuesday	1.06 (0.91, 1.23)	1.18 (0.98, 1.42)	0.90 (0.70, 1.14)	0.83 (0.71, 0.98)
Wednesday	1.00 (0.84, 1.18)	0.90 (0.73, 1.11)	1.01 (0.77, 1.33)	2.01 (1.67, 2.43)
Thursday	1.01 (0.87, 1.18)	1.15 (0.95, 1.39)	0.96 (0.75, 1.22)	0.88 (0.75, 1.03)
Friday	1.21 (1.02, 1.43)	1.08 (0.87, 1.35)	1.33 (1.01, 1.75)	1.08 (0.91, 1.29)
Saturday	1.05 (0.88, 1.25)	1.14 (0.91, 1.41)	1.14 (0.87, 1.49)	0.94 (0.78, 1.13)
Time of post				
8 am to 5 pm	Ref	Ref	Ref	Ref
6 am to 8 am	0.68 (0.57, 0.81)	0.91 (0.74, 1.12)	0.69 (0.53, 0.90)	0.62 (0.52, 0.74)
After 5 pm	0.72 (0.58, 0.89)	0.91 (0.72, 1.14)	0.85 (0.64, 1.13)	0.61 (0.50, 0.73)
Communication technique				
Instructive/call-to-action	Ref	Ref	Ref	Ref
Emotional	1.18 (1.00, 1.39)	1.00 (0.81, 1.24)	0.58 (0.45, 0.75)	1.03 (0.84, 1.27)
Informative	1.05 (0.85, 1.30)	0.90 (0.69, 1.17)	1.00 (0.72, 1.41)	0.98 (0.77, 1.24)
Target behavior				
Other	Ref	Ref	Ref	Ref
Act	0.87 (0.69, 1.08)	1.23 (0.92, 1.64)	1.11 (0.77, 1.58)	0.36 (0.28, 0.45)
Drink	0.96 (0.73, 1.25)	1.53 (1.09, 2.15)	0.84 (0.56, 1.27)	0.32 (0.24, 0.42)
Eat	0.81 (0.64, 1.01)	1.14 (0.85, 1.53)	0.92 (0.65, 1.31)	0.47 (0.36, 0.60)
Interactions with paid or organic^d				
Post type				
Paid link	1.49 (0.98, 2.26)	1.37 (0.83, 2.25)	0.69 (0.36, 1.29)	0.83 (0.53, 1.28)
Paid video	0.15 (0.09, 0.23)	0.32 (0.19, 0.53)	0.25 (0.13, 0.48)	0.46 (0.29, 0.74)
Time of post				
Paid 6 am to 8 am	1.45 (0.94, 2.24)	NS ^e	NS	NS
Paid after 5 pm	1.62 (1.06, 2.48)	NS	NS	NS
Paid or organic/communication technique interaction				
Paid emotional	NS	NS	NS	0.64 (0.43, 0.96)
Paid informative	NS	NS	NS	0.62 (0.34, 1.10)

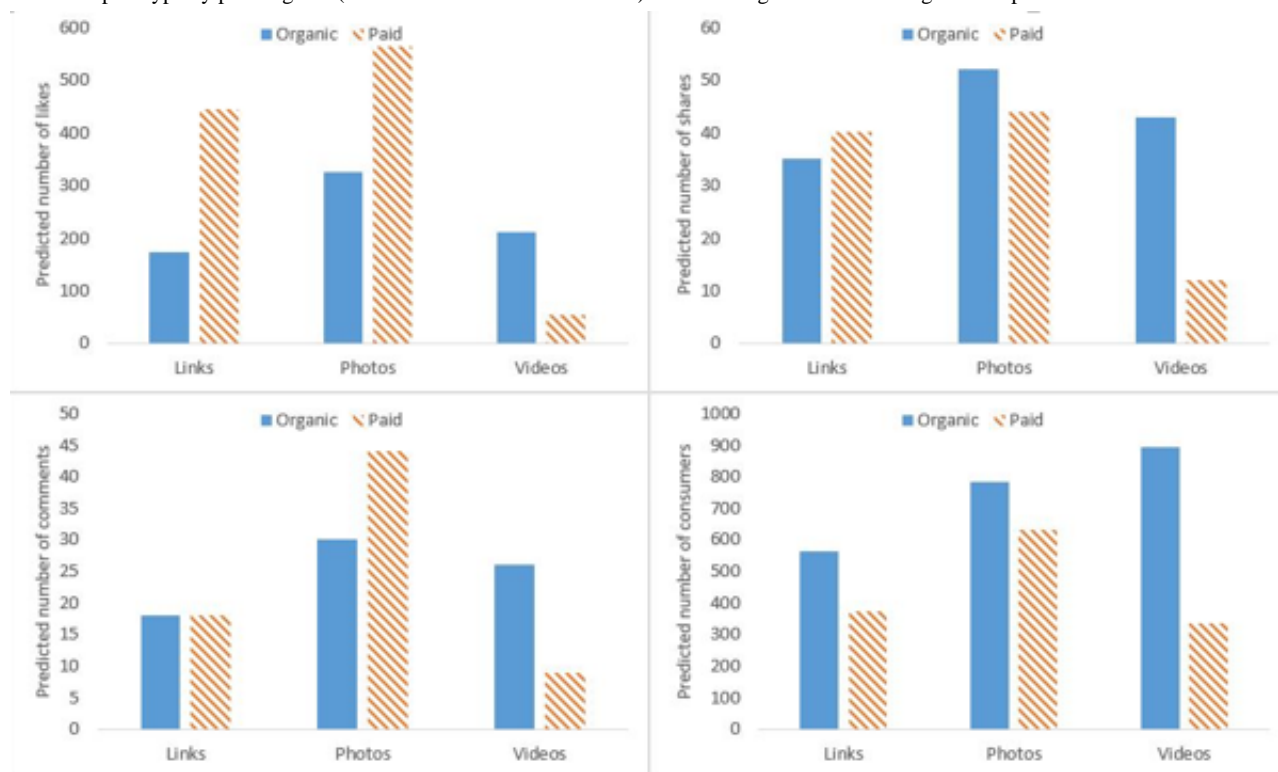
^aIRR: incident rate ratio.^bRef: reference category.

^cPost day is in comparison to the mean of all days.

^dOnly two-way interactions that were significant for at least one outcome are shown. Where the overall test of the interaction was nonsignificant, it was dropped from the final model.

^eNS: nonsignificant.

Figure 1. Predicted number of likes, shares, comments, and consumers by paid/organic status and post type, adjusting for reach. Note: marginal means calculated for post type by paid/organic (mean values for other covariates) based on negative binomial regressions presented in Table 5.



Posts made on Monday received 27% fewer likes, 36% fewer shares, and 28% fewer post consumers compared to the mean, whereas posts on a Tuesday received 17% fewer post consumers. On the other hand, posts on a Wednesday received more post consumers and posts on a Friday received more likes and shares. A significant interaction ($P=.045$) between organic/paid and time of post indicated that paying for posts before 8 am and after 5 pm had a greater incremental effect on likes than paying for posts between those hours. Posts made before 8 am received fewer comments and post consumers compared to posts made between 8 am and 5 pm irrespective of whether the post was paid or organic (ie, the interaction was nonsignificant). Similarly, posts made after 5 pm received fewer post consumers. The communication technique did not influence likes, shares, and comments, with the exception of emotional posts receiving fewer comments than instructive/call-to-action posts. However, the effect of paying for a post on post consumers differed across the three different types of communication techniques ($P=.049$), such that the effect was decremental on emotive posts but not for information posts relative to instructive/call-to-action posts. Finally, drink posts received significantly more shares compared to other posts (by 53%), but act, drink, and eat posts all received between 53% and 68% fewer post consumers compared to other posts.

Page-Level Data

Final time series models for all outcomes included only paid impressions, organic impressions, and viral reach, with all other initially included variables nonsignificant. There were three exceptions to this: organic impressions were nonsignificant in the model predicting weekly engaged male users, viral reach was nonsignificant in the model predicting weekly engaged female users, and TARPs was marginal ($P=.07$) in the model for engaged female users (Table 6).

In all models except weekly engaged males, organic impressions predicted a higher number of new fans and engaged users, compared to paid impressions. Viral reach similarly predicted a higher number of new fans and engaged users compared to paid impressions, but usually not as high as organic impressions. Organic impressions, compared to paid impressions, were considerably more influential for female users than for male users.

For the ITS analyses, none of the trend variables were significant in any of the models (Table 7). As may be expected given that the change in campaign strategy did not seem to change the trend in engagement either acutely or over time, the effect of paid and organic impressions and viral reach were similar in these subgroups to that seen in the models with the full sample and not including these trend terms.

Table 6. Time series results (beta coefficients with 95% CI) showing significant factors in the number of new weekly fans and engaged users (overall and by gender).

Per 10,000...	Weekly new fans, β (95% CI)			Weekly engaged users, β (95% CI)		
	Overall	Male	Female	Overall	Male	Female
Paid impressions	8.81 (7.55, 10.09)	1.97 (1.72, 2.21)	6.40 (5.38, 7.42)	68.53 (54.94, 82.12)	7.51 (6.67, 8.35)	12.67 (7.68, 17.67)
Organic impressions	58.02 (45.30, 70.74)	5.72 (2.70, 8.75)	58.82 (48.74, 68.90)	337.06 (212.22, 461.90)	NS	225.86 (160.47, 291.25)
Viral reach	22.05 (15.42, 28.67)	4.75 (3.19, 6.31)	11.02 (4.47, 17.56)	394.27 (241.86, 546.68)	28.69 (22.83, 34.57)	NS
TARPs ^a	NS ^b	NS	NS	NS	NS	-1.97 (-4.06, 0.12)

^aTARPs: Target Audience Rating Points.

^bNS: nonsignificant.

Table 7. Interrupted time series results showing significant factors in the number of new weekly fans and engaged users (by gender).

Predictors	Weekly new fans, β (95% CI)		Weekly engaged users, β (95% CI)	
	Male aged 35-54	Female aged 25-54	Male aged 35-54	Female aged 25-54
Per 10,000 paid impressions	0.96 (0.49, 1.43)	4.75 (4.05, 5.46)	3.54 (1.80, 5.28)	7.93 (5.41, 10.44)
Per 10,000 organic impressions	NS ^a	21.05 (13.20, 28.91)	NS	99.91 (69.04, 130.78)
Per 10,000 viral reach	4.84 (2.60, 7.08)	12.33 (8.12, 16.55)	25.33 (16.79, 33.87)	31.13 (8.03, 54.23)
Overall trend	-0.01 (-1.32, 1.31)	-0.08 (-8.13, 9.38)	0.19 (-5.55, 5.94)	1.30 (-3.56, 6.16)
Trend change	-0.69 (-5.81, 4.43)	0.63 (-8.13, 9.38)	-11.32 (-33.95, 11.30)	12.12 (-55.73, 79.98)
Level change	20.03 (-76.55, 116.63)	-67.91 (-150.34, 14.52)	99.08 (-297.01, 495.17)	-423.72 (-978.75, 131.31)

^aNS: nonsignificant.

Discussion

Principal Findings

This study examined usage of and engagement with the MHN Facebook page, identifying influential factors at both the post and page level. We found that paying for posts significantly increases reach of posts, but that the effect was not the same across post characteristics, most notably post type. At the same time, we found that organic impressions predicted higher engagement with the MHN page compared to paid impressions, particularly for female users. Together, these findings provide an important insight into the relative value of paid and organic posts: paying for posts is useful in increasing the reach of a page but the content itself must be engaging to capitalize on word-of-mouth marketing through organic reach. In addition, we found no association between television advertising and engagement with the page, suggesting that future campaigns should consider the role of Facebook within broader mass media campaigns and how different channels can complement one another to amplify campaign effects.

Our post-level results showed that paying for posts dramatically increased their reach. This is important because, as the hierarchy of effects predicts, exposure to a message is the first step in bringing about the desired change in behavior [38]. However, the time series analyses clearly showed that organic impressions and viral reach were of critical importance in driving engagement, especially among women. This is likely due to the

very high-level of trust placed in peer-to-peer communication [7] and that women are more likely to engage with health on social networking sites such as Facebook [39]. Collectively, our findings suggest that effective engagement through Facebook requires both maximizing the reach of posts through paid boosts and delivering content that users want to engage with and share in order to capitalize on word-of-mouth marketing [8]. However, how to strike a balance between the two is as yet unclear [40]. Current evidence shows that users will share content when they perceive it will be of benefit to their social network and where the risk of reputational damage is low [41], but what makes public health content “sharable” needs further investigation. This includes understanding why these results are strongest in women.

We also found that the effect of paying for posts on engagement was not the same across the different post types. Specifically, the effect of being paid on video posts appeared to be detrimental once reach was adjusted for, unlike photo and link posts. This may be due to videos requiring more effort on behalf of the user in that they need to watch and, usually, listen for an extended period. The increased effort may then mean that users will more readily scroll past a video if it does not immediately grab their attention, especially considering they will generally react negatively to obvious advertising [42]. When coupled with the fact that Facebook seems to give preferential treatment to videos in its algorithm compared to other post types [4], this finding highlights the need to weigh this preferential treatment against potential audience resistance. Public health agencies

must therefore give careful consideration of how best to use videos within their campaigns on Facebook. This is particularly important given recent changes to the Facebook algorithm, particularly a promise to prioritize content generated by friends and family (ie, organic content) [43].

Day and time of post appear to have had some influence on engagement, with posts made on Mondays generally leading to lower engagement, whereas Fridays led to higher engagement. This finding might reflect users readying themselves for the working week and for the weekend, respectively. That is, on Monday users are focusing on the “serious” tasks of work, subsequently spending less time on Facebook, whereas on Friday they are preparing for more social events and activities of the weekend, reflecting a key motivation for using social media [44]. In addition, posts made outside of working hours generally led to lower levels of engagement, which was unexpected given usage patterns show the most popular times to look at social media are first thing in the morning and in the evening [2]. It is also partly in conflict with a Canadian study that found a negative association between posts made during working hours and engagement, although that study also found a negative association between engagement and posts made after work [23]. Our finding might reflect the fact that more content from larger international markets (eg, the United States and Europe) would be posted at these times, meaning the MHN content would face more competition for users’ attention, but this would not explain the Canadian finding. Alternatively, these seemingly contradictory findings suggest that the more effective time of post might vary depending on the topic of the post.

Other post characteristics, however, appeared to be less influential. That emotional posts did not generate higher levels of engagement is of particular note and largely in line with a previous study [4]. This is surprising given that these types of messages have been shown to be more effective on other media channels [45] and are often presented as being more engaging on social media [46]. The question then is whether emotional appeals are simply not what users want when engaging with health on Facebook, page administrators are not delivering content of sufficient quality, or content is not appealing to the “right” emotions. It was also noteworthy that specific behaviors generally did not generate more (or less) engagement. The exception to this was drink posts receiving more shares, suggesting that users find this content to be more novel, relevant, and interesting [47]. Further research is needed to explore these characteristics in more detail, underscoring the importance of evaluating Facebook campaigns and disseminating the results.

With regards to the page-level analyses, we found that there was no link between Facebook engagement and television advertising, in contrast to the Tips evaluation [27,28]. This is likely because the MHN television advertisements do not specifically mention a Facebook page, but rather direct people to the MHN website that also does not invite visitors to follow the campaign on Facebook. That means that the Facebook page essentially operates independently from the other campaign elements because the only way users can find the page is by searching for it within Facebook or through incidental exposure to MHN content on Facebook. It is likely that stronger linkages between the campaign components would lead to greater

engagement with the Facebook page. However, it is unclear how best to synergize the campaign components, highlighting the need for robust evaluations of all components of mass media campaigns within public health. In addition, we found no evidence that the campaign narrowing its target audience led to any changes in the demographic profile of users who engaged with the Facebook page. This might be because the change in target audience occurred late in our analysis period and more time is needed to see an effect. Alternatively, it may have been because the content did not change appreciably or did not change in the right way to appeal to the new target audience. Campaign managers must therefore consider the role of each channel within a mass media campaign so that they complement one another. Some corporate brands, for instance, use Facebook as a way to associate particular events and values with their brand, as opposed to using it simply as another channel to sell their product [48]. Comprehensive formative and process evaluation would help to address these issues and help to bring about stronger linkages between the different campaign elements. However, formative and process evaluation are frequently overlooked and underreported in campaign evaluations [49].

A major limitation of our study is that we were limited to one campaign Facebook page covering just one health issue (overweight and obesity); tests with more pages that address different health issues are needed to strengthen our findings and increase their generalizability. In addition, our results should only be considered in relation to Facebook, rather than as relevant to other social media platforms given the reasons for using different platforms varies [47,50]. Our post-level analysis was also limited by a relatively small sample size of only 392 posts; more posts would have given us greater power to detect differences between the post characteristics. Finally, our interpretation of the results is based on the assumption that generating engagement is a necessary precursor to population-level impacts but, as yet, there is little evidence available to support this assumption within public health [51]. Outside of Facebook, there is suggestive evidence that skin cancer prevention messages disseminated on Twitter increased knowledge and reduced preference for a tan [52], but the impact of social media-disseminated messaging on health otherwise remains unknown. Investigating this link should be a priority for research, especially as recent changes in media consumption habits have necessitated a rethink in the relative value of different communication channels within mass media campaigns [53].

Conclusion

Our study shows the importance of paying to boost the reach of posts on Facebook while also demonstrating the value of maximizing organic reach, particularly in relation to videos. Therefore, page administrators should give careful consideration to their marketing strategy on Facebook as sole reliance on paid or organic posts could undercut the ability of a page to generate engagement and potentially influence health at a population level. Further, our results highlight the need for campaign managers to think strategically about the role of different campaign channels and how they can amplify and complement one another. These results also underscore the importance of ongoing evaluation of campaigns on social media, especially

on Facebook where the algorithm determining who sees what, when, and how often is adjusted regularly.

Acknowledgments

The authors would like to thank Michelle Maxwell, Beverley Lloyd, and Michael Collins from the New South Wales (NSW) Ministry of Health, and Dheepa Jeyapalan and Amy Vassallo from the Prevention Research Collaboration. The NSW Ministry of Health provided funding to the Prevention Research Collaboration to evaluate physical activity and nutrition programs, including the Make Healthy Normal campaign.

Conflicts of Interest

None declared.

References

1. Facebook Newsroom. Company Info 2017 URL: <https://newsroom.fb.com/company-info/> [accessed 2018-02-20] [WebCite Cache ID 6z4yMEMKn]
2. Sensis. 2015. Sensis social media report May 2015: how Australian people and businesses are using social media URL: https://www.sensis.com.au/assets/PDFdirectory/Sensis_Social_Media_Report_2015.pdf [accessed 2018-02-20] [WebCite Cache ID 6z4yZgk4]
3. Benetoli A, Chen TF, Aslani P. Consumer health-related activities on social media: exploratory study. *J Med Internet Res* 2017 Oct 13;19(10):e352 [FREE Full text] [doi: [10.2196/jmir.7656](https://doi.org/10.2196/jmir.7656)] [Medline: [29030326](https://pubmed.ncbi.nlm.nih.gov/29030326/)]
4. Kite J, Foley BC, Grunseit AC, Freeman B. Please like me: Facebook and public health communication. *PLoS One* 2016;11(9):e0162765 [FREE Full text] [doi: [10.1371/journal.pone.0162765](https://doi.org/10.1371/journal.pone.0162765)] [Medline: [27632172](https://pubmed.ncbi.nlm.nih.gov/27632172/)]
5. Park H, Rodgers S, Stemmle J. Health organizations' use of Facebook for health advertising and promotion. *J Interact Advertising* 2011 Sep;12(1):62-77. [doi: [10.1080/15252019.2011.10722191](https://doi.org/10.1080/15252019.2011.10722191)]
6. Kaplan A, Haenlein M. Two hearts in three-quarter time: how to waltz the social media/viral marketing dance. *Bus Horizons* 2011 May;54(3):253-263. [doi: [10.1016/j.bushor.2011.01.006](https://doi.org/10.1016/j.bushor.2011.01.006)]
7. Nielsen. 2013. Under the influence: consumer trust in advertising URL: <http://www.nielsen.com/us/en/insights/news/2013/under-the-influence-consumer-trust-in-advertising.html> [accessed 2018-02-20] [WebCite Cache ID 6z4yeCvXy]
8. Groeger L, Buttle F. Word-of-mouth marketing influence on offline and online communications: Evidence from case study research. *J Mark Commun* 2013 Jun 10;20(1-2):21-41. [doi: [10.1080/13527266.2013.797736](https://doi.org/10.1080/13527266.2013.797736)]
9. Thackeray R, Neiger BL, Smith AK, Van Wagenen SB. Adoption and use of social media among public health departments. *BMC Public Health* 2012 Mar 26;12:242 [FREE Full text] [doi: [10.1186/1471-2458-12-242](https://doi.org/10.1186/1471-2458-12-242)] [Medline: [22449137](https://pubmed.ncbi.nlm.nih.gov/22449137/)]
10. McGee M. Marketing Land. 2013. EdgeRank is dead: Facebook's News Feed algorithm now has close to 100K weight factors URL: <https://marketingland.com/edgerank-is-dead-facebook-news-feed-algorithm-now-has-close-to-100k-weight-factors-55908> [accessed 2018-05-01] [WebCite Cache ID 6z4yys2vW]
11. Dooley JA, Jones SC, Iverson D. Using Web 2.0 for health promotion and social marketing efforts: lessons learned from Web 2.0 experts. *Health Mark Q* 2014 May;31(2):178-196. [doi: [10.1080/07359683.2014.907204](https://doi.org/10.1080/07359683.2014.907204)] [Medline: [24878406](https://pubmed.ncbi.nlm.nih.gov/24878406/)]
12. McNab C. What social media offers to health professionals and citizens. *Bull World Health Organ* 2009 Aug;87(8):566 [FREE Full text] [Medline: [19704998](https://pubmed.ncbi.nlm.nih.gov/19704998/)]
13. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. *Health Promot Pract* 2013 Jan;14(1):15-23. [doi: [10.1177/1524839911405850](https://doi.org/10.1177/1524839911405850)] [Medline: [21558472](https://pubmed.ncbi.nlm.nih.gov/21558472/)]
14. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *J Med Internet Res* 2013;15(4):e85 [FREE Full text] [doi: [10.2196/jmir.1933](https://doi.org/10.2196/jmir.1933)] [Medline: [23615206](https://pubmed.ncbi.nlm.nih.gov/23615206/)]
15. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. *J Am Med Inform Assoc* 2015 Jan;22(1):243-256. [doi: [10.1136/amiainjnl-2014-002841](https://doi.org/10.1136/amiainjnl-2014-002841)] [Medline: [25005606](https://pubmed.ncbi.nlm.nih.gov/25005606/)]
16. Woolley P, Peterson M. Efficacy of a health-related Facebook social network site on health-seeking behaviors. *Soc Mark Quart* 2012 May 02;18(1):29-39. [doi: [10.1177/1524500411435481](https://doi.org/10.1177/1524500411435481)]
17. Platt T, Platt J, Thiel DB, Kardia SL. Facebook advertising across an engagement spectrum: a case example for public health communication. *JMIR Public Health Surveill* 2016 May 30;2(1):e27 [FREE Full text] [doi: [10.2196/publichealth.5623](https://doi.org/10.2196/publichealth.5623)] [Medline: [27244774](https://pubmed.ncbi.nlm.nih.gov/27244774/)]
18. Swindle TM, Ward WL, Whiteside-Mansell L. Facebook: the use of social media to engage parents in a preschool obesity prevention curriculum. *J Nutr Educ Behav* 2018 Jan;50(1):4.e1-10.e1. [doi: [10.1016/j.jneb.2017.05.344](https://doi.org/10.1016/j.jneb.2017.05.344)] [Medline: [29325661](https://pubmed.ncbi.nlm.nih.gov/29325661/)]

19. Williams G, Hamm MP, Shulhan J, Vandermeer B, Hartling L. Social media interventions for diet and exercise behaviours: a systematic review and meta-analysis of randomised controlled trials. *BMJ Open* 2014 Feb;4(2):e003926 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2013-003926](#)] [Medline: [24525388](#)]
20. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelandotte C. Are health behavior change interventions that use online social networks effective? A systematic review. *J Med Internet Res* 2014 Feb 14;16(2):e40 [[FREE Full text](#)] [doi: [10.2196/jmir.2952](#)] [Medline: [24550083](#)]
21. Balatsoukas P, Kennedy CM, Buchan I, Powell J, Ainsworth J. The role of social network technologies in online health promotion: a narrative review of theoretical and empirical factors influencing intervention effectiveness. *J Med Internet Res* 2015 Jun 11;17(6):e141 [[FREE Full text](#)] [doi: [10.2196/jmir.3662](#)] [Medline: [26068087](#)]
22. Beasley S, Leavy J, Pettigrew S, Croakey. 2017. Selling your message on Facebook: insights from an insider URL: <https://croakey.org/selling-your-health-message-on-facebook-insights-from-an-insider/> [accessed 2018-05-01] [[WebCite Cache ID 6z4ywwPCY](#)]
23. Card KG, Lachowsky N, Hawkins BW, Jollimore J, Baharuddin F, Hogg RS. Predictors of Facebook user engagement with health-related content for gay, bisexual, and other men who have sex with men: content analysis. *JMIR Public Health Surveill* 2018 Apr 06;4(2):e38 [[FREE Full text](#)] [doi: [10.2196/publichealth.8145](#)] [Medline: [29625953](#)]
24. Klassen KM, Borleis ES, Brennan L, Reid M, McCaffrey TA, Lim MS. What people “like”: analysis of social media strategies used by food industry brands, lifestyle brands, and health promotion organizations on Facebook and Instagram. *J Med Internet Res* 2018 Jun 14;20(6):e10227 [[FREE Full text](#)] [doi: [10.2196/10227](#)] [Medline: [29903694](#)]
25. Thackeray R, Neiger BL, Hanson CL, McKenzie JF. Enhancing promotional strategies within social marketing programs: use of Web 2.0 social media. *Health Promot Pract* 2008 Oct;9(4):338-343. [doi: [10.1177/1524839908325335](#)] [Medline: [18936268](#)]
26. Duke J, Hansen H, Kim A, Curry L, Allen J. The use of social media by state tobacco control programs to promote smoking cessation: a cross-sectional study. *J Med Internet Res* 2014 Jul 10;16(7):e169 [[FREE Full text](#)] [doi: [10.2196/jmir.3430](#)] [Medline: [25014311](#)]
27. Ayers JW, Althouse BM, Emery S. Changes in Internet searches associated with the “Tips from Former Smokers” campaign. *Am J Prev Med* 2015 Jun;48(6):e27-e29 [[FREE Full text](#)] [doi: [10.1016/j.amepre.2015.03.015](#)] [Medline: [25998929](#)]
28. Shafer PR, Davis KC, Patel D, Rodes R, Beistle D. Association between media dose, ad tagging, and changes in web traffic for a national tobacco education campaign: a market-level longitudinal study. *J Med Internet Res* 2016;18(2):e39 [[FREE Full text](#)] [doi: [10.2196/jmir.5343](#)] [Medline: [26887959](#)]
29. Davis KC, Shafer PR, Rodes R, Kim A, Hansen H, Patel D, et al. Does digital video advertising increase population-level reach of multimedia campaigns? Evidence from the 2013 Tips From Former Smokers campaign. *J Med Internet Res* 2016 Sep 14;18(9):e235 [[FREE Full text](#)] [doi: [10.2196/jmir.5683](#)] [Medline: [27627853](#)]
30. Allom V, Jongenelis M, Slevin T, Keightley S, Phillips F, Beasley S, et al. Comparing the cost-effectiveness of campaigns delivered various combinations of television and online media. *Front Public Health* 2018;6:83 [[FREE Full text](#)] [doi: [10.3389/fpubh.2018.00083](#)] [Medline: [29629366](#)]
31. Regional TAM, Ozdam, Nielsen. Nielsen. 2016. Australian multi-screen report: quarter 3 URL: <https://shopnielsen.worldsecuresystems.com/reports/consumer/australian-multi-screen-report-q3-2016> [accessed 2018-05-01] [[WebCite Cache ID 6z4yzvFaY](#)]
32. Kite J, Gale J, Grunseit A, Bellew W, Li V, Lloyd B, et al. Impact of the Make Healthy Normal mass media campaign (Phase 1) on knowledge, attitudes and behaviours: a cohort study. *Aust N Z J Public Health* 2018 Apr 12. [doi: [10.1111/1753-6405.12779](#)] [Medline: [29644787](#)]
33. Dunlop SM, Perez D, Cotter T. The natural history of antismoking advertising recall: the influence of broadcasting parameters, emotional intensity and executional features. *Tob Control* 2014 May;23(3):215-222. [doi: [10.1136/tobaccocontrol-2011-050256](#)] [Medline: [23143868](#)]
34. Hutchinson MK, Holtman MC. Analysis of count data using poisson regression. *Res Nurs Health* 2005 Oct;28(5):408-418. [doi: [10.1002/nur.20093](#)] [Medline: [16163676](#)]
35. Dickey DA, Bell WR, Miller RB. Unit roots in time series models: tests and implications. *Am Stat* 1986 Feb;40(1):12-26. [doi: [10.1080/00031305.1986.10475349](#)]
36. StataCorp. Stata Time Series Reference Manual. College Station, TX: Stata Press; 2013.
37. Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. *J Clin Pharm Ther* 2002 Aug;27(4):299-309. [Medline: [12174032](#)]
38. Barry TE, Howard DJ. A review and critique of the hierarchy of effects in advertising. *Int J Advert* 2015 Mar 02;9(2):121-135. [doi: [10.1080/02650487.1990.11107138](#)]
39. Prestin A, Vieux SN, Chou WS. Is online health activity alive and well or flatlining? Findings from 10 Years of the Health Information National Trends Survey. *J Health Commun* 2015 Jul;20(7):790-798. [doi: [10.1080/10810730.2015.1018590](#)] [Medline: [26042588](#)]
40. King R, Racherla P, Bush V. What we know and don't know about online word-of-mouth: a review and synthesis of the literature. *J Interact Mark* 2014 Aug;28(3):167-183 [[FREE Full text](#)] [doi: [10.1016/j.intmar.2014.02.001](#)]

41. Buttle F, Groeger L. Who says what to whom in what channel? A rules theoretic perspective on word-of-mouth marketing. J Mark Manag 2017 May 19;33(13-14):1035-1059. [doi: [10.1080/0267257X.2017.1325390](https://doi.org/10.1080/0267257X.2017.1325390)]
42. Wojdyski BW, Evans NJ. Going native: effects of disclosure position and language on the recognition and evaluation of online native advertising. J Advertising 2015 Dec 14;45(2):157-168. [doi: [10.1080/00913367.2015.1115380](https://doi.org/10.1080/00913367.2015.1115380)]
43. Shields M. Bus Insider Australia. 2018. A bunch of publishers are about to find out that Facebook's algorithm thinks their video content is worthless URL: <https://www.businessinsider.com.au/publishers-are-about-to-find-out-that-facebook-thinks-their-video-content-is-worthless-2018-1?r=US&IR=T> [accessed 2018-02-22] [WebCite Cache ID 6z4z6qJ7F]
44. Whiting A, Williams D. Why people use social media: a uses and gratifications approach. Qualitative Mrkt Res 2013 Aug 30;16(4):362-369. [doi: [10.1108/QMR-06-2013-0041](https://doi.org/10.1108/QMR-06-2013-0041)]
45. Sims M, Langley T, Lewis S, Richardson S, Szatkowski L, McNeill A, et al. Effectiveness of tobacco control television advertisements with different types of emotional content on tobacco use in England, 2004-2010. Tob Control 2016 Jan;25(1):21-26 [FREE Full text] [doi: [10.1136/tobaccocontrol-2013-051454](https://doi.org/10.1136/tobaccocontrol-2013-051454)] [Medline: [25037155](https://pubmed.ncbi.nlm.nih.gov/25037155/)]
46. Socialbakers. 2017. The essentials to building great content URL: <https://www.socialbakers.com/social-media-content/studies/essentials-to-building-great-content/> [accessed 2018-05-01] [WebCite Cache ID 6z4z7n0ZR]
47. Syn SY, Oh S. Why do social network site users share information on Facebook and Twitter? J Inform Sci 2015 May 19;41(5):553-569. [doi: [10.1177/0165551515585717](https://doi.org/10.1177/0165551515585717)]
48. Freeman B, Kelly B, Baur L, Chapman K, Chapman S, Gill T, et al. Digital junk: food and beverage marketing on Facebook. Am J Public Health 2014 Dec;104(12):e56-e64. [doi: [10.2105/AJPH.2014.302167](https://doi.org/10.2105/AJPH.2014.302167)] [Medline: [25322294](https://pubmed.ncbi.nlm.nih.gov/25322294/)]
49. Kite J, Grunseit A, Bohn-Goldbaum E, Bellew B, Carroll T, Bauman A. A systematic search and review of adult-targeted overweight and obesity prevention mass media campaigns and their evaluation: 2000-2017. J Health Commun 2018 Jan;23(2):207-232. [doi: [10.1080/10810730.2018.1423651](https://doi.org/10.1080/10810730.2018.1423651)] [Medline: [29336684](https://pubmed.ncbi.nlm.nih.gov/29336684/)]
50. Davenport S, Bergman S, Bergman J, Fearington ME. Twitter versus Facebook: exploring the role of narcissism in the motives and usage of different social media platforms. Comput Hum Behav 2014 Mar;32:212-220 [FREE Full text] [doi: [10.1016/j.chb.2013.12.011](https://doi.org/10.1016/j.chb.2013.12.011)]
51. Freeman B, Potente S, Rock V, McIver J. Social media campaigns that make a difference: what can public health learn from the corporate sector and other social change marketers? Public Health Res Pract 2015 Mar 30;25(2):e2521517 [FREE Full text] [doi: [10.17061/phrp2521517](https://doi.org/10.17061/phrp2521517)] [Medline: [25848735](https://pubmed.ncbi.nlm.nih.gov/25848735/)]
52. Gough A, Hunter RF, Ajao O, Jurek A, McKeown G, Hong J, et al. Tweet for behavior change: using social media for the dissemination of public health messages. JMIR Public Health Surveill 2017 Mar 23;3(1):e14 [FREE Full text] [doi: [10.2196/publichealth.6313](https://doi.org/10.2196/publichealth.6313)] [Medline: [28336503](https://pubmed.ncbi.nlm.nih.gov/28336503/)]
53. Niederdeppe J. Meeting the challenge of measuring communication exposure in the digital age. Commun Method Meas 2016 Apr 20;10(2-3):170-172. [doi: [10.1080/19312458.2016.1150970](https://doi.org/10.1080/19312458.2016.1150970)]

Abbreviations

IRR: incident rate ratio
ITS: interrupted time series analysis
MHN: Make Healthy Normal
NSW: New South Wales
TARPs: Target Audience Rating Points

Edited by T Sanchez; submitted 25.05.18; peer-reviewed by J Lei, P Shafer, L Laranjo; comments to author 23.07.18; revised version received 16.08.18; accepted 04.09.18; published 14.01.19.

Please cite as:

Kite J, Grunseit A, Li V, Vineburg J, Berton N, Bauman A, Freeman B
 Generating Engagement on the Make Healthy Normal Campaign Facebook Page: Analysis of Facebook Analytics
 JMIR Public Health Surveill 2019;5(1):e11132
 URL: <https://publichealth.jmir.org/2019/1/e11132/>
 doi:[10.2196/11132](https://doi.org/10.2196/11132)
 PMID:[31344679](https://pubmed.ncbi.nlm.nih.gov/31344679/)

©James Kite, Anne Grunseit, Vincy Li, John Vineburg, Nathan Berton, Adrian Bauman, Becky Freeman. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 14.01.2019. 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.

Review

Considering the Impact of Social Media on Contemporary Improvement of Australian Aboriginal Health: Scoping Review

Troy Walker^{1*}, BAppSc, MClChiro, MHumNut; Claire Palermo^{1*}, BSc, MNutr&Diet, MPH, PhD; Karen Klassen¹, BHumNutr&Diet, PhD

Be Active Sleep Eat, Department of Nutrition, Dietetics and Food, Monash University, Notting Hill, Australia

*these authors contributed equally

Corresponding Author:

Troy Walker, BAppSc, MClChiro, MHumNut

Be Active Sleep Eat

Department of Nutrition, Dietetics and Food

Monash University

Level 1

264 Ferntree Gully Road

Notting Hill, 3168

Australia

Phone: 61 9902 4270

Fax: 61 3 9902 4278

Email: troy.walker@monash.edu

Abstract

Background: Social media may have a significant role in influencing the present and future health implications among Australian Aboriginal and Torres Strait Islander people, yet there has been no review of the role of social media in improving health.

Objective: This study aims to examine the extent of health initiatives using social media that aimed to improve the health of Australian Aboriginal communities.

Methods: A scoping review was conducted by systematically searching databases CINAHL Plus; PubMed; Scopus; Web of Science, and Ovid MEDLINE in June 2017 using the terms and their synonyms “Aboriginal” and “Social media.” In addition, reference lists of included studies and the Indigenous HealthInfonet gray literature were searched. Key information about the social media intervention and its impacts on health were extracted and data synthesized using narrative summaries.

Results: Five papers met inclusion criteria. All included studies were published in the past 5 years and involved urban, rural, and remote Aboriginal or Torres Strait Islander people aged 12-60 years. No studies reported objective impacts on health. Three papers found that social media provided greater space for sharing health messages in a 2-way exchange. The negative portrayal of Aboriginal people and negative health impacts of social media were described in 2 papers.

Conclusions: Social media may be a useful strategy to provide health messages and sharing of content among Aboriginal people, but objective impacts on health remain unknown. More research is necessary on social media as a way to connect, communicate, and improve Aboriginal health with particular emphasis on community control, self-empowerment, and decolonization.

(*JMIR Public Health Surveill* 2019;5(1):e11573) doi:[10.2196/11573](https://doi.org/10.2196/11573)

KEYWORDS

aboriginal; Indigenous; social media; health; Web-based

Introduction

The need for evidence-based strategies to improve the health of Australia's Aboriginal and Torres Strait Islander people (hereafter referred to as Aboriginal) is essential to bridging the 10-year gap in life expectancy [1]. Addressing behaviors perpetuating chronic disease linked to diet and lifestyle and the

underlying determinants is complex and, thus, require multifaceted solutions, including screening, assessment, and treatment, support for behavioral change, and changes to the environment to promote healthy choices [2,3]. Having Aboriginal communities at the center of the design and delivery of health-related programs is well established to improve outcomes [4,5].

Social marketing applies marketing principles to disease prevention programs to facilitate health behavioral change [6]. An increase in the use of the internet and portable devices and apps has increased the use of social media as an avenue for social marketing. Social media is any Web-based communication dedicated to participant-based input, interaction, content-sharing, and collaboration. Social media is increasingly being used to try and improve health across the entire population. A recent systematic review of the use and advantages of social media for health communication identified its increasing use and potential to improve health outcomes. Yet, the literature on its benefits and application in Aboriginal populations has not been systematically explored; understanding its potential to improve health in Aboriginal populations is important as the use of some social media in remote areas has been reported as 20% higher than the national average [7]. In addition, some evidence suggests that much of the media portrayal of Aboriginal people is negative and may lead to poorer health outcomes [8]. Racial vilification, where the collective trauma of Aboriginal people is publicized, triggers painful reminders of colonialism [9]. Moreover, sexually explicit content is readily available to the youth of illegal age [3]. However, other data suggest that the ability of social media to support the creation and sharing of content and networking provides opportunities for health messages to be conveyed to a wider social network [10]. While there is some evidence to support the role of social media to promote and improve health in Aboriginal people, there is little evidence of its effect.

Of the evidence that exists, it appears social media may provide a contemporary conduit for Aboriginal people's expression of culture and the ability to access novel ways of health-related knowledge, learning, and engagement among one another and the wider community [11]. Little evidence includes the impact or effect of social media to change behavior or cultural norms [12]. Thus, there is a need to investigate the role of social media in delivering messages related to health for Aboriginal people and its impact on health outcomes.

This study aims to examine the extent of health initiatives using social media that aimed to improve the health of Australian Aboriginal communities.

Methods

Study Conception

To conceptualize outcomes relative to our question, we undertook a scoping review of the potential breadth of health implications that social media may have on Aboriginal Australian's health and well-being. A systematic approach, informed by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines and Population, Intervention, Comparison and Outcomes, was used to construct the research question and search terms [13,14]. A scoping review approach was selected on the basis of the paucity of evidence in the field; this methodology was chosen to summarize what was known and identify gaps based on the guidelines of scoping reviews developed by Arksey and O'Malley [15] and advanced by Levac [16].

Our methodology included searching and reviewing the literature to examine the extent and type of work being undertaken in relation to using social media to improve health in Australian Aboriginal communities and summarize this work and identify gaps. A team of 3 researchers were involved in the totality of the process, 1 Aboriginal Australian and 2 non-Aboriginal Australians.

Search Strategy

The databases CINAHL Plus, PubMed, Scopus, Web of Science, and Ovid Medical Literature Analysis and Retrieval System Online were searched in June 2017 using the terms and their synonyms "Aboriginal" and "Social media" (Textbox 1); this was followed with citation snowballing from relevant systematic reviews and included full-text papers whereby reference lists were scanned. In addition, the first 5 pages of the Aboriginal health evidence repository website Australian Indigenous Health InfoNet was searched for gray literature in September 2017 using the same search terms.

Data Management

In August 2017, results were exported to Covidence software (Covidence Systematic Review Software, Veritas Health Innovation) [17]. Titles and abstracts from searchers were screened by 2 authors (TW and CP). Differences of opinion were resolved through consensus discussions, and where agreement could not be reached, a third author (KK) was brought in to resolve.

Inclusion and Exclusion Criteria

All study designs were included. Published and unpublished studies were included; however, guidelines, protocols, opinion pieces, conference abstracts, and review papers were excluded. Systematic reviews were excluded; however, the reference lists of these studies were searched for relevant papers. Included studies must have reported participants who identified as Aboriginal or Torres Strait Islander, as well as some form of evaluative judgment on the role of social media on improving health.

The intervention or phenomena of interest was social media. Outcomes of interest were the acknowledgment, betterment, or detriment of Aboriginal-related health. In this instance, Aboriginal health and well-being were defined as any potential effects that improved or impaired any element of health, recognizing Aboriginal people's broad conceptualization of health [18]. Papers were excluded where studies were not specific to health outcomes in Aboriginal people and, concomitantly, where there was an absence of the use of social media in conjunction with a focus on improving health outcomes.

Data Analysis

Data extracted included author, date, location, sample size, and demographics (if known), as well as interventions, potential outcomes, and findings related to the study aims. In addition, notes on whether there was Aboriginal involvement in the study were recorded and the impacts of social networking sites on Australian Aboriginal health and well-being specifically summarized.

Textbox 1. The list of databases and all search terms used undertaken in June 2017.

Databases used:

- Ovid Medical Literature Analysis and Retrieval System Online
- Scopus
- CINAHL Plus
- PubMed
- Web of Science

Search terms used:

- aborigi* OR indigen* OR “oceanic ancestry group” OR “first australian” OR “Torres strait islander*” AND
- “social media” OR “social networking site” OR “social network* website*” OR “online social network” OR “online network” (Truncation)

Owing to the small number of included studies, this information was used to understand where the evidence currently exists and inform gaps, rather than a synthesis of findings. Narrative summaries of the qualitative studies included comparing and contrasting social media interventions across studies, as well as potential outcomes from each study to assist in informing a summary of the role of social media in improving the health in Australian Aboriginal communities as is typical of scoping reviews.

Quality Assessment

Studies were assessed for quality by 2 authors (CP and TW) using the Critical Appraisal Skills Programme tool [19]. Studies were scored out of 8 criterion points based on the quality assessment. Studies with a score of >4 (out of 8) were considered good quality; studies with a score of 4 were considered neutral quality and studies with a score of <4 were considered poor quality.

Results

Study Criteria

The initial search revealed 301 studies, which after duplicates were removed, leaving 234 for screening (Figure 1). Screening 234 titles and abstracts left 25 full-text papers for full-text screening. The full text of the remaining 25 papers was assessed, and a further 22 papers were excluded because of not meeting both health-related and social media-related outcomes, or because of not being an empirical study where some evaluation of the intervention was undertaken. Citation snowballing found an additional 7 papers, and following full-text assessment, 1 was included. Searching the gray literature in HealthInfoNet produced 3 papers relevant for assessment. Of the 3 included, 2 more were excluded; 1 was excluded owing to the inability to retrieve further information through contact with the author, whereas the second was excluded given the absence of health and related outcomes.

Included Study Features and Quality Assessment

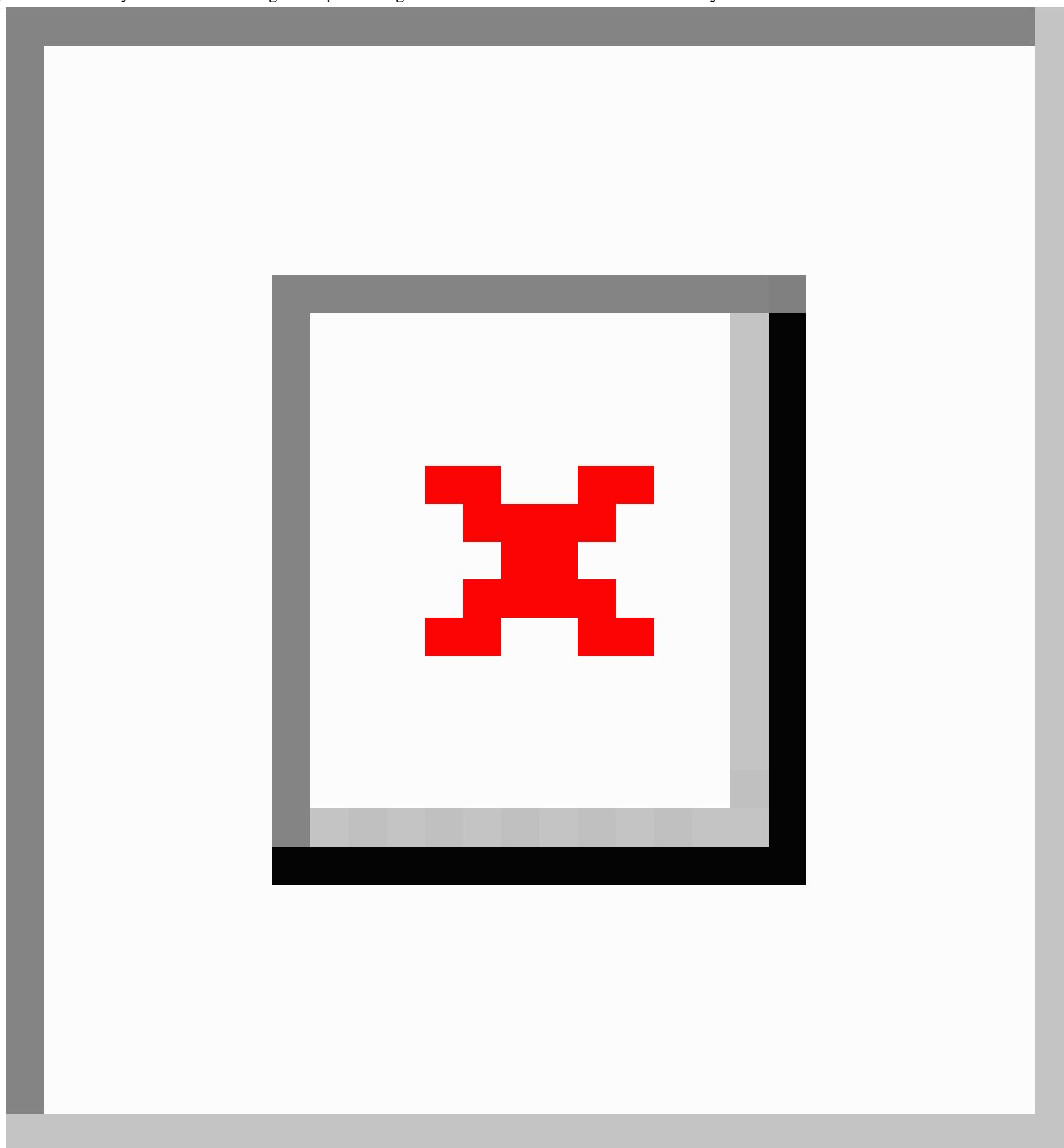
While only 5 studies were found, a narrative summary of these studies was deemed appropriate to assist in understanding the role of social media in improving the health of Aboriginal Australians and guide future research in this area. Of the 5 included papers, 3 used qualitative approaches [11,20,21] and

2 used multiple methods [22,23]. Based on the quality assessment tool used, from the 5 studies, 2 were of neutral quality, 2 were of poor quality, and 1 was of good quality (Multimedia Appendix 1).

All papers were published within the past 5 years, from 2013 to 2017. The studies involved Aboriginal people aged 12-60 years and from both males and females. The settings of the social media campaigns were initiated in urban, rural, and remote Australian locations where they allowed for more widespread involvement. Study durations ranged from 1 day (15 hours) to the present day with ongoing reporting. Participant numbers varied between 28 and 346 people. A key feature and strength of all studies were that Aboriginal people were part of the research project and either involved in gathering experiential data, forming researcher-community partnerships or the research writing itself [24]. No studies reported objective impacts on health.

Social media was used as a tool to enhance social support in all studies, whereby community members were connected online. Social support occurred between social media users or users and Aboriginal health organizations by linking real-world events with Web-based conversations and in improving awareness of access to offline social and emotional support. In addition, social media was used to disseminate information more widely outside the study population as social media was proposed to provide a platform for reaching a broader audience.

One study found that age was associated with social media use for health [21]. Older Aboriginal groups often found using social media for health more complex and, in some cases, having detrimental health outcomes [21], whereas younger groups were more readily receptive to using social media for their health and well-being [11,20,21]. All studies mentioned the need for more time for participants to become familiar with utilizing social media for it to have an impact on health; the reason acknowledged was the relative infancy of social media use and Aboriginal health within Australia [21,23]. All studies showed improved health, which included exercise, nutrition, family, mental health, suicide, death, and grieving. Furthermore, all outlined the need for future social media health campaigns to consider current Australian Aboriginal health culture and perspective [11,20-23].

Figure 1. The study selection flow diagram representing the selection of studies included in the systematic literature review.

A repeated theme that appeared in 3 papers was that social media provided greater space for sharing health messages in a 2-way exchange [11,20,23]. One paper noted the increased awareness and self-empowerment of Aboriginal people in governing their own health after applying one particular social media campaign [11]. Another study showed that when the aim was to increase the quality and duration of Aboriginal people's lives, an emphasis on sport and promotion of physical activity using social media as a medium was well received based on the overall participation and positive feedback. When this approach was combined with other health behaviors, such as quitting smoking, or decreasing the consumption of added sugar and sugar-sweetened beverages, more positive qualitative responses were apparent [22].

Negative health impacts were described in 2 papers using social media, where it was perceived to represent Aboriginal people in a poor light relative to health-related conditions [11,20]. One study specifically noted it could be inadvertently disrespectful by displaying death notices where elder Aboriginal people were unable to use or access social media; this included learning of illness, deaths, and funeral services belatedly in the family through Facebook rather than in-person [21]. Another study outlined that the consistent focus on the health implications, including chronic diseases like diabetes, obesity, and mental health, was potentially negative and labeled Aboriginal people into a deficit position [20]. Respondents in these 2 studies voiced concern with the negative images portrayed in all forms of media of Aboriginal people regarding their health [11,20].

Discussion

Principal Findings

This study aimed to examine the extent of health initiatives using social media that aimed to improve the health of Australian Aboriginal communities; it found 5 studies that evaluated the impact of a range of social media strategies on health or well-being. Social media provided a space for providing social support, sharing health-promoting messages, and increasing awareness and self-efficacy of Aboriginal people in governing their own health. The cocreation of social media content with Aboriginal people and concepts of both self and community empowerment that aimed to improve health appeared to be well received based on the participation and positive feedback.

Literature is scarce regarding the use of social media as a conduit in promoting the health of Australian Aboriginal people. To the best of our knowledge, this is the first scoping review using a systematic approach to evaluating the evidence of health initiatives using social media that aimed to improve the health of Australian Aboriginal communities. A consistent and apparent theme was the concept of a healing and self-empowering dialogue among Aboriginal people. These themes, while often termed in a variety of different ways, centered around end users, researchers, and funders working together to construct contemporary ways to refine, expand, and improve Aboriginal health using multiple platforms of social media. Common alternate names used included, but were not limited to, cocreation, self-determination, 2-way communication, and self-design [11,20,25,26]. Most studies focused on the positive elements of improving Aboriginal health; this is in contrast to much of the previous literature, which framed their research around “disease” and the problems associated with the disease rather than “health.” Other work has investigated social media and its role in racial vilification [9]. The examples analyzed in this study show that social media has significant negative and detrimental impacts on Aboriginal people as they are reminded of colonization. However, the authors acknowledge that their findings highlight the potential vehicle of social media to have conversations that promote change [9]. In addition, a recent study has found that Australian Aboriginal people interact about their health using social media [27]. Our review highlighted that research that addresses and evaluates decolonization and self-empowerment will be more likely to improve Aboriginal health outcomes [11,20,28]. Sharing health information online may gather traction and community capital among Aboriginal communities when using positive messages related to diet, exercise, or smoking rather than threatening approaches frequently used in health media campaigns [27]. When there is an online sense of community support, with a particular focus on self-empowering language that promotes and encourages making better choices related to Aboriginal health and well-being, participation in social media may increase; this area shows promise for more work, given its positive reception and popularity among Aboriginal people [11]. More evaluation is warranted with framing “health” positively to improve Aboriginal health and its associated outcomes.

Social media was used as a platform for social support in most of the included studies. As social and emotional well-being and community connectivity are important for Aboriginal people, enhanced access to social support networks is important for enabling behavioral change [20,29-31]. Social media, through its increased reach could enhance and enlarge support networks; this is important for all Australian Aboriginal communities where access to support may be limited. In addition, information dissemination of public health messages and increasing awareness of access to support and health care can be enhanced for those living in remote communities [32]. The unfavorable findings within the included studies was that social media could be perceived to represent the health of Aboriginal people negatively [20,21,23] or conjure up emotion when learning about funeral services, death, and grieving on social media rather than in-person [21]; these are important considerations for the future use of social media in Aboriginal communities. Likewise, other work has shown that social media may heighten and increase conflict and violence among feuding families [33]. As social media can be used to increase reach for health messages, it can also be used to amplify stigma, racism, and bullying by more widely spreading negative messages. Social media can be used to propagate stigma, and this has been observed in many stigmatized health conditions such as mental health and Alzheimer’s disease [34,35]. Important lessons were learned from #IHMayDay social media strategy as concerns were prospectively raised about the detrimental impacts of negative framing and participants were urged to engage positively throughout the day [20]; this negative potential of social media must be considered for future interventions.

Respecting and appreciating traditional customs of Indigenous groups in building scientific evidence for Indigenous people has been called for in other work [36-38]. The impact of racism on psychological health and the overall negative approach taken by the portrayal of Maori people in all forms of media has been previously highlighted [38]; this fault is noted as a result of the adaptation to recent colonization. A recent systematic review of social marketing targeting Indigenous people across the world found that social marketing interventions primarily used television and radio advertising and appeared to confront health issues of Indigenous populations around the world despite not maximizing all elements of social marketing [39]. These findings together provide evidence for the need to consider social media as strategies to improve the health of Australia’s Aboriginal people, acknowledging the need to use positive health messaging and portray these communities using a strengths-based focus.

Limitations

This study is limited to social media and does not include other social marketing campaigns. Studies only focused on Australian Aboriginal populations and may not be relevant to other Indigenous populations across the world using other platforms for social marketing beyond social media. However, this scoping review has highlighted the lack of studies that actually examine the impact of health-related social media activities in Aboriginal people. While inferences are made toward the perceived or self-reported impact on health or well-being, there was no actual objective measurement in any of the included studies. There is

a need for work examining the impact of social media on actual health outcomes.

Conclusions

Understanding the potential for social media to improve health and well-being in Australian Aboriginal communities is important for researchers, public health professionals, and policy makers. Our scoping review found that there is potential for social media to provide a space for sharing health-promoting messages and increase awareness and self-efficacy of Aboriginal people in governing their own health and for social support.

The cocreation of social media content involving Aboriginal people with the aim to improve health appears to influence participation when framed in a positive health context or form of self-empowerment. However, not all social media approaches are positively associated with Aboriginal people, and some negative health relationships still exist and require further exploration. There is a need for the development and implementation of cocreated messages with the Australian Aboriginal community delivered over social media and the subsequent measurement of its impact on health outcomes.

Conflicts of Interest

None declared.

Multimedia Appendix 1

A summary of included studies on the impact of social media on Aboriginal health outcomes.

[DOCX File, 18KB - [publichealth_v5i1e11573_app1.docx](#)]

References

1. Australian Institute of Health and Welfare. Mortality and life expectancy of Indigenous Australians: 2008 to 2012. Canberra: AIHW; Sep 05, 2014:1-60.
2. Vos T, Barker B, Begg S, Stanley L, Lopez AD. Burden of disease and injury in Aboriginal and Torres Strait Islander Peoples: the Indigenous health gap. *Int J Epidemiol* 2009 Apr;38(2):470-477. [doi: [10.1093/ije/dyn240](#)] [Medline: [19047078](#)]
3. Rice E, Haynes E, Royce P, Thompson SC. Social media and digital technology use among Indigenous young people in Australia: a literature review. *Int J Equity Health* 2016 May 25;15(1):81 [FREE Full text] [doi: [10.1186/s12939-016-0366-0](#)] [Medline: [27225519](#)]
4. Dwyer J, O'Donnell K, Lavioli J, Marlina U, Sullivan P. The Overburden Report: Contracting for Indigenous Health Services. Darwin: Cooperative Research Centre for Aboriginal Health; 2009:1-78.
5. Panaretto KS, Wenitong M, Button S, Ring IT. Aboriginal community controlled health services: leading the way in primary care. *Med J Aust* 2014 Jun 16;200(11):649-652. [Medline: [24938346](#)]
6. Grier S, Bryant CA. Social marketing in public health. *Annu Rev Public Health* 2005;26:319-339. [doi: [10.1146/annurev.publhealth.26.021304.144610](#)] [Medline: [15760292](#)]
7. Callinan T. National Indigenous Television. 2014. Remote Indigenous Australians rely on Facebook to stay in touch URL: <https://www.sbs.com.au/nitv/nitv-news/article/2014/08/26/remote-indigenous-australians-rely-facebook-stay-touch> [accessed 2018-11-30] [WebCite Cache ID 74JChEFAQ]
8. Stoneham M, Goodman J, Daube MA. The Portrayal of Indigenous Health in Selected Australian Media. *IIPJ* 2014 Jan;5(1):1-13. [doi: [10.18584/iipj.2014.5.1.5](#)]
9. Carlson B, Jones LV, Harris M, Quezada N, Frazer R. Trauma, Shared Recognition and Indigenous Resistance on Social media. *AJIS* 2017 Jul 19;21:21. [doi: [10.3127/ajis.v21i0.1570](#)]
10. Moorhead S, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *J Med Internet Res* 2013 Apr 23;15(4):e85 [FREE Full text] [doi: [10.2196/jmir.1933](#)] [Medline: [23615206](#)]
11. McPhail-Bell K, Appo N, Haymes A, Bond C, Brough M, Fredericks B. Deadly Choices empowering Indigenous Australians through social networking sites. *Health Promot Int* 2018 Oct 01;33(5):770-780. [doi: [10.1093/heapro/dax014](#)] [Medline: [28387801](#)]
12. Brusse C, Gardner K, McAullay D, Dowden M. Social media and mobile apps for health promotion in Australian Indigenous populations: scoping review. *J Med Internet Res* 2014 Dec 10;16(12):e280 [FREE Full text] [doi: [10.2196/jmir.3614](#)] [Medline: [25498835](#)]
13. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015 Dec 02;350:g7647 [FREE Full text] [doi: [10.1136/bmj.g7647](#)] [Medline: [25555855](#)]
14. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak* 2007 Jun 15;7:16 [FREE Full text] [doi: [10.1186/1472-6947-7-16](#)] [Medline: [17573961](#)]

15. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology* 2005 Feb;8(1):19-32. [doi: [10.1080/1364557032000119616](https://doi.org/10.1080/1364557032000119616)]
16. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci* 2010 Sep 20;5:69 [FREE Full text] [doi: [10.1186/1748-5908-5-69](https://doi.org/10.1186/1748-5908-5-69)] [Medline: [20854677](https://pubmed.ncbi.nlm.nih.gov/20854677/)]
17. Covidence. 2017. URL: <https://www.covidence.org/about-us> [accessed 2018-12-04] [WebCite Cache ID 74PF6aHFM]
18. Rowley KG, Daniel M, Skinner K, Skinner M, White GA, O'Dea K. Effectiveness of a community-directed 'healthy lifestyle' program in a remote Australian aboriginal community. *Aust N Z J Public Health* 2000 Apr;24(2):136-144. [Medline: [10790932](https://pubmed.ncbi.nlm.nih.gov/10790932/)]
19. CASP Checklists. Middleway, Oxford, UK: CASP; 2017. URL: <https://casp-uk.net/casp-tools-checklists/> [accessed 2018-12-06] [WebCite Cache ID 74SgyuQ1G]
20. Sweet M, Geia L, Dudgeon P, McCallum K. #IHMayDay: tweeting for empowerment and social and emotional wellbeing. *Australas Psychiatry* 2015 Dec;23(6):636-640. [doi: [10.1177/1039856215609762](https://doi.org/10.1177/1039856215609762)] [Medline: [26487324](https://pubmed.ncbi.nlm.nih.gov/26487324/)]
21. Carlson B, Frazer R. It's like Going to a Cemetery and Lighting a Candle?: Aboriginal Australians, Sorry Business and social media. *AlterNative: An International Journal of Indigenous Peoples* 2015;11(3):211-224 [FREE Full text]
22. Berends L, Halliday R. Centre for Health and Social Research. 2016. Sport and Health Social Marketing Strategy (Deadly Sport Gippsland): Evaluation URL: <http://gippsport.com.au/wp-content/uploads/2014/09/DSG-Final-Evaluation-Report-June-2016.pdf> [accessed 2018-11-30] [WebCite Cache ID 74JDZ4Amj]
23. Carlson B, Farrelly T, Frazer R. Mediating tragedy: Facebook, aboriginal peoples and suicide. *Australasian Journal of Information Systems* 2015;19:1-15 [FREE Full text]
24. Humphery K. Dirty questions: Indigenous health and 'Western research'. *Aust N Z J Public Health* 2001 Jun;25(3):197-202. [Medline: [11494986](https://pubmed.ncbi.nlm.nih.gov/11494986/)]
25. Trees K. Mobile media: communicating with and by Indigenous youth about alcohol. *Australian Aboriginal Studies* 2015:97-106.
26. Sweet MA. Social media: new links for Indigenous health. *Med J Aust* 2013 Jul 08;199(1):18. [Medline: [23829246](https://pubmed.ncbi.nlm.nih.gov/23829246/)]
27. Hefler M, Kerrigan V, Henryks J, Freeman B, Thomas DP. Social media and health information sharing among Australian Indigenous people. *Health Promot Int* 2018 Apr 17. [doi: [10.1093/heapro/day018](https://doi.org/10.1093/heapro/day018)] [Medline: [29672684](https://pubmed.ncbi.nlm.nih.gov/29672684/)]
28. Sweet MA, Dudgeon P, McCallum K, Ricketson MD. Decolonising practices: can journalism learn from health care to improve Indigenous health outcomes? *Med J Aust* 2014 Jun 16;200(11):626-627. [Medline: [24938333](https://pubmed.ncbi.nlm.nih.gov/24938333/)]
29. Sweet M, Pearson L, Dudgeon P. @Indigenousx: A Case Study of Community-Led Innovation in Digital Media. *Media International Australia* 2013 Nov;149(1):104-111. [doi: [10.1177/1329878X1314900112](https://doi.org/10.1177/1329878X1314900112)]
30. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. *Health Psychol Rev* 2016 Sep;10(3):277-296 [FREE Full text] [doi: [10.1080/17437199.2016.1151372](https://doi.org/10.1080/17437199.2016.1151372)] [Medline: [26854092](https://pubmed.ncbi.nlm.nih.gov/26854092/)]
31. Latkin C, Knowlton AR. Social Network Assessments and Interventions for Health Behavior Change: A Critical Review. *Behav Med* 2015;41(3):90-97 [FREE Full text] [doi: [10.1080/08964289.2015.1034645](https://doi.org/10.1080/08964289.2015.1034645)] [Medline: [26332926](https://pubmed.ncbi.nlm.nih.gov/26332926/)]
32. Cooper A. The Use of Online Strategies and Social Media for Research Dissemination in Education. *EPAA* 2014 Sep 01:1:22. [doi: [10.14507/epaa.v22n88.2014](https://doi.org/10.14507/epaa.v22n88.2014)]
33. Kral I. Shifting perceptions, shifting identities: Communication technologies and the altered social, cultural and linguistic ecology in a remote indigenous context. *Aust J Anthropol* 2014 Jun 01;25(2):171-189. [doi: [10.1111/taja.12087](https://doi.org/10.1111/taja.12087)]
34. Oscar N, Fox PA, Croucher R, Wernick R, Keune J, Hooker K. Machine Learning, Sentiment Analysis, and Tweets: An Examination of Alzheimer's Disease Stigma on Twitter. *J Gerontol B Psychol Sci Soc Sci* 2017 Sep 01;72(5):742-751. [doi: [10.1093/geronb/gbx014](https://doi.org/10.1093/geronb/gbx014)] [Medline: [28329835](https://pubmed.ncbi.nlm.nih.gov/28329835/)]
35. Athanasopoulou C, Suni S, Hätönen H, Apostolakis I, Lionis C, Välimäki M. Attitudes towards schizophrenia on YouTube: A content analysis of Finnish and Greek videos. *Inform Health Soc Care* 2016;41(3):307-324. [doi: [10.3109/17538157.2015.1008485](https://doi.org/10.3109/17538157.2015.1008485)] [Medline: [25710275](https://pubmed.ncbi.nlm.nih.gov/25710275/)]
36. Stanley L, Swaim RC, Kaholokula JK, Kelly KJ, Belcourt A, Allen J. The Imperative for Research to Promote Health Equity in Indigenous Communities. *Prev Sci* 2017 Nov 06. [doi: [10.1007/s11121-017-0850-9](https://doi.org/10.1007/s11121-017-0850-9)] [Medline: [29110278](https://pubmed.ncbi.nlm.nih.gov/29110278/)]
37. Walker J, Lovett R, Kukutai T, Jones C, Henry D. Indigenous health data and the path to healing. *Lancet* 2017 Nov 04;390(10107):2022-2023. [doi: [10.1016/S0140-6736\(17\)32755-1](https://doi.org/10.1016/S0140-6736(17)32755-1)] [Medline: [29115232](https://pubmed.ncbi.nlm.nih.gov/29115232/)]
38. Nairn R, Pega F, McCreanor T, Rankine J, Barnes A. Media, racism and public health psychology. *J Health Psychol* 2006 Mar;11(2):183-196. [doi: [10.1177/1359105306061179](https://doi.org/10.1177/1359105306061179)] [Medline: [16464918](https://pubmed.ncbi.nlm.nih.gov/16464918/)]
39. Kubacki K, Szablewska N. Social marketing targeting Indigenous peoples: a systematic review. *Health Promot Int* 2017 Sep 23. [doi: [10.1093/heapro/dax071](https://doi.org/10.1093/heapro/dax071)] [Medline: [29036628](https://pubmed.ncbi.nlm.nih.gov/29036628/)]

Edited by G Eysenbach; submitted 17.07.18; peer-reviewed by J Pendergrass, J Du; comments to author 08.10.18; revised version received 18.10.18; accepted 23.10.18; published 05.02.19.

Please cite as:

Walker T, Palermo C, Klassen K

Considering the Impact of Social Media on Contemporary Improvement of Australian Aboriginal Health: Scoping Review

JMIR Public Health Surveill 2019;5(1):e11573

URL: <http://publichealth.jmir.org/2019/1/e11573/>

doi: [10.2196/11573](https://doi.org/10.2196/11573)

PMID: [30720442](https://pubmed.ncbi.nlm.nih.gov/30720442/)

©Troy Walker, Claire Palermo, Karen Klassen. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 05.02.2019. 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

Spatial Access and Willingness to Use Pre-Exposure Prophylaxis Among Black/African American Individuals in the United States: Cross-Sectional Survey

Bisola O Ojikutu^{1*}, MD, MPH; Laura M Bogart^{2*}, PhD; Kenneth H Mayer^{3*}, MD; Thomas J Stopka^{4*}, PhD; Patrick S Sullivan^{5*}, MPH, DVM, PhD; Yusuf Ransome^{6*}, DrPH

¹Department of Medicine, Brigham and Women's Hospital, Division of Global Health Equity, Harvard Medical School, Boston, MA, United States

²RAND Corporation, Santa Monica, CA, United States

³The Fenway Institute, Boston, MA, United States

⁴Department of Public Health and Community Medicine, Tufts University School of Medicine, Boston, MA, United States

⁵Department of Epidemiology, Emory Rollins School of Public Health, Atlanta, GA, United States

⁶Department of Social and Behavioral Sciences, Yale School of Public Health, New Haven, CT, United States

* all authors contributed equally

Corresponding Author:

Bisola O Ojikutu, MD, MPH

Department of Medicine, Brigham and Women's Hospital

Division of Global Health Equity

Harvard Medical School

75 Francis Street

Boston, MA, 02115

United States

Phone: 1 9172872710

Email: bojikutu@partners.org

Abstract

Background: Uptake of pre-exposure prophylaxis (PrEP) among black individuals in the United States is low and may be associated with the limited availability of clinics where PrEP is prescribed.

Objective: We aimed to determine the association between spatial access to clinics where PrEP is prescribed and willingness to use PrEP.

Methods: We identified locations of clinics where PrEP is prescribed from AIDSVu.org and calculated the density of PrEP clinics per 10,000 residents according to the ZIP code. Individual-level data were obtained from the 2016 National Survey on HIV in the Black Community. We used multilevel modelling to estimate the association between willingness to use PrEP and clinic density among participants with individual-level (HIV risk, age, gender, education, income, insurance, doctor visit, census region, urban/rural residence) and ZIP code-level (%poverty, %unemployed, %uninsured, %black population, and density of health care facilities) variables.

Results: All participants identified as black/African American. Of the 787 participants, 45% were men and 23% were found to be at high risk based on the self-reported behavioral characteristics. The mean age of the participants was 34 years (SD 9), 54% of participants resided in the South, and 26% were willing to use PrEP. More than one-third (38%) of the sample had to drive more than 1 hour to access a PrEP provider. Participants living in areas with higher PrEP clinic density were significantly more willing to use PrEP (one SD higher density of PrEP clinics per 10,000 population was associated with 16% higher willingness [adjusted prevalence ratio=1.16, 95% CI: 1.03-1.31]).

Conclusions: Willingness to use PrEP was associated with spatial availability of clinics where providers prescribe PrEP in this nationally representative sample of black African Americans.

(*JMIR Public Health Surveill* 2019;5(1):e12405) doi:[10.2196/12405](https://doi.org/10.2196/12405)

KEYWORDS

barriers; black individuals; HIV prevention; HIV services; pre-exposure prophylaxis; PrEP; racial disparities; spatial access

Introduction

Black Americans comprise 44% of the nearly 40,000 individuals diagnosed with HIV in the United States in 2016 [1]. Pre-exposure prophylaxis (PrEP) in the form of tenofovir/emtricitabine is highly effective for the prevention of HIV [2,3]. Recent data suggest that although the need for PrEP is highest among black individuals, the rate of PrEP uptake is disproportionately low in this group compared to white Americans [4-6].

Individual-level barriers to PrEP use among black individuals include lack of awareness, low perceived HIV risk, and safety concerns [7-10]. PrEP also requires a prescription, and black individuals report a usual source of health care less often than white individuals [11]. Among black individuals, PrEP use may also be limited due to mistrust of the health care system, which is a result of contemporary and historical experiences of racism and discrimination [12,13].

Beyond these barriers, structural factors or aspects of one's environment that are out of one's immediate control, such as spatial access, may limit PrEP uptake [14-17]. Spatial access or proximity to services has been explored as a structural barrier to HIV services [18,19]. However, few studies have explored spatial access to PrEP. The purpose of this study was to determine the association between proximity to PrEP-prescribing clinics and willingness to use PrEP among black individuals in the United States.

Methods

Individual-Level Data

The National Survey on HIV in the Black Community was a cross-sectional survey administered to black individuals (aged 18 to 50 years) through Knowledge Networks, which is a probability-based, online, nonvolunteer Web panel [20]. Panel members were recruited from randomly selected addresses obtained from the US Postal Delivery Sequence File. All surveys were completed via email. Households without internet service were provided internet access.

To develop the survey, cognitive interviews were conducted with a convenience sample of black individuals, and 64 questions, one per screen, were included, some of which had automatic skip logic. A back button could be used to change responses. Once the survey was completed, the participant could not reenter the survey to submit additional surveys. A pretest was conducted with 26 cases and reviewed for accuracy. No personal information was collected. The median completion time was 13 minutes. No cookies were used, and duplication was not allowed in the panel.

This study was approved by Boston Children's Hospital (Boston, MA). Informed consent was obtained prior to administration. Data were collected from February 12 to April 17, 2016. A US \$5 online gift card was offered to all participants. All 1969 black participants in the panel were sampled. Of the 896 (45.50%) who provided consent, 868 (96.88%) were eligible and completed the survey. Only completed surveys were included in this analysis. The sample included individuals reporting high

and low HIV risk. HIV-positive individuals were excluded. Post-stratification weighting was performed using sociodemographic benchmarks from the March 2016 supplement of the Current Population Survey [21] to ensure that estimates were representative of adults living in households in the United States in 2016 [22].

Outcome

Willingness to use PrEP was ascertained by selecting yes, no, or maybe to item "If a pill that could prevent transmission of HIV from an HIV positive sex partner to an uninfected partner were available I would take it." Responses were collapsed to obtain the risk ratio (yes vs no/maybe).

Based on prior research, the following covariates were selected: age (continuous); gender (male, female, or transgender male or female); HIV risk (more than one sex partner and no condom use in the last 3 months, anal sex with more than one partner and no condom use in the last 3 months, men having sex with men, history of a sexually transmitted infection in the last 3 months, drug use in the last 30 days, transgender individuals, and transactional sex work); education (less than high school, high school graduate/General Education Development [GED], college, or higher); income; insurance status (insured or uninsured); visits to a doctor (<12 months or >12 months); metropolitan statistical area (urban or rural); and region.

Exposure

We obtained the locations of PrEP-prescribing clinics from AIDSVu.org [23] on September 29, 2017, for 760 unique ZIP codes and geocoded the addresses in ArcMap 10.4 [24], which produced 173 locations across 127 distinct ZIP Codes. The development and validation of the database of PrEP-prescribing locations has been described elsewhere [25,26]. We selected a random sample of clinics (10%) and confirmed the time of start of prescriptions to ensure the exposure date preceded collection of our outcome variable (willingness to use PrEP). We calculated the density of PrEP clinics per 10,000 residents according to the ZIP code by using the Census 2010 denominators [27] and the number of PrEP clinics per square mile using ArcMap 10.4 [24].

ZIP Code Covariates

ZIP code-level covariates from the American Community Survey 2007-2011 (%living in poverty, %unemployed, %black population) and %uninsured from the American Community Survey 2008-2012 (unavailable for 2007-2011) were downloaded from the American Fact Finder Website [22]. We also adjusted for the density of clinics, community centers, and hospitals in 2016, which were retrieved from ERSI Business Analyst 2016 using North American Industry Classification System codes (621111 and 621112 for doctor offices, 621498 for community health care centers, and 622110 for hospitals) [23,28]. Kernel densities for each variable per ZIP code were created; we used principal component analysis to create one composite score, since the variables were highly correlated (Pearson $r > 0.84$). Next, we created the variable driving distance (in miles) to the nearest PrEP clinic from the population-weighted centroid of all US ZIP codes within a 1-hour buffer, which incorporates street networks. Driving

distance was measured from each geocentroid to the nearest PrEP-prescribing clinic. We limited the distance calculation to a 1-hour maximum to limit calculations that included crossing state lines. The variable was coded into four equal categories based on quartile distributions of drive times below 60 miles or <1 hour driving time, and the remainder of the sample was included within a fifth category, which served as the reference group. Drive time of more than 1 hour was chosen as the reference category, because a large proportion of the sample fell outside the 1-hour mark, which we assumed would be a barrier to accessing PrEP. Geospatial analyses were conducted using ArcMap 10.4 [24], and statistical analyses were performed using STATA 14.1 (Stata Corp., College Station, TX).

Data Analysis

We merged the ZIP code data of individuals and subsequently conducted a multilevel, multivariable analysis to estimate the association between willingness to use PrEP and PrEP clinic density while adjusting for individual and ZIP code covariates in one block. Adjusted prevalence ratios (APR) were calculated along with the 95% CIs. Associations were considered significant at $P < .05$.

Results

We included 787 participants and 700 distinct ZIP codes in the multilevel analysis. Among the participants, 45% were male and 23% were at high risk for HIV infection based on self-reported behavioral characteristics. The mean age of participants was 34 years (SD 9), 54% resided in the South, and 26% were willing to use PrEP. Among high-risk participants, 40.8% were willing to use PrEP. The mean number of PrEP clinics per ZIP code was 1.73 (SD 0.64), the density per 10,000 people was 0.07 (SD 0.22), and 38% of the sample had to drive more than 1 hour to access PrEP. Participants living in areas with higher PrEP clinic density were significantly more willing to use PrEP: 1 SD higher density of PrEP clinics per 10,000 people was associated with 16% higher willingness (APR=1.16, 95% CI=1.03-1.31). Participants with a high school diploma or GED were less likely to be willing to use PrEP than participants without such education levels (APR=0.60, 95% CI=0.37-0.99). Self-reported high HIV risk (APR=1.70, 95% CI=1.27-2.27) and residence in the West compared to Northeast (APR=2.04, 95% CI=1.06-3.93) were significantly associated with higher likelihoods of willingness to use PrEP (Table 1 and Table 2).

Table 1. Individual-level characteristics and multivariable associations with willingness to use pre-exposure prophylaxis (PrEP), 2016.

Individual-level characteristics included in the model (N=787)	Value	Adjusted prevalence ratio	95% CI
Age (years), mean (SD)	34 (9.03)	1.00	0.98-1.01
Gender, n (%)^a			
Male	308 (45.0)	1	Reference
Female	479 (55.0)	1.16	0.87-1.54
Education, n (%)			
Lower than high school	58 (11.0)	1	Reference
High school diploma/GED ^b	169 (34.0)	0.60	0.37-0.99
College or higher	560 (55.0)	0.74	0.46-1.19
Income, n (%)			
<US \$10,000	154 (16.0)	1	Reference
US \$10,000-\$39,999	266 (29.0)	0.89	0.60-1.32
US \$40,000-\$99,999	281 (41.0)	0.80	0.51-1.27
≥US \$100,000	86 (15.0)	0.67	0.34-1.34
Insurance status, n (%)			
Currently insured	670 (83.0)	1.20	0.75-1.92
Not insured	117 (17.0)	1	Reference
Doctor visit, n (%)			
≤12 months	619 (77.0)	1	Reference
>12 months or never	168 (23.0)	0.79	0.51-1.21
HIV risk, n (%)			
Yes	194 (23.0)	1.70	1.27-2.27
No	593 (77.0)	1	Reference
Census region, n (%)			
Northeast	142 (19.0)	1	Reference
Midwest	163 (18.0)	1.52	0.75-3.07
South	389 (53.0)	1.45	0.77-2.74
West	93 (11.0)	2.04	1.06-3.93
Metropolitan statistical area, n (%)			
Urban	729 (91.0)	1	Reference
Rural	58 (9.0)	0.72	0.38-1.37
Willingness to use PrEP^c, n (%)			
Yes	212 (26.0)	— ^d	—
No/maybe	575 (74.0)	—	—

^aPercentages are weighted.^bGED: General Education Development.^cPrEP: pre-exposure prophylaxis.^dNot available.

Table 2. ZIP code–level characteristics and multivariable associations with willingness to use pre-exposure prophylaxis (PrEP), 2016.

ZIP code–level variables included in model	Value	Adjusted prevalence ratio	95% CI
Density of PrEP ^a clinics per 10,000 people, mean (SD)	0.00 (0.23)	1.16	1.03-1.31
Driving distance to PrEP clinic from population centroid (miles)^b, mean (SD)	10.76 (11.77)	— ^c	—
0-2.05	124 (15.76)	0.58	0.33-1.03
2.06-7.19	131 (16.65)	0.70	0.43-1.11
7.20-16.68	122 (15.50)	0.63	0.36-1.11
16.69-57.27	114 (14.49)	0.70	0.45-1.10
>1-hour drive time (reference) ^d	296 (37.61)	1	—
Density of doctors and outpatient clinics, mean (SD)	0.21 (0.24)	—	—
Density of CHCs ^e , mean (SD)	0.00 (0.00)	—	—
Density of hospitals, mean (SD)	0.01 (0.01)	—	—
Density of clinics/CHCs/hospital composites ^f , mean (SD)	—	1.04	0.80-1.33
Black/African American ^g (%), mean (SD)	35.89 (28.52)	1.05	0.87-1.26
Unemployed ^f (%), mean (SD)	11.42 (5.32)	1.05	0.84-1.31
Living in poverty ^g (%), mean (SD)	20.02 (11.15)	0.98	0.75-1.27
Uninsured ^f (%), mean (SD)	16.65 (6.79)	1.04	0.87-1.24

^aPrEP: pre-exposure prophylaxis.

^bThis is a 5-level variable.

^cNot available.

^dDriving time is approximately 1 minute per mile. Therefore, >1-hour drive time is equivalent to 60 miles. This is the reference group.

^eCHC: community health center.

^fA composite variable created through principal components analysis using density of doctors/outpatient clinics, community health centers, and hospitals; higher scores indicate greater density.

^gMean weighted percentage.

Discussion

Spatial proximity is a critical determinant of access to health services [29,30]. Our findings indicate that spatial access to PrEP-prescribing clinics is associated with greater willingness to use PrEP among black individuals. These findings provide additional evidence that access to PrEP must expand to increase uptake. In this study, spatial access may have led to greater willingness to use PrEP because of increased awareness through formal advertising, informal neighborhood networks, or direct knowledge of someone taking PrEP. Social capital and a general “culture of health” may also impact willingness and have been positively associated with use of HIV services [15,31].

Importantly, we found that nearly 40% of the sample would need to drive for >1 hour to access PrEP. Studies have shown that transportation barriers have a significant impact on health outcomes, particularly among disadvantaged individuals [32,33]. Regarding HIV care and treatment, travel time has been found to be a barrier to retention in care [34,35].

Based on our findings, an increase in the number of PrEP-prescribing providers in areas where access is currently

limited would increase the use of PrEP. Recent studies have reported that many providers are unfamiliar with PrEP and have concerns about its safety and risk compensation [36-38]. Interventions that educate providers about PrEP are critical. Novel interventions such as the use of navigators or online prescriptions may also be necessary to increase uptake of PrEP.

This study has several limitations. Although the PrEP-prescribing database has undergone validation [25], sites may have been missed. In addition, the survey did not measure actual PrEP use. However, willingness to use PrEP provides a reasonable measure of potential uptake. Distance to PrEP sites could have been calculated from participants' addresses; however, we only had access to participants' ZIP codes. Although the sample size was modest, these data are nationally representative and weighted to reflect the population composition of black individuals in the United States.

In conclusion, this study demonstrates that black individuals with higher spatial access to PrEP-prescribing clinics were more willing to use this intervention. Scaling up of PrEP prescription at clinics in areas where black individuals reside is necessary to increase access to PrEP.

Acknowledgments

We gratefully acknowledge the participants in the National Survey on HIV in the Black Community and our National Advisory Committee. We would also like to thank Felton Earls, MD, for his leadership in survey development. This publication was possible because of help from the Harvard University Center for AIDS Research (CFAR), an NIH-funded program (P30 AI060354 to KHM and BOO). Additional investigator funding was received from the National Institute of Mental Health (NIMH) K23MH107316 (BOO), NIMH K01MH111374 (YR), and NIMH P30MH058107 (KHM).

Conflicts of Interest

None declared.

References

- Centers for Disease Control and Prevention. HIV Surveillance Report, 2016. 2017 Nov. URL: <https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-2016-vol-28.pdf> [accessed 2018-10-01] [WebCite Cache ID 72qfgz01y]
- Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, et al. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. *N Engl J Med* 2010 Dec 30;363(27):2587-2599 [FREE Full text] [doi: [10.1056/NEJMoa1011205](https://doi.org/10.1056/NEJMoa1011205)] [Medline: [21091279](https://pubmed.ncbi.nlm.nih.gov/21091279/)]
- Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, et al. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. *N Engl J Med* 2012 Aug 2;367(5):399-410 [FREE Full text] [doi: [10.1056/NEJMoa1108524](https://doi.org/10.1056/NEJMoa1108524)] [Medline: [22784037](https://pubmed.ncbi.nlm.nih.gov/22784037/)]
- Eaton LA, Matthews DD, Driffin DD, Bukowski L, Wilson PA, Stall RD, POWER Study Team. A Multi-US City Assessment of Awareness and Uptake of Pre-exposure Prophylaxis (PrEP) for HIV Prevention Among Black Men and Transgender Women Who Have Sex with Men. *Prev Sci* 2017 Dec;18(5):505-516 [FREE Full text] [doi: [10.1007/s11121-017-0756-6](https://doi.org/10.1007/s11121-017-0756-6)] [Medline: [28101813](https://pubmed.ncbi.nlm.nih.gov/28101813/)]
- Smith D, Van Handel M, Grey J. By race/ethnicity, Blacks have the highest number needed PrEP in the United States, 2015. 2018 Presented at: Conference on Retroviruses and Opportunistic Infections; March 4–7, 2018; Boston, Massachusetts URL: <http://www.croiconference.org/sessions/raceethnicity-blacks-have-highest-number-needing-prep-united-states-2015>
- Mera R, McCallister S, Palmer B, Mayer G, Magnuson D, Rawlings K. 21st International AIDS Conference, Durban, Abstract TUAX0105LB. 2016. FTC/TDF (Truvada) for HIV pre-exposure prophylaxis (PrEP) utilization in the United States URL: <http://programme.aids2016.org/Abstract/Abstract/10159> [accessed 2018-10-01] [WebCite Cache ID 72qgr28MT]
- Aaron E, Blum C, Seidman D, Hoyt MJ, Simone J, Sullivan M, et al. Optimizing Delivery of HIV Preexposure Prophylaxis for Women in the United States. *AIDS Patient Care STDS* 2018 Jan;32(1):16-23 [FREE Full text] [doi: [10.1089/apc.2017.0201](https://doi.org/10.1089/apc.2017.0201)] [Medline: [29323558](https://pubmed.ncbi.nlm.nih.gov/29323558/)]
- Philbin MM, Parker CM, Parker RG, Wilson PA, Garcia J, Hirsch JS. The Promise of Pre-Exposure Prophylaxis for Black Men Who Have Sex with Men: An Ecological Approach to Attitudes, Beliefs, and Barriers. *AIDS Patient Care STDS* 2016 Dec;30(6):282-290 [FREE Full text] [doi: [10.1089/apc.2016.0037](https://doi.org/10.1089/apc.2016.0037)] [Medline: [27220036](https://pubmed.ncbi.nlm.nih.gov/27220036/)]
- Kwakwa HA, Bessias S, Sturgis D, Walton G, Wahome R, Gaye O, et al. Engaging United States Black Communities in HIV Pre-exposure Prophylaxis: Analysis of a PrEP Engagement Cascade. *J Natl Med Assoc* 2018 Oct;110(5):480-485. [doi: [10.1016/j.jnma.2017.12.006](https://doi.org/10.1016/j.jnma.2017.12.006)] [Medline: [30129509](https://pubmed.ncbi.nlm.nih.gov/30129509/)]
- Ojikutu BO, Bogart LM, Higgins-Biddle M, Dale SK, Allen W, Dominique T, et al. Facilitators and Barriers to Pre-Exposure Prophylaxis (PrEP) Use Among Black Individuals in the United States: Results from the National Survey on HIV in the Black Community (NSHBC). *AIDS Behav* 2018 Feb 21. [doi: [10.1007/s10461-018-2067-8](https://doi.org/10.1007/s10461-018-2067-8)] [Medline: [29468493](https://pubmed.ncbi.nlm.nih.gov/29468493/)]
- Orgera K, Artiga S. Henry J Kaiser Family Foundation. Disparities in health and health care: five key questions and answers URL: <https://www.kff.org/disparities-policy/issue-brief/disparities-in-health-and-health-care-five-key-questions-and-answers/> [accessed 2018-10-01] [WebCite Cache ID 72qhhAWBm]
- Cahill S, Taylor SW, Elsesser SA, Mena L, Hickson D, Mayer KH. Stigma, medical mistrust, and perceived racism may affect PrEP awareness and uptake in black compared to white gay and bisexual men in Jackson, Mississippi and Boston, Massachusetts. *AIDS Care* 2017 Nov;29(11):1351-1358 [FREE Full text] [doi: [10.1080/09540121.2017.1300633](https://doi.org/10.1080/09540121.2017.1300633)] [Medline: [28286983](https://pubmed.ncbi.nlm.nih.gov/28286983/)]
- Brooks RA, Allen VC, Regan R, Mutchler MG, Cervantes-Tadeo R, Lee S. HIV/AIDS conspiracy beliefs and intention to adopt preexposure prophylaxis among black men who have sex with men in Los Angeles. *Int J STD AIDS* 2018 Dec;29(4):375-381. [doi: [10.1177/0956462417727691](https://doi.org/10.1177/0956462417727691)] [Medline: [28853676](https://pubmed.ncbi.nlm.nih.gov/28853676/)]
- Parkhurst JO. Structural approaches for prevention of sexually transmitted HIV in general populations: definitions and an operational approach. *J Int AIDS Soc* 2014;17:19052 [FREE Full text] [Medline: [25204872](https://pubmed.ncbi.nlm.nih.gov/25204872/)]
- Ransome Y, Kawachi I, Dean LT. Neighborhood Social Capital in Relation to Late HIV Diagnosis, Linkage to HIV Care, and HIV Care Engagement. *AIDS Behav* 2017 Mar;21(3):891-904 [FREE Full text] [doi: [10.1007/s10461-016-1581-9](https://doi.org/10.1007/s10461-016-1581-9)] [Medline: [27752875](https://pubmed.ncbi.nlm.nih.gov/27752875/)]

16. Ransome Y, Kawachi I, Braunstein S, Nash D. Structural inequalities drive late HIV diagnosis: The role of black racial concentration, income inequality, socioeconomic deprivation, and HIV testing. *Health Place* 2016 Dec;42:148-158 [FREE Full text] [doi: [10.1016/j.healthplace.2016.09.004](https://doi.org/10.1016/j.healthplace.2016.09.004)] [Medline: [27770671](https://pubmed.ncbi.nlm.nih.gov/27770671/)]
17. Bronfenbrenner U. *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, MA: Harvard University Press; 1981.
18. Goswami ND, Schmitz MM, Sanchez T, Dasgupta S, Sullivan P, Cooper H, et al. Understanding Local Spatial Variation Along the Care Continuum: The Potential Impact of Transportation Vulnerability on HIV Linkage to Care and Viral Suppression in High-Poverty Areas, Atlanta, Georgia. *J Acquir Immune Defic Syndr* 2016 May 01;72(1):65-72 [FREE Full text] [doi: [10.1097/QAI.0000000000000914](https://doi.org/10.1097/QAI.0000000000000914)] [Medline: [26630673](https://pubmed.ncbi.nlm.nih.gov/26630673/)]
19. Dasgupta S, Kramer MR, Rosenberg ES, Sanchez TH, Sullivan PS. Spatial Accessibility to HIV Providers in Atlanta, Georgia. *AIDS Res Hum Retroviruses* 2015 May;31(5):473-474 [FREE Full text] [doi: [10.1089/aid.2014.0365](https://doi.org/10.1089/aid.2014.0365)] [Medline: [25844663](https://pubmed.ncbi.nlm.nih.gov/25844663/)]
20. Knowledge Networks. URL: <http://www.tessexperiments.org/data/zip/Simas675/Knowledge%20Networks%20Methodology.pdf> [WebCite Cache ID 72tjwMxgL]
21. United States Census Bureau. Current Population Survey (CPS) URL: <https://www.census.gov/programs-surveys/cps/data-detail.html> [accessed 2019-01-09] [WebCite Cache ID 75IrupvJJ]
22. United States Census Bureau - American FactFinder. 2017. URL: <https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml> [accessed 2018-10-01] [WebCite Cache ID 72qiDyW34]
23. AIDSvu. URL: <https://aidsvu.org/> [accessed 2018-10-01] [WebCite Cache ID 72qfzCpkR]
24. ESRI. Redlands, CA; 2017. ArcGIS Blog URL: <https://www.esri.com/arcgis-blog/products/product/mapping/using-and-citing-esri-data/> [accessed 2018-12-14] [WebCite Cache ID 74fM8jeeh]
25. Siegler AJ, Wirtz S, Weber S, Sullivan PS. Developing a Web-Based Geolocated Directory of HIV Pre-Exposure Prophylaxis-Providing Clinics: The PrEP Locator Protocol and Operating Procedures. *JMIR Public Health Surveill* 2017 Sep 06;3(3):e58 [FREE Full text] [doi: [10.2196/publichealth.7902](https://doi.org/10.2196/publichealth.7902)] [Medline: [28877865](https://pubmed.ncbi.nlm.nih.gov/28877865/)]
26. Siegler A, Bratcher A, Weiss K, Mouhanna F, Ahlschlager L, Sullivan P. Location location location: an exploration of disparities in access to publicly listed pre-exposure prophylaxis clinics in the United States. *Ann Epidemiol* 2018 Dec;28(12):858-864 [FREE Full text] [doi: [10.1016/j.annepidem.2018.05.006](https://doi.org/10.1016/j.annepidem.2018.05.006)] [Medline: [30406756](https://pubmed.ncbi.nlm.nih.gov/30406756/)]
27. United States Census Bureau. URL: <https://www.census.gov/data/datasets/2010/dec/summary-file-1.html>. Accessed [accessed 2018-12-17] [WebCite Cache ID 74juim6FV]
28. United States Census Bureau. North American Industry Classification System URL: <https://www.census.gov/eos/www/naics/> [accessed 2018-12-17] [WebCite Cache ID 74jv4vb8n]
29. Fishman J, McLafferty S, Galanter W. Does Spatial Access to Primary Care Affect Emergency Department Utilization for Nonemergent Conditions? *Health Serv Res* 2018 Dec;53(1):489-508. [doi: [10.1111/1475-6773.12617](https://doi.org/10.1111/1475-6773.12617)] [Medline: [27859257](https://pubmed.ncbi.nlm.nih.gov/27859257/)]
30. Huang Y, Meyer P, Jin L. Neighborhood socioeconomic characteristics, healthcare spatial access, and emergency department visits for ambulatory care sensitive conditions for elderly. *Prev Med Rep* 2018 Dec;12:101-105 [FREE Full text] [doi: [10.1016/j.pmedr.2018.08.015](https://doi.org/10.1016/j.pmedr.2018.08.015)] [Medline: [30233997](https://pubmed.ncbi.nlm.nih.gov/30233997/)]
31. Ransome Y, Galea S, Pabayo R, Kawachi I, Braunstein S, Nash D. Social Capital is Associated With Late HIV Diagnosis: An Ecological Analysis. *J Acquir Immune Defic Syndr* 2016 Dec 01;73(2):213-221 [FREE Full text] [doi: [10.1097/QAI.0000000000001043](https://doi.org/10.1097/QAI.0000000000001043)] [Medline: [27632146](https://pubmed.ncbi.nlm.nih.gov/27632146/)]
32. Locatelli SM, Sharp LK, Syed ST, Bhansari S, Gerber BS. Measuring Health-related Transportation Barriers in Urban Settings. *J Appl Meas* 2017;18(2):178-193 [FREE Full text] [Medline: [28961153](https://pubmed.ncbi.nlm.nih.gov/28961153/)]
33. Erskine NA, Waring ME, McManus DD, Lessard D, Kiefe CI, Goldberg RJ. Barriers to Healthcare Access and Long-Term Survival After an Acute Coronary Syndrome. *J Gen Intern Med* 2018 Jul 11. [doi: [10.1007/s11606-018-4555-y](https://doi.org/10.1007/s11606-018-4555-y)] [Medline: [29998434](https://pubmed.ncbi.nlm.nih.gov/29998434/)]
34. Ridgway JP, Almirol EA, Schmitt J, Schuble T, Schneider JA. Travel Time to Clinic but not Neighborhood Crime Rate is Associated with Retention in Care Among HIV-Positive Patients. *AIDS Behav* 2018 Sep;22(9):3003-3008. [doi: [10.1007/s10461-018-2094-5](https://doi.org/10.1007/s10461-018-2094-5)] [Medline: [29600423](https://pubmed.ncbi.nlm.nih.gov/29600423/)]
35. Eberhart MG, Voytek CD, Hillier A, Metzger DS, Blank MB, Brady KA. Travel distance to HIV medical care: a geographic analysis of weighted survey data from the Medical Monitoring Project in Philadelphia, PA. *AIDS Behav* 2014 Apr;18(4):776-782. [doi: [10.1007/s10461-013-0597-7](https://doi.org/10.1007/s10461-013-0597-7)] [Medline: [24141487](https://pubmed.ncbi.nlm.nih.gov/24141487/)]
36. Krakower DS, Mayer KH. The role of healthcare providers in the roll out of preexposure prophylaxis. *Curr Opin HIV AIDS* 2016 Jan;11(1):41-48 [FREE Full text] [doi: [10.1097/COH.0000000000000206](https://doi.org/10.1097/COH.0000000000000206)] [Medline: [26417953](https://pubmed.ncbi.nlm.nih.gov/26417953/)]
37. Petroll AE, Walsh JL, Owczarzak JL, McAuliffe TL, Bogart LM, Kelly JA. PrEP Awareness, Familiarity, Comfort, and Prescribing Experience among US Primary Care Providers and HIV Specialists. *AIDS Behav* 2017 May;21(5):1256-1267. [doi: [10.1007/s10461-016-1625-1](https://doi.org/10.1007/s10461-016-1625-1)] [Medline: [27885552](https://pubmed.ncbi.nlm.nih.gov/27885552/)]
38. Calabrese SK, Magnus M, Mayer KH, Krakower DS, Eldahan AI, Hawkins LAG, et al. Support your client at the space they're in: HIV Pre-Exposure Prophylaxis (PrEP) Prescribers' Perspectives on PrEP-Related Risk Compensation. *AIDS Patient Care STDS* 2017 Apr;31(4):196-204 [FREE Full text] [doi: [10.1089/apc.2017.0002](https://doi.org/10.1089/apc.2017.0002)] [Medline: [28414261](https://pubmed.ncbi.nlm.nih.gov/28414261/)]

Abbreviations

APR: adjusted prevalence ratios

CHC: community health center

GED: General Education Development

PrEP: pre-exposure prophylaxis

Edited by G Eysenbach; submitted 03.10.18; peer-reviewed by L Dean; comments to author 25.10.18; revised version received 08.11.18; accepted 08.11.18; published 04.02.19.

Please cite as:

Ojikutu BO, Bogart LM, Mayer KH, Stopka TJ, Sullivan PS, Ransome Y

Spatial Access and Willingness to Use Pre-Exposure Prophylaxis Among Black/African American Individuals in the United States: Cross-Sectional Survey

JMIR Public Health Surveill 2019;5(1):e12405

URL: <http://publichealth.jmir.org/2019/1/e12405/>

doi: [10.2196/12405](https://doi.org/10.2196/12405)

PMID: [30714945](https://pubmed.ncbi.nlm.nih.gov/30714945/)

©Bisola O Ojikutu, Laura M Bogart, Kenneth H Mayer, Thomas J Stopka, Patrick S Sullivan, Yusuf Ransome. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 04.02.2019. 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

Fraud Detection Protocol for Web-Based Research Among Men Who Have Sex With Men: Development and Descriptive Evaluation

April M Ballard^{1,2}, MPH; Trey Cardwell¹, MPH; April M Young^{1,3}, PhD, MPH

¹Department of Epidemiology, College of Public Health, University of Kentucky, Lexington, KY, United States

²Department of Environmental Health, Emory University, Atlanta, GA, United States

³Center on Drug and Alcohol Research, University of Kentucky, Lexington, KY, United States

Corresponding Author:

April M Young, PhD, MPH

Department of Epidemiology

College of Public Health

University of Kentucky

111 Washington Avenue

Lexington, KY, 40536

United States

Phone: 1 859 218 2090

Email: april.young@uky.edu

Abstract

Background: Internet is becoming an increasingly common tool for survey research, particularly among “hidden” or vulnerable populations, such as men who have sex with men (MSM). Web-based research has many advantages for participants and researchers, but fraud can present a significant threat to data integrity.

Objective: The purpose of this analysis was to evaluate fraud detection strategies in a Web-based survey of young MSM and describe new protocols to improve fraud detection in Web-based survey research.

Methods: This study involved a cross-sectional Web-based survey that examined individual- and network-level risk factors for HIV transmission and substance use among young MSM residing in 15 counties in Central Kentucky. Each survey entry, which was at least 50% complete, was evaluated by the study staff for fraud using an algorithm involving 8 criteria based on a combination of geolocation data, survey data, and personal information. Entries were classified as fraudulent, potentially fraudulent, or valid. Descriptive analyses were performed to describe each fraud detection criterion among entries.

Results: Of the 414 survey entries, the final categorization resulted in 119 (28.7%) entries identified as fraud, 42 (10.1%) as potential fraud, and 253 (61.1%) as valid. Geolocation outside of the study area (164/414, 39.6%) was the most frequently violated criterion. However, 33.3% (82/246) of the entries that had ineligible geolocations belonged to participants who were in eligible locations (as verified by their request to mail payment to an address within the study area or participation at a local event). The second most frequently violated criterion was an invalid phone number (94/414, 22.7%), followed by mismatching names within an entry (43/414, 10.4%) and unusual email addresses (37/414, 8.9%). Less than 5% (18/414) of the entries had some combination of personal information items matching that of a previous entry.

Conclusions: This study suggests that researchers conducting Web-based surveys of MSM should be vigilant about the potential for fraud. Researchers should have a fraud detection algorithm in place prior to data collection and should not rely on the Internet Protocol (IP) address or geolocation alone, but should rather use a combination of indicators.

(*JMIR Public Health Surveill* 2019;5(1):e12344) doi:[10.2196/12344](https://doi.org/10.2196/12344)

KEYWORDS

fraud; HIV; mobile phone; MSM; Web-based research; questionnaires; technology; validity; Web-based methodology

Introduction

The internet is becoming an increasingly common tool for survey research [1-8]. Online instruments present opportunities

to recruit “hidden” or vulnerable populations that have previously been difficult to reach, including men who have sex with men (MSM), people who use drugs, and transgender people [2,9-13]. Internet-based research allows for anonymity and

decreases barriers for participants. The confidentiality and distance from the researcher provided with Web-based data collection creates a space for participants, especially those who are part of a stigmatized population, to provide information honestly and with less apprehension [2,4,9,14,15]. Furthermore, Web-based methods may be more appropriate and effective in reaching communities that are highly active on the Web, such as MSM who use social and sexual networking websites and dating and sexual mobile phone apps [2]. Internet as a data collection tool also allows researchers to reach geographically dispersed populations [4,9] and minimizes cost to researchers, providing an effective, cost-efficient method to recruit and collect data from hard-to-reach populations [2,3,7,8,14-16].

However, the anonymity of internet-based research that serves as an advantage for data collection also poses challenges that portend data quality and validity. Surveying participants in person allows investigators to better enforce inclusion criteria and prevent duplicate enrollment [2,3,7,9,14,15,17], while the process of avoiding fraudulent enrollment in Web-based research is more difficult. Previous research has focused on 2 different types of invalid entries as follows: (1) ineligible participants who misrepresent themselves to fit eligibility criteria to profit from compensation (ie, misrepresentation fraud) and (2) eligible participants who participate more than once without malefic intent or to receive additional compensation (ie, duplicate fraud) [2,4,7,9,13-19]. The confidentiality and anonymity of Web-based research make it difficult to prevent such entries, posing a threat to data integrity if appropriate data quality protocols are not in place [4,7,13-15,17-20]. Bots and smart software that have the capability to produce human-like data with unique Internet Protocol (IP) addresses can also lead to invalid or fraudulent data [21]. Computer program (ie, bots) and human fraud must both, therefore, be considered in fraud identification strategies. As others have noted [1,2], increasing use of Web-based methodology warrants the development of standards and expectations for Web-based research to ensure data validity and accurate result reporting.

Previous research has focused on the use of a handful of characteristics that can be used to identify fraud (eg, IP address; personal information such as name, address, and email address; response patterns; and timestamps) [2,3,7,9,13-15,17,18,20,22]. However, technology is rapidly advancing, as is the savviness of its users, and approaches previously used to prevent and detect fraud may now lack utility or need to be supplemented with other techniques [2]. For example, the development and increasing use of network devices and smartphones have made it more difficult to utilize IP addresses as they can be masked both intentionally and unintentionally by users [23-26]. The purpose of this analysis was to evaluate fraud detection strategies in a Web-based survey of young MSM and describe new protocols to improve fraud detection in Web-based survey research.

Methods

Study Design

The Men's Health Study was a cross-sectional survey that examined individual- and network-level risk factors for HIV

transmission and substance use among young MSM residing in 15 counties in Central Kentucky (total population size, 743,119) [27]. The eligibility criteria included being aged 18-34 years, being biologically male, having engaged in anal sex with another man during the past 6 months, and residing in Central Kentucky. Qualtrics [28], a Web-based survey service, was used to create a Web-based eligibility screening assessment, consent form, and behavioral and demographic survey. A link to the Qualtrics survey was posted on the study website hosted by WordPress [29], which also described the monetary incentive, the process for completing the screening assessment, consent form, and survey, and provided community resources and contact information for the research staff. Based on pilot-testing among the staff, the screening, consent, and survey were anticipated to require a maximum of 45 minutes to complete, and the survey was anticipated to take between 5 and 30 minutes to complete depending on responses and subsequent skip patterns.

Participants were recruited from February to July 2018 utilizing flyers containing the Web address and a quick response code for the study's website. The flyers were posted in local lesbian, gay, bisexual, transgender, and queer (LGBTQ) venues (eg, bars, adult entertainment stores, and health clinics) and on social media via a study-specific page and young adult LGBTQ groups. Staff also set up booths at 2 local LGBTQ events to disseminate flyers; at one event, the booth allowed space for participants to take the survey and be reimbursed on site. Recruitment also involved peer referral. During the survey and informed consent form, participants were informed that they were allowed to refer up to 3 peers to the study and receive US \$10 per eligible referral who completed the survey. Upon completion of the survey, participants received a message from the staff that contained the survey URL and a referral code number and reminded them that they could earn up to US \$30 for peer referrals; the former allowed the staff to link participants to people they refer. Participants could choose to receive this message from the staff via a short message service (SMS) text message or email, and provided their phone number or email address accordingly.

Participants who completed the survey had the option of declining an incentive or receiving a US \$25 e-gift card or mailed payment and provided details either on their email address or mailing address. Participants were informed that the same method of delivery chosen for their US \$25 survey incentive would be used for their peer referral incentives. Those who were recruited at local LGBTQ events were also given the option to receive their US \$25 incentive in-person as payment and, then, were asked if they would like to receive their peer referral incentives as an e-gift card or mailed payment. All procedures were approved by the University of Kentucky's Institutional Review Board.

Fraud Prevention Strategies

Prior to launching the Web-based survey, 4 strategies were implemented to prevent fraud. First, the informed consent form stated that participants should not take the survey more than once, and if this occurred, incentives would not be received more than once. Second, participants could not enter their email or mailing address to receive compensation until the very end of the survey, making the process to receive incentives more

time-consuming in an effort to deter fraud. Third, to prevent participants from taking the survey more than once, either as a duplicate entry or to try to determine the eligibility criteria and subsequently misrepresent themselves to meet those criteria, the “prevent ballot box stuffing” option in Qualtrics was activated. This option places a cookie on the browser when a response is submitted so that if the survey link is clicked on again from the same device and browser, Qualtrics would detect the cookie and prevent the person from entering the survey [30]. Fourth, an option was also activated in Qualtrics to prevent indexing, which blocks search engines from finding the survey and presenting it in search results [30], thereby reducing the likelihood of fraudulent or inapposite participation. Of note, validation using GeoIP location, an estimate of a person’s location based on the IP address, was not used to block survey access because GeoIP location is not always accurate. Smartphones, remote access tools (eg, virtual private network), and network address translators (NATs) used by organizations or companies that assign public addresses to a computer inside a private network for security purposes can result in less accurate or concealed GeoIP estimates [23–26]. In addition, the survey did not use a tool like a Completely Automated Public Turing Test to tell Computers and Humans Apart, which could have reduced fraud by requiring respondents to verify that they were not robots [30]. However, previous research has demonstrated that Completely Automated Public Turing Test to tell Computers and Humans Apart tool verification is not failproof and can be passed despite an entry being invalid or fraud [21].

Fraud Criteria and Detection Strategies

Each survey that was at least 50% complete was manually checked by the study staff utilizing an extensive protocol (Figure 1) to detect fraud based on 8 criteria (described below). A point system was used with each criterion having an assigned point value of 1 or 2, and the total was used to classify surveys as fraudulent, potentially fraudulent, or valid. Data characteristics that were considered to possibly indicate fraud but could also have a reasonable justification were assigned 1 point so that they could be flagged as suspicious and confirmed through correspondence with participants. Data characteristics that were considered to be strongly and independently indicative of fraud were assigned 2 points. A 2-point threshold was, therefore, used for classification as fraud. The measures listed below were used in the fraud detection algorithm.

Geolocation

The survey software logged approximate latitude and longitude locations for participants by comparing IP addresses to a location database [23,24,30]. These data are typically accurate to the city level within the United States. Outside of the United States, data are only accurate at the country level [30].

Geolocation outside of the study area was valuable for assisting with fraud detection but could not be used as a stand-alone fraud detection mechanism because only an external IP address is displayed when a device is connected to a virtual private network or NAT. This causes all devices to have identical IP addresses and geolocation [24,26]. Additionally, research staff

found that surveys they knew to be local and valid (ie, those completed at local LGBTQ events) would frequently have geolocations outside of the study area, sometimes even outside of the state of Kentucky. Surveys at LGBTQ events were completed on smartphones, which can display different IP addresses within minutes because of network proxies within the carrier’s network. This results in inaccurate geolocation based on IP addresses [25].

Geolocation based on IP addresses was, therefore, not used in isolation as a fraud detection mechanism but was considered in the context of other indicators of location. In-person recruitment events and mailing address for those who requested payment incentives allowed the study staff to verify the location of participants. If the survey was completed at an in-person event or if the incentive was requested to be mailed to an address in the study area, the IP address geolocation outside of the study area was not used as an indicator of fraud. However, if the IP address geolocation was outside of the study area, the incentive requested was an e-gift card, and the survey was not completed in-person at a recruitment event, then the survey received 1 point toward the fraud detection point total.

Phone Number

Participants were asked to provide telephone numbers if they consented to be contacted for future research and as a method to receive a referral code. The staff searched phone numbers on the Web using Whitepages.org to identify instances in which the numbers were invalid or corresponded with local businesses or organizations, in which case, the survey received 1 point [15]. Phone numbers were kept separate from all survey data to ensure participants’ privacy and anonymity.

Names

When the survey was initially launched, participants were asked to give their name only if they consented to be contacted for future research or they requested payment through the mail for their incentive. To be able to investigate fraudulent behavior among e-gift card recipients, researchers began requesting participants name when e-gift cards were selected as an incentive. In addition, names were sometimes contained in or associated with email addresses (eg, john.doe@email.com) entered to receive e-gift card incentive, referral code, or community resources. Therefore, names were given up to 3 times—for incentive delivery, through names linked to or contained in email addresses, and in the consent to future research section. This allowed researchers to cross-reference names up to 3 times. If names did not match within a survey, participants were assigned 1 point toward fraud.

Email Address

Email addresses were examined to detect potential fraud. As done in a previous study [21], addresses with alternating letters and numbers (eg, a12bcd34e@email.com) were suspected by researchers to be fake email accounts (ie, created by a Bot program or by a human trying to misrepresent themselves). Surveys associated with such email addresses received 1 point toward fraud.

Figure 1. Web-based survey fraud detection algorithm.

Each entry was initially reviewed using the following criteria and assigned a fraud score (points given in brackets):

- a) [1 point] Was the IP address location outside of the study area AND unable to be verified through other means (i.e., a mailing address for mailed incentives OR completion in-person at a study recruitment event)? (n=164)
- b) [1 point] Is the provided phone number a business or organization number rather than personal number? (n=94)
- c) [1 point] Are names within the entry (i.e., in consent form, email address, incentive mailing address) different? (n=43)
- d) [1 point] Is the email address provided unusual (ex. a12bcd34e@email.com)? (n=37)
- e) [2 points] Does the first and last name AND one or more other personal information items (i.e., phone number, email address, and/or physical address) match previous entries'? (n=13)
- f) [1 point] Do two or more of the following match previous entries' information: phone number, email address, and/or physical address? (n=6)
- g) [2 points] Does the date of birth AND one or more personal information items (i.e., phone number, email address, and/or physical address) match that of previous entries'? (n=5)
- h) [1 point] Was the duration of the survey less than 5 minutes? (n=3)

All entries that received a score of 1 point were categorized as potentially fraudulent. All entries that received a score of 2 points or more were categorized as fraudulent. All entries that received a score of 0 points were categorized as valid.

Personal Information

Personal information, including name, phone number, date of birth, email address, and mailing address, was compared across surveys to detect fraud. If 2 personal information items (ie, phone number, email address, and mailing address) matched a previously completed survey, the survey received 1 point toward their fraud total. Those with matching information but different names and date of birth could represent phone sharing and cohabitation, so surveys that contained contact information that had been entered previously but different names and dates of birth were contacted by the study staff to verify their identity as a unique participant (further described in *Fraud Categorization* below). Those with 1 or more matching personal information items and the same first and last name as a previous survey received 2 points as this was seen as a definite sign fraud. Additionally, surveys with, at least, one matching personal information item and the same date of birth as a previous survey received 2 points.

Survey Duration

Qualtrics collected timestamps for the screening, consent form, and survey sections. For each section, Qualtrics recorded the time each section began and ended. Of note, there was no option to save survey progress, preventing completion across multiple sessions. If the survey was completed in <5 minutes, the entry received 1 point.

Fraud Categorization

Figure 2 displays data on survey entries and fraud categorization. Surveys were classified as *fraudulent* if the total entry point value added to ≥ 2 . Depending on what contact information was provided, fraudulent surveys received a phone call, SMS text message, or email stating, "You recently completed a survey for a health study online. However, we detected that your survey entry was fraudulent. If you think this is a mistake, please contact us", with the intent of deterring further fraudulent behavior. If participants did not respond, their survey was considered invalid and they did not receive compensation. No responses were received from those categorized as fraudulent.

Surveys were classified as *potential fraud* if the total point value added to 1. Potentially fraudulent participants received a phone call, SMS text message, or an email stating, "Thank you for completing the UK Health Study online survey. We have been experiencing fraud in the study and your survey entry has been flagged as suspicious. We sincerely apologize for the inconvenience if this was an error. Please contact us by calling xxx-xxx-xxxx to confirm that you did indeed complete a survey and we will send your \$25 incentive." This message was meant to deter those who were indeed fraudulent but also provide an opportunity for valid submissions to verify legitimacy. They were asked to confirm personal information such as date of birth, mailing address, or email address to validate their survey. If a participant responded and was unable to provide verification,

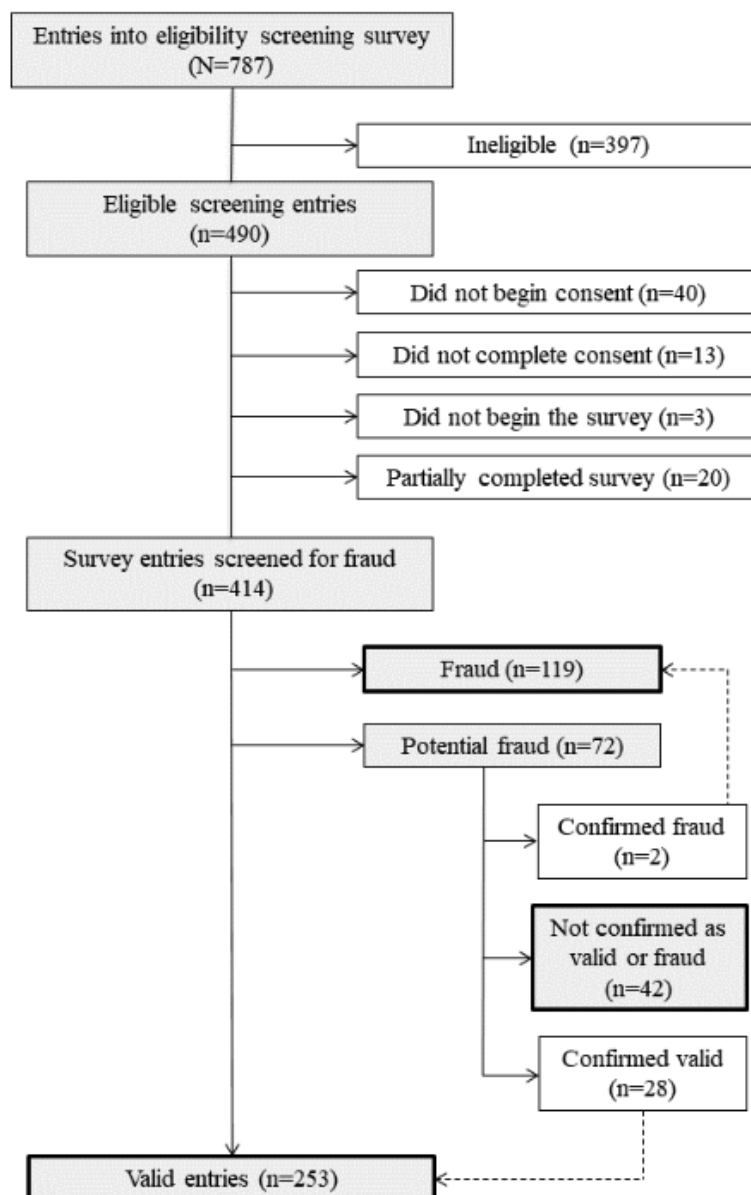
the survey was categorized as fraudulent. If a participant did not respond, the survey remained as potentially fraudulent.

Surveys were classified as *valid* if they met none of the fraud detection criteria or if they were originally classified as potentially fraudulent but contacted study staff to confirm the validity of their survey entry.

Analyses

Descriptive analyses were performed using SAS software, version 9.4 (SAS Institute; Cary, NC, USA) to describe the number and percentage who violated each fraud detection criterion among those that were classified as fraud, potential fraud, and valid.

Figure 2. Flowchart of survey entries and fraud categorization.



Results

Overall, 62.3% (490/787) of the completed screening questionnaires were eligible. Of 490 who were eligible, 437 (89.2%) completed the consent and 408 (51.2%) completed the survey with an average survey duration of 19 minutes and 41 seconds. In addition, there were 26 partial surveys, of which 18 were missing several measures used for fraud detection and a majority of survey responses and, therefore, were not included in analyses presented below or in the final set of valid responses. Six partial surveys were >50% complete and were included in

analyses, making 414 the final number of surveys in need of evaluation for fraud.

Of 414 surveys, 117 (28.3%) were initially categorized as fraud, 72 (17.4%) were categorized as potential fraud, and 225 (54.3%) were considered valid. Of 72 surveys originally classified as potential fraud, 2 participants responded and admitted to living outside of the study area, 42 did not respond, and 28 were able to verify their survey and were reclassified as valid (see Figure 2). Therefore, of the 414 surveys, the final categorization resulted in 119 (28.7%) surveys identified as fraud, 42 (10.1%) as potential fraud, and 253 (61.1%) as valid for the final sample

and analyses (see Table 1). Of the 164 who had geolocation outside of the study area and were not mailed payment or recruited at a local event, 23 (14.0%) had locations outside of

the United States, 118 (72.0%) in states other than Kentucky, and 23 (14.0%) within Kentucky but outside of the 15-county study area (Table 1).

Table 1. The number of surveys that violated the fraud detection measure by category.

Fraud detection measure	Total (n=414), n (%)	Missing, n (%)	Fraud, (n=119) ^a , n (%)	Potential fraud, (n=42) ^a , n (%)	Valid (n=253) ^a , n (%)
Geolocation outside of study area based on IP ^b address and unable to be confirmed through mailing address or participating at a local event	164 (39.6)	0 (0.0)	109 (91.6)	34 (81)	21 (8)
Phone number was a local business or organization phone number	94 (22.7)	93 (22.5) ^c	94 (79.0)	0 (0.0)	0 (0.0)
Mismatching names within entry	43 (10.4)	71 (17.1) ^c	37 (31.1)	2 (4.8)	4 (1.6)
Unusual email address	37 (8.9)	98 (23.7) ^c	34 (28.6)	1 (2.4)	2 (0.8)
First and last name AND one or more other personal items match other previous entry	13 (3.1)	16 (3.9) ^c	13 (10.9)	0 (0.0)	0 (0.0)
Two or more personal items match other previous entry	6 (1.4)	67 (16.2) ^c	3 (2.5)	3 (7.1)	0 (0.0)
Date of birth AND one or more other personal items match other previous entry	5 (1.2)	5 (1.2) ^c	5 (4.2)	0 (0.0)	0 (0.0)
Survey duration <5 minutes	3 (0.7)	0 (0.0)	0 (0.0)	2 (4.8)	1 (0.4)

^aSample size represents final categorizations (ie, reclassification of some that were initially classified as potential fraud as valid or fraudulent based on the verification with participants).

^bIP: internet protocol.

^cParticipants were not required to give personal information (ie, phone number, name, email address) if they did not want to be contacted about future research opportunities, declined their incentive, or if they did not complete the section of the survey describing the referral process. If a personal item was missing needed for a measure, the variable was considered missing.

Figure 1 lists each of the 8 fraud detection criterion used with the assigned point value and the number of surveys that fit that criterion. Surveys had the potential to violate 8 criteria that created a point scale from 0 to 10 points. The actually observed score range was 0-5 points. Data showed that 59.4% (246/414) of all entries had geolocation outside of the study area. However, 33.3% (82/246) of the surveys that had ineligible geolocations belonged to participants who *were* in eligible locations (as verified by their request to mail payment to an address within the study area or participation at a local event). The second most frequently violated criterion was an invalid phone number (94/414, 22.7%), followed by mismatching names within a survey (43/414, 10.4%) and unusual email addresses (37/414, 8.9%). Less than 5% (18/414) had some combination of personal information items matching that of a previous survey (ie, criteria e, f, and g in Table 1). Of the 414 surveys, 3 were below the minimum time threshold established by the study staff, completing it in <5 minutes.

Of note, many surveys violated more than one criterion. Of the 164 entries that violated the geolocation criterion, 94 (57.3%) also had invalid phone numbers, 35 (21.3%) had mismatching names within the survey, and 34 (20.7%) had unusual email addresses. Of the 94 surveys with invalid phone numbers, 29 (30.9%) had mismatching names within the survey and 28 (29.7%) had unusual email addresses. In total, 42.3% (175/414) of surveys violated some combination of the geolocation, phone number, mismatching names, or the email address criteria. The detection measure of matching name and personal identifier

data were not independently useful in detecting fraud. A total of 10.9% (13/119) surveys had first and last names and, at least, one other personal items that matched that of previous entries, but all of these also violated some other fraud criteria. Similarly, all surveys that had a date of birth (5/119) that matched a previous survey's date of birth also violated some other fraud criteria.

Of 119 surveys classified as fraudulent, 109 (91.6%) had geolocations outside the study area, 94 (79.0%) provided business or organization phone numbers rather than personal phone numbers, 37 (31.1%) provided mismatching names within the surveys, and 34 (28.6%) had unusual email addresses. In total, 93.3% (111/119) of fraudulent surveys violated at least one of the following criteria: geolocation, phone number, mismatching names, or the email address; 86.6% (103/119) violated at least 2 of those 4 criteria.

IP addresses were not used in the fraud detection algorithm because analyses demonstrated that they would have limited utility. Overall, 23.0% (95/414) of the surveys had an IP address identically matching that of at least one other survey entry; however, 13.7% (13/95) of those with matching IP addresses were surveys completed at in-person local events and were observed by the staff as being unique individuals. Of the remaining 82 surveys, 47 (57.3%) were classified as valid surveys, 12 (14.6%) were potential fraud, and 23 (28.1%) were fraudulent.

Discussion

Principal Findings

In this Web-based study of young MSM, a robust fraud detection algorithm involving geolocation, phone number, email address, name, and other personal data revealed that 28.7% (119/414) of all eligible survey entries were fraudulent. The majority of fraudulent surveys involved participants whose IP address had geolocation outside of the study area, invalid phone numbers (ie, local business phone numbers), mismatching names within the survey, and unusual email addresses (ie, a12bcd34e@email.com). These findings are consistent with prior studies' conclusions that fraud can be pervasive in Web-based research [5,7,9,13-22], but extends previous studies by outlining a step-by-step fraud detection strategy that does not rely on IP address and geolocation, which were identified in this study to have substantial limitations.

Previous studies have used IP address to detect duplicate entries [2,9,13-15,17,19] but this was not used in this study because (1) Qualtrics' ballot box stuffing prevention feature was used to prevent multiple entries from the same IP addresses and (2) matching IP addresses may not necessarily indicate fraud but occur in instances where >1 eligible participant lives in the same household, shares devices, or is accessing the survey from a communal space (eg, library or other venues with public Wi-Fi access). Furthermore, IP address to identify duplicate entries was not found to be appropriate based on the increasing use of smartphones, remote access tools, and NATs [23-26]. Remote access tools and NATs present the same IP address across users [26] and present challenges to researchers in differentiating between users. On the other hand, smartphones can present different IP addresses within minutes based on the mobile tower the signal is being routed through [26]; this allows individuals to intentionally or unintentionally obscure their IP address and geolocation, therein limiting the usefulness of the measure. In addition, the increasing threat of bots and smart software that produce human-like data with unique IP addresses further limit this measure as a tool to detect fraud [21]. In this study, the majority of repeat IP addresses were valid entries, suggesting that IP address should be used with caution as a measure for fraud detection.

Geolocation linked to IP address to assist in assessing geographic eligibility (ie, residence in the 15-county study area) was used, but in combination with other indicators of fraud because of the inaccuracy of geolocation data that can be introduced by smartphones, remote access tools, NATs, bots, and smart software [21,23-26]. One out of 3 surveys that had ineligible geolocations based on the IP address belonged to participants who *were* in eligible locations, as verified by their request to mail payment to an address within the study area or by the staff observing their participation at a local event. Thus, like IP address, geolocation used for fraud detection should not be used in isolation.

Other measures that were found to be useful in fraud detection were invalid phone numbers, mismatching names within a survey, and unusual email addresses. Invalid phone numbers included a variety of local businesses and organizations,

including manufacturing companies, health care organizations, professional associations, and university offices. Surprisingly, names at different parts of the survey (ie, consent form, contact information, and survey incentive information) were entered inconsistently. Distinct email address patterns that emerged among fraudulent surveys included addresses that began with a numeric value, switched between numbers and letters throughout the address, at least, 3 times, and did not appear to include initials or a name. Unexpectedly, the detection measure of matching name and personal identifier data across surveys, which would indicate duplicate entries, was not useful in detecting unique cases of fraud; surveys that violated these criteria also violated other criteria. These findings, however, may only be representative of this sample and a holistic approach should be considered for monitoring and detecting fraud.

Respondents were contacted to verify entries that were suspicious but not blatantly fraudulent and utilized messages tailored to fraud category when recontacting respondents. Similar to previous research among MSM [21], none of the individuals who received the message tailored to fraudulent cases responded, possibly indicating that these were, indeed, fraud. Of those who received the message tailored to potentially fraudulent cases, 38.9% (28/72) responded with reasonable explanations such as using a partner's phone number or email address or providing a physical address verifying geographic eligibility and were reclassified as valid. Interestingly, 2 of the potentially fraudulent cases that were contacted verified their fraud, explaining that they were indeed outside of the study area and were confused about geographic eligibility criteria. In addition, real-time monitoring of fraud was helpful in deterring subsequent fraud. A majority of fraudulent entries were submitted within the first month of data collection, but the frequency dramatically decreased after staff began calling, SMS text messaging, and emailing targeted messages. Communicating directly with fraudulent participants may help deter continued invalid entries from being submitted. Future studies should consider monitoring fraud in real time and verifying suspicious surveys with respondents if possible.

Limitations

While the fraud detection algorithm used in the study was robust, there were limitations. For example, participants' name and phone numbers could be left blank and were left blank by 3.9% (16/414) and 22.5% (93/414) of surveys, respectively. Other data were available to assist in fraud detection in these cases (ie, geolocation, email address, physical address, date of birth, or survey duration); however, those who did not provide name and phone information had a greater likelihood of not being detected as fraudulent. In addition, virtual phone systems that allow for the quick creation of phone numbers may limit the ability to use phone numbers as a fraud detection algorithm, as respondents can create local numbers and enter them in the survey. Furthermore, certain email address providers display names associated with email addresses when messages are sent while others do not; those that display names allowed the staff to better detect name inconsistency within surveys and fraud.

Conclusions

This study suggests that researchers conducting Web-based surveys of MSM should be vigilant about the potential for fraud. Researchers should have a fraud detection algorithm in place prior to data collection to ensure that (1) data needed for fraud detection are being collected; (2) the informed consent document can describe that surveys will be evaluated for fraud and what the consequences are for incentives; and (3) fraud can be monitored in real time. The latter allows the staff to send messages to those flagged as fraudulent and potentially fraudulent and deter subsequent fraud and avoid misclassification. Importantly, in research involving populations engaged in stigmatized or illegal behavior, researchers should take extra precautions to ensure that messages sent to respondents do not disclose the focus of the study or eligibility criteria in case the message is intercepted by an unintended recipient. In addition, as discussed by Sullivan et al [2], researchers should be mindful that rapid evolutions in technology could impact the utility and relevance of previously published protocols and methodologies. Therefore, continued innovation in fraud prevention and detection for Web-based research will be necessary, as will the development of automated or semiautomated algorithms for detecting and mitigating fraud [2,5,21]. The latter is especially important given that manual approaches, such as the one used in this study, can be labor- and time-intensive and may be difficult to implement in studies

with large samples sizes. Of note, manual approaches may continue to be important for small Web-based studies on novel topics [21].

In research involving geographic eligibility criteria (ie, residence in a certain area), automatic algorithms embedded in Web-based survey tools that exclude people based on the geolocation should be used with caution, as geolocations are often inaccurate. Furthermore, in studies with geographic inclusion criteria, the entry of a phone number with a local area code should not be assumed to be a local person, as the entry of local business or organization numbers can be common. Providing the option of receiving an incentive through mailed payment rather than only by e-gift card proved to be a useful tool in fraud detection, not only for the evaluation of whether the mailing address was within the study area for those who opted for mailed payments but also because the staff could deter fraud by contacting respondents who violated multiple fraud criteria and ask for a mailing address for the incentive rather than automatically sending an e-gift card. In addition, the collection of name data in multiple places throughout the survey, as well as email address, phone number, mailing address, and geolocation based on the IP address were, in combination, useful in assisting with fraud detection.

Future research should consider similar algorithms and should publish algorithms used for fraud detection to improve replicability and to benefit the field of online research.

Acknowledgments

This research was supported by the National Institute on Drug Abuse (NIH NIDA R03 DA039740). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res* 2004 Dec 29;6(3):e34 [[FREE Full text](#)] [doi: [10.2196/jmir.6.3.e34](#)] [Medline: [15471760](#)]
2. Sullivan PS, Grey JA, Simon Rosser BR. Emerging technologies for HIV prevention for MSM: what we have learned, and ways forward. *J Acquir Immune Defic Syndr* 2013 Jun 01;63 Suppl 1:S102-S107 [[FREE Full text](#)] [doi: [10.1097/QAI.0b013e3182949e85](#)] [Medline: [23673879](#)]
3. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. *J Med Internet Res* 2002;4(2):E13 [[FREE Full text](#)] [doi: [10.2196/jmir.4.2.e13](#)] [Medline: [12554560](#)]
4. Mustanski BS. Getting wired: Exploiting the internet for the collection of valid sexuality data. *Journal of Sex Research* 2001 Nov;38(4):292-301. [doi: [10.1080/00224490109552100](#)]
5. Pequegnat W, Rosser BRS, Bowen AM, Bull SS, DiClemente RJ, Bockting WO, et al. Conducting Internet-based HIV/STD prevention survey research: considerations in design and evaluation. *AIDS Behav* 2007 Jul;11(4):505-521. [doi: [10.1007/s10461-006-9172-9](#)] [Medline: [17053853](#)]
6. van Gelder MMHJ, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? *Am J Epidemiol* 2010 Dec 01;172(11):1292-1298. [doi: [10.1093/aje/kwq291](#)] [Medline: [20880962](#)]
7. Choi SH, Mitchell J, Lipkus I. Lessons Learned From an Online Study with Dual-smoker Couples. *Am J Hlth Behav* 2017 Jan 01;41(1):61-66. [doi: [10.5993/AJHB.41.1.6](#)]
8. Ekman A, Litton J. New times, new needs; e-epidemiology. *Eur J Epidemiol* 2007;22(5):285-292. [doi: [10.1007/s10654-007-9119-0](#)] [Medline: [17505896](#)]
9. Grey JA, Konstan J, Iantaffi A, Wilkerson JM, Galos D, Rosser BRS. An updated protocol to detect invalid entries in an online survey of men who have sex with men (MSM): how do valid and invalid submissions compare? *AIDS Behav* 2015 Oct;19(10):1928-1937 [[FREE Full text](#)] [doi: [10.1007/s10461-015-1033-y](#)] [Medline: [25805443](#)]

10. Theunissen K, Hoebe C, Kok G, Crutzen R, Kara-Zaitri C, de Vries N, et al. A Web-Based Respondent Driven Sampling Pilot Targeting Young People at Risk for Chlamydia Trachomatis in Social and Sexual Networks with Testing: A Use Evaluation. *Int J Environ Res Public Health* 2015 Aug 20;12(8):9889-9906 [FREE Full text] [doi: [10.3390/ijerph120809889](https://doi.org/10.3390/ijerph120809889)] [Medline: [26308015](https://pubmed.ncbi.nlm.nih.gov/26308015/)]
11. Wilkerson JM, Shenk JE, Grey JA, Simon Rosser BR, Noor SW. Recruitment Strategies of Methamphetamine-Using Men Who Have Sex with Men into an Online Survey. *J Subst Use* 2015;20(1):33-37 [FREE Full text] [doi: [10.3109/14659891.2013.868936](https://doi.org/10.3109/14659891.2013.868936)] [Medline: [25642143](https://pubmed.ncbi.nlm.nih.gov/25642143/)]
12. Bauermeister JA, Zimmerman MA, Johns MM, Glowacki P, Stoddard S, Volz E. Innovative recruitment using online networks: lessons learned from an online study of alcohol and other drug use utilizing a web-based, respondent-driven sampling (webRDS) strategy. *J Stud Alcohol Drugs* 2012 Sep;73(5):834-838 [FREE Full text] [Medline: [22846248](https://pubmed.ncbi.nlm.nih.gov/22846248/)]
13. Miner MH, Bocking WO, Romine RS, Raman S. Conducting Internet Research With the Transgender Population: Reaching Broad Samples and Collecting Valid Data. *Soc Sci Comput Rev* 2012 May 01;30(2):202-211 [FREE Full text] [doi: [10.1177/0894439311404795](https://doi.org/10.1177/0894439311404795)] [Medline: [24031157](https://pubmed.ncbi.nlm.nih.gov/24031157/)]
14. Bauermeister JA, Pingel E, Zimmerman M, Couper M, Carballo-Diéguez A, Strecher VJ. Data Quality in web-based HIV/AIDS research: Handling Invalid and Suspicious Data. *Field methods* 2012 Aug 01;24(3):272-291 [FREE Full text] [doi: [10.1177/1525822X12443097](https://doi.org/10.1177/1525822X12443097)] [Medline: [23180978](https://pubmed.ncbi.nlm.nih.gov/23180978/)]
15. Bowen AM, Daniel CM, Williams ML, Baird GL. Identifying multiple submissions in Internet research: preserving data integrity. *AIDS Behav* 2008 Nov 1;12(6):964-973 [FREE Full text] [doi: [10.1007/s10461-007-9352-2](https://doi.org/10.1007/s10461-007-9352-2)] [Medline: [18240015](https://pubmed.ncbi.nlm.nih.gov/18240015/)]
16. Schmidt WC. World-Wide Web survey research: Benefits, potential problems, and solutions. *Behavior Research Methods, Instruments, & Computers* 1997 Jun;29(2):274-279. [doi: [10.3758/BF03204826](https://doi.org/10.3758/BF03204826)]
17. Quach S, Pereira JA, Russell ML, Wormsbecker AE, Ramsay H, Crowe L, et al. The good, bad, and ugly of online recruitment of parents for health-related focus groups: lessons learned. *J Med Internet Res* 2013 Nov 14;15(11):e250 [FREE Full text] [doi: [10.2196/jmir.2829](https://doi.org/10.2196/jmir.2829)] [Medline: [24231040](https://pubmed.ncbi.nlm.nih.gov/24231040/)]
18. Teitcher JEF, Bocking WO, Bauermeister JA, Hoefer CJ, Miner MH, Klitzman RL. Detecting, preventing, and responding to “fraudsters” in internet research: ethics and tradeoffs. *J Law Med Ethics* 2015;43(1):116-133 [FREE Full text] [doi: [10.1111/jlme.12200](https://doi.org/10.1111/jlme.12200)] [Medline: [25846043](https://pubmed.ncbi.nlm.nih.gov/25846043/)]
19. Kramer J, Rubin A, Coster W, Helmuth E, Hermos J, Rosenbloom D, et al. Strategies to address participant misrepresentation for eligibility in Web-based research. *Int J Methods Psychiatr Res* 2014 Mar;23(1):120-129 [FREE Full text] [doi: [10.1002/mpr.1415](https://doi.org/10.1002/mpr.1415)] [Medline: [24431134](https://pubmed.ncbi.nlm.nih.gov/24431134/)]
20. Konstan J, Rosser B, Ross M, Stanton J, Edwards W. The Story of Subject Naught: A Cautionary but Optimistic Tale of Internet Survey Research. *Journal of Computer-Mediated Communication* 2005 Jan;10(2):A. [doi: [10.1111/j.1083-6101.2005.tb00248.x](https://doi.org/10.1111/j.1083-6101.2005.tb00248.x)]
21. Dewitt J, Capistrant B, Kohli N, Rosser B, Mitteldorf D, Merengwa E, et al. Addressing Participant Validity in a Small Internet Health Survey (The Restore Study): Protocol and Recommendations for Survey Response Validation. *JMIR Res Protoc* 2018 Apr 24;7(4):e96 [FREE Full text] [doi: [10.2196/resprot.7655](https://doi.org/10.2196/resprot.7655)] [Medline: [29691203](https://pubmed.ncbi.nlm.nih.gov/29691203/)]
22. Nosen E, Woody S. Online Surveys: Effects of Research Design Decisions on Rates of Invalid Participation and Credibility. *Graduate Student Journal of Psychology* 2008;10:3-14.
23. Poese I, Uhlig S, Kaafar MA, Donnet B, Gueye B. IP geolocation databases. *SIGCOMM Comput. Commun. Rev* 2011 Apr 15;41(2):53. [doi: [10.1145/1971162.1971171](https://doi.org/10.1145/1971162.1971171)]
24. IP Addresses Beginners Guide. 2011 Mar. URL: <https://www.icann.org/en/system/files/files/ip-addresses-beginners-guide-04mar11-en.pdf> [WebCite Cache ID 72cAunLnS]
25. Balakrishnan M, Mohamed I, Ramasubramanian V. Where's that Phone?: Geolocating IP Addresses on 3G Networks. 2009 Presented at: Proceedings of the 9th ACM SIGCOMM conference on Internet measurement ACM Nov; November 04 - 06, 2009; Chicago, Illinois, USA p. 294-300. [doi: [10.1145/1644893.1644928](https://doi.org/10.1145/1644893.1644928)]
26. Zhang L. A retrospective view of network address translation. *IEEE Network* 2008 Sep;22(5):8-12. [doi: [10.1109/MNET.2008.4626226](https://doi.org/10.1109/MNET.2008.4626226)]
27. U.S. Census Bureau. 2012. 2010 Census of Population Housing, Population Housing Unit Counts, CPH-2-19, Kentucky URL: <https://www.census.gov/prod/cen2010/cph-2-19.pdf> [accessed 2018-09-22] [WebCite Cache ID 72cBBCjbc]
28. Qualtrics. Qualtrics URL: <https://www.qualtrics.com/> [accessed 2018-09-22] [WebCite Cache ID 72cBHUVmv]
29. Wordpress. Wordpress URL: <https://wordpress.com/> [accessed 2018-09-22] [WebCite Cache ID 72cBLQurN]
30. Snow J. Qualtrics. 2011. The Complete Research Suite: A step-by-step guide to using Qualtrics URL: https://www.qualtrics.com/docs/TheCompleteResearchSuite_print.pdf [accessed 2018-09-22] [WebCite Cache ID 72cBYDREe]

Abbreviations

MSM: men who have sex with men

IP: Internet Protocol

NAT: network address translator

LGBTQ: lesbian, gay, bisexual, transgender, and queer

SMS: short message service

Edited by G Eysenbach; submitted 27.09.18; peer-reviewed by BS Rosser, G Kok; comments to author 25.10.18; revised version received 02.11.18; accepted 05.11.18; published 04.02.19.

Please cite as:

Ballard AM, Cardwell T, Young AM

Fraud Detection Protocol for Web-Based Research Among Men Who Have Sex With Men: Development and Descriptive Evaluation
JMIR Public Health Surveill 2019;5(1):e12344

URL: <http://publichealth.jmir.org/2019/1/e12344/>

doi: [10.2196/12344](https://doi.org/10.2196/12344)

PMID: [30714944](https://pubmed.ncbi.nlm.nih.gov/30714944/)

©April M Ballard, Trey Cardwell, April M Young. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 04.02.2019. 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

Barriers to Implementation of Perinatal Death Audit in Maternity and Pediatric Hospitals in Jordan: Cross-Sectional Study

Yousef Khader^{1*}, ScD; Mohammad Alyahya², PhD; Anwar Batieha^{1*}, PhD

¹Department of Public Health and Community Medicine, Jordan University of Science and Technology, Irbid, Jordan

²Department of Health Management and Policy, Jordan University of Science and Technology, Irbid, Jordan

* these authors contributed equally

Corresponding Author:

Yousef Khader, ScD

Department of Public Health and Community Medicine

Jordan University of Science and Technology

Al-Ramtha-Amman Street

Irbid, 22110

Jordan

Phone: 962 0796802040

Email: yaskhader@just.edu.jo

Abstract

Background: Perinatal death audit is a feasible and cost-effective quality improvement tool that helps to improve the quality of health care and reduce perinatal deaths. Perinatal death audit is not implemented in almost all hospitals in Jordan.

Objective: This study aimed to assess health professionals' attitude toward perinatal death auditing and determine the main barriers for effective implementation of perinatal death auditing as perceived by health professionals in Jordanian hospitals.

Methods: A cross-sectional study was conducted among health professionals in 4 hospitals in Jordan. All physicians (pediatricians and obstetricians) and nurses working in these hospitals were invited to participate in the study. The study questionnaire assessed the attitude of health professionals toward perinatal death audit and assessed barriers for implementation of perinatal death audit in their hospitals.

Results: This study included a total of 84 physicians and 218 nurses working in the 4 selected maternity hospitals. Only 35% (29/84) of physicians and 36.2% (79/218) of nurses reported that perinatal death audit would help to improve the quality of prenatal health care services to a great or very great extent. Lack of time was the first-mentioned barrier for implementing perinatal death audit by both physicians (35/84, 42%) and nurses (80/218, 36.7%). Almost the same proportions of health professionals reported inadequate patient information being documented in hospital records as a barrier. Lack of a health information system was the third-mentioned barrier by health professionals. Fear of having conflicts with the family of the dead baby was reported by almost one-third of physicians and nurses. Only 28% (23/83) of physicians and 16.9% (36/213) of nurses reported that they would like to be involved in perinatal death audit in their health facilities.

Conclusions: Health professionals in Jordan had poor attitude toward perinatal death audit. The main barriers for implementing perinatal death audit in Jordanian hospitals were lack of time, inadequate patient information being documented in hospital records, and lack of health information systems.

(*JMIR Public Health Surveill* 2019;5(1):e11653) doi:[10.2196/11653](https://doi.org/10.2196/11653)

KEYWORDS

perinatal death; quality of health care; cause of death; Jordan

Introduction

Perinatal death audit is a feasible and cost-effective quality improvement tool that helps to improve the quality of health care and reduce perinatal deaths [1]. Perinatal death audits are implemented to generate accurate perinatal data, determine

medical and nonmedical causes of perinatal deaths, identify appropriate interventions to address these causes, and improve the quality of services. Countries that have implemented perinatal death audit have achieved a significant decrease in perinatal deaths. Studies have shown that perinatal death audit was associated with a 30% reduction in perinatal mortality [2].

Perinatal death audits have been used widely in high-income countries [3,4]. However, they are less frequently used in low- and middle-income countries where 98% of perinatal deaths take place [5]. Few studies have assessed the attitudes of health care providers about perinatal death audits, the challenges facing their implementation, and the barriers for implementing perinatal death audits. One study [6] showed that physicians have positive attitudes toward the death audit and reported that it is a good quality-of-care indicator in the hospital, which are valuable and necessary to improve the quality of health services. However, inadequate patient information in hospital records, lack of time for health care providers, high turnover of health professionals, fear of blame, lack of a national policy for perinatal audit, and lack of training were identified as important barriers for implementing death audit [7-9].

In the late 1990s, the neonatal mortality rate in Jordan fell from 19 to 15 per 1000 live births and remained relatively constant as Jordan transitioned into the new millennium [10]. Jordan is one of the many countries in the world that failed to achieve the Millennium Development Goal 4's target [11-13]. This is particularly because of the lack of effective planning and monitoring of health services. On the other hand, the United Nations International Children's Emergency Fund (UNICEF)-funded study "Perinatal and Neonatal (PNN) Mortality study in Jordan" [13] showed that a large proportion (74%) of neonatal deaths were preventable and only 37% of neonatal deaths received optimal health care. The study highlighted the need for strengthening the essential newborn care and improving the quality of maternal and neonatal health care. Other studies in Jordan showed that barriers to access maternal and neonatal care, low social status of women, poverty, and inequality were major determinants of PNN deaths especially in rural and remote areas in Jordan [14,15]. Moreover, the influx of Syrian refugees has a negative impact on economic, social, and health development and has stressed the country's health system [16].

Previous studies in Jordan strongly recommended establishing and implementing perinatal death audits in hospitals to improve the quality of services and decrease perinatal deaths. In Jordan, most hospitals have perinatal death review committees. However, these committees are not functional, and perinatal death audits have not yet been implemented in health facilities. To build an effective perinatal death audit, it is important to understand health professionals' attitude toward perinatal death audit and their perception of barriers and challenges for proper perinatal death audit implementation. Therefore, this study aimed to assess health professionals' attitude toward perinatal death auditing and determine the main barriers for effective implementation of perinatal death auditing as perceived by health professionals in pediatric hospitals in Jordan. Moreover, the study aimed to determine whether health professionals' attitude and their perception of the main barriers differ according to gender, profession, and years of experience.

Methods

Study Design

A cross-sectional study was conducted in 3 public hospitals and 1 teaching hospital, 3 hospitals from the north (Al-Mafraq Pediatrics Hospital, Princess Rahma Hospital, and King Abdullah University Hospital) and 1 from the south of Jordan (Al-Karak Hospital). Princess Rahma Pediatric Hospital (120 beds) and Al-Mafraq Pediatrics Hospital (108 beds) are the only pediatric referral public hospitals in Jordan. Out of the 2 teaching hospitals in Jordan, we selected King Abdullah University Hospital (683 beds), which is affiliated with Jordan University of Science and Technology (JUST) and serves approximately 1 million inhabitants in the north of Jordan. Al-Karak Hospital is the largest and the main hospital that provides pediatric services in the south of Jordan, with 125 beds. All physicians (pediatricians and obstetricians) and nurses working in these hospitals were invited to participate in the study, and those who agreed were interviewed using face-to-face structured interview by trained nurses. The study was approved by the institutional review board at JUST.

Study Questionnaire

The first part of the study questionnaire collected information about health professionals' age, gender, and years of experience. The second part of the questionnaire included 2 main questions to assess the perception of health professionals about perinatal death review: "To what extent a perinatal death review committee would help to improve the quality of prenatal healthcare services?" and "To what extent a perinatal death review committee would help to reduce perinatal deaths?" The possible responses for each question were "not at all," "to a small extent," "to some extent," "to a moderate extent," "to a great extent," and "to a very great extent." For the purpose of analysis, the responses "to a great extent" and "to a very great extent" were pooled together in 1 category to indicate great extent.

The third part of the questionnaire assessed the barriers for implementation of perinatal death audit in their hospitals. Health professionals were presented with a list of potential barriers that were identified from the relevant literature [17-29]. Moreover, they were asked to report any barrier that is not mentioned in the list.

Statistical Analysis

Data were analyzed using IBM SPSS version 20. Data were presented using percentages for categorical variables and means and SDs for continuous variables. The differences between proportions were tested using chi-square test. A *P* value of less than .05 was considered statistically significant.

Results

Participants' Characteristics

This study included a total of 84 physicians and 218 nurses working in the 4 selected maternity hospitals. More than half (44/84, 53%) of physicians were females, and almost all nurses were females. Their age ranged from 17 to 62 years, with a

mean (SD) of 31.1 (6.5) years. Their years of experience ranged from 1 to 35 years, with a mean (SD) of 7.6 (6.5) years and median of 5.0 years. All selected hospitals had a nonfunctional perinatal deaths review committee.

Attitude Toward Perinatal Death Audit

Only 35% (29/84) of physicians and 36.2% (79/218) of nurses reported that perinatal death audit would help to improve the quality of prenatal health care services to a great or very great extent (Table 1). Similarly, 32% (27/84) of physicians and 38.5% (84/218) of nurses stated that perinatal death audit would help to reduce perinatal deaths (Table 1). However, 12% (10/84)

of physicians and 21.6% (47/218) of nurses reported that perinatal death audit would not help to improve the quality of prenatal health care services, and an almost similar proportion reported that the perinatal death audit would not help to reduce perinatal deaths.

Table 2 shows the participants' responses on whether perinatal death audit would help to improve the quality of prenatal health care services and reduce perinatal deaths to a great or very great extent according to gender, profession, and years of experience. The attitude of health professionals toward perinatal death auditing did not differ significantly according to gender, profession, and years of experience.

Table 1. Health professionals' attitude toward perinatal death audits in maternity and pediatric hospitals in Jordan.

Attitude toward perinatal death audit	Health professionals		Total (N=302), n (%)
	Physicians (n=84), n (%)	Nurses (n=218), n (%)	
The extent to which a perinatal death audit would improve the quality of prenatal health care services			
Not at all	10 (12)	47 (21.6)	57 (18.9)
To a small extent	7 (8)	13 (6.0)	20 (6.6)
To some extent	12 (14)	29 (13.3)	41 (13.6)
To a moderate extent	26 (31)	50 (22.9)	76 (25.2)
To a great extent	20 (24)	64 (29.4)	84 (27.8)
To a very great extent	9 (11)	15 (6.9)	24 (7.9)
The extent to which a perinatal death audit would help to reduce perinatal deaths			
Not at all	10 (12)	40 (18.3)	50 (16.6)
To a small extent	16 (19)	13 (6.0)	29 (9.6)
To some extent	7 (8)	25 (11.5)	32 (10.6)
To a moderate extent	24 (29)	56 (25.7)	80 (26.5)
To a great extent	22 (26)	69 (31.7)	91 (30.1)
To a very great extent	5 (6)	15 (6.9)	20 (6.6)

Table 2. Participants' responses on whether perinatal death audit would help to improve the quality of prenatal health care services and reduce perinatal deaths to a great or very great extent according to gender, profession, and years of experience.

Variable	Perinatal death audit would help to improve the quality of prenatal health care services			Perinatal death audit would help to reduce perinatal deaths		
	Total	n (%)	P value	Total	n (%)	P value
Gender			.30			.38
Male	42	12 (29)		42	13 (31)	
Female	258	95 (36.8)		258	98 (38.0)	
Profession			.78			.30
Physician	84	29 (35)		84	27 (32)	
Nurse	218	79 (36.2)		218	84 (38.5)	
Years of experience			.72			.62
≤5	158	58 (36.7)		158	56 (35.4)	
>5	144	50 (34.7)		144	55 (38.2)	

The Main Barriers for Implementation of Perinatal Death Audit

The main barriers for implementing perinatal death audit in the hospitals were lack of time, inadequate patient information being documented in hospital records, lack of health information system, and fear of having problems with the family of the dead baby (Table 3). Lack of time was the first-mentioned barrier by both physicians (35/84, 42%) and nurses (80/218, 36.7%) for implementing perinatal death audit. Almost the same proportions of health professionals reported inadequate patient information being documented in hospital records as a barrier. Lack of a health information system was the third-mentioned barrier by health professionals. Fear of having conflicts with the family of the dead baby was reported by almost one-third of physicians and nurses. Fear of legal problems and the sensitivity between the concerned physicians and nurses were reported as barriers for effective implementation of perinatal death audit by 25.8% (78/302) and 23.2% (70/302) of health professionals, respectively. Having difficulties in ensuring confidentiality and not trained to conduct perinatal death were reported by almost

one-tenth of physicians and nurses. Health professionals' frequent turnover was the least-mentioned barrier. Physicians were significantly more likely to report "not trained to conduct perinatal death review" as a barrier compared with nurses (17/84, 20% vs 17/218, 7.8%; P value=.004). The participants' responses in regard to the perceived barriers for effective implementation of perinatal death audits in hospitals in Jordan did not differ significantly according to gender and years of experience.

Intention to Be Involved in Perinatal Death Audit

Only 28% (23/83) of physicians and 16.9% (36/213) of nurses reported that they would like to be involved in perinatal death audit in their health facilities. More than half of the physicians (48/83, 58%) and 63.8% (136/213) of nurses stated that they would probably like to be involved in perinatal death audit if it is implemented in their health facilities. The intention of health professionals to be involved in perinatal death audit did not differ significantly according to gender, profession, and years of experience (Table 4).

Table 3. The main barriers for effective implementation of perinatal death audits in hospitals in Jordan as perceived by health professionals.

Main barriers for effective implementation of perinatal death audits	Health professionals		Total (N=302), n (%)	P value
	Physicians (n=84), n (%)	Nurses (n=218), n (%)		
Lack of time	35 (42)	80 (36.7)	115 (38.1)	.48
Inadequate patient information being documented in hospital records	34 (41)	77 (35.3)	111 (36.8)	.45
Lack of health information system	33 (40)	75 (34.4)	108 (35.8)	.48
Fear of having problems with the family of the dead baby	29 (35)	74 (33.9)	103 (34.1)	.97
Fear of medico-legal problems	25 (30)	53 (24.3)	78 (25.8)	.41
Sensitivity between the concerned physicians and nurses	18 (21)	52 (23.9)	70 (23.2)	.76
Difficulty in ensuring confidentiality	16 (19)	24 (11.0)	40 (13.2)	.10
No need for the death review	7 (8)	30 (13.8)	37 (12.3)	.27
Not trained to conduct perinatal death review	17 (20)	17 (7.8)	34 (11.3)	.004
Health professionals' frequent turnover	5 (6)	8 (3.7)	13 (4.3)	.58

Table 4. The participants' responses to whether they would you like to be involved in perinatal death audit if it is implemented in their health facilities.

Variable	Would you like to be involved in perinatal death audit in your health facility?			P value
	Definitely yes, n (%)	Probably yes, n (%)	No, n (%)	
Gender				.07
Male	14 (33)	22 (52)	6 (14)	
Female	45 (17.9)	160 (63.5)	47 (18.7)	
Profession				.10
Physician	23 (28)	48 (58)	12 (14)	
Nurse	36 (16.9)	136 (63.8)	41 (19.2)	
Years of experience				.35
≤5	28 (17.9)	103 (66.0)	25 (16.0)	
>5	31 (22.1)	81 (57.9)	28 (20.0)	

Discussion

This study showed that only one-third of health professionals had reported that perinatal death audit would help to improve the quality of prenatal health care services to a great or very great extent. As perceived by health professionals, lack of time, inadequate patient information being documented in hospital records, and lack of a health information system were the first 3 mentioned barriers for implementing perinatal death audit. The attitude of health professionals toward perinatal death auditing and their intention to be involved in perinatal death audit did not differ significantly according to gender, profession, and years of experience. The perceived barriers for effective implementation of perinatal death audits in hospitals in Jordan did not differ significantly according to gender and years of experience.

The causes of perinatal deaths in Jordan need to be quickly addressed if the Sustainable Development Goals target is to be met. To increase the survival of babies, it is essential to identify the causes of perinatal deaths and their contributing factors and improve the quality of services. This can be achieved by effective implementation of perinatal death audit in Jordan hospitals. The audit process offers a chance to learn from critical situations in the management of maternity and neonatal care. Health care providers are urged to modify their care to better practice once there is detection about the poor practices that lead to these problems [17]. The capability to respond efficiently to recommendations acknowledged through audits is crucial to reducing deaths.

Over the past few decades, Jordan has made substantial progress in improving maternal, neonatal, and infant health. However, there are still challenges to achieving the third Sustainable Developmental Goal. Existing references indicate that the majority of perinatal deaths are preventable. Jordan is now ready for the next step toward eliminating preventable perinatal deaths. A vital component of any elimination strategy is a continuous surveillance system that not only tracks the number of deaths but also provides information about the underlying contributing factors and how they should be addressed. Stillbirths and neonatal deaths surveillance “J-SANDS” and auditing system is a model of such a system.

Although literature supports the fact that perinatal death audit strongly contributes to the avoidance of perinatal deaths, a relatively small proportion of health professionals (27/84, 32% of physicians and 84/218, 38.5% of nurses) stated that perinatal review audit would help to reduce perinatal deaths. This finding reflects the poor awareness of the value of perinatal death audit among health professionals. The main barriers to perinatal death audit implementation in our study included lack of time for health care providers, inadequate patient information in hospital records, lack of health information system, and fear of having problems with the family of the dead baby. Lack of time by health professionals was the first-mentioned barrier in our study. This barrier has been reported in other studies [18-20]. In a study conducted in Uganda, the majority of respondents reported that the main challenge to conduct death review was heavy workload with fewer staff [21]. To overcome this barrier,

perinatal death audit should be included in job descriptions of health professionals [19]. The management should also support perinatal death review as one of the health professionals’ duties and as a part of their daily work.

Inadequate patient information in hospital records was the second-mentioned barrier. This barrier was also reported as a barrier to completing audit successfully in many studies in Malawi, Tanzania, and Uganda [18,22-25]. Inadequate information hinders the ability of health professionals from assessing the causes of deaths.

Lack of an electronic health information system was the third-mentioned barrier. Many hospitals in Jordan do not have the capacity to process the limited available data to capture deaths, assign causes of deaths, and identify the avoidable factors. One study in Jordan showed that only 14% of neonatal deaths are registered and reported to the Department of Civil Registration because Jordan relies on paper-based systems to register and report births and deaths. None of the hospitals in Jordan report stillbirths. Lack of an electronic health information system and lack of a centralized database for compiling audit results makes data interpretation and identification of avoidable factors difficult to create actionable recommendations. Electronic health information system and centralized database for compiling audit, registering births and deaths, and assigning causes of deaths should be developed and implemented. Electronic platforms may pose an initial additional financial burden, although they may save time and money in the long term [26].

Fear of blame including loss of face among peers and potential legal ramifications have been shown to be important deterrents to conducting perinatal death audit in other studies [27]. To ensure successful implementation, having participants agree to a code of conduct for review meetings, establishing a no-blame environment, and ensuring confidentiality insofar as it is possible contribute to an environment where audit is more likely to be successful [27].

Health professionals’ frequent turnover was the least-mentioned barrier in our study. However, this was shown as an important barrier in other studies [28]. Consistent with other studies, not being trained on perinatal death audit was one of the mentioned barriers [29]. Unlike our studies, previous studies reported other barriers such as the lack of a national policy, strategy, and guidelines for perinatal audit [7]. However, the availability of a policy alone does not guarantee the success of the implementation.

One of the main limitations of this study is that the findings cannot be generalized to all hospitals of Jordan because 2 of the selected hospitals were pediatric hospitals, 1 was a teaching hospital, and the fourth hospital was a public hospital in the south of Jordan. Moreover, our findings are limited only to public and teaching hospitals as we did not include private hospitals. Another limitation is that the sample of health professionals is small to conduct subgroup analysis. In conclusion, health professionals in Jordan had a poor attitude toward perinatal death audit. The main barriers for implementing perinatal death audit in Jordanian hospitals were lack of time, inadequate patient information being documented in hospital

records, lack of a health information system, and fear of having problems with the family of the dead baby. Training activities are needed to increase the awareness of health professionals about the value of perinatal death in improving the quality of

services and perinatal deaths. An electronic health information system and centralized database for compiling audit, registering births and deaths, and assigning causes of deaths should be developed and implemented.

Acknowledgments

The authors would like to acknowledge the International Development Research Centre and UNICEF-Jordan for their support of the implementation research of establishing a perinatal surveillance system in Jordan.

Conflicts of Interest

None declared.

References

1. Papiernik E, Bucourt M, Zeitlin J. Audit of obstetrical practices and prevention of perinatal deaths. *Bull Acad Natl Med* 2005 Jan;189(1):71-84. [Medline: [16119881](#)]
2. Wilkinson D. Perinatal mortality--an intervention study. *S Afr Med J* 1991 May 4;79(9):552-553. [Medline: [2024213](#)]
3. Flenady V, Mahomed K, Ellwood D, Charles A, Teale G, Chadha Y, AustralianNew Zealand Stillbirth Alliance. Uptake of the Perinatal Society of Australia and New Zealand perinatal mortality audit guideline. *Aust N Z J Obstet Gynaecol* 2010 Apr;50(2):138-143. [doi: [10.1111/j.1479-828X.2009.01125.x](#)] [Medline: [20522069](#)]
4. Eskes M, Waelput AJ, Erwich JJ, Brouwers HA, Ravelli AC, Achterberg PW, et al. Term perinatal mortality audit in the Netherlands 2010-2012: a population-based cohort study. *BMJ Open* 2014 Oct 14;4(10):e005652 [FREE Full text] [doi: [10.1136/bmjopen-2014-005652](#)] [Medline: [25763794](#)]
5. Cousens S, Blencowe H, Stanton C, Chou D, Ahmed S, Steinhardt L, et al. National, regional, and worldwide estimates of stillbirth rates in 2009 with trends since 1995: a systematic analysis. *Lancet* 2011 Apr 16;377(9774):1319-1330. [doi: [10.1016/S0140-6736\(10\)62310-0](#)] [Medline: [21496917](#)]
6. Biswas A, Rahman F, Eriksson C, Halim A, Dalal K. Facility death review of maternal and neonatal deaths in Bangladesh. *PLoS One* 2015;10(11):e0141902 [FREE Full text] [doi: [10.1371/journal.pone.0141902](#)] [Medline: [26540233](#)]
7. Armstrong CE, Lange IL, Magoma M, Ferla C, Filippi V, Ronsmans C. Strengths and weaknesses in the implementation of maternal and perinatal death reviews in Tanzania: perceptions, processes and practice. *Trop Med Int Health* 2014 Sep;19(9):1087-1095 [FREE Full text] [doi: [10.1111/tmi.12353](#)] [Medline: [25039579](#)]
8. Hofman JJ, Mohammed H. Experiences with facility-based maternal death reviews in northern Nigeria. *Int J Gynaecol Obstet* 2014 Aug;126(2):111-114 [FREE Full text] [doi: [10.1016/j.ijgo.2014.02.014](#)] [Medline: [24834852](#)]
9. Kongnyuy EJ, Mlava G, van den Broek N. Facility-based maternal death review in three districts in the central region of Malawi: an analysis of causes and characteristics of maternal deaths. *Womens Health Issues* 2009;19(1):14-20. [Medline: [19111783](#)]
10. World Health Organization. 2006. Neonatal and perinatal mortality. Country, regional and global estimates URL: <http://www.who.int/sorry/> [accessed 2019-01-28] [WebCite Cache ID 75IKVmQl2]
11. Rajaratnam JK, Marcus JR, Flaxman AD, Wang H, Levin-Rector AD, Dwyer L, et al. Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970-2010: a systematic analysis of progress towards Millennium Development Goal 4. *Lancet* 2010 Jun 5;375(9730):1988-2008. [doi: [10.1016/S0140-6736\(10\)60703-9](#)] [Medline: [20546887](#)]
12. UNDP 2010; reports by UN agencies, funds and programmes. 2010 Mar 19. Millennium Development Goals Report 2010, Keeping the Promise URL: <http://undocs.org/en/A/RES/65/1> [accessed 2019-01-28] [WebCite Cache ID 75IQiVjOO]
13. Batieha AM, Khader YS, Berdzuli N, Chua-Oon C, Badran EF, Al-Sheyab NA, et al. Level, causes and risk factors of neonatal mortality, in Jordan: results of a national prospective study. *Matern Child Health J* 2016 May;20(5):1061-1071. [doi: [10.1007/s10995-015-1892-x](#)] [Medline: [26645614](#)]
14. Sivasubramaniam P, Quinn CE, Blevins M, Al Hajjara A, Khuri-Bulos N, Faouri S, et al. Neonatal outcomes of infants admitted to a large government hospital in Amman, Jordan. *Glob J Health Sci* 2015 Jan 14;7(4):217-234 [FREE Full text] [doi: [10.5539/gjhs.v7n4p217](#)] [Medline: [25946927](#)]
15. Khader YS, Batieha A, Al Fursan RK, Al-Hader R, Hijazi SS. Rate of teenage pregnancy in Jordan and its impact on maternal and neonatal outcomes. *Int J Adolesc Med Health* 2017 Jul 26. [doi: [10.1515/ijamh-2017-0075](#)] [Medline: [28782350](#)]
16. Bouchghoul H, Hornez E, Duval-Arnould X, Philippe HJ, Nizard J. Humanitarian obstetric care for refugees of the Syrian war. The first 6 months of experience of Gynécologie Sans Frontières in Zaatari Refugee Camp (Jordan). *Acta Obstet Gynecol Scand* 2015 Jul;94(7):755-759. [doi: [10.1111/aogs.12638](#)] [Medline: [25817053](#)]

17. Jamtvedt G, Young JM, Kristoffersen DT, Thomson O'Brien MA, Oxman AD. Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2003 Jun 13;6(3):CD000259. [doi: [10.1002/14651858.CD000259](https://doi.org/10.1002/14651858.CD000259)] [Medline: [12917891](https://pubmed.ncbi.nlm.nih.gov/12917891/)]
18. Richard F, Ouédraogo C, Zongo V, Ouattara F, Zongo S, Gruénais ME, et al. The difficulty of questioning clinical practice: experience of facility-based case reviews in Ouagadougou, Burkina Faso. *BJOG* 2009 Jan;116(1):38-44 [FREE Full text] [doi: [10.1111/j.1471-0528.2008.01741.x](https://doi.org/10.1111/j.1471-0528.2008.01741.x)] [Medline: [18503575](https://pubmed.ncbi.nlm.nih.gov/18503575/)]
19. Belizán M, Bergh AM, Cilliers C, Pattinson RC, Voce A, Synergy Group. Stages of change: a qualitative study on the implementation of a perinatal audit programme in South Africa. *BMC Health Serv Res* 2011 Sep 30;11:243 [FREE Full text] [doi: [10.1186/1472-6963-11-243](https://doi.org/10.1186/1472-6963-11-243)] [Medline: [21958353](https://pubmed.ncbi.nlm.nih.gov/21958353/)]
20. Bakker W, van den Akker T, Mwagomba B, Khukulu R, van Elteren M, van Roosmalen J. Health workers' perceptions of obstetric critical incident audit in Thyolo District, Malawi. *Trop Med Int Health* 2011 Oct;16(10):1243-1250 [FREE Full text] [doi: [10.1111/j.1365-3156.2011.02832.x](https://doi.org/10.1111/j.1365-3156.2011.02832.x)] [Medline: [21767335](https://pubmed.ncbi.nlm.nih.gov/21767335/)]
21. Agaro C, Beyeza-Kashesya J, Waiswa P, Sekandi JN, Tusiime S, Anguzu R, et al. The conduct of maternal and perinatal death reviews in Oyam District, Uganda: a descriptive cross-sectional study. *BMC Womens Health* 2016 Dec 14;16:38 [FREE Full text] [doi: [10.1186/s12905-016-0315-5](https://doi.org/10.1186/s12905-016-0315-5)] [Medline: [27418127](https://pubmed.ncbi.nlm.nih.gov/27418127/)]
22. Namazzi G, Waiswa P, Nakakeeto M, Nakibuuka VK, Namutamba S, Najjemba M, et al. Strengthening health facilities for maternal and newborn care: experiences from rural eastern Uganda. *Glob Health Action* 2015 Mar 31;8:24271. [doi: [10.3402/gha.v8.24271](https://doi.org/10.3402/gha.v8.24271)] [Medline: [25843496](https://pubmed.ncbi.nlm.nih.gov/25843496/)]
23. Moxon SG, Lawn JE, Dickson KE, Simen-Kapeu A, Gupta G, Deorari A, et al. Inpatient care of small and sick newborns: a multi-country analysis of health system bottlenecks and potential solutions. *BMC Pregnancy Childbirth* 2015;15(Suppl 2):S7 [FREE Full text] [doi: [10.1186/1471-2393-15-S2-S7](https://doi.org/10.1186/1471-2393-15-S2-S7)] [Medline: [26391335](https://pubmed.ncbi.nlm.nih.gov/26391335/)]
24. Owolabi H, Ameh CA, Bar-Zeev S, Adaji S, Kachale F, van den Broek N. Establishing cause of maternal death in Malawi via facility-based review and application of the ICD-MM classification. *BJOG* 2014 Sep;121(Suppl 4):95-101 [FREE Full text] [doi: [10.1111/1471-0528.12998](https://doi.org/10.1111/1471-0528.12998)] [Medline: [25236641](https://pubmed.ncbi.nlm.nih.gov/25236641/)]
25. Pattinson R, Kerber K, Waiswa P, Day LT, Mussell F, Asiruddin SK, et al. Perinatal mortality audit: counting, accountability, and overcoming challenges in scaling up in low- and middle-income countries. *Int J Gynaecol Obstet* 2009 Oct;107(Suppl 1):S113-S121. [doi: [10.1016/j.ijgo.2009.07.011](https://doi.org/10.1016/j.ijgo.2009.07.011)] [Medline: [19815206](https://pubmed.ncbi.nlm.nih.gov/19815206/)]
26. Lewis G. The cultural environment behind successful maternal death and morbidity reviews. *BJOG* 2014 Sep;121(Suppl 4):24-31 [FREE Full text] [doi: [10.1111/1471-0528.12801](https://doi.org/10.1111/1471-0528.12801)] [Medline: [25236630](https://pubmed.ncbi.nlm.nih.gov/25236630/)]
27. Pearson L, deBernis L, Shoo R. Maternal death review in Africa. *Int J Gynaecol Obstet* 2009 Jul;106(1):89-94. [doi: [10.1016/j.ijgo.2009.04.009](https://doi.org/10.1016/j.ijgo.2009.04.009)] [Medline: [19428010](https://pubmed.ncbi.nlm.nih.gov/19428010/)]
28. Khanam R, Khan M, Halim M, Begum K, Jahan S. Facility and community based maternal death review in Bangladesh. *Bangladesh J Obstet Gynaecol* 1970 Jan 1;24(1):18-21. [doi: [10.3329/bjog.v24i1.6322](https://doi.org/10.3329/bjog.v24i1.6322)]
29. Dumont A, Gaye A, De BL, Chaillet N, Landry A, Delage J. Facility-based maternal death reviews: effects on maternal mortality in a district hospital in Senegal. *Bull World Health Organ* 2006;84(3):218-224.

Abbreviations

JUST: Jordan University of Science and Technology

PNN: perinatal and neonatal

UNICEF: United Nations International Children's Emergency Fund

Edited by T Sanchez; submitted 21.07.18; peer-reviewed by Y Halasa, A Harste; comments to author 12.11.18; revised version received 01.12.18; accepted 07.12.18; published 06.03.19.

Please cite as:

Khader Y, Alyahya M, Batieha A

Barriers to Implementation of Perinatal Death Audit in Maternity and Pediatric Hospitals in Jordan: Cross-Sectional Study

JMIR Public Health Surveill 2019;5(1):e11653

URL: <http://publichealth.jmir.org/2019/1/e11653/>

doi: [10.2196/11653](https://doi.org/10.2196/11653)

PMID: [30839277](https://pubmed.ncbi.nlm.nih.gov/30839277/)

©Yousef Khader, Mohammad Alyahya, Anwar Batieha. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 06.03.2019. 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: 2016 Key Indicators Report

Maria Zlotorzynska¹, MPH, PhD; Patrick Sullivan¹, DVM, PhD; Travis Sanchez¹, MPH, DVM

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

Corresponding Author:

Maria Zlotorzynska, MPH, PhD

Department of Epidemiology

Rollins School of Public Health

Emory University

1518 Clifton Rd NE

Atlanta, GA, 30322

United States

Phone: 1 4047278799

Email: mzlotor@emory.edu

Abstract

The American Men's Internet Survey (AMIS) is an annual Web-based behavioral survey of men who have sex with men (MSM) living in the United States. This Rapid Surveillance Report describes the fourth cycle of data collection (September 2016 through February 2017; AMIS 2016). The key indicators are the same as previously reported for AMIS (December 2013 through May 2014, AMIS 2013; November 2014 through April 2015, AMIS 2014; and September 2015 through April 2016, AMIS 2015). The AMIS survey methodology has not substantively changed since AMIS 2015. MSM were recruited from a variety of websites using banner advertisements and email blasts. Additionally, participants from AMIS 2015 who agreed to be recontacted for future research were emailed a link to the AMIS 2016 survey. Men were eligible to participate if they were ≥ 15 years old, resided in the United States, provided a valid US zone improvement plan code, and reported ever having sex with a man or identified as gay or bisexual. We examined demographic and recruitment characteristics using multivariable regression modeling ($P < .05$) stratified by participants' self-reported HIV status. The AMIS 2016 round of data collection resulted in 10,166 completed surveys from MSM representing every US state, Puerto Rico, Guam, and the US Virgin Islands. Participants were mainly non-Hispanic white, over the age of 40 years, living in the Southern United States and urban areas, and recruited from general social networking websites. Self-reported HIV prevalence was 10.80% (1098/10,166). Compared to HIV-negative/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 (75.77% vs 65.88%, $P < .001$) and more likely to have had anal sex without a condom with a serodiscordant or unknown-status partner (33.24% vs 16.06%, $P < .001$). The reported use of marijuana, methamphetamines, and other illicit substances in the past 12 months was higher among HIV-positive participants than among HIV-negative/unknown-status participants (28.05% vs 24.99%, 11.48% vs 2.16%, and 27.60% vs 18.22%, respectively; all $P < .001$). Most HIV-negative/unknown-status participants (79.93%, 7248/9068) reported ever having a previous HIV test, and 56.45% (5119/9068) 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/unknown-status participants (70.86% vs 40.13% and 24.04% vs 8.97%, respectively; both $P < .001$).

(*JMIR Public Health Surveill* 2019;5(1):e11313) doi:[10.2196/11313](https://doi.org/10.2196/11313)

KEYWORDS

bisexual; gay; HIV; homosexual; internet; MSM; STD; surveillance; survey

Introduction

The American Men's Internet Survey (AMIS) is an annual online behavioral survey of men who have sex with men (MSM), living in the United States. AMIS was developed to produce

timely data from large-scale monitoring of behavior trends among MSM recruited online. 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]. The methods and previous AMIS cycle data (AMIS 2013, AMIS 2014, and AMIS 2015) have been previously published [2-4].

This supplemental report updates previous information with data collected in AMIS 2016. Methods in the AMIS 2016 did not change from the previously published methods, unless otherwise noted. An in-depth analysis and discussion of multiyear trends for indicators reported herein has been published and includes data for the first four cycles of AMIS (AMIS 2013 through AMIS 2016) [5].

Methods

Recruitment and Enrollment

Similar to the prior year's recruitment process, AMIS participants were recruited through convenience sampling from a variety of websites using banner advertisements or email blasts to website members (hereafter referred to generically as "ads"). For AMIS 2016, data were collected from September 2016 through February 2017. The survey was not incentivized. Data on the number of clicks on all banner ads were obtained directly from the websites. Men who clicked on the ads were taken directly to the survey website hosted on a secure server administered by SurveyGizmo (Boulder, Colorado). Participants were also recruited by emailing participants from the previous cycle of AMIS (AMIS 2015) who consented to be recontacted for future studies. To be eligible for the survey, participants had to be ≥ 15 years of age, consider themselves as male, reside in the United States, and report that they either had oral or anal sex with a man at least once in the past or identify as gay or bisexual (hereafter referred to as men who have sex with men [MSM]). Persons who were < 15 years of age 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 online survey immediately. The full questionnaire for AMIS 2016 is presented in [Multimedia Appendix 1](#).

Several data-cleaning steps were performed on the raw dataset of eligible responses to obtain the final analysis dataset. First, deduplication of survey responses was performed in the same manner as in previous AMIS cycles [2-5]. Briefly, the demographic data for near-complete ($> 70\%$) survey responses with nonunique internet protocol addresses were compared, and responses that showed 100% match for all characteristics were considered to be duplicate responses. Only the observation with the highest survey completion was retained. The dataset was then limited to those surveys deemed successful (ie, observations with no missing values for the first question of at least two consecutive sections). Finally, the dataset was restricted to include participants who reported having oral or anal sex in the past 12 months and provided a valid US zone improvement plan (ZIP) code. ZIP codes were validated in same manner as in AMIS 2015 [4]. 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 [6]. 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 [7]. ZIP codes that could not be found were classified as invalid.

Measures and Analyses

For AMIS 2016 analyses, participants were categorized as either AMIS 2015 participants who took the survey again or new participants from the website/app based on target audience and purpose: gay social networking ($n=2$), gay general interest ($n=1$), general social networking ($n=3$), and geospatial social networking ($n=2$). Recruitment outcomes and demographic characteristics for the AMIS 2015 participants are presented in [Tables 1](#) and [2](#), and thereafter, they are recategorized according to their original source of recruitment. We do not provide the names of the websites/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 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 [8]. Population density was defined in the same manner as in AMIS 2015 and was based on the National Center for Health Statistics Rural-Urban classification scheme for counties [9]. The self-reported HIV status was categorized as HIV positive or HIV negative/unknown status, consistent with surveillance reports produced by the NHBS system [8]. Three 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 2015 [4].

The analysis methods for AMIS 2016 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 between 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/ethnicity, age group, NHBS city residency, and type of recruitment website. Metropolitan statistical areas included in the NHBS system in 2016 were Atlanta, GA; Baltimore, MD; Boston, MA; Chicago, IL; Dallas, TX; Denver, CO; Detroit, MI; Houston, TX; Los Angeles, CA; Miami, FL; Nassau-Suffolk, NY; New Orleans, LA; New York City, NY; Newark, NJ; Philadelphia, PA; San Diego, CA; San Francisco, CA; San Juan, PR; Seattle, WA; and Washington, DC. HIV testing behaviors were only examined among those who did not report that they were HIV positive, and these data were presented by participant characteristics. Multivariable logistic regression results are presented as Wald chi-square P values to denote an independently significant difference in the behavior for each subgroup compared to a reference group. Statistical significance was set at $P < .05$.

Table 1. Recruitment outcomes for the American Men's Internet Survey, United States, 2016.

Recruitment outcomes	Total	Gay social networking (n=2)	General gay interest (n=1)	General social networking (n=3)	Geospatial social networking (n=2)	AMIS ^a 2015 participants
Clicked ad (N)	147,143	4162	557	58,917	83,507	Not applicable
Screened ^b , n (%)	51,876 (35.26)	2877 (69.13)	181 (32.50)	39,281 (66.67)	8137 (9.74)	1400
Ineligible^c, n (%)	23,173 (44.67)	564 (19.60)	147 (81.22)	19,271 (49.06)	3039 (37.35)	152 (10.86)
Not ≥15 years of age ^d	16,643 (71.82)	438 (77.66)	91 (61.90)	13,572 (70.43)	2441 (80.32)	101 (66.45)
Not male ^d	19,079 (82.33)	511 (90.60)	94 (63.95)	15,641 (81.16)	2704 (88.98)	129 (84.87)
Not MSM ^e ever or not identifying as gay/bisexual ^d	22,282 (96.16)	549 (97.34)	99 (67.35)	18,721 (97.15)	2790 (91.81)	123 (80.92)
Nonresident ^d	18,989 (81.94)	471 (83.51)	144 (97.96)	15,383 (79.82)	2845 (93.62)	146 (96.05)
Eligible ^c , n (%)	28,703 (55.33)	2313 (80.40)	34 (18.78)	20,010 (50.94)	5098 (62.65)	1248 (89.14)
Consented ^f , n (%)	20,583 (71.71)	1716 (74.19)	28 (82.35)	13,776 (68.85)	3928 (77.05)	1135 (90.95)
Unduplicated ^g , n (%)	18,038 (87.64)	1604 (93.47)	23 (82.14)	11,876 (86.21)	3501 (89.13)	1034 (91.10)
Success ^h , n (%)	11,636 (64.51)	1242 (77.43)	14 (60.87)	7594 (63.94)	1870 (53.41)	916 (88.59)
MSM in past 12 months ⁱ , n (%)	10,222 (87.85)	1165 (93.80)	13 (92.86)	6443 (84.84)	1756 (93.90)	845 (92.25)
Valid ZIP ^j code ^k , n (%)	10,166 (99.45)	1160 (99.57)	13 (100.00)	6401 (99.35)	1750 (99.66)	842 (99.64)

^aAMIS: American Men's Internet Survey.

^bProportion of total participants who clicked on the ad, including those who started the screening questionnaire.

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

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

^eMSM: men who have sex with men.

^fProportion of eligible participants.

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

^hProportion of unduplicated participants. Success in deduplication removes participants who did not pass the test for survey completeness.

ⁱProportion of successes.

^jZIP: zone improvement plan.

^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.

Table 2. Characteristics of men who have sex with men in the American Men's Internet Survey by recruitment type, United States, 2016.

Participant characteristics	Recruitment type						<i>P</i> value ^b
	Total	Gay social networking (n=2)	General gay interest (n=1)	General social networking (n=3)	Geospatial social networking (n=2)	AMIS ^a 2015 participants	
Race/ethnicity, n (%)							<.001
Black, non-Hispanic	879 (8.65)	35 (3.02)	0 (0.00)	664 (10.37)	141 (8.06)	39 (4.63)	
Hispanic	1311 (12.90)	73 (6.29)	2 (15.38)	852 (13.31)	286 (16.34)	98 (11.64)	
White, non-Hispanic	7073 (69.58)	986 (85.00)	8 (61.54)	4274 (66.77)	1157 (66.11)	648 (76.96)	
Other or multiple races	903 (8.88)	66 (5.69)	3 (23.08)	611 (9.55)	166 (9.49)	57 (6.77)	
Age (years), n (%)							<.001
15-24	2718 (26.74)	27 (2.33)	6 (46)	2268 (35.43)	271 (15.49)	146 (17.34)	
25-29	1693 (16.65)	34 (2.93)	2 (15)	1223 (19.11)	265 (15.14)	169 (20.07)	
30-39	1414 (13.91)	92 (7.93)	1 (8)	719 (11.23)	455 (26.00)	147 (17.46)	
≥40	4341 (42.70)	1007 (86.81)	4 (31)	2191 (34.23)	759 (43.37)	380 (45.13)	
Region, n (%)							<.001
Northeast	1879 (18.48)	235 (20.26)	2 (15.38)	1185 (18.51)	306 (17.49)	151 (17.93)	
Midwest	1988 (19.56)	232 (20.00)	4 (30.77)	1352 (21.12)	233 (13.31)	167 (19.83)	
South	4055 (39.89)	422 (36.38)	4 (30.77)	2506 (39.15)	800 (45.71)	323 (38.36)	
West	2240 (22.03)	271 (23.36)	3 (23.08)	1355 (21.17)	410 (23.43)	201 (23.87)	
US-dependent areas	4 (0.04)	0 (0)	0 (0)	3 (0.05)	1 (0.06)	0 (0)	
NHBS^c city resident, n (%)							<.001
Yes	4224 (41.55)	500 (43.10)	7 (53.85)	2291 (35.81)	1085 (62.00)	341 (40.50)	
No	5942 (58.45)	660 (56.90)	6 (46.15)	4110 (64.21)	665 (38.02)	501 (59.50)	
Population density^d, n (%)							<.001
Urban	4288 (42.18)	436 (37.59)	5 (38.46)	2538 (39.65)	944 (53.94)	365 (43.35)	
Suburban	2200 (21.64)	316 (27.24)	6 (46.15)	1298 (20.28)	409 (23.37)	171 (20.31)	
Small/medium metropolitan	2790 (27.44)	305 (26.29)	1 (7.69)	1929 (30.14)	305 (17.43)	250 (29.69)	
Rural	884 (8.70)	103 (8.88)	1 (7.69)	633 (9.89)	91 (5.20)	56 (6.65)	
Self-reported HIV status, n (%)							<.001
Positive	1098 (10.80)	125 (10.78)	2 (15.38)	619 (9.67)	263 (15.03)	89 (10.57)	
Negative	7089 (69.73)	828 (71.38)	9 (69.23)	4283 (66.91)	1291 (73.77)	678 (80.52)	
Unknown	1979 (19.47)	207 (17.84)	2 (15.38)	1499 (23.42)	196 (11.20)	75 (8.91)	
Total (N)	10,166	1160	13	6401	1750	842	

^aAMIS: American Men's Internet Survey.^bChi-square test for difference in characteristics between recruitment types.^cNHBS: National HIV Behavioral Surveillance.^dThe National Center for Health Statistics urban/rural category could not be assigned for four participants living in US territories.

Results

AMIS 2016 was performed from September 2016 through February 2017 and resulted in 147,143 persons clicking on the ads and landing on the study's recruitment page (Table 1). Most persons who clicked on the ads were from geospatial networking websites (83,507/147,143; 56.75%). Of the 4513 participants

who completed the AMIS 2015 survey and were emailed links to the AMIS 2016 survey, 31.02% (1400) clicked on the link. Just over one-third (35.26%) of all of participants who landed on the study page started the screening process and 55.33% of them were eligible. The most common reason for ineligibility was not ever having male-male sex or not identifying as gay or bisexual. Almost three-quarters (71.71%) of participants who

were eligible consented to participate in the survey. A total of 2545 (12.36%) surveys were likely from duplicate participants. Among unduplicated surveys, almost two-thirds (64.51%) were considered successful. Most successful surveys were from men who reported having sex with another man in the past 12 months (87.85%). Almost all these surveys (10,166/10,222; 99.45%) provided a valid US ZIP code. Overall, the completion rate was 6.9%, with an analytical sample consisting of 10,166 surveys from 147,143 clicks.

Over two-thirds (7073/10,166; 69.58%) of the participants included in this report were non-Hispanic white, and less than half were ≥ 40 years of age (4341/10,166; 42.70%); 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 from each of the 28 states and the District of Columbia (Figure 1). About 4 in 10 (4224/10,166; 41.55%) participants resided in an NHBS city and about the same proportion (4288/10,166; 42.18%) lived in an urban county. Overall, 10.80% (1098/10,166) of participants were HIV positive, 69.73% were HIV negative (7089/10,166),

and 19.74% (1979/10,166) had an unknown HIV status. All participant characteristics differed significantly based on the recruitment source (Table 2).

Most participants reported having anal sex without a condom with another man in the past 12 months (Table 3). Compared to HIV-negative/unknown-status participants, those who were HIV positive were significantly more likely to report anal intercourse without a condom (adjusted odds ratio [aOR]=1.79, 95% CI: 1.53-2.08), including with male partners who were of discordant or unknown status (aOR=2.76, 95% CI: 2.38-3.19). Stratified by the serostatus group, anal intercourse without a condom differed significantly by race/ethnicity (HIV-negative/unknown-status participants only), age group (HIV-negative/unknown-status participants only), and recruitment website (HIV-positive and HIV-negative/unknown-status participants). Anal intercourse without a condom with partners of discordant or unknown HIV status differed significantly by race/ethnicity, age, and recruitment website for both HIV-positive participants and HIV-negative/unknown-status participants.

Figure 1. Number of men who have sex with men who participated in the American Men's Internet Survey by state, 2016.

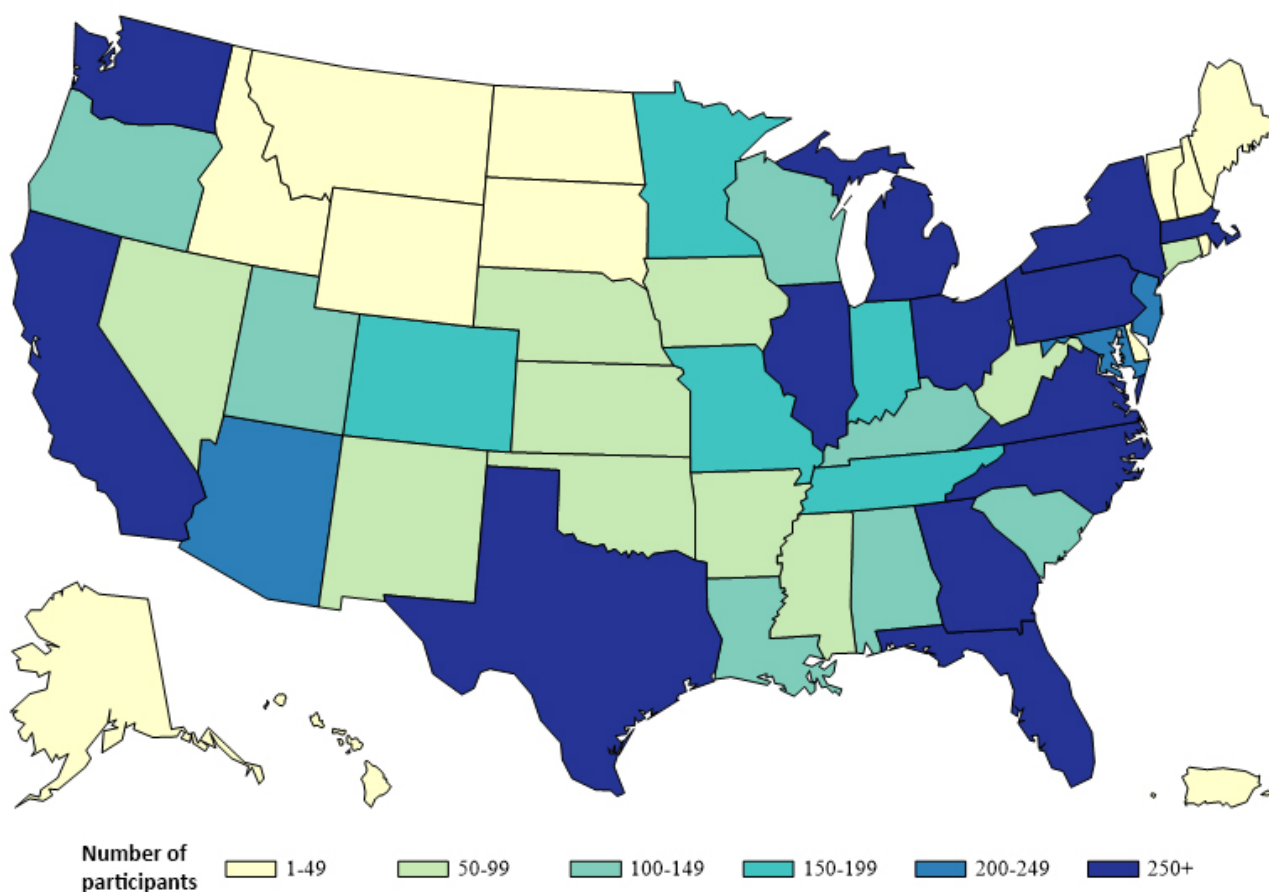


Table 3. Sexual Behaviors with male partners of men who have sex with men in the American Men's Internet Survey, United States, 2016.

Participant characteristics	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 discordant or unknown HIV status	
		n (%)	P value ^a	n (%)	P value ^a
HIV positive	1098	832 (75.77)	<.001 ^b	365 (33.24)	<.001 ^b
Race/ethnicity					
Black, non-Hispanic	248	175 (70.56)	.17	76 (30.65)	.70
Hispanic	130	96 (73.85)	.23	34 (26.15)	.001
White, non-Hispanic	632	492 (77.85)	Ref ^a	218 (34.49)	Ref ^a
Other or multiple races	88	69 (78.41)	.34	37 (42.05)	.05
Age (years)					
15-24	59	50 (84.75)	.37	21 (35.59)	.87
25-29	97	80 (82.47)	.78	48 (49.48)	.008
30-39	186	159 (85.48)	.16	76 (40.86)	.41
≥40	756	543 (71.83)	Ref ^a	220 (29.10)	Ref ^a
NHBS^c city resident					
Yes	568	429 (75.53)	.78	195 (34.33)	.43
No	530	403 (76.04)	Ref ^a	170 (32.08)	Ref ^a
Recruitment type					
Gay social networking	139	102 (73.38)	.45	58 (41.73)	.004
General gay interest ^d	<5	N/A ^e	N/A	N/A	N/A
General social networking	651	471 (72.35)	Ref ^a	193 (29.65)	Ref ^a
Geospatial social networking	303	255 (84.16)	.007	111 (36.63)	.47
HIV negative or unknown status	9068	5974 (65.88)	Ref ^a	1456 (16.06)	Ref ^a
Race/ethnicity					
Black, non-Hispanic	631	388 (61.49)	.002	116 (18.38)	.08
Hispanic	1181	833 (70.53)	<.001	218 (18.46)	.12
White, non-Hispanic	6441	4231 (65.69)	Ref ^a	1010 (15.68)	Ref ^a
Other or multiple races	815	522 (64.05)	.28	112 (13.74)	.009
Age (years)					
15-24	2659	1694 (63.71)	<.001	417 (15.68)	.87
25-29	1596	1202 (75.31)	<.001	285 (17.86)	.03
30-39	1228	892 (72.64)	.002	232 (18.89)	.09
≥40	3585	2186 (60.98)	Ref ^a	522 (14.56)	Ref ^a
NHBS^c city resident					
Yes	3656	2436 (66.63)	>.99	630 (17.23)	.14
No	5412	3538 (65.37)	Ref ^a	826 (15.26)	Ref ^a
Recruitment type					
Gay social networking	1119	626 (55.94)	<.001	210 (18.77)	.007
General gay interest	73	50 (68.49)	.76	10 (13.70)	.38
General social networking	6125	4026 (65.73)	Ref ^a	911 (14.87)	Ref ^a

Participant characteristics	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 discordant or unknown HIV status	
		n (%)	P value ^a	n (%)	P value ^a
Geospatial social networking	1731	1257 (72.62)	<.001	321 (18.54)	.35

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

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

^cNHBS: National HIV Behavioral Surveillance.

^dTo prevent identification, data for groups with less than five participants are not presented.

^eN/A: not applicable.

Over one-quarter (308/1098; 28.05%) of HIV-positive participants reported using marijuana in the past 12 months (Table 4). Compared to HIV-negative/unknown-status participants, HIV-positive participants were significantly more likely to report use of marijuana (aOR=1.63, 95% CI: 1.40-1.90), methamphetamines (aOR=5.53, 95% CI: 4.30-7.11), and other illicit substances in the past 12 months (aOR=2.15, 95% CI: 1.84-2.52). Among HIV-positive participants, the use of marijuana did not vary significantly for any participant characteristic, but the use of methamphetamines varied significantly by race/ethnicity, age, and recruitment site. In this group, the use of other illicit substances varied significantly by race/ethnicity, age, residence in an NHBS city, and recruitment site. Use of marijuana, methamphetamines, and other illicit substances differed significantly by race/ethnicity and age among HIV-negative/unknown-status participants. Additionally, the use of marijuana and other illicit substances differed significantly by residence in an NHBS city, and the use of other illicit substances differed significantly by recruitment site among HIV-negative/unknown-status participants.

HIV testing behaviors were examined among participants who were not HIV positive (Table 5). Most participants (7248/9068; 79.93%) were previously tested for HIV infection, and just over

half (5119/9068; 56.45%) were tested in the past 12 months. HIV testing behavior, both ever tested and tested in the past 12 months, differed significantly by race/ethnicity, age, residence in an NHBS city, and type of recruitment website.

Compared to HIV-negative/unknown-status participants, HIV-positive participants were significantly more likely to report testing for sexually transmitted infection (STI) (aOR=3.62, 95% CI: 3.13-4.19) and STI diagnosis (aOR=3.25, 95% CI: 2.74-3.87) in the past 12 months (Table 6). The most common STI diagnosis among HIV-positive participants was syphilis (164/1098; 14.94%), followed by gonorrhea (125/1098; 11.38%) and chlamydia (108/1098; 9.84%). Gonorrhea was the most common STI diagnosis among HIV-negative/unknown-status participants (456/9068; 5.03%), followed by chlamydia (412/9068; 4.54%) and syphilis (226/9068; 2.49%). Among HIV-negative/unknown-status participants, STI testing differed significantly by race/ethnicity and age. Among both HIV-positive and HIV-negative/unknown-status participants, STI testing differed significantly by NHBS city residence and type of recruitment website, and STI diagnosis differed significantly by age, NHBS city residence, and type of recruitment website.

Table 4. Substance use behaviors of men who have sex with men in the American Men's Internet Survey, United States, 2016.

Participant characteristics	Participants (N)	Substance use behaviors in the past 12 months					
		Used marijuana		Used methamphetamines		Used other substance(s)	
		n (%)	P value ^a	n (%)	P value ^a	n (%)	P value ^a
HIV positive	1098	308 (28.05)	<.001 ^b	126 (11.48)	<.001 ^b	303 (27.60)	<.001 ^b
Race/ethnicity							
Black, non-Hispanic	248	66 (26.61)	.52	16 (6.45)	.02	52 (20.97)	.04
Hispanic	130	40 (30.77)	.98	22 (16.92)	.12	44 (33.85)	.21
White, non-Hispanic	632	176 (27.85)	Ref ^a	75 (11.87)	Ref ^a	183 (28.96)	Ref ^a
Other or multiple races	88	26 (29.55)	.96	13 (14.77)	.45	24 (27.27)	.64
Age (years)							
15-24	59	22 (37.29)	.29	5 (8.47)	.51	21 (35.59)	.60
25-29	97	33 (34.02)	.80	10 (10.31)	.72	41 (42.27)	.03
30-39	186	71 (38.17)	.18	33 (17.74)	.03	65 (34.95)	.96
≥40	756	182 (24.07)	Ref ^a	78 (10.32)	Ref ^a	176 (23.28)	Ref ^a
NHBS^c city resident							
Yes	568	174 (30.63)	.10	78 (13.73)	.11	175 (30.81)	.04
No	530	134 (25.28)	Ref ^a	48 (9.06)	Ref ^a	128 (24.15)	Ref ^a
Recruitment type							
Gay social networking	139	36 (25.90)	.98	17 (12.23)	.93	34 (24.46)	.56
General gay interest ^d	<5	N/A ^e	N/A	N/A	N/A	N/A	N/A
General social networking	651	170 (26.11)	Ref ^a	50 (7.68)	Ref ^a	153 (23.50)	Ref ^a
Geospatial social networking	303	100 (33.00)	.35	57 (18.81)	.007	113 (37.29)	.01
HIV negative or unknown status	9068	2266 (24.99)	Ref ^a	196 (2.16)	Ref ^a	1652 (18.22)	Ref ^a
Race/ethnicity							
Black, non-Hispanic	631	126 (19.97)	.006	8 (1.27)	.049	86 (13.63)	.001
Hispanic	1181	344 (29.13)	.04	37 (3.13)	.03	275 (23.29)	<.001
White, non-Hispanic	6441	1597 (24.79)	Ref ^a	128 (1.99)	Ref ^a	1152 (17.89)	Ref ^a
Other or multiple races	815	199 (24.42)	.33	23 (2.82)	.12	139 (17.06)	.27
Age (years)							
15-24	2659	886 (33.32)	<.001	37 (1.39)	.03	548 (20.61)	.001
25-29	1596	500 (31.33)	<.001	35 (2.19)	.60	388 (24.31)	.18
30-39	1228	322 (26.22)	.74	42 (3.42)	.04	270 (21.99)	<.001
≥40	3585	558 (15.56)	Ref ^a	82 (2.29)	Ref ^a	446 (12.44)	Ref ^a
NHBS^c city resident							
Yes	3656	1024 (28.01)	<.001	94 (2.57)	.48	779 (21.31)	<.001
No	5412	1242 (22.95)	Ref ^a	102 (1.88)	Ref ^a	873 (16.13)	Ref ^a
Recruitment type							
Gay social networking	1119	201 (17.96)	.71	37 (3.31)	.13	150 (13.40)	.68
General gay interest	73	16 (21.92)	.55	1 (6.25)	.48	12 (16.44)	.52
General social networking	6125	1572 (25.67)	Ref ^a	90 (1.47)	Ref ^a	1068 (17.44)	Ref ^a

Participant characteristics	Participants (N)	Substance use behaviors in the past 12 months					
		Used marijuana		Used methamphetamines		Used other substance(s)	
		n (%)	P value ^a	n (%)	P value ^a	n (%)	P value ^a
Geospatial social networking	1731	473 (27.33)	.11	68 (3.93)	.08	419 (24.21)	<.001

^aWald chi-square from multivariable logistic regression comparing behavior (yes vs no) between groups with specific characteristics and a reference (Ref) group.

^bWald chi-square from multivariable logistic regression comparing behavior (yes vs no) between HIV-positive participants and HIV-negative or unknown-serostatus participants. Model controlled for race/ethnicity, age, National HIV Behavioral Surveillance system residency, and website type.

^cNHBS: National HIV Behavioral Surveillance.

^dTo prevent identification, data for groups with less than five participants are not presented.

^eN/A: not applicable.

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, 2016.

Participant characteristics	Participants (N)	HIV testing behaviors			
		HIV tested ever		HIV tested in past 12 months	
		n (%)	P value ^a	n (%)	P value ^a
Race/ethnicity					
Black, non-Hispanic	631	563 (89.22)	<.001	440 (69.73)	<.001
Hispanic	1181	916 (77.56)	.09	710 (60.12)	.68
White, non-Hispanic	6441	5134 (79.71)	Ref ^a	3512 (54.53)	Ref ^a
Other or multiple races	815	635 (77.91)	.30	457 (56.07)	.02
Age (years)					
15-24	2659	1521 (57.20)	<.001	1207 (45.39)	<.001
25-29	1596	1393 (87.28)	.004	1040 (65.16)	<.001
30-39	1228	1129 (91.94)	<.001	829 (67.51)	<.001
≥40	3585	3205 (89.40)	Ref ^a	2043 (56.99)	Ref ^a
NHBS^b city resident					
Yes	3656	3112 (85.12)	<.001	2356 (64.44)	<.001
No	5412	4136 (76.42)	Ref ^a	2763 (51.05)	Ref ^a
Recruitment type	1119	926 (82.75)			
Gay social networking	73	66 (90.41)	<.001	577 (51.56)	<.001
General gay interest	6125	4676 (76.34)	.18	44 (60.27)	.93
General social networking	1731	1563 (90.29)	Ref ^a	3193 (52.13)	Ref ^a
Geospatial social networking	1119	926 (82.75)	<.001	1292 (74.64)	<.001
Total	9068	7248 (79.93)		5119 (56.45)	

^aWald chi-square from multivariable logistic regression comparing behavior (yes vs no) between groups with specific characteristics and a reference (Ref) group.

^bNHBS: National HIV Behavioral Surveillance.

Table 6. Sexually transmitted infection testing and diagnosis of men who have sex with men participants in the American Men's Internet Survey, United States, 2016.

Participant characteristics	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	1098	778 (70.86)	<.001 ^c	264 (24.04)	<.001 ^c
Race/ethnicity					
Black, non-Hispanic	248	178 (71.77)	.40	67 (27.02)	.30
Hispanic	130	105 (80.77)	.24	39 (30.00)	.82
White, non-Hispanic	632	426 (67.41)	Ref ^b	134 (21.20)	Ref ^b
Other or multiple races	88	69 (78.41)	.40	24 (27.27)	.94
Age (years)					
15-24	59	47 (79.66)	.52	16 (27.12)	.90
25-29	97	80 (82.47)	.28	37 (38.14)	.03
30-39	186	152 (81.72)	.45	67 (36.02)	.26
≥40	756	499 (66.01)	Ref ^b	144 (19.05)	Ref ^b
NHBS^d city resident					
Yes	568	426 (75.00)	.03	166 (29.23)	.004
No	530	352 (66.42)	Ref ^b	98 (18.49)	Ref ^b
Recruitment website type					
Gay social networking	139	92 (66.19)	.44	29 (20.86)	.90
General gay interest ^e	<5	N/A ^f	N/A	N/A	N/A
General social networking	651	439 (67.43)	Ref ^b	124 (19.05)	Ref ^b
Geospatial social networking	303	242 (79.87)	.004	110 (36.30)	<.001
HIV negative or unknown status	9068	3639 (40.13)	Ref ^b	813 (8.97)	Ref ^b
Race/ethnicity					
Black, non-Hispanic	631	313 (49.60)	.01	77 (12.20)	.20
Hispanic	1181	566 (47.93)	.16	154 (13.04)	.08
White, non-Hispanic	6441	2409 (37.40)	Ref ^b	498 (7.73)	Ref ^b
Other or multiple races	815	351 (43.07)	.35	84 (10.31)	.59
Age (years)					
15-24	2659	935 (35.16)	<.001	216 (8.12)	.21
25-29	1596	848 (53.13)	<.001	213 (13.35)	<.001
30-39	1228	617 (50.24)	.01	155 (12.62)	.23
≥40	3585	1239 (34.56)	Ref ^b	229 (6.39)	Ref ^b
NHBS^d city resident					
Yes	3656	1762 (48.19)	<.001	438 (11.98)	<.001
No	5412	1877 (34.68)	Ref ^b	375 (6.93)	Ref ^b
Recruitment website type					
Gay social networking	1119	338 (30.21)	<.001	55 (4.92)	.05
General gay interest	73	27 (36.99)	.35	5 (6.85)	.53
General social networking	6125	2264 (36.96)	Ref ^b	469 (7.66)	Ref ^b

Participant characteristics	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	1731	1001 (57.83)	<.001	284 (16.41)	<.001

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

^bWald chi-square from multivariable logistic regression comparing behavior (yes vs no) between groups with specific characteristics and a reference (Ref) group.

^cWald chi-square from multivariable logistic regression comparing behavior (yes vs no) between HIV-positive participants and HIV-negative or unknown-serostatus participants. Model controlled for race/ethnicity, age, National HIV Behavioral Surveillance system residency, and website type.

^dNHBS: National HIV Behavioral Surveillance.

^eTo prevent identification, data for groups with less than five participants are not presented.

^fN/A: not applicable.

Acknowledgments

The study was funded by a grant from the MAC AIDS Fund and by 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. It was reviewed and handled by an independent editor.

Multimedia Appendix 1

American Men's Internet Survey 2016 questionnaire.

[[PDF File \(Adobe PDF File\), 450KB - publichealth_v5i1e11313_app1.pdf](#)]

References

1. Finlayson TJ, Le B, Smith A, Bowles K, Cribbin M, Miles I, Centers for Disease ControlPrevention (CDC). 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 Surveill Summ* 2011 Oct 28;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. 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 12;22(8):2413-2425. [doi: [10.1007/s10461-018-2168-4](#)] [Medline: [29948340](#)]
6. Office of Policy Development and Research (PD&R). HUD USPS ZIP Code Crosswalk Files URL: https://www.huduser.gov/portal/datasets/usps_crosswalk.html [accessed 2018-12-14] [[WebCite Cache ID 74fO5kWxl](#)]
7. Look Up a ZIP Code. USPS URL: https://tools.usps.com/go/ZipLookupAction_input [accessed 2018-12-14] [[WebCite Cache ID 74fOESDbx](#)]
8. Centers for Disease Control and Prevention. HIV Surveillance Special Report 15. 2016. 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-hssr-nhbs-msm-2014.pdf> [accessed 2018-06-14] [[WebCite Cache ID 74fOKs6SB](#)]
9. Ingram DD, Franco SJ. 2013 NCHS Urban-Rural Classification Scheme for Counties. *Vital Health Stat* 2 2014 Apr(166):1-73 [[FREE Full text](#)] [Medline: [24776070](#)]

Abbreviations

aOR: adjusted odds ratio
AMIS: American Men's Internet Survey
MSM: men who have sex with men
NHBS: National HIV Behavioral Surveillance
STI: sexually transmitted infection
ZIP: zone improvement plan

Edited by A Lansky; submitted 16.06.18; peer-reviewed by S Tregear, D Callander; comments to author 14.08.18; revised version received 28.11.18; accepted 30.11.18; published 20.02.19.

Please cite as:

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;5(1):e11313

URL: <http://publichealth.jmir.org/2019/1/e11313/>

doi: [10.2196/11313](https://doi.org/10.2196/11313)

PMID: [30785405](https://pubmed.ncbi.nlm.nih.gov/30785405/)

©Maria Zlotorzynska, Patrick Sullivan, Travis Sanchez. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 20.02.2019. 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

Applying National Estimates of Adults With Indications for Pre-Exposure Prophylaxis to Populations of Men Who Have Sex With Men and People Who Inject Drugs in Colorado: Modeling Study

Jennifer A Donnelly¹, MPH; Thomas T Deem², MSN, FNP-BC; Megan A Duffy¹, MPH; Anita K Watkins¹, MPH; Alia A Al-Tayyib³, PhD; Daniel J Shodell¹, MD, MPH; Mark Thrun⁴, MD; Sarah E Rowan³, MD

¹Colorado Department of Public Health and Environment, Denver, CO, United States

²Metro Community Provider Network, Denver, CO, United States

³Denver Public Health, Denver Health Medical Center, Denver, CO, United States

⁴Gilead Sciences, Foster City, CA, United States

Corresponding Author:

Sarah E Rowan, MD
Denver Public Health
Denver Health Medical Center
605 Bannock Street
Denver, CO, 80204
United States
Phone: 1 303 602 6392
Email: Sarah.Rowan@dhha.org

Abstract

Background: Oral pre-exposure prophylaxis (PrEP) is a highly effective option for HIV prevention. To realize the full benefit of PrEP at the population level, uptake must reach those at the greatest risk of HIV acquisition. Guidance published by Centers for Disease Control and Prevention (CDC) suggests that the number of individuals with indications for PrEP is 1.1-1.2 million nationally based on survey data of key populations and local transmission patterns. We applied these estimates at state and county levels to determine the number of individuals who might benefit from PrEP locally and compared our estimates to CDC-published estimates for Colorado.

Objective: This analysis aimed to produce estimates of key populations with indications for PrEP in Colorado as a whole and by county type. These estimates will be used for public health strategic planning for HIV prevention goals at the state and county jurisdictional levels.

Methods: Colorado population estimates were obtained from the state demography office, which utilizes US decennial census data and input from county and local agencies to forecast the population. We limited our analysis to adults aged 18-59 years to be consistent with CDC methodology for PrEP estimates. We performed a literature review to define the best population-level percentages to determine numbers of HIV-negative men who have sex with men (MSM) and people who inject drugs (PWID) in Colorado. These percentages were applied to the state and to each county by its rural-urban designation. Finally, CDC-derived percentages of MSM and PWID with indications for PrEP were applied to these estimates to determine numbers of MSM and PWID who may benefit from PrEP use.

Results: In 2017, 3,252,648 adults aged 18-59 years were living in Colorado. By applying published estimates of percentages of men who had sex with other men in the past 12 months, we determined that 41,353-49,624 adult males could be considered sexually active MSM. We estimated that 9758-13,011 adults aged 18-59 years were likely to have injected drugs in the past 12 months. By accounting for numbers of people living with HIV in those categories and applying the CDC PrEP percentages of MSM and PWID with indications for PrEP nationally, we estimated that 8792-12,528 MSM and PWID in Colorado had indications for PrEP; this number is smaller than that estimated by CDC, although within the lower CI limit.

Conclusions: By employing a simple framework consisting of census data, literature review, population estimates, and national estimates for PrEP indicators, we derived estimates for potential PrEP use in our state. Statewide estimates of key populations

by state and county type will enable health officials to set informed goals and track progress toward optimizing PrEP uptake. This formula may be applicable to other states with similar epidemics and resources.

(*JMIR Public Health Surveill* 2019;5(1):e11113) doi:[10.2196/11113](https://doi.org/10.2196/11113)

KEYWORDS

HIV prevention; key population estimates; men who have sex with men; pre-exposure prophylaxis; people who inject drugs

Introduction

Like many states, Colorado observed a decline in HIV diagnosis rates between 2005 and 2015, leading regional stakeholders to consider the possibility of ending the HIV epidemic in the state [1]. The Joint United Nations Programme on HIV/AIDS 90-90-90 movement further propelled that work by introducing international population-level goals for the diagnosis and care of people living with HIV [2]. Colorado is close to reaching the goals of 90-90-90, yet the number of individuals who are newly diagnosed with HIV annually has stabilized and in some regions increased in the past few years, marking a change from the prior decade of declining HIV rates [3]. It is clear that to complement the prevention benefits of optimal treatment for people living with HIV, pre-exposure prophylaxis (PrEP) for people at risk for acquiring HIV is also a necessary tool to end new transmissions and propel efforts to end the epidemic [4].

Unlike the parameters of the 90-90-90 initiative, targets for PrEP use have not been well established at the local level. The National HIV/AIDS Strategy recommends a 500% increase in PrEP prescriptions by 2020, though this is currently considered a developmental indicator, expected to be modified as additional information becomes available [5]. In their paper in the *Morbidity and Mortality Weekly Report*, Smith et al analyzed survey data from the National Health and Nutrition Examination Survey, National Survey on Drug Use and Health, and National HIV Behavioral Surveillance System and concluded that nationally, approximately 1.2 million individuals in the United States had the Centers for Disease Control and Prevention (CDC)-recommended indications for PrEP [6]. They determined that 24.7% of men who had sex with men (MSM) in the past 12 months, 18.5% of people who injected drugs (PWID) in the past 12 months, and 0.4% of sexually active heterosexual adults had indications for PrEP. A subsequent estimate published in 2018 refined the national estimate to reflect regional differences in HIV transmission risk groups and to identify the number of black, Hispanic or Latino, and white individuals in each transmission group with indications for PrEP in each state. Using this method, CDC reported that 24,310 (95%CI 13,480-44,430) individuals in Colorado have indications for PrEP [7]. To test this estimate and to better understand the need for PrEP in Colorado as a whole and in each county, we utilized a variety of population-level data sources to determine numbers of HIV-negative MSM and PWID in the state who are likely to have indications for PrEP. These estimates were obtained by applying population-level percentages of MSM and PWID to the state as a whole and by calculating county-level estimates

based on each county's rural-urban designation and the percentage of estimated MSM and PWID in each county type. We then applied national CDC-derived percentages of MSM and PWID with PrEP indications to our MSM and PWID estimates to determine the number of individuals with indications for PrEP in Colorado. Our goal was to develop a formula to derive estimates of PrEP indications that could lead to timely, precise, and actionable goals for PrEP uptake at state and local levels.

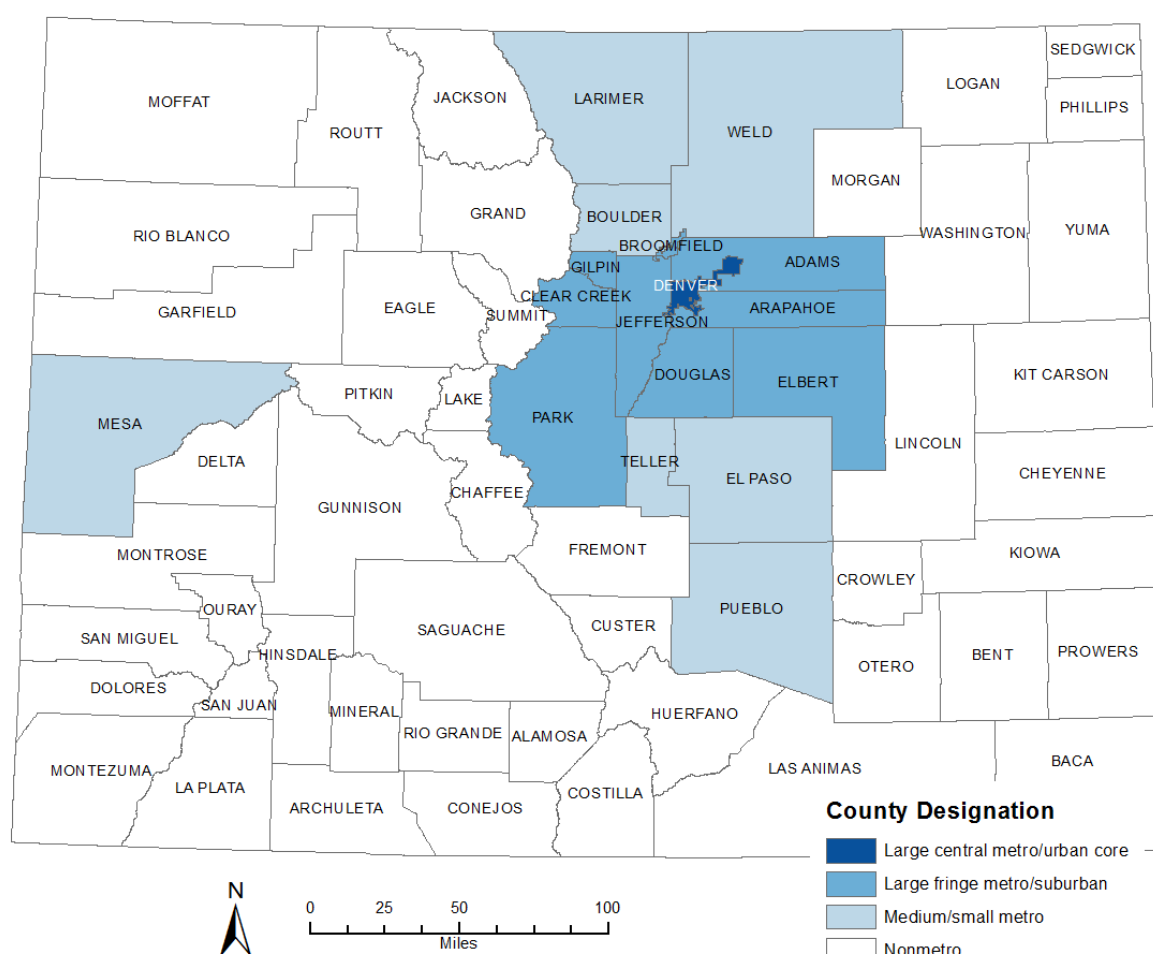
Methods

Colorado Population Estimates

Census data for the total population of Colorado was obtained from the Colorado State Demography website, which uses the most recent decennial US Census Bureau data and input from county and local agencies to estimate and forecast the population for intercensal years to forecast population at the state and county levels [8]. Numbers of adults aged 18-59 years were extracted from the population totals as was the distribution by sex and the geographic distribution of adults by county. We aggregated Colorado counties using the 2013 National Center for Health Statistics Urban-Rural Classification Scheme for Counties in the following groupings: large central metro/urban core, large fringe metro/suburban, medium/small metro, and nonmetro [9] (see [Figure 1](#)). Forecasted 2017 county population data were used to determine total county population, sex, and age stratifications [8].

Estimates for Men Who Have Sex With Men and People Who Inject Drugs Not Known to Be Living With HIV

We conducted a literature review to determine the most relevant and accurate percentages of adult populations that were likely to fall into the MSM and PWID categories. To reflect populations at increased risk of HIV and to remain consistent with the selection criteria used by Smith and colleagues in their first national PrEP estimates, for our final calculations, we selected references that included estimates of proportions of men who had had sex with other men in the past 12 months and proportions of adults who had injected drugs in the past 12 months [6]. We then applied the national and regional estimates from the literature to the Colorado adult male and overall adult populations. Once the estimated numbers of MSM and PWID in Colorado were calculated, we utilized state HIV surveillance data to subtract the number of individuals known to be living with HIV from each group to determine the potential number of individuals at increased risk of HIV acquisition [3].

Figure 1. Map of Colorado counties by urbanicity designation.

Textbox 1. Formulas to determine the number of individuals with indications for pre-exposure prophylaxis (PrEP) by the men who have sex with men (MSM) and people who inject drugs (PWID) subgroups.

- Estimated number of MSM with PrEP indications=[(Adult male population×Estimated percentage MSM in the past 12 months)–HIV-positive MSM]×Estimated percentage of MSM with PrEP indications (24.7%)
- Estimated Number of PWID with PrEP indications=[(Adult population×Estimated percentage of PWID in the past 12 months)–HIV-positive PWID]×Estimated percentage of PWID with PrEP indications (18.5%)

Individuals with a history of both male-male sex and injection drug use were subtracted from the MSM estimates. The exercise was repeated at the county level by applying varying percentages of individuals estimated to have had male-male sex or injected drugs in the past 12 months by county type (urban-rural designation) and then subtracting the number of MSM or PWID living with HIV in each county type from the total.

Estimates for Men Who Have Sex With Men and People Who Inject Drugs With Indications for Pre-Exposure Prophylaxis

Using CDC-derived percentages of individuals in the MSM and PWID categories with indications for PrEP, we calculated the number of individuals in these categories by applying

percentages to Colorado as a whole and by summing estimates for each county type [6] (see [Textbox 1](#) for the complete formula).

Results

Colorado Population Estimates

The number of adults aged 18-59 years living in Colorado in 2017 was 3,252,648. Of those, 50.58% (1,654,138/3,252,648) were men. Overall, 52.66% (1,713,125/3,252,648) of the population lived in counties classified as large central metro/urban core or large fringe metro/suburban[8] (see [Table 1](#) for the distribution of adults by county urbanicity type).

Table 1. Distribution of Colorado population in 2017 by urbanicity.

Urbanicity type	Counties	Male population	Female population	State population
State, n (%)	64 (100.00)	1,654,138 (100.00)	1,598,510 (100.00)	3,252,648 (100.00)
Urbanicity^a, n (%)				
Large central metro/urban core	1 (1.56)	228,148 (13.79)	220,601 (13.80)	448,749 (13.79)
Large fringe metro/suburban	9 (14.06)	633,801 (38.32)	630,575 (39.45)	1,264,376 (38.87)
Medium/small metro	7 (10.93)	585,451 (35.39)	568,440 (35.56)	1,153,892 (35.47)
Nonmetro	47 (73.43)	206,737 (12.50)	178,894 (11.19)	385,631 (11.85)

^aCounties assigned to urbanicity in accordance with the 2013 National Center for Health Statistics urban-rural classification scheme.

Estimates for Men Who Have Sex With Men and People Who Inject Drugs Not Known to Be Living With HIV

Our literature review yielded 4 publications that characterized the percentage of the given populations of males that could be considered MSM, 3 of which included estimates for male-male sexual activity in the past 12 months [10-13]. The review yielded 2 publications that described population proportions of PWID, 1 of which specifically characterized the percentage of populations with a history of injection drug use in the past 12 months [11,14] (see Table 2 for details of the reviewed publications). The estimates most applicable to our analysis were described by Oster et al [11], who suggested that 2.5% of the male population nationally and 3% of the males in the Western United States had a history of sex with men in the past 12 months. At the county level, estimates for recent male-male sexual activity ranged from 1.1% of adult males in nonmetro counties to 4.4% of adult males in large central metro counties [11]. We compared these findings with estimates produced by Grey et al [10], who suggested that 2.4% of the adult male population nationally had a history of sex with men in the past 12 months and that 3.8% of men in Colorado had had sex with a man in the past 5 years. Male-male sexual activity in Colorado in the past 12 months was not described in the study by Grey et al. County-level estimates ranged from 1% to 6.8% in nonmetro and large central metro counties, respectively [11].

Oster et al [11] also estimated that 0.3% of the adult population nationally and 0.4% of the adult population in the Western United States had a history of injection drug use in the past 12 months. Estimates of recent injection drug use at the county urbanicity levels ranged from 0.3% in the large central metro counties to 0.5% in the nonmetro counties [11].

By applying the national and regional MSM and PWID percentages by Oster et al [11] to the Colorado population as a whole, regardless of county type, we determined that statewide, 41,353-49,624 males aged 18-59 years were likely to be MSM in the past 12 months, depending on the whether we applied the

national MSM estimates (lower estimate) or Western US MSM estimates (higher estimate). Using the national and Western US estimates for PWID, we determined that 9758-13,011 individuals (males and females aged 18-59 years) were likely to have injected drugs in the past 12 months, also with the higher estimate derived from estimates for the Western United States rather than nationally. After accounting for individuals living with HIV in those 2 categories, we determined that 33,199-41,470 MSM and 9098-12,351 PWID were eligible for the PrEP indications analysis.

When we applied estimates of percentages of individuals with MSM and PWID behavior in the past 12 months by county urbanicity type to the number of adults ages 18-59 years living in each county type in Colorado, we determined that 36,354 males were estimated to have had male-male sex in the past 12 months and 10,143 individuals were likely to have injected drugs in the past 12 months. After accounting for MSM and PWID living with HIV in Colorado, we estimated that 28,200 MSM and 9483 PWID were eligible for the PrEP indications analysis.

Estimates for Men Who Have Sex With Men and People Who Inject Drugs With Indications for Pre-Exposure Prophylaxis

Applying published estimates for the proportions of MSM and PWID with CDC-recommended indications for PrEP, we determined that the MSM population with indications for PrEP ranged from 8200 to 10,243 males statewide. We estimated that 1683-2285 PWID had indications for PrEP statewide. By applying the formula for PrEP indications using the county-level MSM and PWID estimates, we determined that 6965 MSM in Colorado were likely to have indications for PrEP and 1827 PWID were likely to have indications for PrEP (see Table 3 for the estimated numbers of MSM and PWID with indications for PrEP statewide and by county type). Majority of MSM with PrEP indications were located in a large central metro county (Denver) or in large fringe metro counties, while the number of PWID with PrEP indications was more evenly distributed throughout the state (see Figures 2 and 3).

Table 2. Literature reviewed for population estimates [10-14].

Study and behavioral characteristic	Time frame	US population percentage ^a , n (%)	Colorado or Western US population percentage, n (%)	Population age (years)	Geographic distribution
Oster et al, 2015					
MSM ^b	Lifetime	5,933,000 (4.7)	1,696,000 (5.7)	≥13	National, regional, and county level
MSM	Past 12 months	3,156,000 (2.5)	893,000 (3.0)	≥13	National, regional, and county level
PWID ^c	Lifetime	5,949,000 (2.3)	1,980,000 (3.3)	≥13	National, regional, and county level
PWID	Past 12 months	776,000 (0.3)	240,000 (0.4)	≥13	National, regional, and county level
Grey et al, 2016: MSM	Past 5 years	4,503,080 (3.9)	73,357 (3.8)	≥18	National, state, county, and core-based statistical areas
Purcell et al, 2012					
MSM	Lifetime	8,476,848 (6.9)	N/A ^d	≥13	National
MSM	Past 5 years	4,791,262 (3.9)	N/A	≥13	National
MSM	Past 12 months	3,562,733 (2.9)	N/A	≥13	National
Lieb et al, 2011: MSM	Lifetime	12,986 (6.9)	N/A	≥18	State level
Lansky et al, 2014: PWID	Lifetime	6,612,488 (2.6)	N/A	≥13	National level

^aMSM: percentage of adult males; PWID: percentage of all adults.

^bMSM: men who have sex with men.

^cPWID: people who inject drugs.

^dN/A: not applicable.

Table 3. Estimated number of individuals with indications for pre-exposure prophylaxis in Colorado.^a

Key population	Statewide (United States based) ^a	Statewide (Western United States based)	Statewide (sum of county-level estimates)	Large central metro	Large fringe metro	Medium/small metro	Nonmetro
Men who have sex with men, n	8200	10,243	6965	1536	3311	1676	442
People who inject drugs, n	1683	2285	1827	207	669	610	341
Total, n	9683	12,528	8792	1743	3980	2286	783

^aThis table presents only “n” values. All estimates were derived by applying the CDC percentages for PrEP in MSM (24.7%) and PWID (18.5%) across all estimates.

Figure 2. Map of Colorado counties with numbers of men who have sex with men (MSM), including MSM that also inject drugs (MSM/PWID) with indications for pre-exposure prophylaxis (PrEP).

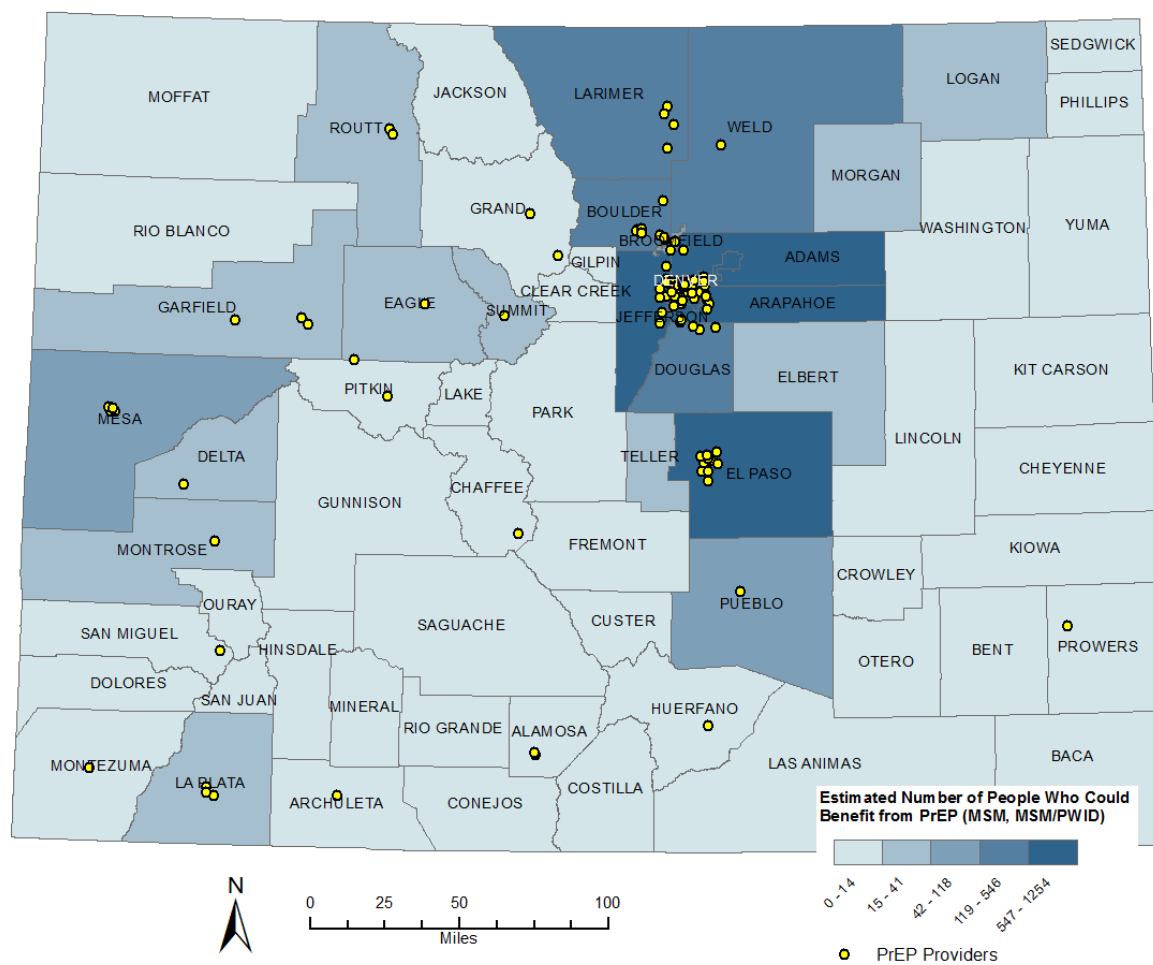
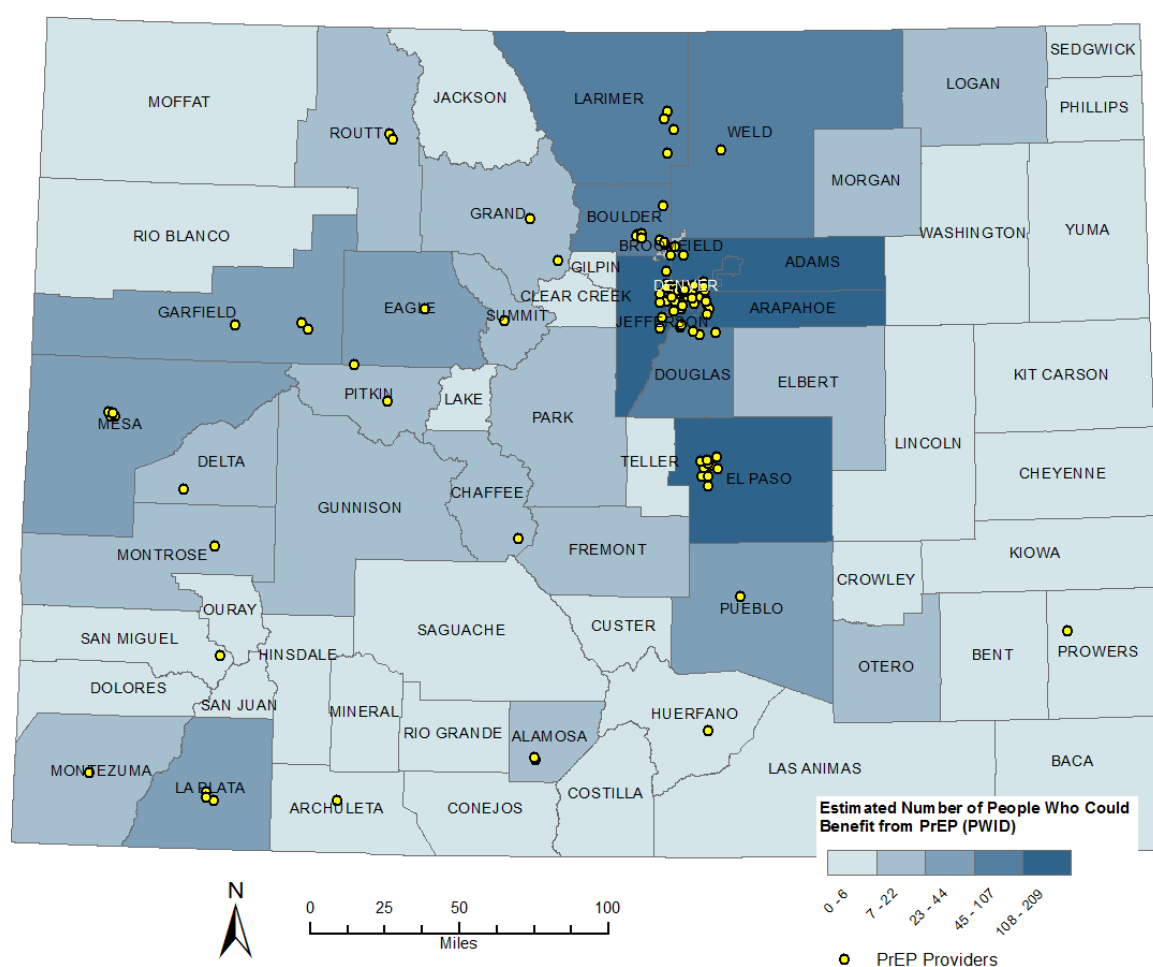


Figure 3. Map of Colorado counties with numbers of people who inject drugs (PWID) with indications for pre-exposure prophylaxis (PrEP).

Discussion

Principal Findings

Of the 3,252,648 individuals aged 18-59 years residing in Colorado in 2017, we determined that 8792-12,528 MSM and PWID were likely to have indications for PrEP as described in the 2015 CDC guidelines for PrEP use. Of those, approximately 81%-82% were MSM, and 65% lived in large central or large fringe metro areas. By target population, 70% of MSM lived in large central or large fringe metro areas and 50% of PWID lived in large central or large fringe metro areas, reflecting the more rural distribution of injection drug use. The county-level distribution of PrEP for MSM was similar to the distribution of HIV among MSM in Colorado, in which 70%-75% of the MSM living with HIV resided in the 5-county Denver Metro area [3].

As has been observed with national PrEP estimates, the number of individuals estimated to have an indication for PrEP in Colorado was close to the number of individuals living with HIV in the state [6]. Some authors have suggested that an alternate method for estimating PrEP need could employ HIV diagnoses as a reference point [15]. In 2016, approximately 8500 MSM, 1400 MSM-PWID, and 500 PWID were known to be living with HIV in Colorado, which are similar number overall to the estimates for PrEP we present here [3]. However, the number of non-MSM-PWID living with HIV in Colorado

is significantly lower than the number of individuals we estimated were PWID with indications for PrEP.

A more nuanced approach to using HIV diagnosis data to estimate numbers of individuals with indications for PrEP at the local level was recently published by the CDC. This approach relies on the ratio of the percentage of PWID diagnosed with HIV to the percentage of MSM diagnosed with HIV in a given area. This ratio is further refined by applying race and ethnicity data [16]. This approach accounts for heterogeneity in transmission risk factors regionally. Interestingly, the published estimates for MSM with PrEP indications in Colorado were significantly higher than those we present here. This is partly due to the weighted proportion of HIV diagnoses in Colorado who are MSM, but more importantly, this more recent publication relied on much larger MSM estimates based on the report of sexual activity with men in the past 5 years as opposed to that in the past 12 months, as is in the PrEP guidelines [7]. This significant methodological difference is the key to understanding the variation in estimates. Given that sexual activity varies for people over time, especially in the light of advances in PrEP and understanding of treatment as prevention, having both estimates gives a broader picture of the potential for PrEP use in the state. This alternative methodological approach is limited in states with lower numbers of HIV diagnoses annually, where the transmission category distribution may vary significantly year to year. The recent CDC

publication also does not describe estimates for numbers of individuals with PrEP indications at the county level, which are crucial for local resource allocation and planning, in particular as it relates to support for PrEP clinical services in underserved counties (see [Figures 2 and 3](#)). The method we present here enables jurisdictions to generate targeted PrEP estimates based on more timely local data, which can then be compared to nationally generated state-level estimates as they become available.

Limitations

Our analysis is subject to several limitations. Most notably, we did not include estimates of heterosexual individuals with indications for PrEP. We initiated this process using similar methodology as that employed for MSM and PWID estimates but deemed the estimates likely to be inaccurate as the epidemiology of HIV in Colorado is heavily skewed toward MSM with a much lower percentage of individuals living with HIV in the state being heterosexual than is the case nationally. Similarly, we were not able to estimate the prevalence of transgender people in Colorado or subsequent numbers of transgender individuals with indications for PrEP. The inclusion of these 2 populations would make the analysis richer and more complete and will be the focus of future efforts at the state and local health department levels.

The analysis is also limited by our reliance on national and regional estimates of sexual behavior and injection drug use, which may or may not be accurate for Colorado. In particular, Colorado has been heavily affected by the opioid epidemic and may have a significantly higher number of PWID than presented here [17]. Also, both the behavioral estimates obtained from the papers by Grey et al and Oster et al as well as the PrEP indication estimates by Smith et al rely on data from the National Health and Nutrition Examination Survey, which is limited to individuals who are not institutionalized or homeless [6,10,11]. This exclusion is likely to lead to an underestimation of the prevalence of recent injection drug use, thereby leading to an underestimate of the number of PWID with indications for PrEP. A revised estimate of the prevalence of PWID that accounts for

homeless and incarcerated populations would be of great benefit. Finally, as noted by Smith and colleagues, as sample sizes get smaller, estimation is more unstable [6]. Therefore, the estimates we have presented have been used specifically for program planning purposes and are limited in generalizability.

Implications and Final Summary

Although simple in its methodology, this exercise is a practical means to estimate the need for PrEP at the state and local levels. To our knowledge, this is the first instance of a state- and county-level application of national estimates. Additional methodologies using surveys and more precise population-level statistics have been employed in other jurisdictions as alternate approaches to estimating PrEP demand or PrEP targets [18-20]. However, to obtain preliminary estimates, especially for states with relatively smaller epidemics or for whom resources for population-level HIV prevention analyses are more limited, we offer this approach as a feasible option that provides immediate, actionable estimates that can be quickly revised as new estimates for key populations become available.

As Colorado and its individual metro areas develop strategic plans to end the HIV epidemic, measurable targets for PrEP uptake help direct efforts to the most relevant populations and regions [1]. While the estimates for PrEP indication vary when derived by county type compared to the statewide estimates we produced and compared to the most recent estimates for Colorado from CDC, taken as a whole, these estimates provide a range of numbers that can serve as targets for PrEP use, possibly in a stepwise manner starting with more conservative estimates and increasing our targets as we build demand and gain capacity for PrEP provision. At this time, our state health department is conducting an analysis of insurance claims data to determine the approximate number of PrEP prescriptions filled in 2017. This will serve as a starting point for measuring progress toward optimal PrEP uptake. Ongoing education and financial support for PrEP programs will be crucial to ensuring that this highly effective intervention reaches all individuals for whom it could be beneficial.

Conflicts of Interest

None declared.

References

1. Colorado Department of Public Health and Environment. Denver, CO: Colorado Department of Public Health and Environment; 2016 Sep 16. Colorado HIV/AIDS Strategy 2017-2021 URL: https://drive.google.com/file/d/17OmT6VGzTx6LgusDDK-dgHJ_W5OStuRI/view [accessed 2018-05-16]
2. Granich R, Williams B, Montaner J, Zuniga JM. 90-90-90 and ending AIDS: necessary and feasible. *Lancet* 2017 Dec 22;390(10092):341-343. [doi: [10.1016/S0140-6736\(17\)31872-X](https://doi.org/10.1016/S0140-6736(17)31872-X)] [Medline: [28745591](https://pubmed.ncbi.nlm.nih.gov/28745591/)]
3. CDPHE HIV Surveillance Program. Colorado Department of Public Health and Environment. Denver, CO: CDPHE DCEED STI/HIV and Viral Hepatitis Branch; 2017 Dec 15. HIV Surveillance Semiannual Report 4th Quarter 2016 URL: https://drive.google.com/file/d/1rtZjIci3bQAZ4OcZkz49mNFrXoUSnpv_/view [accessed 2018-05-18]
4. Fauci A, Marston HD. Ending the HIV-AIDS Pandemic--Follow the Science. *N Engl J Med* 2015 Dec 03;373(23):2197-2199. [doi: [10.1056/NEJMp1502020](https://doi.org/10.1056/NEJMp1502020)] [Medline: [26624554](https://pubmed.ncbi.nlm.nih.gov/26624554/)]
5. Indicator Supplement. Washington, DC: The White House Office of National AIDS Policy; 2016 Dec 01. National HIV/AIDS Strategy for the United States: Updated to 2020 URL: <https://files.hiv.gov/s3fs-public/nhas-indicators-supplement-dec-2016.pdf> [accessed 2018-11-16] [[WebCite Cache ID 73ybVkJru](#)]

6. Smith D, Van Handel M, Wolitski R, Stryker J, Hall H, Prejean J, et al. Vital Signs: Estimated Percentages and Numbers of Adults with Indications for Preexposure Prophylaxis to Prevent HIV Acquisition--United States, 2015. *MMWR Morb Mortal Wkly Rep* 2015 Nov 27;64(46):1291-1295 [FREE Full text] [doi: [10.15585/mmwr.mm6446a4](https://doi.org/10.15585/mmwr.mm6446a4)] [Medline: [26606148](https://pubmed.ncbi.nlm.nih.gov/26606148/)]
7. Smith D, Van Handel M, Grey J. Estimates of adults with indications for HIV pre-exposure prophylaxis by jurisdiction, transmission risk group, and race/ethnicity, United States, 2015. *Ann Epidemiol* 2018 May 18. [doi: [10.1016/j.annepidem.2018.05.003](https://doi.org/10.1016/j.annepidem.2018.05.003)] [Medline: [29941379](https://pubmed.ncbi.nlm.nih.gov/29941379/)]
8. Colorado State Demography Office. Colorado Totals for Colorado Counties. Denver, CO: Colorado Department of Local Affairs; 2018 Jan 01. Colorado Population Data URL: <https://demography.dola.colorado.gov/> [accessed 2018-11-16] [WebCite Cache ID 73ybrbxdz]
9. Ingram D, Franco SJ. Vital Health Statistics Series 2 Number 166. Atlanta, GA: Centers for Disease Control and Prevention; 2014 Apr. 2013 NCHS Urban-Rural Classification Scheme for Counties URL: https://www.cdc.gov/nchs/data/series/sr_02/sr02_166.pdf [accessed 2018-11-16] [WebCite Cache ID 73yccySwz]
10. Grey J, Bernstein K, Sullivan P, Purcell D, Chesson H, Gift T, et al. Estimating the Population Sizes of Men Who Have Sex With Men in US States and Counties Using Data From the American Community Survey. *JMIR Public Health Surveill* 2016;2(1):e14 [FREE Full text] [doi: [10.2196/publichealth.5365](https://doi.org/10.2196/publichealth.5365)] [Medline: [27227149](https://pubmed.ncbi.nlm.nih.gov/27227149/)]
11. Oster AM, Sternberg M, Lansky A, Broz D, Wejnert C, Paz-Bailey G. Population Size Estimates for Men who Have Sex with Men and Persons who Inject Drugs. *J Urban Health* 2015 Aug;92(4):733-743 [FREE Full text] [doi: [10.1007/s11524-015-9970-3](https://doi.org/10.1007/s11524-015-9970-3)] [Medline: [26115985](https://pubmed.ncbi.nlm.nih.gov/26115985/)]
12. Purcell D, Johnson C, Lansky A, Prejean J, Stein R, Denning P, et al. Estimating the population size of men who have sex with men in the United States to obtain HIV and syphilis rates. *Open AIDS J* 2012;6:98-107 [FREE Full text] [doi: [10.2174/1874613601206010098](https://doi.org/10.2174/1874613601206010098)] [Medline: [23049658](https://pubmed.ncbi.nlm.nih.gov/23049658/)]
13. Lieb S, Fallon SJ, Friedman SR, Thompson DR, Gates GJ, Liberti TM, et al. Statewide estimation of racial/ethnic populations of men who have sex with men in the U.S. *Public Health Rep* 2011;126(1):60-72 [FREE Full text] [doi: [10.1177/003335491112600110](https://doi.org/10.1177/003335491112600110)] [Medline: [21351396](https://pubmed.ncbi.nlm.nih.gov/21351396/)]
14. Lansky A, Johnson C, Oraka E, Sionean C, Joyce M, DiNenno E, et al. Estimating the Number of Heterosexual Persons in the United States to Calculate National Rates of HIV Infection. *PLoS One* 2015;10(7):e0133543 [FREE Full text] [doi: [10.1371/journal.pone.0133543](https://doi.org/10.1371/journal.pone.0133543)] [Medline: [26214309](https://pubmed.ncbi.nlm.nih.gov/26214309/)]
15. Siegler AM, Giler-Mera R, Yeung H, Jones J, Guest J, Kramer M, et al. Distribution of Active PrEP Prescriptions and the PrEP-to-Need Ratio, US Q2. 2018 Mar 07 Presented at: Conference on Retroviruses and Opportunistic Infections (CROI) 2018; March 4-7, 2018; Boston, MA.
16. Smith DV, Van Handel M, Grey JA. By Race/Ethnicity, Blacks have Highest Number Needing PrEP in the United States, 2015. 2018 Mar 04 Presented at: Conference on Retroviruses and Opportunistic Infections (CROI); March 3-7, 2018; Boston, MA.
17. Heroin Response Work Group. Heroin in Colorado: Preliminary Assessment. Denver, CO: Heroin Response Work Group; 2017 Apr 01. URL: https://www.colorado.gov/pacific/sites/default/files/PW_ISVP_PDO_Heroin-in-Colorado-2017.pdf [accessed 2018-11-16] [WebCite Cache ID 73ydgpbUn]
18. Guido A, Packard S. Assessing the Feasibility of Pre-Exposure Prophylaxis for HIV Prevention in Maricopa County, Arizona: An Estimate of the PrEP-Eligible Population.: AIDS Education Training Center Program Pacific; 2016 Mar 01. URL: <https://bit.ly/2RsMmZP> [accessed 2018-11-16] [WebCite Cache ID 73ye2U9dk]
19. Keen P, Gray R, Telfer B, Guy R, Schmidt H, Whittaker B, et al. The 2016 HIV diagnosis and care cascade in New South Wales, Australia: meeting the UNAIDS 90-90-90 targets. *Journal of the International AIDS Society* 2018 Apr 20;21(4):1-6.
20. Zablotska I, Selvey C, Guy R, Price K, Holden J, Schmidt H, EPIC-NSW study group. Expanded HIV pre-exposure prophylaxis (PrEP) implementation in communities in New South Wales, Australia (EPIC-NSW): design of an open label, single arm implementation trial. *BMC Public Health* 2018 Dec 02;18(1):210 [FREE Full text] [doi: [10.1186/s12889-017-5018-9](https://doi.org/10.1186/s12889-017-5018-9)] [Medline: [29394918](https://pubmed.ncbi.nlm.nih.gov/29394918/)]

Abbreviations

CDC: Centers for Disease Control and Prevention

MSM: men who have sex with men

PrEP: pre-exposure prophylaxis

PWID: people who inject drugs

Edited by T Sanchez; submitted 31.05.18; peer-reviewed by A Lansky, E Rosenberg, D Purcell; comments to author 28.06.18; revised version received 07.08.18; accepted 04.09.18; published 16.01.19.

Please cite as:

Donnelly JA, Deem TT, Duffy MA, Watkins AK, Al-Tayyib AA, Shodell DJ, Thrun M, Rowan SE

Applying National Estimates of Adults With Indications for Pre-Exposure Prophylaxis to Populations of Men Who Have Sex With Men and People Who Inject Drugs in Colorado: Modeling Study

JMIR Public Health Surveill 2019;5(1):e11113

URL: <http://publichealth.jmir.org/2019/1/e11113/>

doi: [10.2196/11113](https://doi.org/10.2196/11113)

PMID: [30664481](https://pubmed.ncbi.nlm.nih.gov/30664481/)

©Jennifer A Donnelly, Thomas T Deem, Megan A Duffy, Anita K Watkins, Alia A Al-Tayyib, Daniel J Shodell, Mark Thrun, Sarah E Rowan. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 16.01.2019. 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

Programmatic Mapping to Estimate Size, Distribution, and Dynamics of Key Populations in Kosovo

Dafina Gexha Bunjaku^{1*}, MHM; Edona Deva^{2*}, MSci; Luljeta Gashi^{1*}, MSci; Pranvera Kaçaniku-Gunga^{1*}, MD; Carly A Comins^{3*}, MPH; Faran Emmanuel^{4*}, MBBS, MSci

¹National Institute of Public Health, Department of Epidemiology, Pristina, Kosovo, Albania

²Community Development Fund, Kosovo, Albania

³Bloomberg School of Public Health, Department of Epidemiology, Johns Hopkins University, Baltimore, MD, United States

⁴Centre for Global Public Health, University of Manitoba, Winnipeg, MB, Canada

* all authors contributed equally

Corresponding Author:

Carly A Comins, MPH

Bloomberg School of Public Health

Department of Epidemiology

Johns Hopkins University

610 N Wolfe St

Baltimore, MD, 21205

United States

Phone: 1 4105024274

Email: ccomins1@jhu.edu

Abstract

Background: The burden of an HIV epidemic in Kosovo lies among the key populations (KPs) of female sex workers (FSWs), men who have sex with men (MSM), and people who inject drugs (PWIDs). The available interventions for KPs are fragmented and lack sufficient and appropriate granularity of information needed to develop large-scale outreach programs.

Objective: The aim of this study was to estimate the size and distribution of these populations to create evidence for developing action plans for HIV prevention.

Methods: The programmatic mapping approach was used to collect systematic information from key informants, including geographic and virtual locations in 26 municipalities of Kosovo between February to April 2016. In level 1, information was gathered about KPs' numbers and locations through 1537 key informant interviews within each municipality. Level 2 involved validating these spots by conducting another 976 interviews with KPs congregating at those spots. Population size estimates were calculated for each spot, and finally a national-level estimate was developed, which was corrected for duplication and overlaps.

Results: Of the estimated 6814 MSM (range: 6445 to 7117), nearly 4940 operate through the internet owing to the large stigma and discrimination against same-sex relationships. Geo-based MSM (who operate through physical spots) congregate at a few spots with large spot sizes (13.3 MSM/spot). Three-fourths of the MSM are distributed in 5 major municipalities. Fridays and Saturdays are the peak days of operation; however, the number only increases by 5%. A significant number are involved in sex work, that is, provide sex to other men for money. PWIDs are largely geo-based; 4973 (range: 3932 to 6015) PWIDs of the total number of 5819 (range: 4777 to 6860) visit geographical spots, with an average spot size of 7.1. In smaller municipalities, they mostly inject in residential locations. The numbers stay stable during the entire week, and there are no peak days. Of the 5037 (range: 4213 to 5860) FSWs, 20% use cell phones, whereas 10% use websites to connect with clients. The number increases by 25% on weekends, especially in larger municipalities where sex work is mostly concentrated. Other than a few street-based spots, most spots are establishments run by pimps, which is reflective of the highly institutionalized, structured, and organized FSW network.

Conclusions: This study provides valuable information about the population size estimates as well as dynamics of each KP, which is the key to developing effective HIV prevention strategies. The information should be utilized to develop microplans and effectively provide HIV prevention services to various KPs.

(*JMIR Public Health Surveill* 2019;5(1):e11194) doi:[10.2196/11194](https://doi.org/10.2196/11194)

KEYWORDS

HIV; program development; Kosovo; prevention and control; vulnerable populations; geographic mapping; epidemiology

Introduction

Background

Kosovo is the smallest country in the Balkans region of Europe and, in terms of registered HIV cases, has the smallest HIV epidemic in the World Health Organization (WHO) Regional Office for Europe (WHO/Europe) [1,2]. A total of 97 HIV infections in Kosovo have been registered since the first reported case in 1986, and 41 AIDS-related deaths have been reported until 2015 [3]. The possibility of a high proportion of undiagnosed infections makes it difficult to estimate the overall HIV prevalence accurately and to confirm whether HIV incidence has remained stable. As no estimation process for people living with HIV has been carried out to date, based on current testing figures, it is difficult to make an estimate with any degree of accuracy [4].

Despite the small epidemic size, it is likely that the highest burden of HIV lies among various key populations (KPs), especially in men who have sex with men (MSM) and people who inject drugs (PWIDs), as seen in many neighboring countries [5]. The previous integrated biological and behavioral surveys (IBBS) rounds showed low numbers of HIV cases; however, one needs to consider that data were collected at only a couple of sites, and the study recruited a very small sample size mostly through convenience sampling [6,7]. HIV rates were low, but Hepatitis C rates among PWIDs remains a cause for concern, as it is a predictor of how widespread the HIV could become among PWIDs [8].

To date, the HIV epidemic in Kosovo has remained insignificant, yet there is a strong potential for its growth, particularly among MSM and PWIDs [9]. The experience working with these populations has demonstrated that effective interventions built around a population focus not only protect and engage members of these communities but also make a major contribution to averting a wider epidemic [10,11]. Substantial progress has been made to meet the strategic objectives and subobjectives of Kosovo HIV/AIDS strategy for 2009 to 2014 [12]. However, to improve the response to the new HIV/AIDS strategic targets for the period 2017 to 2021 [13], future interventions need to focus on KPs and should be based on WHO recommendations concerning KPs [14]. So far, the available interventions for KPs are fragmented, lack sufficient and appropriate coverage, and need a much more scientific base for program development. From a program planning perspective, it is essential to first quantify the size of KPs, understand their subtypes, and identify locations where they can be found [15].

Objectives

On the basis of these principles, this study aimed to develop size estimates of KPs in Kosovo and provide the HIV prevention program with grassroot-level information to improve program design and coverage. The key objectives were to quantify the size of KPs, understand their subtypes, and identify locations where they can be found. Data from this study generated

evidence for developing action plans for HIV prevention interventions tailored to KPs' needs to scale up the HIV response.

Methods

Programmatic Mapping Overview

In epidemiological terms, the mapping methodology employed within this study resembled a cross-sectional survey to identify KPs in the epidemiological context of person, place, and time. The study was conducted in 26 of the 38 municipalities in Kosovo. The remaining municipalities were excluded because of security issues or political unrest. Data were collected from February to April 2016. During the preparation phase, KP communities and networks were engaged to discuss and develop procedures that ensured no harm was caused to the KP. This engagement with KPs was critical and consistent, with the approach of *nothing for us without us* through an assessment of the risks and benefits of programmatic mapping. Other important elements of the preparation phase included recruitment of field teams, training of the teams comprehensively in a 4-day training workshop on all aspects of field work, and pretesting and finalization of data collection tools. Other than the core training, the field teams were also provided a 2-day refresher training in between the 2 phases of field data collection.

Geo-Mapping Approach

The broad methodological approach included 2 sequential phases. Level 1 (L1) was a systemic process of gathering information from secondary key informants on geo-locations where KPs congregate to meet sexual partners or inject drugs. Secondary key informants are people who are not KPs themselves but are knowledgeable of the presence and operations of KPs, for example, bar tenders, cab drivers, pimps, and massage parlor workers etc. A total of 1537 interviews were conducted in L1. The entire study area was divided geographically into smaller data collection units referred to as *zones*. The basic geo-administrative divisions in Kosovo called municipalities were considered as zones for this study. Data were collected in 26 out of 38 municipalities in Kosovo, as municipalities having political unrest and/or security issues were excluded. The municipalities of Prishtina and Prizren were further divided into 5 and 3 zones subsequently to a total number of 32 study zones where data were collected. The number of interviews within each municipality was kept between 35 and 60, based on population density and the geographic span of each zone. The primary outcome of L1 was the development of a spot list for each KP type within each zone. The following phase, level 2 (L2), included validation of spots identified in L1 for KP presence and activity as well as an in-depth profiling of each spot. L2 involved conducting interviews with KP informants, that is, FSWs, MSM, and PWIDs operating at the spots listed during L1 mapping. All respondents interviewed were older than 18 years and provided informed consent. Level 2 is considered to be the *validation stage*, where sites/spots mentioned by KIs in L1 are validated, and new spots are

identified. Most spots for MSM and PWIDs were validated, whereas for FSW spots, the validation process was conducted only in spots where access was possible and a social mobilizer was available to facilitate the validation process. Residential spots were also not validated owing to issues of confidentiality. Members of the sex worker community were employed as *social mobilizers* who facilitated access and entry at various sex work spots. Social mobilizers connected with network operators or venue managers (eg, brothel, bars, and night clubs etc) and, after their consent, met the KP members operating at those venues to conduct a short interview. For data collection, teams of at least two (an interviewer and a social mobilizer) visited the identified hotspot and randomly selected a member of the KP present at the spot to conduct the interview. The social mobilizer facilitated the process by identifying and encouraging the KP members to participate in the study. The interviewer then formally conducted the interview after taking informed consent. The information was noted on an L2 form, which collected information on the type of spot, number of KP members who operate at the spot on a usual or peak day and obtained more specific information on the timings and days of operation. It also inquired the number of spots each KP member utilized to adjust for duplications. A total number of 976 spots were validated in L2. This included 255 spots for FSWs, 226 spots for MSM, and 495 spots for PWIDs.

Web Mapping Approach for Men Who Have Sex With Men

Geographic mapping captures only the visible segment of KPs who operate at publicly accessible venues. Consultations with the MSM community revealed that it was important to account for the less visible/frequently active MSM who meet their clients/partners exclusively by other means (eg, through internet websites and mobile phone-based geo-social networking apps (GNA) [16,17]. Geographic mapping of MSM spots was therefore supplemented with virtual mapping of MSM who find same-sex partners primarily through these virtual sites.

The process involved working closely with the MSM community and making a list of all such MSM dating websites and mobile phone-based GNAs as the first step. These sites included global sites, county-specific sites, as well as Facebook and WhatsApp groups specific to the country. A total of 2 virtual mappers (VMs) who were knowledgeable about the use of such websites and apps were hired from the MSM community. A user profile was created at all active sites to be able to browse the websites/apps regularly. VMs logged in at each active internet site/GNA site 3 times a day for a period of 2 weeks to gather information about the total number of registered users, the number online at the particular time of visit, and the number of new registrations each day. To further understand these Web operations, the VMs contacted various virtual users randomly, introducing the research and encouraging them to participate in the study. Those who agreed to participate were invited for a face-to-face interview at a time and place of their convenience. During the interview, information was gathered on all websites/apps the MSM are registered with, multiple registrations with different identity documents at each site, the number of contacts made at each website, as well as use of geo-locations to look at overlap between virtual and geo-spaces.

Although a high nonresponse was noticed (nearly one-thirds of the MSM contacted refused to participate), a sample of 84 MSM who operate through these sites were interviewed.

Data Management and Analysis

All data collected in the field were entered in a database developed in Microsoft Excel. The team leaders, along with the data manager, were responsible for all aspects of data quality and consistency. All spots were provided a unique code, and data were checked for consistency (eg, missing information, spelling mistakes, and outliers etc) by the field supervisors before forms were submitted to the data management team. Data entry was done at the project office in Pristina, under the supervision of a database supervisor.

KP size estimates were calculated for each spot. As a first step, all reported minimum and maximum KP estimates were summed up to get a range of estimate for each spot, which were rolled up into municipality estimates, while adjusting for duplication, that is, multiple spots visited by the same KP and also for invisible KPs who do not come to spots, using the following formula, where E_i =the adjusted estimated number of the KP, si =crude estimate of KPs for the spot; p_i =proportion of population who visit more than 1 spot in a day, and m_i =average number of spots visited in a day.

$$E_i = si(1 - p_i) + [1]$$

The municipality estimates were aggregated into national-level estimates after adjusting for invisibility factor. Finally, average spot sizes (number of KPs per spot) were determined as well as the density of each KP per 1000 adult men or women was also calculated. The adult male and female population of Kosovo was used as a denominator [18], whereas the numerator was the estimate produced by this study.

Ethical Considerations

This appraisal was designed to meet international ethical protocols by taking effective measures to avoid risk, protect individuals' rights, and ensure safety of all study participants. Ethical approval for the study was taken from the Ethical Review Board at the Ministry of Health, Kosovo. In addition, specific measures were taken to ensure the safety of the field teams and KPs. Data were collected only in municipalities which were deemed safe and received support from the local police. The KP community was given the power to make decisions on how this project was implemented, and all their concerns and suggestions were duly incorporated in the research protocol. Recruitment of participants was conducted only after describing the study procedures and obtaining informed consent. A nonidentifying coding system was used to track study data while assuring nondisclosure of participants' identities. All participants were provided information on existing services and were linked to the health and social services that are available for this community.

Results

Estimated Numbers of Key Populations

A total number of 17,670 KP members were identified in Kosovo. The largest KP found within the country was MSM,

followed by PWIDs and FSWs. Aside from a difference in size estimation, stark differences were found surrounding the operational dynamics of each population, the type and location of spots frequented, and the alternative methods for seeking partners and/or engaging in risk activity.

This study estimated a total of 5037 (range: 4213 to 5860) FSWs in Kosovo (Table 1). Although an estimated number of 4163 out of 5037 (82.6%) FSWs operate through geographic spots, nearly one-fifth (1058) are clandestine and remain unseen. These FSWs do not come to geo-physical locations but use other forms of contact to connect with their clients, that is, cell phones, internet hookups, or through personal contacts with pimps etc. Approximately, a tenth of the FSWs use the internet to find clients, but a significant overlap with geo-spots and cell phone hookups was reported. PWIDs were estimated at 5819 (range: 4777 to 6860). A larger number, that is, 4974 of 5819 were geo-spot-based, spread over 847 spots. Although this study identified no females in PWIDs, they were probably a part of the unseen PWIDs who operate undercover, not frequenting geographic locations. MSM were found to be the largest KP in Kosovo, with a total estimate of 6814 (range: 6445 to 7117). A smaller number of MSM operated at geo-spots (1874 MSMs congregate at 141 geographical spots), whereas a much larger number of MSM (ie, 4940 out of 6814) reported to operate through internet sites and mobile apps. Interviews conducted at L2 confirmed that Facebook is the most popular networking site, whereas Grinder and Planet Romeo are the most used mobile apps. A significant proportion of MSM provide sexual

services to other men in return for money and can thus be regarded as male sex workers (MSWs). An estimated number of 731 (range: 595 to 865) MSWs were found to be distributed throughout the country, mostly operating through geo-spots. Results of the analysis which calculated KP density (number of FSWs per 1000 adult females and number of MSM per 1000 adult males, etc) showed approximately 9 FSWs per 1000 adult females, 10 PWIDs per 1000 adult men, and 12 MSM per 1000 adult men in Kosovo.

Municipality Distribution of Key Populations in Kosovo

Figure 1 shows the distribution of FSWs, PWIDs, and MSM in various municipalities of Kosovo. Ferizaj, Prizren, Prishtinë, and Gjilan were municipalities with the highest number of FSWs having 16% (806/5037), 13.1 % (660/5037), 9.9% (499/5037), and 8.8% (443/5037) of the estimated FSWs, respectively. More than half of the municipalities in Kosovo had an insignificant number of FSWs (less than 2% of the total FSWs). The distribution of PWIDs varied significantly by municipality. Nearly half of the PWIDs concentrate in 3 municipalities of Prishtinë, Ferizaj, and Prizren accounting for 24.5%, 15.2%, and 9.6%, respectively. MSM seem to concentrate in fewer municipalities with Prishtinë and Prizren, reporting the highest number of MSM, that is, 2613 and 1277 MSM, respectively. In Prishtinë, of the 2613 MSM identified, 84.8% (2215/2613) MSM were Web-based, whereas in Prizren, 58.2% (743/1276) were Web-based. Other municipalities with higher numbers of MSM included Mitrovicë, Gjakovë, and Pejë.

Table 1. Estimated number of female sex workers (FSWs), people who inject drugs (PWIDs), and men who have sex with men (MSM) in Kosovo in 2016.

Characteristics	FSWs ^a	PWIDs ^b	MSM ^c
Estimated number ^d	5037 (4213-5860)	5819 (4777-6860)	6814 (6445-7117)
Estimated number at geo-spots	4163 (3482-4843)	4974 (3932-6015)	1874 (1570-2177)
Estimated number of nongeo-spot/Web-based	1058	845	4940
Estimated number of male sex workers	— ^e	—	731 (595-865)
Number of geo-spots	790	847	141
Adult population	566560 ^f	568903 ^g	568903 ^g
Number of KPs ^h per 1000 adults	8.9 per adult females	10.2 per adult males	12.0 per adult males

^aFSWs: female sex workers.

^bPWIDs: people who inject drugs.

^cMSM: men who have sex with men.

^dThe sum of geo-spot-based and nongeo-spot-based estimates might not add up to the total estimated number because of adjustments made in the final estimate to account for duplication.

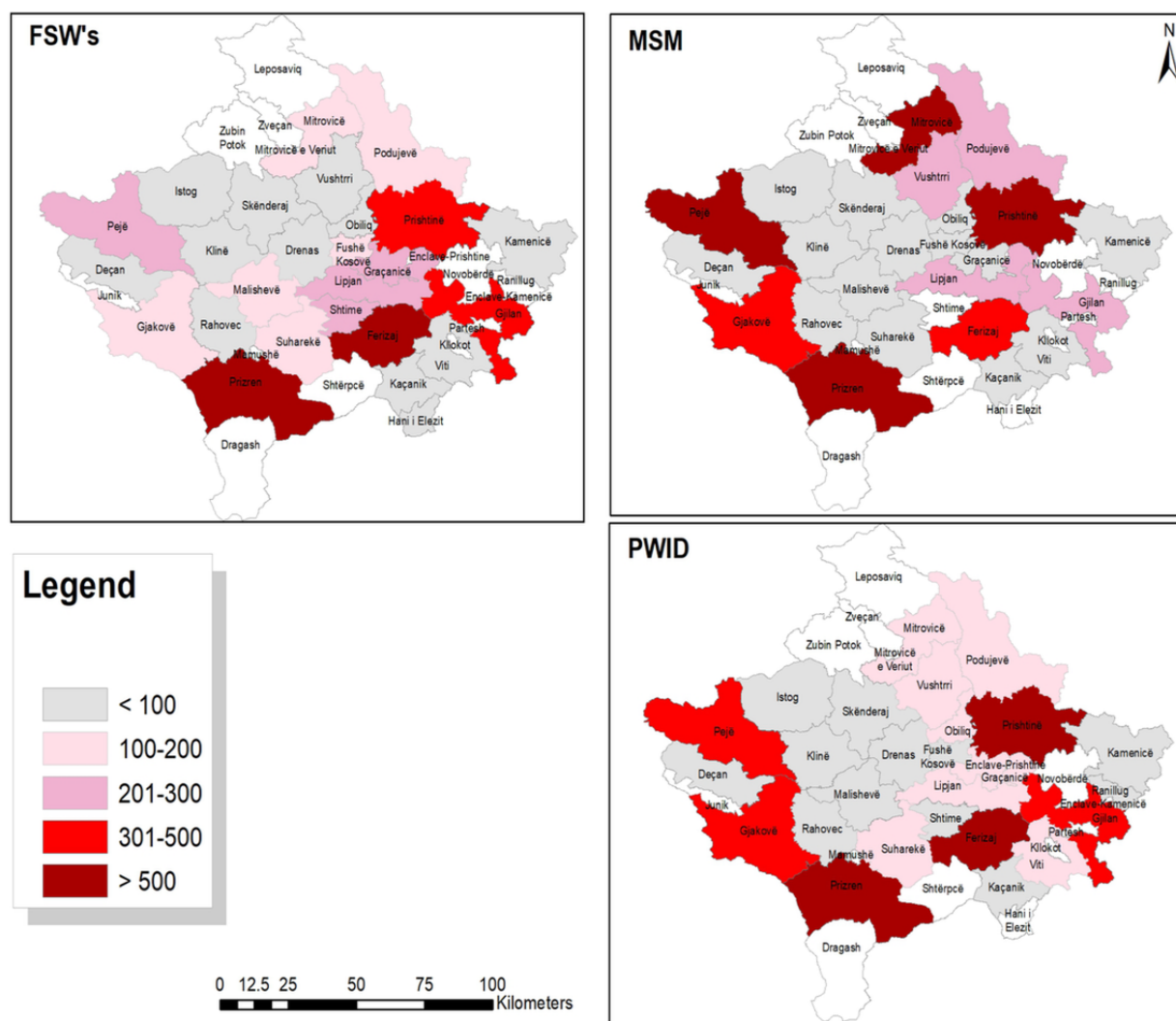
^eNot applicable.

^fAdult female population.

^gAdult male population.

^hKPs: key populations.

Figure 1. Distribution of female sex workers (FSWs), men who have sex with men (MSM), and people who inject drugs (PWIDs) by municipalities in Kosovo in 2016.



Spot Typology and Operational Dynamics

Table 2 shows the spot typologies and operational dynamics of FSWs, PWIDs, and MSM hotspots of Kosovo. FSWs mostly operate through geo-spots having an average spot size of 5.3 FSWs per spot, with significant differences across municipalities. The peak days of activity for FSWs was reported to be Friday and Saturday. Nearly three fourth FSWs operate in evenings, whereas half of them operate at night as well. More than two-thirds of FSWs find clients at restaurants with or without music and coffee shops, where they work as hostesses. The spots for PWIDs were different from FSW spots. These included abandoned buildings, establishments, public transport

stops or parks, streets, and some residential buildings as well. Street spot was the largest spot typology followed by abandoned buildings, which accounted for 41.1% (2044/4974) and 31.5% (1567/4974) of PWIDs, respectively. A much smaller proportion of MSM is visible on physical spots. Over half of MSM spots comprised street spots/parks and bus stop spots etc. Although the average spot size of MSM is large (13.3 MSM per spot), great variation existed across municipalities and by spot type. The average spot size for restaurant/coffee shop spots was high, with an estimated 19.3 MSM per spot, whereas the spot size for street spots was 11.3. Although peak days for MSM was reported to be on weekends, that is, Fridays and Saturdays, all days of the week reported substantial MSM activity.

Table 2. Spot typologies and operational dynamics for various key populations in Kosovo in 2016.

Mapping	FSWs ^a		PWIDs ^b		MSMs ^c	
	Number of spots	Estimates	Number of spots	Estimates	Number of spots	Estimates
Estimated number	790	4163	847	4974	141	1874
Spot size	5.3	5.3	5.9	5.9	13.3	13.3
Type of spots						
Abandoned buildings	0	0	240	1567	0	0
Hotel/motel/guest house	38	238	155	804	3	22
Public transport stop	13	62	42	261	63	709
Residential place	4	16	63	298	11	95
Restaurant with live music	225	1326	0	0	0	0
Restaurant/coffee shop	330	1584	0	0	18	348
Salons/casino/shops	66	273	0	0	20	318
Streets/open spaces/park	114	664	347	2044	26	382
Peak day of operation	Friday, Saturday		All week		Friday, Saturday	
Peak time of operation, (n/N)						
Morning (before 12 noon)	11.4 (29/481)		71.0 (431/607)		12.1 (16/132)	
Afternoon (12-5 PM)	41.2 (105/481)		87.2 (529/607)		44.7 (59/132)	
Evening (5-9 PM)	51.8 (132/481)		57.8 (351/607)		59.8 (79/132)	
Night (9 PM onward)	78.0 (199/481)		20.9 (127/607)		39.4 (52/132)	

^aFSWs: female sex workers.^bPWIDs: people who inject drugs.^cMSM: men who have sex with men.

Discussion

Summary

Although there have been some efforts in Kosovo in the past to estimate KP size and understand their operational dynamics [6,7,19], this study has significantly contributed to enhance our knowledge. Not only were we able to estimate the size of various KPs in Kosovo, we also identified geo-locations where HIV transmission may be maximum and developed a pragmatic typology of KPs and spots, which improved our understanding of these populations. There is enough evidence provided to understand the distribution and structure of KPs as well as know their size and operational dynamics, which is central to develop a prevention response to halt the progression of HIV and AIDS [11,15].

Our study followed a simple community-led approach, ensuring active leadership and involvement of the KPs themselves in validating estimates. A total number of 6814 MSM, 5819 PWIDs, and 5037 FSWs were identified in Kosovo. Other than FSW estimates, the estimated numbers for both MSM and PWIDs are much lower than what have been calculated through previous research [6,7,19]. The previous estimates were derived from IBBS data using multiplier technique methods. IBBS samples were collected from Prishtina and Prizren alone, which were extrapolated to develop national estimates for MSM and PWIDs. Our study found that the number of KPs in large urban centers, that is, Pristina, Ferizaj, and Prizren, was many folds

in comparison with smaller municipalities, and thus extrapolating rates from urban centers to smaller municipalities could have led to an overestimation of the size of PWIDs and MSM in Kosovo. In contrast to previous research, the estimates derived from this study followed a comprehensive exercise based on data collected at the grassroot level, led by KPs. While calculating these estimates, we adjusted for KPs visiting multiple spots and also for hidden KPs who do not come to spots. Finally, we triangulated the results with program data from service delivery programs and verified it with nongovernmental organizations (NGOs) working with these populations. The triangulation process included carefully looking at monitoring data from the programs and conducting discussion with peer workers and program managers to validate various key spots and number of KPs in various municipalities.

This research revealed in-depth information on the structure, subculture, and operational dynamics of each KP and found striking differences between them. This study identified 7 different types of geographic spots where FSWs in Kosovo congregate, find sexual partners, or engage in sexual activities. FSWs in Kosovo are centered around restaurants/coffee shops with live music and hotels/motels etc with a number of typologies involved; each having its own operational dynamics and prevention needs. Nearly all FSWs are a part of a wider but hidden network that functions undercover and are largely managed by network operators and pimps. Although sex workers are visible at a few spots, the network that operates them is

extremely organized and well-connected and mostly keeps the sex work clandestine. This nature of sex work in Kosovo is reflective of the illegality of sex work and high prevalence of sex trafficking into and through the country [20].

PWIDs represent a small proportion of an overall large population of drug users in Kosovo [19]. Unlike FSWs, PWIDs have a much wider distribution across most municipalities surveyed. Injecting drugs typically occurs in abandoned places/houses, open spaces, streets, parks etc and mostly happen during times when these places are not frequented by people. Abandoned buildings were also found to be locations where PWIDs mostly purchase drugs and inject. A predominate percentage of PWIDs visit geo-spots after noon and continue injecting through the entire day. As injecting is an everyday phenomenon, there were no peak days reported, and the estimated number of PWIDs did not change. Most injecting occurs at homes, and street-based drug injecting is not very common; thus, PWIDs are also covert and not visible. Major cities such as Pristina, Ferizaj, etc were identified as more open societies, whereas Peja, Podujeva, Skenderaj, and Gjakova were identified as very stigmatized, and PWIDs are not prepared to publicly identify themselves. Nonetheless, because of the overall drug use patterns and availability of injectable drugs in Kosovo [19], PWIDs make up a large KP identified in this mapping activity. MSM have distinct features, which makes them notably different from PWIDs and FSWs. Owing to the large stigma and discrimination faced generally, MSM concentrate in a few municipalities, that is, Prishtinë, Prizren, Mitrovicë, among others, which are considered urban and more open-minded. The higher number of MSM in these areas are representative of the higher number of students and MSM-friendly establishments such as night clubs. Most MSM operate discreetly through internet websites or mobile phones and are not visible to the general community in most municipalities, as is seen in many other countries globally [16,21]. After connecting with other peer members, they would go to either public spots or discreet locations, for example, homes, abandoned buildings, etc and usually meet after the sunset and during the late hours of the night. During discussions, MSM emphasized that they do not feel safe during the night hours, especially in the small cities, as they might be identified easier. Among MSM who do visit geographic spots, a large majority do not meet within their own municipality/city but travel to neighboring cities for physical contact.

Overall, this study has provided valuable information about the operational typologies and dynamics of these populations, which is the key to developing effective HIV prevention strategies. As part of utilization of the results, the knowledge gained from this study could be used to strategize target regions and towns where provision of services would be most effective and cost beneficial [22]. Of the key strengths of this approach, 1 lies not only in its development of estimates but also in providing a consequential distribution of KP members at different spots. Thus, larger spots with a high number of KP sizes should be the focus of prevention programs and could be the hubs of service delivery [23]. It is difficult to fully comprehend the extent and organizational dimensions of sex work or same sex without a long engagement and trust-building period with KPs. With such numbers of KPs reported, there is a need to continue a focused HIV prevention program for these populations.

Limitations

There were few limitations the study encountered regarding the geographic-spot validation process. Security issues and prohibited entry surrounding PWIDs and FSW spots inhibited the validation of various spots. Moreover, misclassification of exposures may also have occurred leading to an over or under representation of the study population. This could have resulted in an over classification of drug users as PWIDs or under classification of MSM who only engage in sexual activity with men occasionally or experimentally; thus, various drug users who share the same spots as PWIDs could have been included in the estimated PWIDs numbers, and MSM who do not engage regularly in MSM activities might have been missed in the overall estimates of MSM.

Conclusions

Overall, this study has shown that a substantial number of KPs are present in Kosovo, comprising of MSM, PWIDs, and FSWs. Each KP has a unique geographic distribution and operational dynamics by which they seek partners and/or engage in risk activity. Despite all limitations, this study's findings can guide program planners to develop appropriate HIV program implementation strategies and enhance coverage. Knowledge gained can be used to develop macroplans to strategically target regions/town and microplans to strategically deliver services, as well as to develop and strengthen the structural components of the HIV prevention programs within Kosovo.

Acknowledgments

The authors would like to thank members of KPs, NGOs, and data collection staff for their engagement and commitment. They would also like to thank The Global Fund and Community Development Fund for financial support, and also Muhammad Zubair for developing ArcGIS maps and formatting of this manuscript.

Conflicts of Interest

None declared.

References

1. Godinho J, Jaganjac N, Eckertz D, Renton A, Novotny T. World Bank. Washington, D.C; 2005. HIV/AIDS in the Western Balkans Priorities for Early Prevention in a High-Risk Environment URL: <http://documents.worldbank.org/curated/en/>

- [956161468251378385/pdf/3401810HIV0AIDS0in0the0Western0Balkans.pdf](#) [accessed 2019-02-12] [WebCite Cache ID 7686Tg3up]
2. Kosovo Agency of Statistics. Pristina, Kosovo; 2017. Statistical Yearbook of the Republic of Kosovo for 2017 URL: <http://ask.rks-gov.net/en/kosovo-agency-of-statistics/add-news/statistical-yearbook-of-the-republic-of-kosovo-for-2017> [accessed 2019-02-13] [WebCite Cache ID 76AZUO08J]
 3. UNAIDS. Republic of Kosovo: Ministry of Health; 2015. Kosovo Narrative Report 2015 URL: http://www.unaids.org/sites/default/files/country/documents/KOSOVO_narrative_report_2015.pdf [accessed 2019-02-12] [WebCite Cache ID 7686zRyv9]
 4. Burrows D, Manukyan A, Musa M, APM Global Health. WHO Regional Office for Europe. Copenhagen, Denmark; 2015. Review of the HIV Strategic Plan in Kosovo (in accordance with United Nations Security Council Resolution 1244 (1999)) URL: http://www.euro.who.int/_data/assets/pdf_file/0004/270094/Final-report-Review-of-the-HIV-strategic-plan-in-Kosovo.pdf?ua=1 [accessed 2019-02-12] [WebCite Cache ID 7687AjGAK]
 5. Pachankis JE, Hatzenbuehler ML, Hickson F, Weatherburn P, Berg RC, Marcus U, et al. Hidden from health: structural stigma, sexual orientation concealment, and HIV across 38 countries in the European MSM internet Survey. *AIDS* 2015 Jun 19;29(10):1239-1246. [doi: [10.1097/QAD.0000000000000724](https://doi.org/10.1097/QAD.0000000000000724)] [Medline: [26035323](#)]
 6. National Institute of Public Health of Kosova. Community Development Fund. Pristina, Kosovo; 2014. HIV Integrated Behavioral and Biological Surveillance Surveys-Kosovo 2014: People who Inject Drugs in Prishtina and Prizren Men who have Sex with Men in Prishtina URL: http://csgd-ks.org/wp-content/uploads/2016/09/English-Kosovo_Report_PWID_MSM.pdf
 7. National Institute of Public Health of Kosova. Community Development Fund. Pristina, Kosovo; 2014. HIV Integrated Behavioral and Biological Surveillance Surveys-Kosovo 2014: Female Sex Workers in Ferizaj URL: <http://kcdf.org/> [accessed 2019-02-12] [WebCite Cache ID 7688Hg2mF]
 8. Akbarzadeh V, Mumtaz GR, Awad SF, Weiss HA, Abu-Raddad LJ. HCV prevalence can predict HIV epidemic potential among people who inject drugs: mathematical modeling analysis. *BMC Public Health* 2016 Dec 3;16(1):1216 [FREE Full text] [doi: [10.1186/s12889-016-3887-y](https://doi.org/10.1186/s12889-016-3887-y)] [Medline: [27912737](#)]
 9. Burrows D, Manukyan A, Musa M, APM Global Health. Copenhagen, Denmark: WHO Regional Office for Europe; 2014. Review of the HIV programme in Kosovo (in accordance with United Nations Security Council Resolution 1244 (1999)) URL: http://www.euro.who.int/_data/assets/pdf_file/0003/270093/Final-report-Review-of-the-HIV-programme-in-Kosovo.pdf [accessed 2019-02-12] [WebCite Cache ID 7688XWoU6]
 10. Coates TJ, Richter L, Caceres C. Behavioural strategies to reduce HIV transmission: how to make them work better. *Lancet* 2008 Aug 23;372(9639):669-684 [FREE Full text] [doi: [10.1016/S0140-6736\(08\)60886-7](https://doi.org/10.1016/S0140-6736(08)60886-7)] [Medline: [18687459](#)]
 11. Anderson S, Cherutich P, Kilonzo N, Cremin I, Fecht D, Kimanga D, et al. Maximising the effect of combination HIV prevention through prioritisation of the people and places in greatest need: a modelling study. *Lancet* 2014 Jul 19;384(9939):249-256. [doi: [10.1016/S0140-6736\(14\)61053-9](https://doi.org/10.1016/S0140-6736(14)61053-9)] [Medline: [25042235](#)]
 12. Ministry of Health, Republic of Kosovo. Pristina, Kosovo; 2009. Health Sector Strategy 2010 - 2014 URL: http://mei-ks.net/repository/docs/ANNEX_10_-_Health_V.pdf [accessed 2019-02-12] [WebCite Cache ID 7688IQmNw]
 13. Ministry of Health, Republic of Kosovo. Pristina, Kosovo; 2016 Nov. Health Sector Strategy 2017 - 2021 URL: http://www.kryeministri-ks.net/repository/docs/Strategjia sektoriale e shendetesise final-nentor_2016_ENG.pdf [accessed 2019-02-12] [WebCite Cache ID 7688w3EWH]
 14. World Health Organization. 2014. Consolidated guidelines on HIV prevention, diagnosis, treatment and care for key populations URL: <https://www.who.int/hiv/pub/guidelines/keypopulations/en/> [accessed 2019-01-27] [WebCite Cache ID 75jxoU1iv]
 15. Emmanuel F, Isac S, Blanchard JF. Using geographical mapping of key vulnerable populations to control the spread of HIV epidemics. *Expert Rev Anti Infect Ther* 2013 May;11(5):451-453. [doi: [10.1586/eri.13.33](https://doi.org/10.1586/eri.13.33)] [Medline: [23627850](#)]
 16. Wei C, Lim SH, Guadamuz TE, Koe S. Virtual versus physical spaces: which facilitates greater HIV risk taking among men who have sex with men in East and South-East Asia? *AIDS Behav* 2014 Aug;18(8):1428-1435 [FREE Full text] [doi: [10.1007/s10461-013-0628-4](https://doi.org/10.1007/s10461-013-0628-4)] [Medline: [24077974](#)]
 17. Holloway IW, Pulsipher CA, Gibbs J, Barman-Adhikari A, Rice E. Network influences on the sexual risk behaviors of gay, bisexual and other men who have sex with men using geosocial networking applications. *AIDS and Behavior* 2015 Jun;19(Suppl 2):112-122. [doi: [10.1007/s10461-014-0989-3](https://doi.org/10.1007/s10461-014-0989-3)] [Medline: [25572832](#)]
 18. Kosovo Agency of Statistics. Pristina, Kosovo; 2016. Statistical Yearbook of the Republic of Kosovo 2016 URL: <http://ask.rks-gov.net/media/2577/statistical-yearbook-2016-ang.pdf> [accessed 2019-02-12] [WebCite Cache ID 768967N1A]
 19. UNICEF. Prishtina, Kosovo; 2009. Psychoactive substance use in Kosovo: Rapid Assessment and Response with youth, injecting drug users and prisoners URL: https://www.unicef.org/kosovoprogramme/Press_release_on_RAR_eng.pdf [accessed 2018-05-31] [WebCite Cache ID 6zpn2Mrqs]
 20. Wildt R. Preliminary study on sex trafficking and prostitution market dynamics in Kosovo. Utrecht University In cooperation with the Kosovo Population Foundation. URL: <https://www.legislationline.org/documents/id/18618>. [accessed 2019-02-24]

21. Beyrer C, Baral SD, van Griensven F, Goodreau SM, Chariyalertsak S, Wirtz AL, et al. Global epidemiology of HIV infection in men who have sex with men. *Lancet* 2012 Jul 28;380(9839):367-377 [FREE Full text] [doi: [10.1016/S0140-6736\(12\)60821-6](https://doi.org/10.1016/S0140-6736(12)60821-6)] [Medline: [22819660](https://pubmed.ncbi.nlm.nih.gov/22819660/)]
22. Bertozzi SM, Laga M, Bautista-Arredondo S, Coutinho A. Making HIV prevention programmes work. *The Lancet* 2008 Sep;372(9641):831-844. [doi: [10.1016/S0140-6736\(08\)60889-2](https://doi.org/10.1016/S0140-6736(08)60889-2)] [Medline: [18687457](https://pubmed.ncbi.nlm.nih.gov/18687457/)]
23. Wand H, Ramjee G. Targeting the hotspots: investigating spatial and demographic variations in HIV infection in small communities in South Africa. *J Int AIDS Soc* 2010 Oct 27;13:41 [FREE Full text] [doi: [10.1186/1758-2652-13-41](https://doi.org/10.1186/1758-2652-13-41)] [Medline: [20979651](https://pubmed.ncbi.nlm.nih.gov/20979651/)]

Abbreviations

FSW: female sex worker
GNA: geo-social networking apps
IBBS: integrated biological and behavioral survey
KP: key population
L1: Level 1
L2: Level 2
MSW: male sex worker
MSM: men who have sex with men
NGO: nongovernmental organization
PWIDs: people who inject drugs
VM: virtual mapper
WHO: World Health Organization

Edited by T Sanchez; submitted 31.05.18; peer-reviewed by J Zhao, A Altaf, A Abdul-Quader; comments to author 31.07.18; revised version received 23.09.18; accepted 09.11.18; published 05.03.19.

Please cite as:

Gexha Bunjaku D, Deva E, Gashi L, Kaçaniku-Gunga P, Comins CA, Emmanuel F
Programmatic Mapping to Estimate Size, Distribution, and Dynamics of Key Populations in Kosovo
JMIR Public Health Surveill 2019;5(1):e11194
URL: <http://publichealth.jmir.org/2019/1/e11194/>
doi: [10.2196/11194](https://doi.org/10.2196/11194)
PMID: [30835241](https://pubmed.ncbi.nlm.nih.gov/30835241/)

©Dafina Gexha Bunjaku, Edona Deva, Luljeta Gashi, Pranvera Kaçaniku-Gunga, Carly A. Comins, Faran Emmanuel. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 05.03.2019. 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.

Viewpoint

The Continuing Value of CD4 Cell Count Monitoring for Differential HIV Care and Surveillance

Brian Rice¹, BSc, MSc, PhD; Andrew Boule², MSc, PhD, MBchB; Sandra Schwarcz³, MPH, MD; Amir Shroufi⁴, MPhil, ChB, MB; George Rutherford³, AM, MD; James Hargreaves¹, MSc, PhD

¹Faculty of Public Health and Policy, Public Health, Environments and Society, London School of Hygiene & Tropical Medicine, London, United Kingdom

²Faculty of Health Sciences, Department of Public Health & Family Medicine, University of Cape Town, Cape Town, South Africa

³Global Health Sciences, University of California San Francisco, San Francisco, CA, United States

⁴Médecins Sans Frontières, Cape Town, South Africa

Corresponding Author:

Brian Rice, BSc, MSc, PhD

Faculty of Public Health and Policy

Public Health, Environments and Society

London School of Hygiene & Tropical Medicine

15-17 Tavistock Place

London, WC1H 9SH

United Kingdom

Phone: 44 2079272567

Email: brian.rice@lshtm.ac.uk

Abstract

The move toward universal provision of antiretroviral therapy and the expansion of HIV viral load monitoring call into question the ongoing value of CD4 cell count testing and monitoring. We highlight the role CD4 monitoring continues to have in guiding clinical decisions and measuring and evaluating the epidemiology of HIV. To end the HIV/AIDS epidemic, we require strategic information, which includes CD4 cell counts, to make informed clinical decisions and effectively monitor key surveillance indicators.

(*JMIR Public Health Surveill* 2019;5(1):e11136) doi:[10.2196/11136](https://doi.org/10.2196/11136)

KEYWORDS

CD4; HIV; differential care; antiretroviral therapy; surveillance; monitoring

The CD4 cell count has been the principal basis for assessing an HIV-infected person's level of immunosuppression and for timing initiation of antiretroviral therapy (ART) [1]. In 2015, the World Health Organization (WHO) recommended starting ART at any CD4 count, regardless of the clinical symptoms or conditions [2]. These guidelines and subsequent studies argue that, for clinical purposes, the frequency of CD4 monitoring post-ART initiation can be reduced or ceased when viral load testing is available and patients are suppressed [2-6].

The move toward universal ART, the expansion of viral load monitoring [7-9], and recommendations to reduce or cease CD4 testing post-ART initiation call into question the value of CD4 testing. Decreased support for CD4 testing could potentially not only result in reduced CD4 monitoring among people who have initiated ART, but also unintentionally lead to reduced quality of monitoring at diagnosis and the period up to ART initiation. In this communication, we highlight the continuing

role of CD4 monitoring in guiding clinical decisions and measuring and evaluating the epidemiology of HIV.

The routine collection of CD4 data at diagnosis (baseline) from laboratories and health care facilities (conducting primary or confirmatory HIV tests) continues to provide an assessment of treatment priorities. Importantly, this information remains critical in identifying late diagnosis (as often indicated by a CD4 count of <350 cells/μL) [10-12]. A late diagnosis indicates that a person is at a significantly elevated risk of HIV-related opportunistic infections and mortality, and has been identified as a primary cause of HIV-related deaths in settings where ART is widely and freely available [10,12-13].

Low CD4 counts are triggers for more intensive follow-up and care in differentiated care models. For example, WHO guidelines on advanced disease recommend that people with a CD4 count <100 cells/μL be screened for cryptococcal disease

and managed with fluconazole if asymptomatic, those with a CD4 <200 cells/ μ L receive tailored counselling, and those with a CD4 <350 cells/ μ L receive cotrimoxazole prophylaxis [14]. It is also recommended that people living with HIV with a count <100 cells/ μ L be offered the urinary lipoarabinomannan point-of-care test for tuberculosis [14]. In this way, CD4 is facilitating a shift away from symptom-based tuberculosis screening toward an approach of testing all those at high risk of disease. In Uganda, ART-naïve adults with CD4 counts \leq 250 cells/ μ L are currently screened for tuberculosis and those with CD4 \leq 100 cells/ μ L are also screened for cryptococcal antigen [15], similar to the case in South Africa [16]. In a recent three-country trial, presumptive antimicrobial treatment in patients initiating ART with CD4 counts <100 cells/ μ L resulted in a 30% reduction in 6-month mortality [17]. The routine collection and use of CD4 data have also been shown to be cost effective in promoting clinical outcomes such as disability-adjusted life years averted [18,19]. As loss to follow-up is a common outcome along the HIV care continuum [20], it is important that differentiated care models, informed by CD4 monitoring, also consider persons re-engaging in care.

Clinically mediated CD4 monitoring has also been an important feature of HIV surveillance. At the population level, the prevalence of CD4-defined late diagnoses helps monitor the effectiveness of program efforts for early identification [11,21] and is a WHO linkage to care indicator [21]. The linkage of CD4 data at diagnosis with longitudinal CD4 cell counts up to ART initiation has provided important information on trajectories of CD4 depletion between diagnosis and treatment. This information has been used at international, national, and subnational levels to back calculate from the time of diagnosis to the probable time of infection in order to estimate the incidence of HIV [22-26], estimate the prevalence of undiagnosed HIV [24-28], and assign a probable place of infection [29,30].

In a number of settings, the application of CD4-based models and their analyses are either being expanded or newly adopted. In 2016, a new model incorporating CD4 at or after diagnosis, but before ART, was introduced in the United States to estimate HIV incidence, prevalence, and undiagnosed infections [24]. Among European Union member states, a CD4 back-calculation model, which assigns probable place of HIV infection among migrant populations by estimating the time of infection and comparing it with the time of arrival in the host country, is being promoted to inform prevention programs [29,30]. In addition to informing pre-ART care and policy decisions concerning the use of ART for prevention [31], routine CD4 monitoring in South Africa has recently been used to assess the risk of subsequent loss to follow-up from care [32] and to estimate care cascade measures [33]. Although most of the CD4-based activities cited are focused in middle- and high-income settings, the promotion of HIV case surveillance [34] and the collection of CD4 within these systems will hopefully further expand the application of these methods to low- and middle-income settings, including high-prevalence settings in sub-Saharan Africa.

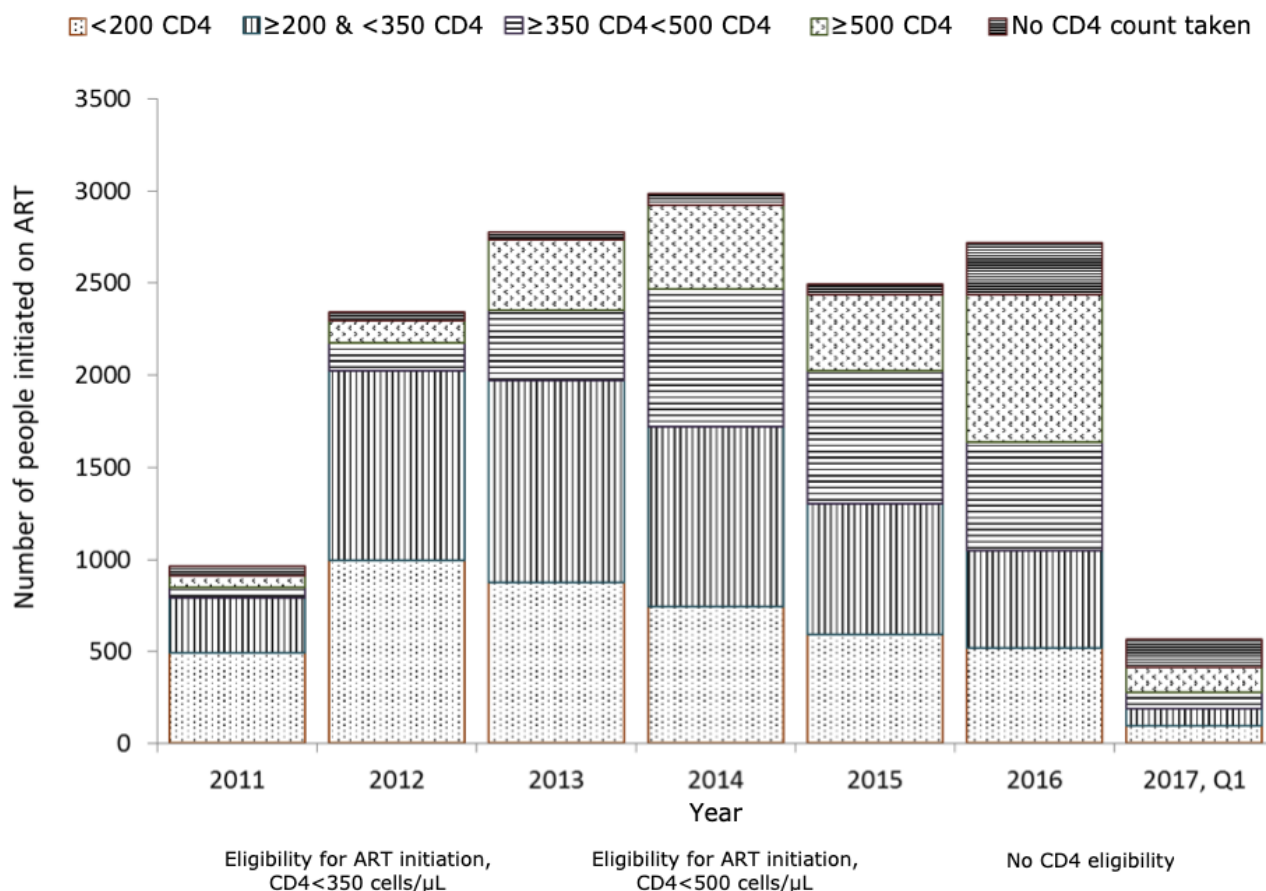
Clinical and surveillance activities relying on CD4 testing will be impaired if testing is reduced or discontinued between diagnosis and ART initiation, or in a setting where viral load testing remains suboptimal. Although it has been suggested that access to viral load monitoring in low-income, high-HIV burden settings may be limited [4], the Joint United Nations Programme on HIV/AIDS in 2016 reported that a number of resource-limited countries have drastically reduced CD4 monitoring in favor of increased viral load testing [35]. In 2018, the President's Emergency Plan for AIDS Relief announced that they will reduce their overall level of support in donor countries for CD4 testing to prioritize viral load testing [36].

Although CD4 monitoring remains essential for the detection and management of HIV-related opportunistic infections such as *Cryptococcus*, signatories of a 2017 advanced HIV position statement claim donor support for CD4 testing at the primary care level has decreased in recent years [37]. The signatories of this statement included Médecins Sans Frontières. Figure 1 presents ART initiation by CD4 count in a Médecins Sans Frontières treatment program in South Africa. The top bar of the figure suggests an increase in recent years in the number and proportion of people not receiving a CD4 test at the time of treatment initiation. The reduction of CD4 monitoring both at and subsequent to diagnosis has also been brought to the attention of a research team carrying out HIV system assessments in resource-limited settings in 2015 and 2016 (R. Harklerode, personal communication, January 2018) [38].

While highlighting the role of continuing CD4 monitoring in informing clinical and epidemiological activities, we remain fully supportive of the expansion of viral load monitoring. In several areas of clinical management, for example, the monitoring of pediatric HIV infection [39], a combination of CD4 and viral load monitoring is essential. As CD4 tests are more affordable than viral load tests in many countries [18,19] and have already been scaled up, we believe CD4 monitoring presents a model of learning for scaling up optimal and affordable viral load testing.

It is inevitable that the role of CD4 monitoring in guiding clinical decisions will become more selective. However, vigilance and oversight are required to ensure that while we reduce reliance on CD4 monitoring in virologically suppressed patients, we retain our capacity to conduct CD4 testing at diagnosis and up to ART initiation. This remains critically important in diagnosing and treating comorbidities, determining whether a person requires an advanced package of screening and care, reducing mortality, and ensuring the continuity of critical data for surveillance activities (such as estimating HIV incidence and undiagnosed infections). Although we do not advocate for routine CD4 monitoring for all, CD4 should continue to guide the clinical management of persons re-engaging in care or remaining in care but failing treatment. Capacity will preferably be retained at the population level; where this is not the case, representative sampling methods should be considered. To end the HIV/AIDS epidemic, we must obtain essential data to make informed clinical decisions and effectively monitor key surveillance indicators.

Figure 1. ART initiations in Eshowe and Mbongolwane, KwaZulu-Natal, stratified by CD4 count, 2011 to 2017. CD4 counts are measured as CD4 cells/ μ L. Q1 was from January to end March 2017. ART: antiretroviral therapy.



Acknowledgments

We wish to thank the MeSH Consortium for their support in this work. The MeSH Consortium is funded by the Bill and Melinda Gates Foundation (BMGF-OPP1120138). The views presented in this paper are those of the authors and do not necessarily reflect the views of the BMGF.

Conflicts of Interest

None declared.

References

1. World Health Organization. 2010. Antiretroviral therapy for HIV infection in adults and adolescents: recommendations for a public health approach - 2010 revision URL: <https://www.who.int/hiv/pub/arv/adult2010/en/> [accessed 2019-02-20] [WebCite Cache ID 76KxkeaYS]
2. World Health Organization. 2015. Guideline on when to start antiretroviral therapy on pre-exposure prophylaxis for HIV URL: <https://www.who.int/hiv/pub/guidelines/earlyrelease-arv/en/> [accessed 2019-02-20] [WebCite Cache ID 76Kxmnur2]
3. Davies M, Ford N, Rabie H, Fatti G, Stinson K, Giddy J, et al. Reducing CD4 Monitoring in Children on Antiretroviral Therapy With Virologic Suppression. *Pediatr Infect Dis J* 2015 Dec;34(12):1361-1364 [FREE Full text] [doi: [10.1097/INF.0000000000000912](https://doi.org/10.1097/INF.0000000000000912)] [Medline: [26379169](https://pubmed.ncbi.nlm.nih.gov/26379169/)]
4. Ford N, Stinson K, Davies M, Cox V, Patten G, Cragg C, et al. Is it safe to drop CD4+ monitoring among virologically suppressed patients: a cohort evaluation from Khayelitsha, South Africa. *AIDS* 2014 Sep 10;28(14):2003-2005. [doi: [10.1097/QAD.0000000000000406](https://doi.org/10.1097/QAD.0000000000000406)] [Medline: [25102088](https://pubmed.ncbi.nlm.nih.gov/25102088/)]
5. Ford N, Meintjes G, Pozniak A, Bygrave H, Hill A, Peter T, et al. The future role of CD4 cell count for monitoring antiretroviral therapy. *The Lancet Infectious Diseases* 2015 Feb;15(2):241-247. [doi: [10.1016/S1473-3099\(14\)70896-5](https://doi.org/10.1016/S1473-3099(14)70896-5)] [Medline: [25467647](https://pubmed.ncbi.nlm.nih.gov/25467647/)]

6. Working Group on Modelling of Antiretroviral Therapy Monitoring Strategies in Sub-Saharan Africa, Phillips A, Shroufi A, Vojnov L, Cohn J, Roberts T, et al. Sustainable HIV treatment in Africa through viral-load-informed differentiated care. *Nature* 2015 Dec 03;528(7580):S68-S76 [FREE Full text] [doi: [10.1038/nature16046](https://doi.org/10.1038/nature16046)] [Medline: [26633768](https://pubmed.ncbi.nlm.nih.gov/26633768/)]
7. Cohen M, Chen Y, McCauley M, Gamble T, Hosseinipour M, Kumarasamy N, HPTN 052 Study Team. Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl J Med* 2011 Aug 11;365(6):493-505 [FREE Full text] [doi: [10.1056/NEJMoa1105243](https://doi.org/10.1056/NEJMoa1105243)] [Medline: [21767103](https://pubmed.ncbi.nlm.nih.gov/21767103/)]
8. Nsanzimana S, Kanters S, Mills E. Towards test and treat strategy for HIV in sub-Saharan Africa. *BMJ* 2015 Dec 29;351:h6839. [doi: [10.1136/bmj.h6839](https://doi.org/10.1136/bmj.h6839)] [Medline: [26714871](https://pubmed.ncbi.nlm.nih.gov/26714871/)]
9. World Health Organization. 2014. Technical and operational considerations for implementing HIV viral load testing: interim technical update URL: <https://www.who.int/hiv/pub/arv/viral-load-testing-technical-update/en/> [accessed 2019-02-20] [WebCite Cache ID 76Ky6Pise]
10. Adler A, Mounier-Jack S, Coker R. Late diagnosis of HIV in Europe: definitional and public health challenges. *AIDS Care* 2009 Mar;21(3):284-293. [doi: [10.1080/09540120802183537](https://doi.org/10.1080/09540120802183537)] [Medline: [19031304](https://pubmed.ncbi.nlm.nih.gov/19031304/)]
11. Antinori A, Coenen T, Costagliola D, Dedes N, Ellefson M, Gatell J, European Late Presenter Consensus Working Group. Late presentation of HIV infection: a consensus definition. *HIV Med* 2011 Jan;12(1):61-64 [FREE Full text] [doi: [10.1111/j.1468-1293.2010.00857.x](https://doi.org/10.1111/j.1468-1293.2010.00857.x)] [Medline: [20561080](https://pubmed.ncbi.nlm.nih.gov/20561080/)]
12. Rice B, Yin Z, Brown A, Croxford S, Conti S, De Angelis D, et al. Monitoring of the HIV Epidemic Using Routinely Collected Data: The Case of the United Kingdom. *AIDS Behav* 2017 Jul;21(Suppl 1):83-90. [doi: [10.1007/s10461-016-1604-6](https://doi.org/10.1007/s10461-016-1604-6)] [Medline: [27832390](https://pubmed.ncbi.nlm.nih.gov/27832390/)]
13. Mocroft A, Lundgren J, Sabin M, Monforte A, Brockmeyer N, Casabona J, Collaboration of Observational HIV Epidemiological Research Europe (COHERE) study in EuroCoord. Risk factors and outcomes for late presentation for HIV-positive persons in Europe: results from the Collaboration of Observational HIV Epidemiological Research Europe Study (COHERE). *PLoS Med* 2013;10(9):e1001510 [FREE Full text] [doi: [10.1371/journal.pmed.1001510](https://doi.org/10.1371/journal.pmed.1001510)] [Medline: [24137103](https://pubmed.ncbi.nlm.nih.gov/24137103/)]
14. World Health Organization. 2017. Guidelines for managing advanced HIV disease and rapid initiation of antiretroviral therapy URL: <http://apps.who.int/iris/bitstream/10665/255884/1/9789241550062-eng.pdf?ua=1> [accessed 2019-02-20] [WebCite Cache ID 76KyGBhIU]
15. Pac L, Horwitz M, Namutebi A, Auerbach BJ, Semeere A, Namulema T, et al. Implementation and operational research: Integrated pre-antiretroviral therapy screening and treatment for tuberculosis and cryptococcal antigenemia. *J Acquir Immune Defic Syndr* 2015 Apr 15;68(5):e69-e76 [FREE Full text] [doi: [10.1097/QAI.0000000000000527](https://doi.org/10.1097/QAI.0000000000000527)] [Medline: [25761234](https://pubmed.ncbi.nlm.nih.gov/25761234/)]
16. Vallabhaneni S, Longley N, Smith M, Smith R, Osler M, Kelly N, et al. Implementation and Operational Research: Evaluation of a Public-Sector, Provider-Initiated Cryptococcal Antigen Screening and Treatment Program, Western Cape, South Africa. *J Acquir Immune Defic Syndr* 2016 Dec 01;72(2):e37-e42 [FREE Full text] [doi: [10.1097/QAI.0000000000000976](https://doi.org/10.1097/QAI.0000000000000976)] [Medline: [26926942](https://pubmed.ncbi.nlm.nih.gov/26926942/)]
17. Hakim J, Musiime V, Szubert A, Mallewa J, Siika A, Agutu C, REALITY Trial Team. Enhanced Prophylaxis plus Antiretroviral Therapy for Advanced HIV Infection in Africa. *N Engl J Med* 2017 Dec 20;377(3):233-245 [FREE Full text] [doi: [10.1056/NEJMoa1615822](https://doi.org/10.1056/NEJMoa1615822)] [Medline: [28723333](https://pubmed.ncbi.nlm.nih.gov/28723333/)]
18. Kahn J, Marseille E, Moore D, Bunnell R, Were W, Degerman R, et al. CD4 cell count and viral load monitoring in patients undergoing antiretroviral therapy in Uganda: cost effectiveness study. *BMJ* 2011 Nov 09;343:d6884. [doi: [10.1136/bmj.d6884](https://doi.org/10.1136/bmj.d6884)] [Medline: [22074713](https://pubmed.ncbi.nlm.nih.gov/22074713/)]
19. Keebler D, Revill P, Braithwaite S, Phillips A, Blaser N, Borquez A, et al. Cost-effectiveness of different strategies to monitor adults on antiretroviral treatment: a combined analysis of three mathematical models. *The Lancet Global Health* 2014 Jan;2(1):e35-e43. [doi: [10.1016/S2214-109X\(13\)70048-2](https://doi.org/10.1016/S2214-109X(13)70048-2)]
20. Keane J, Pharr J, Buttner M, Ezeanolue E. Interventions to Reduce Loss to Follow-up During All Stages of the HIV Care Continuum in Sub-Saharan Africa: A Systematic Review. *AIDS Behav* 2017 Jun;21(6):1745-1754. [doi: [10.1007/s10461-016-1532-5](https://doi.org/10.1007/s10461-016-1532-5)] [Medline: [27578001](https://pubmed.ncbi.nlm.nih.gov/27578001/)]
21. World Health Organisation. 2015. Consolidated strategic information guidelines for HIV in the health sector URL: <https://www.who.int/sorry/> [accessed 2019-02-20] [WebCite Cache ID 76KySVO4Q]
22. Birrell P, Chadborn T, Gill O, Delpech V, De Angelis D. Estimating trends in incidence, time-to-diagnosis and undiagnosed prevalence using a CD4-based Bayesian back-calculation. *Stat Comm Infect Dis* 2012;4(1).
23. Jansson J, Kerr C, Mallitt K, Wu J, Gray R, Wilson D. Inferring HIV incidence from case surveillance with CD4+ cell counts. *AIDS* 2015 Jul 31;29(12):1517-1525. [doi: [10.1097/QAD.0000000000000679](https://doi.org/10.1097/QAD.0000000000000679)] [Medline: [25870983](https://pubmed.ncbi.nlm.nih.gov/25870983/)]
24. Song R, Hall H, Green T, Szwarcwald C, Pantazis N. Using CD4 Data to Estimate HIV Incidence, Prevalence, and Percent of Undiagnosed Infections in the United States. *J Acquir Immune Defic Syndr* 2017 Dec 01;74(1):3-9. [doi: [10.1097/QAI.0000000000001151](https://doi.org/10.1097/QAI.0000000000001151)] [Medline: [27509244](https://pubmed.ncbi.nlm.nih.gov/27509244/)]
25. van Sighem A, Nakagawa F, De Angelis D, Quinten C, Bezemer D, de Coul EO, et al. Estimating HIV Incidence, Time to Diagnosis, and the Undiagnosed HIV Epidemic Using Routine Surveillance Data. *Epidemiology* 2015 Sep;26(5):653-660 [FREE Full text] [doi: [10.1097/EDE.0000000000000324](https://doi.org/10.1097/EDE.0000000000000324)] [Medline: [26214334](https://pubmed.ncbi.nlm.nih.gov/26214334/)]

26. Szwarcwald C. Estimation of the HIV Incidence and of the Number of People Living With HIV/AIDS in Brazil, 2012. *J AIDS Clin Res* 2015;06(03):430. [doi: [10.4172/2155-6113.1000430](https://doi.org/10.4172/2155-6113.1000430)]
27. Lodwick R, Nakagawa F, van Sighem A, Sabin C, Phillips A. Use of surveillance data on HIV diagnoses with HIV-related symptoms to estimate the number of people living with undiagnosed HIV in need of antiretroviral therapy. *PLoS One* 2015;10(3):e0121992 [FREE Full text] [doi: [10.1371/journal.pone.0121992](https://doi.org/10.1371/journal.pone.0121992)] [Medline: [25768925](https://pubmed.ncbi.nlm.nih.gov/25768925/)]
28. Wong N, Wong K, Lee M, Tsang O, Chan D, Lee S. Estimation of the Undiagnosed Intervals of HIV-Infected Individuals by a Modified Back-Calculation Method for Reconstructing the Epidemic Curves. *PLoS One* 2016;11(7):e0159021 [FREE Full text] [doi: [10.1371/journal.pone.0159021](https://doi.org/10.1371/journal.pone.0159021)] [Medline: [27403882](https://pubmed.ncbi.nlm.nih.gov/27403882/)]
29. Rice B, Elford J, Yin Z, Delpech V. A new method to assign country of HIV infection among heterosexuals born abroad and diagnosed with HIV. *AIDS* 2012 Sep 24;26(15):1961-1966. [doi: [10.1097/QAD.0b013e3283578b80](https://doi.org/10.1097/QAD.0b013e3283578b80)] [Medline: [22781226](https://pubmed.ncbi.nlm.nih.gov/22781226/)]
30. European Centre for Disease Prevention and Control. 2013. Migrant health: sexual transmission of HIV within migrant groups in the EU/EEA and implications for effective interventions URL: <https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/Migrant-health-sexual-transmission.pdf> [accessed 2019-02-20] [WebCite Cache ID 76KyfPh0E]
31. McGrath N, Lessells R, Newell M. Time to eligibility for antiretroviral therapy in adults with CD4 cell count > 500 cells/ μ L in rural KwaZulu-Natal, South Africa. *HIV Med* 2015 Sep;16(8):512-518 [FREE Full text] [doi: [10.1111/hiv.12255](https://doi.org/10.1111/hiv.12255)] [Medline: [25959724](https://pubmed.ncbi.nlm.nih.gov/25959724/)]
32. Grimsrud A, Cornell M, Schomaker M, Fox M, Orrell C, Prozesky H, International Epidemiologic Databases to Evaluate AIDS Southern Africa Collaboration (IeDEA-SA). CD4 count at antiretroviral therapy initiation and the risk of loss to follow-up: results from a multicentre cohort study. *J Epidemiol Community Health* 2016 Dec;70(6):549-555 [FREE Full text] [doi: [10.1136/jech-2015-206629](https://doi.org/10.1136/jech-2015-206629)] [Medline: [26700300](https://pubmed.ncbi.nlm.nih.gov/26700300/)]
33. Takuva S, Brown A, Pillay Y, Delpech V, Puren A. The continuum of HIV care in South Africa: implications for achieving the second and third UNAIDS 90-90-90 targets. *AIDS* 2017 Dec 20;31(4):545-552. [doi: [10.1097/QAD.0000000000001340](https://doi.org/10.1097/QAD.0000000000001340)] [Medline: [28121668](https://pubmed.ncbi.nlm.nih.gov/28121668/)]
34. World Health Organisation. 2017. Consolidated guidelines on person-centred HIV patient monitoring and case surveillance Geneva: World Health Organization URL: <https://www.who.int/sorry/> [accessed 2019-02-20] [WebCite Cache ID 76KyqHdPn]
35. UNAIDS. 2016. The need for routine viral load testing: questions and answers URL: http://www.unaids.org/sites/default/files/media_asset/JC2845_en.pdf [accessed 2019-02-20] [WebCite Cache ID 76Kyui7Vk]
36. PEPFAR. PEPFAR 2018 Country Operational Plan Guidance for Standard Process Countries URL: <https://www.pepfar.gov/documents/organization/276459.pdf> [accessed 2019-02-20] [WebCite Cache ID 76KyxpacA]
37. Medecins sans Frontières. Advanced HIV position statement 2017 URL: <https://www.msf.org.za/hiv-position-statement> [accessed 2019-02-20] [WebCite Cache ID 76Kz1xlzg]
38. Harklerode R, Schwarcz S, Hargreaves J, Boule A, Todd J, Xueref S, et al. Feasibility of Establishing HIV Case-Based Surveillance to Measure Progress Along the Health Sector Cascade: Situational Assessments in Tanzania, South Africa, and Kenya. *JMIR Public Health Surveill* 2017 Jul 10;3(3):e44 [FREE Full text] [doi: [10.2196/publichealth.7610](https://doi.org/10.2196/publichealth.7610)] [Medline: [28694240](https://pubmed.ncbi.nlm.nih.gov/28694240/)]
39. U.S. Department of Health and Human Services. Guidelines for the Use of Antiretroviral Agents in Pediatric HIV Infection, 2018 URL: <https://aidsinfo.nih.gov/guidelines/html/2/pediatric-arv/59/clinical-and-laboratory-monitoring-of-pediatric-hiv-infection> [accessed 2019-02-20] [WebCite Cache ID 76Kz4eNXo]

Abbreviations

ART: antiretroviral therapy

WHO: World Health Organization

Edited by S Baral; submitted 25.05.18; peer-reviewed by J Todd, S Xueref; comments to author 23.09.18; revised version received 19.10.18; accepted 02.11.18; published 20.03.19.

Please cite as:

Rice B, Boule A, Schwarcz S, Shroufi A, Rutherford G, Hargreaves J
 The Continuing Value of CD4 Cell Count Monitoring for Differential HIV Care and Surveillance
JMIR Public Health Surveill 2019;5(1):e11136
 URL: <http://publichealth.jmir.org/2019/1/e11136/>
 doi:[10.2196/11136](https://doi.org/10.2196/11136)
 PMID:[30892272](https://pubmed.ncbi.nlm.nih.gov/30892272/)

©Brian Rice, Andrew Boule, Sandra Schwarcz, Amir Shroufi, George Rutherford, James Hargreaves. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 20.03.2019. 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

Population Size Estimation of Venue-Based Female Sex Workers in Ho Chi Minh City, Vietnam: Capture-Recapture Exercise

Giang Le¹, MPH; Nghia Khuu², MD, MA; Van Thi Thu Tieu³, MD; Phuc Duy Nguyen², MPH; Hoa Thi Yen Luong⁴, MPH; Quang Duy Pham², MD, PhD; Hau Phuc Tran², MD, MPH, PhD; Thuong Vu Nguyen², MD, DTM&H, PhD; Meade Morgan⁵, PhD; Abu S Abdul-Quader⁴, PhD

¹Program Development Office, United States Agency for International Development, Vietnam, Hanoi, Viet Nam

²Department for Disease Control and Prevention, Pasteur Institute, Ho Chi Minh City, Viet Nam

³Ho Chi Minh City HIV/AIDS Center, Ho Chi Minh City, Viet Nam

⁴Division of Global HIV and Tuberculosis, United States Centers for Disease Control and Prevention, Hanoi, Viet Nam

⁵Division of Global HIV and Tuberculosis, United States Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:

Giang Le, MPH

Program Development Office

United States Agency for International Development, Vietnam

2 Ngo Quyen Street

Hoan Kiem District

Hanoi,

Viet Nam

Phone: 84 0962032206

Email: letonggiang@gmail.com

Abstract

Background: There is limited population size estimation of female sex workers (FSWs) in Ho Chi Minh City (HCMC)—the largest city in Vietnam. Only 1 population size estimation among venue-based female sex workers (VFSWs) was conducted in 2012 in HCMC. Appropriate estimates of the sizes of key populations are critical for resource allocation to prevent HIV infection.

Objective: The aim of this study was to estimate the population size of the VFSWs from December 2016 to January 2017 in HCMC, Vietnam.

Methods: A multistage capture-recapture study was conducted in HCMC. The capture procedures included selection of districts using stratified probability proportion to size, mapping to identify venues, approaching all VFSWs to screen their eligibility, and then distribution of a unique object (a small pink makeup bag) to all eligible VFSWs in all identified venues. The recapture exercise included equal probability random selection of a sample of venues from the initial mapping and then approaching FSWs in those venues to determine the number and proportion of women who received the unique object. The proportion and associated confidence bounds, calculated using sampling weights and accounting for study design, were then divided by the number of objects distributed to calculate the number of VFSWs in the selected districts. This was then multiplied by the inverse of the proportion of districts selected to calculate the number of VFSWs in HCMC as a whole.

Results: Out of 24 districts, 6 were selected for the study. Mapping identified 573 venues across which 2317 unique objects were distributed in the first capture. During the recapture round, 103 venues were selected and 645 VFSWs were approached and interviewed. Of those, 570 VFSWs reported receiving the unique object during the capture round. Total estimated VFSWs in the 6 selected districts were 2616 (95% CI 2445-3014), accounting for the fact that only 25% (6/24) of total districts were selected gives an overall estimate of 10,465 (95% CI 9782-12,055) VFSWs in HCMC.

Conclusions: The capture-recapture exercise provided an estimated number of VFSWs in HCMC. However, for planning HIV prevention and care service needs among all FSWs, studies are needed to assess the number of sex workers who are not venue-based, including those who use social media platforms to sell services.

(*JMIR Public Health Surveill* 2019;5(1):e10906) doi:[10.2196/10906](https://doi.org/10.2196/10906)

KEYWORDS

population size estimation; venue-based; female sex workers; Ho Chi Minh City; capture-recapture

Introduction

Of the approximately 248,000 people living with HIV in Vietnam in 2017, more than 100,000 people (approximately 42%) are receiving antiretroviral therapy [1,2]. With a national estimated prevalence of about 0.3% among adults, the HIV epidemic in Vietnam remains concentrated among 3 key populations (KPs)—people who inject drugs, men who have sex with men, and female sex workers (FSWs) [2,3].

Vietnam's capacity to further reduce HIV prevalence among the general population largely depends on knowledge and information about KPs. Population size estimates of those affected by HIV/AIDS help central and provincial level policy makers and program administrators understand the scope of the HIV epidemic, help plan appropriate interventions, allocate sufficient resources, and estimate program coverage. In Vietnam, HIV sentinel surveillance (HSS) among KPs is implemented in 20 priority provinces. The HSS provides seroprevalence and limited behavioral data. Population size estimation activity, however, has not been included in the data collection exercise [4,5]. There is limited information on FSW population size in Vietnam. In 2012, capture-recapture method was used to conduct a population size estimation activity among FSWs in Ho Chi Minh City (HCMC). The results showed that there were about 16,500 to 22,500 FSWs in the city [6]. Since then, the total population of the city had increased significantly, which may have led to a higher demand for various services. This necessitated the planning and implementation of another population size estimation activity among FSWs in HCMC. However, we decided to focus this size estimation activity among venue-based female sex workers (VFSWs) who can be reached through different venues (locations) such as restaurants, bars, massage parlors, and karaoke palaces but not all FSWs in the city for a number of reasons. First, high coverage of mobile phone users in Vietnam makes it easy for FSWs to reach their clients without being present at physical venues. Second, unavailability of practical and appropriate methods that can be used to estimate the population size of this population who approach their clients on the Internet, such as a social network platform (eg, Facebook) or messaging apps (eg, Zalo). Between December 2016 and January 2017, the Pasteur Institute in HCMC conducted the population size estimation activity among VFSWs in 6 selected districts of HCMC—the largest city in the country. The US Centers for Disease Control and Prevention and the Pasteur Institute in HCMC approved the study.

Methods

Overview

Before implementation, a meeting was held with FSWs and local stakeholders, including peer outreach workers, to share information about the study. Moreover, a two-sample capture-recapture method was used to estimate the population size of VFSWs in HCMC, Vietnam. VFSWs were defined as all females aged 18 years or older, having sex for money or goods in the last 3 months, residing in the city for the past 6 months, and who operate out of venues such as restaurant, coffee shops, karaoke, bars, spas, and park. Implementation of the capture-recapture method involved the following steps: selection of districts, mapping of venues, developing a list of accessible venues, distribution of unique objects to VFSWs in venues (capture), sampling of venues for recapture, and counting at the recapture sites both the number of VFSWs who received the unique object and the number of VFSWs who did not. After consulting with VFSWs and local service providers, pink makeup bags were chosen to be used as unique objects in this study (hereafter referred to as “unique objects”).

Selection of Districts

In HCMC, there are 24 administrative districts covering approximately about 2100 km², with a total population of about 9 million. In an effort to cover the city as broadly as possible, we used the HCMC Public Security Office (PSO) data and estimated the number of VFSWs that was used as an input of The Joint United Nations Programme on HIV and AIDS (UNAIDS) Estimation and Project Package (EPP) to select the districts where population size estimation activity among VFSWs was implemented.

The HCMC PSO reported that in 2013, there were about 7021 FSWs in the city disaggregated by district [7]. According to UNAIDS EPP input estimates, there are about 20,000 FSWs in HCMC. The number of FSWs reported by the HCMC PSO was used to estimate the proportion of FSWs in HCMC by district. Then, those proportions were applied to the EPP estimations to derive at an estimated number of FSWs in each district. To provide broader geographic coverage of the city, the districts were ranked from highest to lowest (column E, Table 1) and then divided into 3 groups based on the estimates. Group 1 was the top 6 districts that account for 50% of the estimated number of FSWs in the city. The second (7 districts) and third groups (11 districts) accounted for 30% and 20% of the estimated number of FSWs, respectively. To cover at least 25% of the total estimates, 6 districts were selected for the size estimation data collection activity. Moreover, 2 districts were selected at random from each group with equal probability (Table 1).

Table 1. District selection in Ho Chi Minh City for female sex workers.

District (A)	EPP ^a provincial estimate distributed across districts with public security proportion (B), n	Reported FSWs ^b by public security 2014 (C), n	Contribution of each district to EPP estimate (column B) (D), %	Cumulative (E), %	District selected (F)
Binh Thanh	3245	1139	16.22	16.22	Yes
Cu Chi	1419	498	7.09	23.32	No
District 1	1410	495	7.05	30.37	No
Go Vap	1367	480	6.84	37.20	No
Phu Nhuan	1191	418	5.95	43.16	Yes
Binh Tan	1111	390	5.55	48.71	No
Tan Phu	1014	356	5.07	53.78	No
District 10	903	317	4.52	58.30	No
Tan Binh	863	303	4.32	62.61	No
District 9	806	283	4.03	66.64	Yes
Binh Chanh	781	274	3.90	70.55	Yes
District 3	772	271	3.86	74.41	No
Hoc Mon	735	258	3.67	78.08	No
District 11	695	244	3.48	81.56	No
District 5	615	216	3.08	84.63	No
District 7	493	173	2.46	87.10	Yes
District 6	433	152	2.16	89.26	No
Nha Be	393	138	1.97	91.23	No
District 8	393	138	1.97	93.19	No
Thu Duc	345	121	1.72	94.92	No
District 12	319	112	1.60	96.51	No
District 2	251	88	1.25	97.76	No
District 4	234	82	1.17	98.93	Yes
Can Gio	214	75	1.07	100.00	No
Total	20,002	7021	100.00	—	6 districts

^aEPP: Estimation and Project Package.

^bFSWs: female sex workers.

Mapping of Districts

After the districts were selected, 12 field teams—2 for each district—were formed. The field team members included 2 peer outreach workers who had prior experience of conducting outreach in their assigned districts. The field teams were then trained on method, implementation procedures, and purposes of the size estimation exercise. The training focused on how to conduct mapping including observation, documentation, maintaining confidentiality of information, and how to maintain safety in the field.

Before conducting observation and mapping, field team members were asked to generate a list of all potential venues within each of their assigned districts where they believed VFSWs congregated or worked, describe the locations, estimate the number of VFSWs that congregated in each of these venues,

and the recommended times of the day and days of the week when VFSWs could be approached.

After completion of the preliminary list of venues, observation and mapping were conducted in each selected district to identify venues where VFSWs congregated or worked. During the observation, field teams also visited locations or venues within their assigned areas to ensure if any *new potential locations or venues* were added to the list. In addition, the field teams conducted brief informal *interviews* with local business owners (street-side tea and coffee shops), *xe om* (motorbike taxi operators), as well as other business establishments to corroborate the information collected based on observation and mapping.

Development of the List of Venues

On the basis of the mapping exercise, a list of venues was created with the following information: address and description

of the venues and estimated number of VFSWs congregated in each venue by days and times when they were accessible to the field team members. For the venues to be included in capture-recapture exercise, they had to meet the following criteria: (1) could be approached, (2) were safe for the study team and potential participants, and (3) had more VFSWs than the number of people who need to be approached or captured in the standardized period. It was estimated that in 1 hour, a group of 2 field team members could recruit at least 5 VFSWs. Thus, for sampling purposes, only venues that are estimated to have 5 or more VFSWs were selected to develop the list of venues. Those venues with fewer than 5 VFSWs (small venues) were combined with other venues in its proximity (about 5-min walk) and added to the list of venues. Large venues—those with 20 or more VFSWs—were divided into 2 or more venues to make sure on average these had about 5 to 10 VFSWs per venue.

Distribution of Unique Objects (Capture)

On the basis of an initial estimate of 20,000 VFSWs and a sample size for the recapture round (554), we calculated that 2232 objects needed to be distributed to estimate the size of the population with a precision of 33% and incorporating a survey design effect of 2.0 [8]. Before distribution of unique objects, all field team members were trained on how to count potential participants, how to identify VFSWs and assess eligibility before giving them the unique objects, what to say about the unique objects, how to maintain anonymity and confidentiality of VFSWs, and how to maintain safety in the field. After the training, the field teams closely followed the schedules developed by the study team to visit all venues to distribute unique objects to all eligible VFSWs (capture exercise).

After the selection of the venues, the field team went to the venues and approached the venue owners to explain to them the purposes and procedure of the study. Once they agreed, the field team went inside the venues and approached the potential participants. The number of all potential participants during the visits was counted and recorded. Then, they were explained the purposes of the study and were asked the screening questions to determine their eligibility. If they met the selection criteria, they were informed clearly that participation was completely voluntary. The informed consent form was read to them, and verbal consents were obtained before including them in the study. Each eligible VFSW was given 1 unique object. They were also asked to keep the object because they might be asked to identify it at a later date. The object distributor used log forms to keep track of the number of unique objects distributed by date and location. No other information was collected. At the end of the distribution of unique object activity, the total number of objects actually distributed was determined. The capture exercise was completed in 2 weeks from December 5 to 18, 2016.

Recapture

Moreover, 2 weeks after the completion of the capture exercise, a subset of venues was randomly selected with equal probability from the total list of venues used in the capture round. New

field teams were formed by switching the team members within each selected district. Each new field team was assigned to approach and recruit potential participants at venues/locations where they were not assigned in the capture round. At each selected venue, the field team went to the venue and approached the venue owners to explain to them the purposes and procedure of the study. Once they agreed, the field team went inside the venues and counted and recorded all potential VFSWs. All potential VFSWs were approached, verbally consented, and assessed to make sure that they met the selection criteria. Then, they were asked whether they received any object from someone. If they mentioned that they did receive an object during the previous weeks and they could show or correctly describe and identify the unique object distributed from printed pictures showing the unique object and other 4 different objects, the VFSW was considered to be recaptured. VFSWs who mentioned that they never received any unique objects or could not describe and identify correctly the unique object were recorded as new capture in the recapture round. The recapture round was completed in 2 weeks from January 1 to 15, 2017.

Data Analysis

Population size estimates for VFSWs across the 6 sampled districts were calculated using the formula $N=MC/R$, where N is the size estimate, M is the number of unique objects distributed during the capture round, C is the number of VFSWs captured in the recapture round, and R is the number of VFSWs captured on both the capture and recapture rounds [9]. Survey-weighted logit-based 95% CIs for P values were calculated using Taylor-linearized variance estimates to account for the selection of districts and venue, and these were then used to calculate bounds for N . The estimate for N and associated bounds were then multiplied by the inverse of the district sampling fraction to scale from the 6 sampled districts to all 24 districts in the city. The estimates for the estimated proportion of VFSWs who reported receiving an object during the recapture round and CI calculations included weights to account for district sampling probabilities and stratum and district and venue variables to account for the sampling design. The calculations were conducted in Stata version 14 (StataCorp LLC, College Station, TX).

Results

In the capture round, a total of 2317 VFSWs received unique objects across all 573 mapped venues in the 6 selected districts. In the recapture round, 103 venues were selected and 645 VFSWs were interviewed. Of those, 88.67% (2317/2613; 95% CI 76.9% to 94.7%) were found to have received a unique object. Dividing the number of objects distributed by this estimate and bounds gives 2616 (95% CI 2445-3014) as the number of VFSWs in the 6 sampled districts. Furthermore, multiplying by 4 to account for the fact that only 6 of the 24 districts were selected gives an overall estimate of 10,465 (95% CI 9782-12,055) for the size of the VFSWs population in HCMC. Table 2 below provides the information on each of the selected districts as well as the total for HCMC.

Table 2. Summary of capture and recapture for venue-based female sex workers in Ho Chi Minh City.

Stratum	District	Venues identified	Unique objects distributed	Venues in recapture	VFSWs ^a in recapture	VFSWs with unique object in recapture	Percentage of VF-SWs with unique object	Estimated number of VFSWs
High	Binh Thanh	98	438	20	110	96	87.3	502
High	Phu Nhuan	62	397	14	113	88	77.9	510
Middle	Binh Chanh	174	597	28	148	135	91.2	654
Middle	District 9	45	185	9	95	77	81.1	228
Low	District 4	54	211	8	43	43	100.0	211
Low	District 7	130	489	24	136	131	96.3	508
	Total	563	2317	103	645	570	88.6 ^b	2616

^aVFSWs: venue-based female sex workers.

^bPoint estimates weighted to account for the probability of district selection and CIs calculated to account for clustering within venues.

After completing data analysis, a half-day consultation was held to present the findings to the stakeholders in the city and to discuss the challenges and limitation of the findings. The consultation was attended by officials from the HCMC Provincial AIDS Center, senior field workers, and scientific staff from the Pasteur Institute in HCMC and the US Centers for Disease Control and Prevention. The presentations included a brief description of the capture-recapture method, how it was implemented in HCMC, the results, and how to interpret the findings.

Discussion

Principal Findings

The data collection activity was conducted in 6 randomly selected districts to estimate the size of VFSWs aged 18 years or older. We considered 2 options for extrapolating from the 6 sampled districts to the entire city. First was to inflate the results for the 6 districts by a factor of 4 as we sampled one-fourth of the districts. The second was to inflate the size estimate for the 6 districts by the ratio of the number of females aged between 18 and 49 years in the city divided by the number in the 6 districts, based on 2016 census estimate. Using 2016 census estimates, we calculated that the proportion was 4.067 [10]. With this factor of 4.067, the estimated population size of VFSWs in HCMC would be similar to what had been estimated, accounting for the fact that only 25% (6/24) of total districts were selected (10,465).

With the advent of technology and internet, it is possible that over the years, the nature of sex work has changed. However, a significant proportion of FSWs continue to operate through different venues. In this exercise, VFSWs operated at various types of venues such as restaurants, massage parlors, bars/karaoke, cafeterias, and hostels. However, only few locations on streets where FSWs could be found were recorded. The reduction in the number of street venues might be explained by the influence of mobile phone technology, significant increase of internet users in Vietnam, and the presence of police [11]. The modern technology has helped VFSWs to reach their clients without being present at a physical location. Those who continue to operate through the venues represent a segment of

the FSWs' population in HCMC who are somewhat visible and accessible to health care workers. Significant efforts were made to cover all venues where sex workers could be approached during capture and recapture; however, the field staff reported that because of many factors (approachability and security issues), only about 70% of venues in the selected districts during mapping were covered. If the missed venues were similar to those included in the study, then the VFSW population size for HCMC would be estimated at 10,465/0.70 or 14,950. This estimated size of the population is lower than that of a study using a capture-recapture method (19,602; 95% CI 16,590-22,554) but is higher than high the estimate of VFSW in the same study using data from police system (10,595) in 2013 [12,13].

Limitations

There are a number of limitations to this study. First, the findings presented above are limited to those VFSWs in HCMC who are accessible only at selected venues. Findings do not include VFSWs who operate out of venues that were not identified and not included in the list of venues for data collection. In a large city such as HCMC, there are venues the study teams were not able to access, such as *high-class* restaurants, massage parlors, or entertainment venues. Other FSWs who do not operate out of venues were not accessed and could not be included in the estimation either. Second, the estimation does not include FSWs who connect to their client through modern technology such as Web or telephones. Recent figures show that Vietnam has the largest and fastest growing smartphone and internet users in the world, with an estimate of over 45% of the population using smartphones and internet [11]. Therefore, it is possible FSWs are using other methods such as telephones and internet to connect with prospective clients. Third, in Vietnam, sex work is illegal; the public security authorities maintain a relatively close watch on individuals who may be working as sex workers, making the hidden populations even more difficult to reach. That primarily limits the street operations of FSW. However, this may also limit those who operate at various venues. In addition to public scrutiny, the venue owners control the number of FSWs at any venue at a given time. They control their working hours and how and when they can be accessed. An illustration of this was a very low number of VFSWs venues

existed during the mapping, and not all venues could be accessed. Police activity in some of the districts may have influenced the number of VFSWs who could be approached during the size estimation exercise. Fourth, as we recruited those aged 18 years or older, the population size estimates of VFSWs did not include those aged younger than 18 years. Finally, the assumption of *closed population* might have been violated. However, it was minimized by shortening the time of the study (in about 6 weeks) and the time between the 2 samples (2 weeks)

and avoiding to collect data on unusual days (such as Lunar New Year). Despite the potential limitations, the findings of this study are useful for program planning and assessments that highlight the importance of designing appropriate interventions for VFSWs in HCMC. The findings also highlight the important need to investigate possible ways to estimate those who do not operate out of venues, especially those who use various social media platforms; their characteristics; and how they could be reached with appropriate interventions.

Acknowledgments

The authors would like to thank the staff at the Pasteur Institute in Ho Chi Minh City (HCMC) and the HCMC Provincial HIV/AIDS Center. They also appreciate the professional work of the field teams, private business owners, and all respondents who spent their time and effort to support and participate in the study. This study has been supported by the President's Emergency Plan for AIDS Relief through the United States Centers for Disease Control and Prevention under the terms of 5U2GGH001628-02. The findings and conclusions in this study are those of the authors and do not necessarily represent the official position of the funding agencies.

Conflicts of Interest

None declared.

References

1. Vietnam Authority of HIV/AIDS Control. Ministry of Health: Department of HIV/AIDS Prevention and Control. Hanoi, Vietnam: VAAC; 2017. Report on HIV/AIDS prevention and control in the first six months of 2017 and key tasks in the last 6 months of 2017 URL: http://vaac.gov.vn/Cms_Data/Contents/Vaac/Folders/Solieubaocao/Solieu/~contents/4UUJBYQ6FQYNT64M/958-BC-6-thang-2017.pdf [accessed 2018-11-29] [WebCite Cache ID 74Hi5n1yE]
2. Vietnam AOHC. Vietnam HIV/AIDS Estimates and Projections 2017. Hanoi, Vietnam: VAAC; 2018.
3. Vietnam Authority of HIV/AIDS Control. Ministry of Health: Department of HIV/AIDS Prevention and Control. Hanoi, Vietnam: VAAC; 2016 Aug 09. Report on HIV/AIDS prevention and control in the first six months of 2016 and key tasks in the last 6 months of 2016 URL: <http://vaac.gov.vn/solieu/Detail/Bao-cau-cong-tac-phong-chong-HIV-AIDS-6-thang-dau-nam-2016-va-nhiem-vu-trong-tam-6-thang-cuoi-nam-2016> [accessed 2018-04-29] [WebCite Cache ID 6z21ooxpP]
4. Vietnam Authority of HIV/AIDS Control. Report on the Results of HIV sentinel surveillance plus behavioral surveys among high-risk populations in Vietnam (HSS/HSS+) 2016. Hanoi, Vietnam: VAAC; 2017:1.
5. Vietnam Ministry of Health. Circular 09 Guiding on HIV/AIDS surveillance and STI in Vietnam. Hanoi, Vietnam: Vietnam Ministry of Health; 2012. URL: <https://bit.ly/2FWMTgB> [WebCite Cache ID 75VzxMNIIM]
6. Tuan LA, Quang T, Hue N, Tuan N. Estimated number of Female Sex Workers in Ho Chi Minh City. Hanoi: Vietnam National Institute of Hygiene and Epidemiology; 2012.
7. Ho Chi Minh City Dept of Public Security. Distribution of people who inject drugs and female sex workers in Ho Chi Minh City. HCMC: Ho Chi Minh City Dept of Public Security; 2013.
8. World Health Organization. Geneva: World Health Organization; 2017. Biobehavioral survey guidelines for populations at risk for HIV: Supplemental Materials URL: <http://apps.who.int/iris/bitstream/handle/10665/258924/9789241513012-eng.pdf;jsessionid=6333E19627757BC4BEC2A8C803DE8968?sequence=1> [accessed 2018-12-05] [WebCite Cache ID 74R1QOBdU]
9. World Health Organization. Geneva: World Health Organization; 2010. Guidelines on estimating the size of populations most at risk to HIV URL: http://apps.who.int/iris/bitstream/handle/10665/44347/9789241599580_eng.pdf?sequence=1&isAllowed=y
10. General Statistics Office of Vietnam. Hanoi: General Statistics Office of Vietnam; 2011 Feb. Population projection for Vietnam 2009-2049 URL: https://www.gso.gov.vn/default_en.aspx?tabid=617&ItemID=11016 [accessed 2018-04-29]
11. Representasia. Represent Asia News.: Representasia; 2016 May 10. Vietnam's Mobile Market is booming URL: <http://representasianews.com/vietnams-mobile-market-is-booming/> [accessed 2018-04-28] [WebCite Cache ID 6z22ksqAr]
12. Tuan LA, Tuan N, Thanh D, Quang T, Keith S. Multiple approaches to population size estimation of injecting drug users (IDU), female sex workers (VFSW), and men who have sex with men (MSM) in three Vietnamese provinces. San Francisco: The University of California; 2012 Presented at: 19th International AIDS conference; July 22-27, 2012; Washington, DC URL: <http://hivinsite.ucsf.edu/InSite?page=cfaids-12-00>

13. Tuan LA, Quang T, Van T, Tuan N. Population estimation of FSWs in Ho Chi Minh City. Ho Chi Minh City; 2012 Presented at: 5th National HIV/AIDS conference; December 12, 2012; Ho Chi Minh City, Vietnam.

Abbreviations

EPP: Estimation and Project Package

FSW: female sex worker

HCMC: Ho Chi Minh City

HSS: HIV sentinel surveillance

KP: key population

PSO: Public Security Office

UNAIDS: United Nations Programme on HIV and AIDS

VFSW: venue-based female sex worker

Edited by D Prybylski, J Neal; submitted 29.04.18; peer-reviewed by A Hakim, D Gexha Bunjaku; comments to author 23.07.18; revised version received 09.10.18; accepted 31.10.18; published 29.01.19.

Please cite as:

Le G, Khuu N, Tieu VTT, Nguyen PD, Luong HTY, Pham QD, Tran HP, Nguyen TV, Morgan M, Abdul-Quader AS

Population Size Estimation of Venue-Based Female Sex Workers in Ho Chi Minh City, Vietnam: Capture-Recapture Exercise

JMIR Public Health Surveill 2019;5(1):e10906

URL: <http://publichealth.jmir.org/2019/1/e10906/>

doi: [10.2196/10906](https://doi.org/10.2196/10906)

PMID: [30694204](https://pubmed.ncbi.nlm.nih.gov/30694204/)

©Giang Le, Nghia Khuu, Van Thi Thu Tieu, Phuc Duy Nguyen, Hoa Thi Yen Luong, Quang Duy Pham, Hau Phuc Tran, Thuong Vu Nguyen, Meade Morgan, Abu S Abdul-Quader. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 29.01.2019. 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 Population Size of Female Sex Workers in Namibia Using a Respondent-Driven Sampling Adjustment to the Reverse Tracking Method: A Novel Approach

Paul Douglas Wesson^{1*}, PhD; Rajatashuvra Adhikary^{2*}, PhD; Anna Jonas^{3,4}, MPH, MPA; Krysta Gerndt⁵, MPH; Ali Mirzazadeh⁶, PhD, MD, MPH; Frieda Katuta³, NP; Andrew Maher⁵, MPH; Karen Banda⁵, PhD, MPH; Nicholas Mutenda³, MSc; Willi McFarland⁶, MD, MPH, PhD; David Lowrance⁴, MD, MPH; Dimitri Prybylski⁴, PhD, MPH; Sadhna Patel⁴, MPH

¹Center for AIDS Prevention Studies, Division of Prevention Science, University of California, San Francisco, San Francisco, CA, United States

²Strategic Information/Monitoring and Evaluation, WHO India Country Office, Gurugau, Haryana, India

³Directorate of Special Programmes, Response Monitoring & Evaluation Subdivision, Ministry of Health and Social Services, Windhoek, Namibia

⁴Division of Global HIV and Tuberculosis, US Centers for Disease Control and Prevention, Atlanta, GA, United States

⁵Institute for Global Health Sciences, University of California San Francisco, San Francisco, CA, United States

⁶Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, United States

* these authors contributed equally

Corresponding Author:

Paul Douglas Wesson, PhD

Center for AIDS Prevention Studies

Division of Prevention Science

University of California, San Francisco

550 16th Street, 3rd Floor

San Francisco, CA, 94158

United States

Phone: 1 415 476 6288

Email: paul.wesson@ucsf.edu

Abstract

Background: Key populations, including female sex workers (FSWs), are at a disproportionately high risk for HIV infection. Estimates of the size of these populations serve as denominator data to inform HIV prevention and treatment programming and are necessary for the equitable allocation of limited public health resources.

Objective: This study aimed to present the respondent-driven sampling (RDS) adjusted reverse tracking method (RTM; RadR), a novel population size estimation approach that combines venue mapping data with RDS data to estimate the population size, adjusted for double counting and nonattendance biases.

Methods: We used data from a 2014 RDS survey of FSWs in Windhoek and Katima Mulilo, Namibia, to demonstrate the RadR method. Information from venue mapping and enumeration from the survey formative assessment phase were combined with survey-based venue-inquiry questions to estimate population size, adjusting for double counting, and FSWs who do not attend venues. RadR estimates were compared with the official population size estimates, published by the Namibian Ministry of Health and Social Services (MoHSS), and with the unadjusted RTM.

Results: Using the RadR method, we estimated 1552 (95% simulation interval, SI, 1101-2387) FSWs in Windhoek and 453 (95% SI: 336-656) FSWs in Katima Mulilo. These estimates were slightly more conservative than the MoHSS estimates—Windhoek: 3000 (1800-3400); Katima Mulilo: 800 (380-2000)—though not statistically different. We also found 75 additional venues in Windhoek and 59 additional venues in Katima Mulilo identified by RDS participants' responses that were not detected during the initial mapping exercise.

Conclusions: The RadR estimates were comparable with official estimates from the MoHSS. The RadR method is easily integrated into RDS studies, producing plausible population size estimates, and can also validate and update key population maps for outreach and venue-based sampling.

KEYWORDS

population density; social networking; sex workers; vulnerable populations; human immunodeficiency virus

Introduction

Background

Over the past decade, population size estimation (PSE) has become increasingly important in the global fight against the HIV epidemic, particularly among key populations [1,2]. Key populations—for example, men who have sex with men, female sex workers (FSWs), and people who inject drugs—face a disproportionate burden of HIV infection, relative to the general population, because of behavioral risk factors [3-5]. Size estimation for key populations is necessary to estimate the absolute burden of disease by serving as the denominator for the population at risk [6]. Furthermore, PSE for key populations is used to evaluate the reach and coverage of existing services for key populations as well as for program planning and allocation of limited public health resources to most effectively combat local HIV epidemics [7]. PSE of key populations is also necessary for tracking progress toward 90-90-90 diagnosis and treatment targets: by 2020, 90% of people living with HIV (PLWH) knowing their status, 90% of diagnosed PLWH receiving antiretroviral treatment, and 90% of PLWH on treatment being virally suppressed [8]. Without PSE, investigators and policy makers cannot translate these proportional targets into absolute numbers to ensure appropriate resource allocation. Currently, there is no gold standard method to estimate the size of key populations [9]. Instead, researchers must choose from a menu of different PSE methods, each leveraging information about the population size from various data sources while also being vulnerable to different biases [6].

The reverse tracking method (RTM), for example, is an underutilized PSE method that leverages information about the venues visited by a key population to estimate population size [10]. The RTM takes a 2-stage approach to size estimation. During the first stage, researchers map out the venues where the key population can be found (the venue-based sampling frame) and obtain the size for each venue, M_i (ie, how many members of the key population can be found there), typically from at least 3 key informants at each venue or from direct observation by the research team. During the second stage, the researchers return to either all or a sample of the venues from the sampling frame, sampled with probability proportional to the size of the venue, and directly count the number of people who are members of the key population for each venue, N_i . The ratio of N_i/M_i , averaged over all the venues from the second visit, is a correction factor that is then multiplied by the sum of the counts from the first stage to estimate the population size.

Although the RTM creatively leverages venue-based data, the method relies on several strong assumptions. First, the RTM implicitly assumes that all members of the key population can be found at physical venues that are known to the researchers. Second, the method assumes that members of the key population exclusively *belong* to a single venue; the equation does not

adjust for double counting [10]. In addition, to efficiently use resources allocated for research, this venue-based PSE method is best integrated into a study that uses venue-based sampling to find the key population, such as time location sampling (TLS) [11]. However, many surveillance studies of key populations use respondent-driven sampling (RDS), a peer referral-based sampling method [12]. As such, the RTM is inefficient, from a resource perspective, to include as a PSE method for RDS-based surveillance studies of key populations.

Objectives

We sought to modify the RTM for the RDS study design. By including *venue inquiry* questions in the RDS survey to serve as a *virtual second visit* and combining that information with the size of mapped venues (ie, the number of members of the key population found at the mapped venue) obtained before the RDS survey during the formative assessment stage, we successfully developed a novel PSE method that leverages venue information within the RDS context. With additional information from the survey, we adjusted our size estimate for double counting (accounting for people who visit multiple venues) and for people who do not attend any venues, thereby overcoming notable limitations of the original RTM. We refer to this approach as the RDS-adjusted RTM (RadR, pronounced “radar”). In this paper, we describe the RadR method and demonstrate its implementation using FSW data from an integrated biobehavioral surveillance study (IBBSS) in Namibia as a case example.

Methods

Study Design

From September 2012 to June 2014, the Namibian Ministry of Health and Social Services (MoHSS) partnered with the US Centers for Disease Control and Prevention (CDC) and the University of California, San Francisco (UCSF), to conduct a cross-sectional IBBSS of FSWs. The purpose of this study was to measure the prevalence of HIV infection among the FSW population, to assess HIV-related risk, preventive, and health-seeking behaviors, and to estimate the size of the FSW population in selected urban areas. The IBBSS was implemented in 4 sites on the basis of HIV prevalence from sentinel surveillance and the availability of community-based organizations and other organizations providing services to the FSW population. The IBBSS took place in Windhoek (the capital and largest city), Walvis Bay and Swakopmund (neighboring cities on the coast), Oshikango and Oshikati (neighboring towns in the northern region), and Katima Mulilo (a border town in the northeastern region, receiving traffic from Angola, Botswana, Zambia, and Zimbabwe) [13].

Study Subjects

Participants were recruited for the IBBSS using RDS. RDS is a social networking-based sampling and analytic approach

[14,15]. Briefly, the sampling process begins with a selection of *seeds* members of the target population purposively selected by the research team, to often represent the presumed diversity of the target population. Each seed is given a limited number of coupons (eg, 3), to recruit other members of the target population (eg, FSWs) from within their social network to participate in the study. After participating in this study, recruiters are also given a limited number of coupons to recruit people from their social network. These coupons are used to participate in the study and for the researchers to track who was recruited by whom. In addition to the primary incentive that everyone receives for participating in the study, recruiters also receive a secondary incentive for every one of their recruits who successfully participates in the study. This process of recruitment continues until the sample size is reached and equilibrium (ie, when additional recruitment does not substantially change the sample characteristics/proportions) is achieved for selected variables. In theory, with enough waves of recruitment and adjustment for the network recruitment design, the final RDS sample will approximate a representative sample of the target population, independent of the characteristics of the initial seeds. The number and selection of seeds are also guided by practical study implementation considerations (eg, people well connected to the social network to begin and facilitate active recruitment and chosen for diversity with respect to key demographic and behavioral variables). The MoHSS report provides the results of the main objectives of the IBBSS as well as specifics on the implementation of the survey among FSWs [13]. For a more detailed discussion of RDS implementation, we refer the reader to Johnston et al's systematic review of published RDS studies [12] and Gile et al's study on RDS diagnostics [16].

For the Namibia IBBSS, participants were eligible for the study if they met the following criteria: were at least 18 years of age, biologically female, able to speak English, Oshiwambo, Silozi, or Afrikaans, exchanged vaginal, anal, or oral sex for money during the 30 days preceding the IBBSS, and were residents of the study area for at least 6 months preceding the IBBSS. In total, 12 seeds were selected in Windhoek and 8 seeds were selected in Katima Mulilo to represent the diversity of the population with respect to age, cultural and linguistic background, and socioeconomic status. Recruitment was conducted from September 2012 to August 2013 for Windhoek and from October 2013 to June 2014 for Katima Mulilo. The number of coupons distributed to each person ranged from 3 to 11 to facilitate a diverse sample and aid the progress of recruitment in each city.

For the purpose of demonstrating the RadR method, we restricted our analysis to 2 of the 4 IBBSS sites, Windhoek and Katima Mulilo. Unlike the other sites, the Windhoek and Katima Mulilo sites did not combine multiple towns into a single site.

The mapping and count, the first stage of the RTM, took place during a 3-week formative assessment period beginning 3 weeks before the start of the RDS survey. Venues where FSWs congregate and find clients were identified on the basis of key informant interviews and focus group discussion with members of the FSW community who were identified by local nongovernmental organizations (NGOs) providing services to key populations in Namibia. Key informants were asked to

name and identify known or frequented venues or hotspots. A list and map of venues were compiled for each population in each site. As fewer than 30 venues were named in each site, a complete census was taken at all venues identified. Types of venues identified by key informants principally included bars, night clubs, hotels and guesthouses, streets, as well as petrol and service stations. Study teams and individuals familiar with the local context visited the venues for observation and direct counts of the study populations for approximately 5 hours on the peak times indicated by key informants.

Measures

As a *virtual second visit*, we asked each RDS participant to name the venues that she goes to most frequently to find or solicit clients. Specifically, participants were asked, "What is the name of the venue you go to most frequently to find clients?" and "In the past 30 days, how often did you attend this venue?" Participants could list up to 3 venues or respond that they do not go to venues to find clients. Responses were open-ended; participants provided names of the venues they attended most frequently as opposed to selecting from a predetermined list. After data collection, all responses were tabulated and summed by venue.

Data Analysis

The RTM approach to size estimation is calculated according to equation (a) in Figure 1.

\hat{S} is the estimated population size, n is the number of venues visited on the second visit, N_i is the number of people observed at venue i on the second visit, M_i is the number of people observed at venue i on the first visit (or reported by a key informant for venue i), and M is the total number of people observed at all venues on the first visit.

For the RadR method, M is the total number of people observed at all venues on the first visit (from the mapping and enumeration exercise during the formative assessment), M_i is the number of people observed at venue i during the first visit, and N_i is the number of people observed at venue i during the *virtual second visit*. The original RTM equation is insufficient for RDS studies as the virtual second visit (N_i) is taken from the RDS sample that reports attending venue i instead of from the target population that is physically observed at venue i . We therefore augmented the original RTM equation with additional correction factors to standardize to the target population and leverage additional information provided by the RDS survey through inverse probability weights. The augmented equation is given as equation (b) in Figure 1.

C1 (correction factor 1) is the *standardization parameter*, used to standardize the study population to the target population. It is calculated as $1/(R^*/RDS)$, where R^* is the number of RDS respondents who visited a mapped venue (ie, a venue included in the mapping phase), and RDS is the RDS sample size. The standardization parameter assumes that the RDS sample is representative of the target population. Formally, we assume the following equivalency, $n/N = R^*/RDS$, where n is the number of people from the target population who are observed at venues, and N is the number of people in the target population. C2

(correction factor 2), the *visibility parameter*, accounts for the visible, or reachable, population; M is upweighted to account for people who attend venues previously unmapped, and they were therefore not a part of the original sampling frame. This parameter is calculated as $(1-r)/p$, where r is the proportion of the RDS sample that reports not going to any venues and p is the proportion of the RDS sample that reports attending a mapped venue. C3 (correction factor 3), the *hidden population parameter*, accounts for people who do not go to venues. This parameter is calculated as $1/(1-r)$. Finally, C4 (correction factor 4), the *double counting parameter*, accounts for people who go to multiple venues. This parameter is calculated as $1-(s/2)-(2t/3)$, where s is the proportion of venue-attending RDS participants who attend 2 venues, and t is the proportion of venue-attending RDS participants who attend 3 venues. For C1, R^*/RDS is equivalent to p in C2. The equation for RadR then simplifies to equation (c) in Figure 1.

To calculate 95% simulation intervals (SIs), we created probability distributions for the correction factors and resampled from those distributions. Drawing on the simplified RadR equation above, we fit the RDS-weighted values of p , s , and t (point estimates and 95% CI) to beta distributions. We assume a beta distribution here as this family of distributions is flexible and convenient for fitting quantities that are constrained to values between 0 and 1 [17]. Resampling from these distributions 10,000 times, we calculated the values of the simplified correction factors, storing the product, $(1/p)^2 * (1-(s/2)-(2t/3))$, for each iteration, thus creating a distribution of the simplified correction factors. The 2.5 and 97.5 percentile values were then obtained and multiplied by the fixed values portion of the RadR equation to calculate the 95% SI. The calculations are illustrated in Figure 2.

To assess the performance of RadR method, we chose 2 comparisons. First, as RadR shares underlying assumptions and theoretically and practically builds upon the RTM method, we compare its results with the (unadjusted) results of RTM [6,10]. Second, to assess its acceptability and usefulness to policy makers, we also compared the results with the official PSE for FSWs from the MoHSS [13]. These results were adopted following a stakeholder consensus following a *modified Delphi* method [5,6,13,18]. Representatives from the MoHSS, CDC,

the US Agency for International Development, local NGOs working with the FSW population, and FSW population members convened at a stakeholder workshop following data collection for the IBBSS. Each stakeholder provided an initial estimate for the FSW population in the study site on the basis of their experience with the population. These estimates were then allowed to be revised after stakeholders had the chance to discuss the rationale behind their estimates and after seeing the empirical results from several PSE methods that were included in the IBBSS (ie, key informant interview, unique object multiplier, wisdom of the crowd, and literature review, but not the RTM or RadR, which was not available at the time of the stakeholder meeting). The median of the revised stakeholder estimates was presented as the official population size estimate in the MoHSS report [13].

RDS-weighted values were calculated using RDS-A software version 0.42 (Handcock, Fellows, and Gile) [19]. The RDS-II estimator was used to calculate RDS-weighted point estimates and 95% CI [20]. Imputed visibility, a measure of a person's connectedness in the social network, was used in place of network size for RDS-weighted estimates [21]. Self-reported network size may be a biased representation of a person's position and influence in a network because of recall bias, digit preference for round numbers (eg, people reporting a network size of 20 rather than 23), and access to yet unsampled members of the target population at the time of recruitment. Imputed visibility overcomes these potential biases by leveraging self-reported network size, the time during the sampling process during which study participants were recruited, and the number of people study participants were able to recruit. R statistical software version 3.4.1 (R Core Team) was used to estimate 95% SIs [22].

Ethics Approval

The protocol for the main IBBSS received approval from the Research Committee of the Directorate for Policy, Planning and Human Resources of the MoHSS in Windhoek, Namibia, the Committee on Human Research at the UCSF in San Francisco, California, USA, and the Division of Global HIV and Tuberculosis in the CDC, Atlanta, Georgia, USA. All study participants provided verbal informed consent before enrollment in the survey.

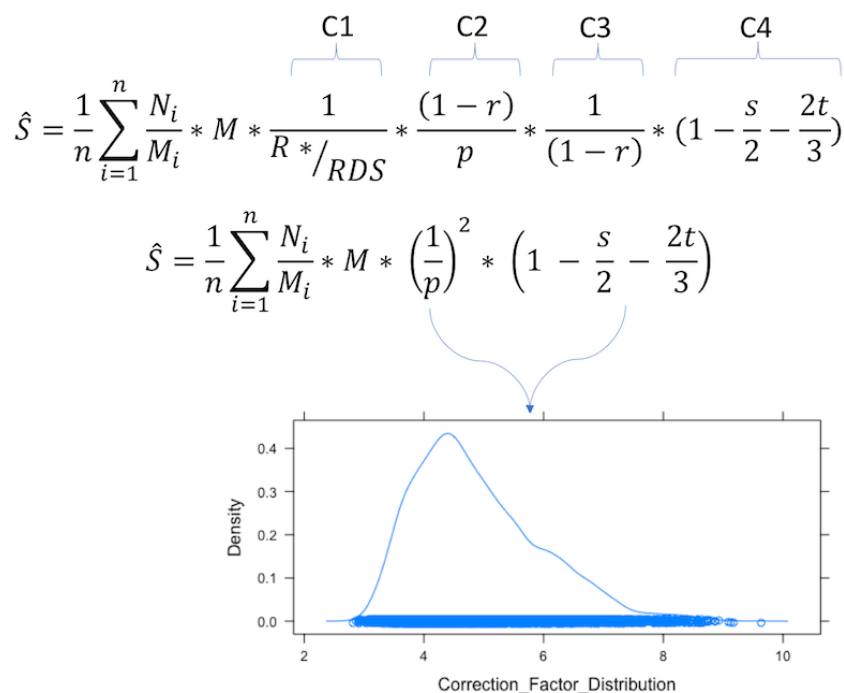
Figure 1. Equations: (a) Reverse Tracking Method, (b) Respondent-driven sampling adjusted Reverse Tracking Method, and (c) simplified Respondent-driven sampling adjusted Reverse Tracking Method.

$$(a) \quad \hat{S} = \frac{1}{n} \sum_{i=1}^n \frac{N_i}{M_i} * M$$

$$(b) \quad \hat{S} = \frac{1}{n} \sum_{i=1}^n \frac{N_i}{M_i} * M * C1 * C2 * C3 * C4$$

$$(c) \quad \hat{S} = \frac{1}{n} \sum_{i=1}^n \frac{N_i}{M_i} * M * \left(\frac{1}{p}\right)^2 * \left(1 - \frac{s}{2} - \frac{2t}{3}\right)$$

Figure 2. Respondent-driven sampling adjusted reverse tracking method equation (complete and simplified). Correction factors are calculated from the simulated distributions, from which the 2.5 percentile and 97.5 percentile are used to calculate 95% simulation intervals. Key: \hat{S} =the estimated population size; n =the number of venues visited on the second visit; N_i =the number of people observed at venue i on the second visit; M_i =the number of people observed at venue i on the first visit; M =the total number of people observed at all venues on the first visit; R^* =the sum of the number of times that a mapped venue is reported from the venue inquiry questions; RDS =the RDS sample size; r =the proportion of the RDS sample that reports not going to any venues; p =the proportion of the RDS sample that report attending a mapped venue; s =the proportion of venue-attending RDS participants who attend two venues; t =the proportion of venue-attending RDS participants who attend three venues.



Results

Sampling/Recruitment

In Windhoek, 10 seeds were initially selected to begin recruitment and 9 additional seeds were added to increase the pace of recruitment. All 19 seeds were productive recruits, resulting in a total of 316 participants sampled over 7 waves of recruitment. Of the total number of coupons distributed in Windhoek, 28.8% (366/1271) were returned by potential participants. In Katima Mulilo, 9 seeds were identified to recruit for the study; however, 1 seed was found to be ineligible. The remaining 8 seeds were productive recruits, resulting in 309 participants sampled over 11 waves of recruitment. Of the total number of coupons distributed in Katima Mulilo, 48.0% (426/887) were returned by potential participants. FSWs were younger in Katima Mulilo (mean age 27.3 years) compared with Windhoek (mean age 30.3 years). The majority of FSWs in both locations had achieved at least a secondary school education. FSWs in Windhoek reported more client partners in the 30 days preceding the interview compared with FSWs in Katima Mulilo. RDS-weighting of the sample indicated that over half of the FSWs in Katima Mulilo are HIV positive (56.8%; 177/309) compared with nearly one-third of FSWs in Windhoek (32.6%; 103/316; [Table 1](#)).

In Windhoek, 31.6% (100/316) of FSWs reported not visiting venues to find clients; 37.6% (122/316) of FSWs also reported that they had visited at least one of the venues previously mapped by the research team during the formative assessment. In comparison, in Katima Mulilo, 17.0% (56/309) of FSWs reported not visiting venues to find clients; 42.3% (127/309) of FSWs also reported that they had visited at least one of the venues previously mapped by the research team during the formative assessment. This percentage increases to nearly 70% if venues are included that were mapped by the research team but no FSWs were observed by the research team during the enumeration exercise. [Figure 3](#) depicts the RDS recruitment tree by site. Each node, representing a participant, is scaled according to venue attendance; large nodes indicate the participant did not visit venues to find clients. The venue inquiry questions identified over 75 previously unmapped venues in Windhoek and 59 previously unmapped venues in Katima Mulilo.

Population Size

Using the RadR method, the FSW population size was estimated at 1552 (95% SI: 1101-2387) in Windhoek, corresponding to roughly 1.8% of the adult female population. The FSW population size in Katima Mulilo was estimated at 453 (95% SI: 336-656), corresponding to roughly 4.9% of the adult female population. [Table 2](#) compares these estimates with the stakeholder consensus and the unadjusted RTM estimate.

Table 1. Demographics and descriptive statistics for respondent-driven sampling sample of female sex workers in Katima Mulilo and Windhoek, Namibia (because of rounding, percentages may not sum to 100%).

RDS ^a participant characteristics	Katima Mulilo		Windhoek	
	n	RDS Adjusted % (95% CI)	n	RDS Adjusted % (95% CI)
Age (years), mean (minimum-maximum)	309	27.3 (18-53)	316	30.3 (18-65)
Education				
Primary/less than primary	89	28.6 (22.9-34.1)	116	36.7 (30.6-42.7)
Secondary	219	71.1 (70.4-71.7)	198	62.6 (61.9-63.3)
Vocational/Technical	0	0	2	0.7 (0.0-6.8)
Client partners during 30 days preceding interview				
<5	186	60.9 (55.1-66.9)	139	42.6 (36.8-48.4)
5-9	88	27.9 (22.9-32.8)	94	30.9 (25.2-36.5)
10-14	19	6.4 (3.6-9.2)	24	7.4 (4.3-10.4)
>15	16	4.8 (2.1-7.4)	59	19.2 (15.0-23.3)
Marital status				
Never married	250	81.6 (77.2-86.1)	276	87.8 (83.5-92.1)
Previously or currently married	59	18.4 (14.0-22.8)	40	12.2 (7.9-16.5)
HIV status^b				
Positive	177	56.8 (50.1-63.4)	103	32.6 (26.7-38.5)
Negative	131	43.0 (36.3-49.6)	206	65.7 (59.7-71.7)
Venues visited to find clients				
I do not go to venues	56	17.0 (11.5-22.4)	100	31.6 (26.1-37.2)
1 venue	32	11.1 (7.2-15.0)	84	27.9 (22.5-33.4)
2 venues	106	35.0 (28.7-40.3)	80	24.8 (19.7-30.0)
3 venues	115	37.4 (31.8-43.1)	51	15.2 (11.0-20.0)
Refuse to Answer	— ^c		1	— ^c
Visited >1 of the mapped venues				
Yes	127	42.3 (35.1-49.5)	122	37.6 (30.2-44.9)
No	182	57.7 (50.5-64.9)	194	62.4 (55.1-69.8)
Correction factors				
C1: standardization parameter		2.36 ^d		2.66
C2: visibility parameter		1.96		1.82
C3: hidden population parameter		1.21		1.46
C4: double counting parameter		0.49		0.67

^aRDS: respondent-driven sampling.^b§:Stratified counts do not sum to N because of indeterminate HIV test results.^cNot applicable.^d95% CI values are not applicable.

Figure 3. Respondent-driven sampling recruitment tree of female sex workers in Katima Mulilo, Namibia and Windhoek, Namibia. Large nodes indicate participants who report not attending venues to find clients.

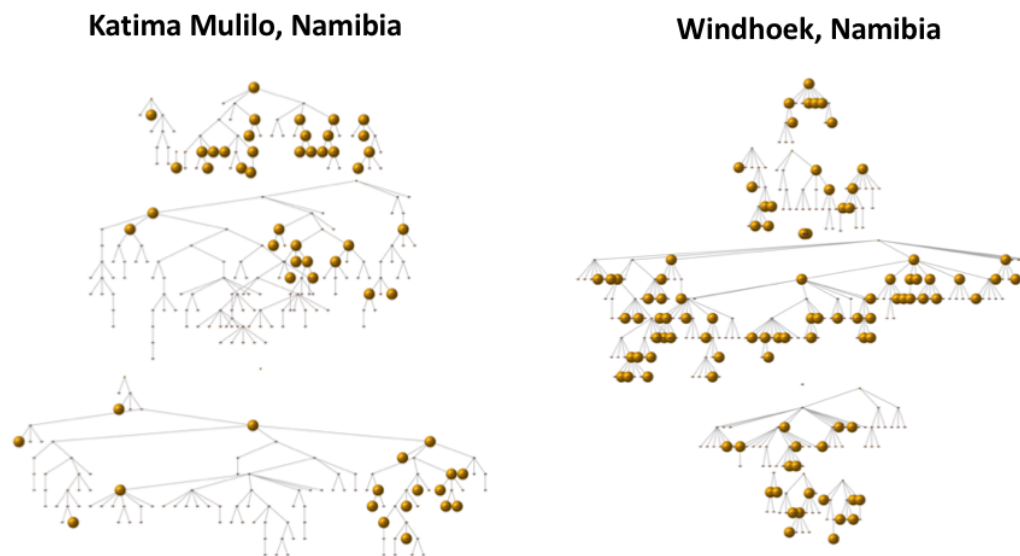


Table 2. Population size estimates of female sex workers by study site and population size estimation method (we use “Acceptable bounds” here as an umbrella term as some methods report 95% CIs, other methods report plausibility bounds, and the respondent-driven sampling adjusted reverse tracking method reports 95% simulation intervals).

Study site and population size estimates method	Estimated number of FSWs ^a (acceptable bounds)	Estimated percentage of adult (15 to 49 years) female population who are FSWs (acceptable bounds)
Windhoek		
Stakeholder consensus ^b	3000 (1800-3400)	2.2 (2.0-3.6)
Reverse tracking method	492 (418-565)	0.56 (0.48-0.64)
RadR ^c	1552 (1101-2387)	1.77 (1.25-2.72)
Katima Mulilo		
Stakeholder consensus	800 (380-2,000)	8.6 (4.1-21.5)
Reverse tracking method	192 (181-203)	2.06 (1.95-2.18)
RadR	453 (336-656)	4.85 (3.60-7.03)

^aFSWs: female sex workers.

^bThe stakeholder consensus was informed by the following population size estimates methods: key informant interview, unique object multiplier, wisdom of the crowd, and literature review.

^cRadR: respondent-driven sampling adjusted reverse tracking method.

Discussion

Principal Results

We estimated the size of the FSW population to be 1552 (95% SI: 1101-2387) in Windhoek and 453 (95% SI: 336-656) in Katima Mulilo, using the RadR method. FSW size estimates were notably larger when using the RadR method compared with the unadjusted RTM (1552 vs 492 in Windhoek and 453 vs 192 in Katima Mulilo). This is expected as the RadR method was conceived to explicitly account for the *hidden* members of the key population in its calculation (ie, those who cannot be

found at physical venues but participated in the RDS survey), whereas the unadjusted RTM relies only on the observable population when estimating the population size. Accounting for the *hidden* members is especially important in research involving key populations as the social marginalization often faced by these groups may result in a sampling bias that may be particularly strong for venue-based study designs. For example, we estimated that over 30% of the FSW population in Windhoek and 17% of the FSW population in Katima Mulilo would not be found at venues (Table 1, Figure 3). Therefore, we believe that incorporating such correction factors as the *hidden population parameter* improves the validity of the RadR

population size estimates compared with the unadjusted RTM estimates.

The RadR estimates are consistent with results from the official FSW size estimates from the stakeholder consensus with respect to having overlapping CIs or SIs. Of note, in Katima Mulilo, there is a substantial difference between the stakeholder consensus estimate and the unadjusted RTM estimate. However, the RadR estimate is in closer agreement with the final stakeholder estimate, illustrating the impact of our correction factors on calculating a more plausible population size estimate. Still, the RadR estimate is slightly more conservative than the stakeholder consensus. This may be because of FSW demographic groups, such as higher income FSWs, who may not be observed in either the census mapping or the RDS study. This stratum of FSWs would then be absent from the RadR calculation, whereas stakeholders who have knowledge of this group may incorporate them into their estimation of the population size.

Strengths

In addition to calculating plausible size estimates for the target population, the RadR method advances the PSE field and key population surveillance in 4 ways. First, the venue inquiry questions serve to validate the existing sampling frame and census mapping. Responses to these questions indicate whether the target population actually attends the venues identified during the formative assessment. In addition, the frequency with which a venue is reported in the survey provides some insight into the popularity of that venue among the target population, assuming that RDS recruits who attend the same venue are not more likely to recruit each other. A second improvement is that the RadR method expands and updates venue-based mapping (such as those used by outreach programs) and potential sampling frames (such as those used in TLS surveys, conventional cluster sampling, and census mapping). Additional venues previously unknown to the research team and not included in the original mapping exercise can be identified. Taken together, these first 2 advancements inform more targeted and efficient mapping, outreach efforts, and venue-based sampling frames. Third, the RadR method further advances the PSE field by leveraging information collected in the RDS survey to account for double counting. The original RTM makes the strong assumption that people exclusively “belong” to 1 venue. This assumption can be evaluated using the venue inquiry questions. If participants attend multiple venues, as was the case in Namibia, this information is collected and used to adjust the size estimate appropriately. Finally, the RadR method advances the PSE field and improves upon the original RTM by accounting for the proportion of the key population that is virtually *invisible* to venue-based sampling as the members in that proportion do not go to venues to find clients. The RDS methodology has often been credited with finding the more hidden members of the key population [23] and generating a more representative sample of the population [24]. The RadR method leverages this quality of the sampling design and the data collected on nonvenue attendance to calculate an inverse-probability weight for venue attendance. The inverse-probability weight adjusts the venue-based population size estimate to also account for the segment of the population

that cannot be found at physical venues (eg, those who find clients or sex partners through social networking websites).

Limitations

The RadR method assumes that the RDS sample is representative of the target population. This assumption is especially necessary for C1, the standardization parameter. If the RDS sample is representative of the target population, then R^*/RDS should reflect the same relationship in the broader target population, that is, the number of people visiting a mapped venue divided by the total number of people in the population. If the representativeness of the RDS sample is of concern, several diagnostic approaches such as bottleneck plots and convergence plots are recommended to evaluate the sample [16]. These diagnostic approaches are available in RDS-A. Nonetheless, RDS has noted limitations in implementation and underlying assumptions that are difficult to prove [11,12,25], and therefore it remains uncertain if it consistently produces a truly representative sample.

We assume that study participants are not more likely to recruit other participants who attend the same venues. Although we were unable to investigate homophily (the likelihood that respondents preferentially recruit others who are similar to themselves on specified characteristics) for specific venues, we did assess this measure for overall venue attendance. Using RDS-A's recruitment homophily function, we explored homophily by attending any venue and attending a mapped venue. In Windhoek, we found evidence for recruitment homophily with respect to reporting a mapped venue (Chi-squared $P=.005$). In Katima Mulilo, we found evidence for recruitment homophily with respect to reporting any venue attendance (Chi-squared $P=.002$) and reporting a mapped venue (Chi-squared $P<.001$). Considering how this might affect the RadR estimates, recruitment homophily would suggest that the RDS sample may not be representative of the underlying target population. This may violate our assumption that $n/N=R^*/RDS$ and impact our estimation of the components for the correction factors. One potential solution may be to use the RDS-I estimator, which is designed to account for patterns of recruitment among subgroups [25-27], and re-estimate p , s , and t , which are used in the simplified RadR formula. Using the RDS-I estimator, we re-estimated the Windhoek FSW population to be 1720 (95% SI: 1198-2640)—compared with the original RadR estimate of 1552 (95% SI: 1101-2,387); we re-estimated the Katima Mulilo FSW population to be 405 (95% SI: 303-574)—compared with the original RadR estimate of 453 (95% SI: 336-656). In this case, our approach to account for recruitment homophily did not result in substantially different population size estimates. However, additional research may be warranted to investigate the impact of recruitment homophily on RadR estimates in other populations.

To reduce survey fatigue, we limited our venue-inquiry questions to, at most, 3 venues. It is possible that FSWs attended more than 3 venues to find clients. Allowing for the inclusion of additional venues could expand the census mapping but may have unpredictable results for the RadR estimate, depending on whether the additional venues mentioned were mapped. Statistical modeling studies may be appropriate to determine

the optimal number of venues inquired about to balance the rewards of additional information with potentially diminishing statistical returns. Investigators may also consider asking participants for the total number of venues visited before inquiring further about the 3 most often visited venues. This additional information could provide better insight into the mobility of the key population among local venues.

Although the RadR method can easily be integrated into the RDS survey with the addition of a few questions, the initial data setup can be labor intensive. Responses to the venue inquiry questions are open-ended, requiring researchers to identify and assess multiple ways of spelling the same venue name and recode these multiple references as the same venue. Researchers must then be cautious of mismatched venue names, that is, different venue names referring to the same venue or similar venue names actually referring to different venues. The potential for mismatched venue names is a limitation in this study. Future studies that implement the RadR method should collaborate with local researchers to confirm the correct matching of venue names. However, following this initial time investment, the PSE calculation is straightforward, and the data can be tabulated to validate (and update) the venue-based sampling frame and census mapping.

Conclusions

Despite these limitations, we found that the RadR method is easily integrated into RDS studies, leveraging already collected

data from a census mapping of venues during the formative assessment stage. In fact, this approach to size estimation could still be used if the census mapping and enumeration took place independently of the formative assessment for the RDS study (eg, client mapping by key population programs). Investigators must consider whether the population enumerated during the separate census mapping is the same population that is being surveyed for the RDS study. Our census mapping took place 3 weeks before the RDS study. Investigators must consider the mobility of the population when determining whether separate census mapping and enumeration can reasonably serve as a *first visit* for the target population before applying the RDS adjustment. If the target population is highly mobile, in the sense that a substantial proportion of the population either left the study site or changed venue attendance behavior in the period between the census mapping and the RDS study, then the approach detailed in this paper would not be appropriate. The RadR method improves upon the unadjusted RTM by further collecting information on multiple venues visited and the proportion of the members of the population who do not visit venues to find clients or sex partners. In addition to calculating plausible size estimates, as demonstrated here, the RadR method directly informs public health (prevention) programming by updating the census mapping and identifying venues where outreach services can take place.

Acknowledgments

PW was funded by a National Institutes of Health postdoctoral training grant (T32 MH19105).

This project was supported in part by the President's Emergency Plan for AIDS Relief (PEPFAR) through the CDC (under the terms of award number PS001468-05).

PEPFAR/CDC/other agency authorship disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the funding agencies. The manuscript required internal review and approval by the funder (CDC) before authorization for publication.

Conflicts of Interest

None declared.

References

1. Schwartländer B, Stover J, Hallett T, Atun R, Avila C, Gouws E, Investment Framework Study Group. Towards an improved investment approach for an effective response to HIV/AIDS. *Lancet* 2011 Jun 11;377(9782):2031-2041. [doi: [10.1016/S0140-6736\(11\)60702-2](https://doi.org/10.1016/S0140-6736(11)60702-2)] [Medline: [21641026](https://pubmed.ncbi.nlm.nih.gov/21641026/)]
2. The Office of the U.S. Global AIDS Coordinator. The United States President's Emergency Plan for AIDS Relief. Washington, DC: United States Agency for International Development; 2014 12. PEPFAR 3.0 - Controlling the Epidemic: Delivering on the Promise of an AIDS-free Generation URL: <https://www.pepfar.gov/documents/organization/234744.pdf> [accessed 2019-02-19] [WebCite Cache ID 76JBipMZ]
3. Githuka G, Hladik W, Mwalili S, Cherutich P, Muthui M, Gitonga J, KAIS Study Group. Populations at increased risk for HIV infection in Kenya: results from a national population-based household survey, 2012. *J Acquir Immune Defic Syndr* 2014 May 01;66(Suppl 1):S46-S56 [FREE Full text] [doi: [10.1097/QAI.0000000000000137](https://doi.org/10.1097/QAI.0000000000000137)] [Medline: [24732821](https://pubmed.ncbi.nlm.nih.gov/24732821/)]
4. Mumtaz G, Riedner G, Abu-Raddad LJ. The emerging face of the HIV epidemic in the Middle East and North Africa. *Curr Opin HIV AIDS* 2014 Mar;9(2):183-191 [FREE Full text] [doi: [10.1097/COH.0000000000000038](https://doi.org/10.1097/COH.0000000000000038)] [Medline: [24445372](https://pubmed.ncbi.nlm.nih.gov/24445372/)]
5. Okal J, Geibel S, Muraguri N, Musyoki H, Tun W, Broz D, et al. Estimates of the size of key populations at risk for HIV infection: men who have sex with men, female sex workers and injecting drug users in Nairobi, Kenya. *Sex Transm Infect* 2013 Aug;89(5):366-371 [FREE Full text] [doi: [10.1136/sextrans-2013-051071](https://doi.org/10.1136/sextrans-2013-051071)] [Medline: [23761166](https://pubmed.ncbi.nlm.nih.gov/23761166/)]

6. Wesson P, Reingold A, McFarland W. Theoretical and empirical comparisons of methods to estimate the size of hard-to-reach populations: a systematic review. *AIDS Behav* 2017 Jul;21(7):2188-2206. [doi: [10.1007/s10461-017-1678-9](https://doi.org/10.1007/s10461-017-1678-9)] [Medline: [28078496](#)]
7. Abdul-Quader AS, Baughman AL, Hladik W. Estimating the size of key populations: current status and future possibilities. *Curr Opin HIV AIDS* 2014 Mar;9(2):107-114. [doi: [10.1097/COH.0000000000000041](https://doi.org/10.1097/COH.0000000000000041)] [Medline: [24393694](#)]
8. Joint United Nations Programme on HIV/AIDS. Joint United Nations Programme on HIV/AIDS. Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS; 2014. 90-90-90: An Ambitious treatment target to help end the AIDS epidemic URL: http://www.unaids.org/sites/default/files/media_asset/90-90-90_en.pdf [accessed 2019-02-19] [WebCite Cache ID 76JBqudjir]
9. UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance. UNAIDS. Geneva, Switzerland: Co-published by World Health Organization and UNAIDS; 2010. Guidelines on Estimating the Size of Populations Most at Risk to HIV URL: http://data.unaids.org/pub/manual/2010/guidelines_popnestimationsize_en.pdf [accessed 2019-02-18] [WebCite Cache ID 76JD5jnVw]
10. Vadivoo S, Gupte M, Adhikary R, Kohli A, Kangusamy B, Joshua V, IBBA Study Team. Appropriateness and execution challenges of three formal size estimation methods for high-risk populations in India. *AIDS* 2008 Dec;22(Suppl 5):S137-S148. [doi: [10.1097/01.aids.0000343772.82647.2d](https://doi.org/10.1097/01.aids.0000343772.82647.2d)] [Medline: [19098474](#)]
11. Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS* 2005 May;19(Suppl 2):S67-S72. [Medline: [15930843](#)]
12. Johnston L, Hakim A, Dittrich S, Burnett J, Kim E, White RG. A systematic review of published respondent-driven sampling surveys collecting behavioral and biologic data. *AIDS Behav* 2016 Dec;20(8):1754-1776. [doi: [10.1007/s10461-016-1346-5](https://doi.org/10.1007/s10461-016-1346-5)] [Medline: [26992395](#)]
13. Ministry of Health and Social Services. Ministry of Health and Social Services. Windhoek, Namibia: Directorate of Special Programmes; Response Monitoring & Evaluation Subdivision; 2015. Results of the Integrated Biological and Behavioral Surveillance Studies among Female Sex Workers in Namibia URL: http://www.aidsinfoonline.org/kpatlas/document/NAM/NAM_2014_IBBS_PSE_FSW.pdf [accessed 2019-02-24] [WebCite Cache ID 76Q2N0UmU]
14. Heckathorn D. Respondent-driven sampling: a new approach to the study of hidden populations. *Soc Probl* 1997 May;44(2):174-199 [FREE Full text] [doi: [10.2307/3096941](https://doi.org/10.2307/3096941)]
15. Heckathorn D. Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. *Soc Probl* 2002 Feb;49(1):11-34 [FREE Full text] [doi: [10.1525/sp.2002.49.1.11](https://doi.org/10.1525/sp.2002.49.1.11)]
16. Gile K, Johnston L, Salganik M. Diagnostics for Respondent-driven Sampling. *J R Stat Soc Ser A Stat Soc* 2015 Jan;178(1):241-269 [FREE Full text] [doi: [10.1111/rssa.12059](https://doi.org/10.1111/rssa.12059)] [Medline: [27226702](#)]
17. Joseph L, Gyorkos T, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *Am J Epidemiol* 1995 Feb 01;141(3):263-272. [Medline: [7840100](#)]
18. Khalid FJ, Hamad FM, Othman AA, Khatib AM, Mohamed S, Ali AK, et al. Estimating the number of people who inject drugs, female sex workers, and men who have sex with men, Unguja Island, Zanzibar: results and synthesis of multiple methods. *AIDS Behav* 2014 Jan;18(Suppl 1):S25-S31. [doi: [10.1007/s10461-013-0517-x](https://doi.org/10.1007/s10461-013-0517-x)] [Medline: [23709254](#)]
19. Handcock MS, Fellows IE, Gile KJ. Hard-to-Reach Population Methods Research Group. Los Angeles, CA; 2014. RDS Analyst: Software for the Analysis of Respondent-Driven Sampling Data, Version 0.42 URL: <http://hpmrg.org> [accessed 2019-02-19] [WebCite Cache ID 76JEWdxLV]
20. Volz E, Heckathorn D. Probability based estimation theory for respondent driven sampling. *J Off Stat* 2008;24:79-97.
21. Johnston L, McLaughlin K, El Rhilani H, Latifi A, Toufik A, Bennani A, et al. Estimating the Size of hidden populations using respondent-driven sampling data: case examples from Morocco. *Epidemiology* 2015 Nov;26(6):846-852 [FREE Full text] [doi: [10.1097/EDE.0000000000000362](https://doi.org/10.1097/EDE.0000000000000362)] [Medline: [26258908](#)]
22. R Core Team. The R Project for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. R: A Language and Environment for Statistical Computing URL: <http://www.r-project.org/> [accessed 2019-02-19] [WebCite Cache ID 76JFYkL9Q]
23. Kendall C, Kerr L, Gondim R, Werneck GL, Macena RH, Pontes MK, et al. An empirical comparison of respondent-driven sampling, time location sampling, and snowball sampling for behavioral surveillance in men who have sex with men, Fortaleza, Brazil. *AIDS Behav* 2008 Jul;12(4 Suppl):S97-104. [doi: [10.1007/s10461-008-9390-4](https://doi.org/10.1007/s10461-008-9390-4)] [Medline: [18389357](#)]
24. McCreesh N, Frost SD, Seeley J, Katongole J, Tarsh MN, Ndunguse R, et al. Evaluation of respondent-driven sampling. *Epidemiology* 2012 Jan;23(1):138-147 [FREE Full text] [doi: [10.1097/EDE.0b013e31823ac17c](https://doi.org/10.1097/EDE.0b013e31823ac17c)] [Medline: [22157309](#)]
25. Gile K, Handcock M. Respondent-driven sampling: an assessment of current methodology. *Sociol Methodol* 2010 Aug;40(1):285-327 [FREE Full text] [doi: [10.1111/j.1467-9531.2010.01223.x](https://doi.org/10.1111/j.1467-9531.2010.01223.x)] [Medline: [22969167](#)]
26. Wirtz A, Mehta S, Latkin C, Zelaya CE, Galai N, Peryshkina A, et al. Comparison of respondent driven sampling estimators to determine HIV prevalence and population characteristics among men who have sex with men in Moscow, Russia. *PLoS One* 2016;11(6):e0155519 [FREE Full text] [doi: [10.1371/journal.pone.0155519](https://doi.org/10.1371/journal.pone.0155519)] [Medline: [27248818](#)]
27. Salganik M, Heckathorn D. Sampling and estimation in hidden populations using respondent-driven sampling. *Sociol Methodol* 2016 Jun 24;34(1):193-240. [doi: [10.1111/j.0081-1750.2004.00152.x](https://doi.org/10.1111/j.0081-1750.2004.00152.x)]

Abbreviations

CDC: Centers for Disease Control and Prevention
FSW: female sex worker
IBBS: integrated biobehavioral surveillance study
MoHSS: Ministry of Health and Social Services
NGO: nongovernmental organization
PLWH: people living with HIV
PSE: population size estimation
RadR: respondent-driven sampling adjusted reverse tracking method
RDS: respondent-driven sampling
RTM: reverse tracking method
SI: simulation interval
TLS: time location sampling
UCSF: University of California, San Francisco

Edited by D Prybylski, J Neal; submitted 03.08.18; peer-reviewed by K McLaughlin, MR Baneshi, H Ruisenor; comments to author 01.11.18; revised version received 27.12.18; accepted 27.01.19; published 14.03.19.

Please cite as:

Wesson PD, Adhikary R, Jonas A, Gerndt K, Mirzazadeh A, Katuta F, Maher A, Banda K, Mutenda N, McFarland W, Lowrance D, Prybylski D, Patel S

Estimating the Population Size of Female Sex Workers in Namibia Using a Respondent-Driven Sampling Adjustment to the Reverse Tracking Method: A Novel Approach

JMIR Public Health Surveill 2019;5(1):e11737

URL: <http://publichealth.jmir.org/2019/1/e11737/>

doi: [10.2196/11737](https://doi.org/10.2196/11737)

PMID: [30869646](https://pubmed.ncbi.nlm.nih.gov/30869646/)

©Paul Douglas Wesson, Rajatashuvra Adhikary, Anna Jonas, Krysta Gerndt, Ali Mirzazadeh, Frieda Katuta, Andrew Maher, Karen Banda, Nicholas Mutenda, Willi McFarland, David Lowrance, Dimitri Prybylski, Sadhna Patel. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 14.03.2019. 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

Population Size Estimations Among Hidden Populations Using Respondent-Driven Sampling Surveys: Case Studies From Armenia

Katherine R McLaughlin¹, PhD; Lisa G Johnston², PhD; Laura J Gamble¹, MSc; Trdat Grigoryan³, MD; Arshak Papoyan³, MD; Samvel Grigoryan³, MD

¹Department of Statistics, Oregon State University, Corvallis, OR, United States

²5% Initiative, Expertise France, Paris, France

³National Center for AIDS Prevention, Yerevan, Armenia

Corresponding Author:

Katherine R McLaughlin, PhD

Department of Statistics

Oregon State University

239 Weniger Hall

Corvallis, OR, 97331

United States

Phone: 1 541 737 3269

Email: katherine.mclaughlin@oregonstate.edu

Abstract

Background: Estimates of the sizes of hidden populations, including female sex workers (FSW), men who have sex with men (MSM), and people who inject drugs (PWID), are essential for understanding the magnitude of vulnerabilities, health care needs, risk behaviors, and HIV and other infections.

Objective: This article advances the successive sampling-population size estimation (SS-PSE) method by examining the performance of a modification allowing visibility to be jointly modeled with population size in the context of 15 datasets. Datasets are from respondent-driven sampling (RDS) surveys of FSW, MSM, and PWID from three cities in Armenia. We compare and evaluate the accuracy of our imputed visibility population size estimates to those found for the same populations through other unpublished methods. We then suggest questions that are useful for eliciting information needed to compute SS-PSE and provide guidelines and caveats to improve the implementation of SS-PSE for real data.

Methods: SS-PSE approximates the RDS sampling mechanism via the successive sampling model and uses the order of selection of the sample to provide information on the distribution of network sizes over the population members. We incorporate visibility imputation, a measure of a person's propensity to participate in the study, given that inclusion probabilities for RDS are unknown and social network sizes, often used as a proxy for inclusion probability, are subject to measurement errors from self-reported study data.

Results: FSW in Yerevan (2012, 2016) and Vanadzor (2016) as well as PWID in Yerevan (2014), Gyumri (2016), and Vanadzor (2016) had great fits with prior estimations. The MSM populations in all three cities had inconsistencies with expert prior values. The maximum low prior value was larger than the minimum high prior value, making a great fit impossible. One possible explanation is the inclusion of transgender individuals in the MSM populations during these studies. There could be differences between what experts perceive as the size of the population, based on who is an eligible member of that population, and what members of the population perceive. There could also be inconsistencies among different study participants, as some may include transgender individuals in their accounting of personal network size, while others may not. Because of these difficulties, the transgender population was split apart from the MSM population for the 2018 study.

Conclusions: Prior estimations from expert opinions may not always be accurate. RDS surveys should be assessed to ensure that they have met all of the assumptions, that variables have reached convergence, and that the network structure of the population does not have bottlenecks. We recommend that SS-PSE be used in conjunction with other population size estimations commonly used in RDS, as well as results of other years of SS-PSE, to ensure generation of the most accurate size estimation.

(JMIR Public Health Surveill 2019;5(1):e12034) doi:[10.2196/12034](https://doi.org/10.2196/12034)

KEYWORDS

population size estimation; respondent-driven sampling; men who have sex with men; female sex workers; people who inject drugs; Armenia

Introduction

Having accurate estimates of the sizes of hidden populations, including female sex workers (FSW), men who have sex with men (MSM), and people who inject drugs (PWID), are essential for understanding the magnitude of vulnerabilities, health care needs, risk behaviors, and HIV and other infections. In addition, population size estimations (PSEs) are used to inform resource allocation to develop programs to support sexual health and well-being, counseling and treatment for drug use, to advance social and economic justice, and to respond to and monitor critical health needs and epidemics. However, measuring a hidden population is extremely challenging and current methods contain numerous biases [1-5]. Given the importance of measuring the sizes of hidden and vulnerable populations, the advancement and continued critical review of current methods are needed [2,6].

Currently, many PSEs of FSW, MSM, and PWID are conducted in conjunction with HIV biobehavioral surveys (BBS) using respondent-driven sampling (RDS) [1-5,7-9]. These surveys are routinely implemented to measure the prevalence of sexual risks, drug use, HIV testing and knowledge, stigma and discrimination, and HIV and other infections among key populations at higher risk of HIV exposure.

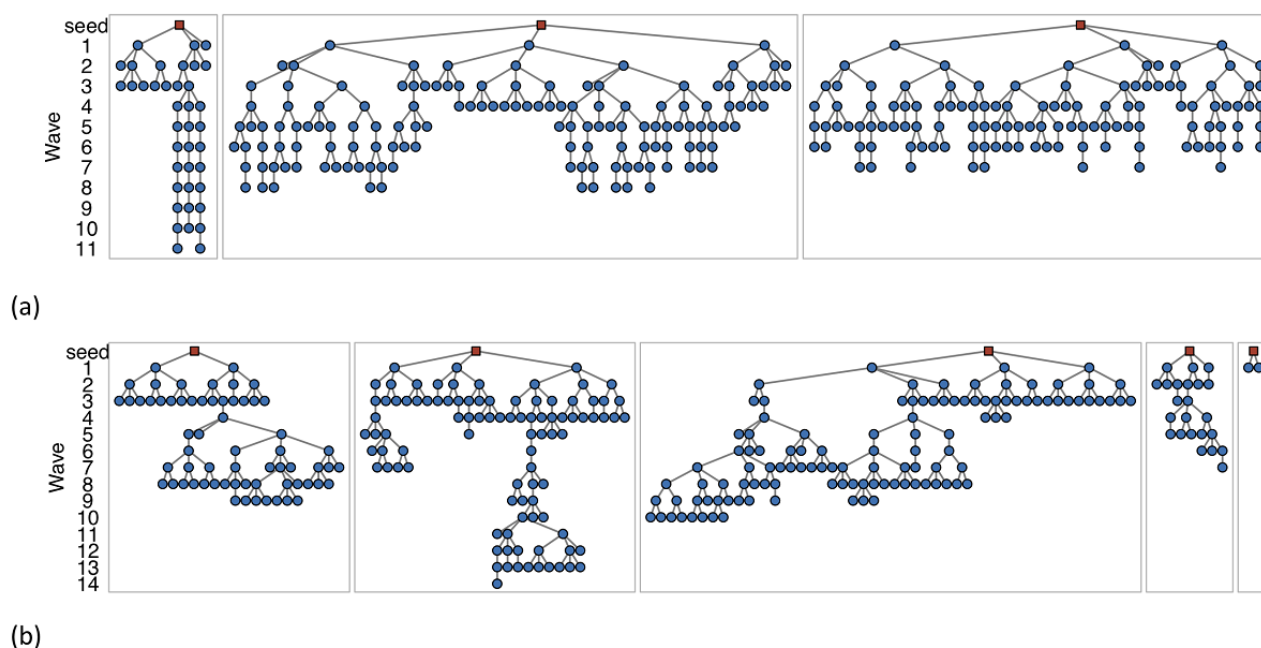
RDS is a probability-based sampling method which, when implemented and analyzed correctly, can yield findings representing the network of the population sampled [10-13]. Sampling begins with a convenience sample of well-networked population members, referred to as *seeds*. Seeds enroll, complete the survey and biological specimen collection, and are provided a fixed number of coupons to recruit members from their social network. All participants provide a measurement of their social network size, or degree, which is the number of people they know, who know them, that are in their social network. For BBS, social networks are described as groups of people who know each other and engage in common behaviors, such as injecting drugs, having anal sex, or exchanging sex for money or goods during a specified time period (eg, 6 months). Coupons, which contain a unique number to manage peer-to-peer recruitment, allow participants to remain anonymous, making it especially acceptable to populations that are stigmatized or practice illegal behaviors. Sampling should result in long recruitment chains, whereby the final sample is not biased by the initial convenience sample of seeds. Data collected using RDS methods are adjusted based on each participant's social network size and other covariates when making inferences about the population to account for the complex sampling process.

The intuitive reasoning is that individuals with larger social network sizes are more likely to be sampled, and be sampled earlier in the RDS process, so their responses need to be weighted accordingly.

One of the PSE methods being commonly used in conjunction with RDS surveys is successive sampling-population size estimation (SS-PSE) [4,5,7,9,14,15], which relies on the successive sampling model for the RDS sampling process [16]. Unlike other PSE methods that use two or more sources of data, such as object and service multipliers and capture-recapture, SS-PSE can be used with data from just one RDS study. In SS-PSE, the order of enrollment and network size of each participant are used to estimate the distribution of population network sizes and the depletion of network size over the study period is used to model the overall population size. More details will be provided in the Methods section.

This article describes the use of SS-PSE in three rounds of BBS, conducted in 2012, 2014, and 2016, among FSW, MSM, and PWID in three cities in Armenia: Yerevan, the capital city (2016 population: 467,087 females and 373,903 males, aged 18 years or older); Gyumri, the second-largest city, located in the northwest of Armenia (2016 population: 49,482 females and 41,535 males, aged 18 years or older); and Vanadzor, the third-largest city, located in the north of Armenia (2016 population: 26,052 females and 28,962 males, aged 18 years or older) [17]. Roughly 43% of the country's population live in these areas. RDS recruitment chains for selected populations are shown in Figure 1 and the complete set are provided in Multimedia Appendix 1. We advance the SS-PSE methodology by examining the performance of a modification allowing visibility, a measure of a person's propensity to participate in the study, to be jointly modeled with population size in the context of 15 datasets of FSW, MSM, and PWID populations from Armenia. Visibility [18] is modeled because inclusion probabilities for RDS are unknown and social network sizes, often used as a proxy for inclusion probability, are subject to biases and measurement errors from self-reported study data. For example, self-reported social network sizes may be an inaccurate measure of inclusion probability due to heaping or rounding [16]; intentional misreporting, perhaps to minimize one's connection to a stigmatized population [19-21]; and unintentional misreporting, perhaps due to memory recall bias [22-24]. We compare and evaluate the quality of our imputed visibility PSEs to those found for the same populations through other unpublished methods. In addition, we provide guidelines and caveats to improve the implementation of SS-PSE for real data.

Figure 1. Respondent-driven sampling (RDS) recruitment chains for selected populations: (a) female sex workers (FSW), Yerevan 2016 and (b) men who have sex with men (MSM), Yerevan 2014. Seeds are indicated by red squares and the waves of recruitment are shown vertically.



Methods

Survey

Standard RDS methods were used to recruit FSW, MSM, and PWID in 2012, 2014, and 2016 in Yerevan as well as in 2016 in Gyumri and Vanadzor [10]. Network size questions were structured based on the eligibility criteria of each sampled population: FSW were women who received money in exchange for sexual intercourse in the previous three months; MSM were men, including transgender women in 2012, 2014, and 2016, who had anal sex with another man in the previous 12 months; and PWID were people who injected drugs for nonmedical purposes in the previous three months. All participants were 18 years of age or older and residents of the survey location. Seeds or persons with a valid coupon who presented to a survey location were screened for eligibility and underwent informed consent. No one refused to enroll, despite having to consent to both the biological and behavioral parts of the survey. Enrollees were then interviewed by a trained interviewer, provided HIV pretest counseling, and underwent a venous blood draw for laboratory testing for HIV and other infections. Following the blood draw, each respondent received a set number of coupons—no more than three—along with recruitment instructions on how to recruit eligible peers. The different number of coupons to distribute and different target sample sizes—100 in Gyumri and Vanadzor and 300 in Yerevan—reflect differences in population size and connectedness, as well as anticipated speed of recruitment, identified during formative research. To maintain respondents' confidentiality, unique identification codes were used to link behavioral and biological data and to track who recruited whom. Respondents received primary compensation of AMD 4000 (Armenian Dram) in 2012 and 2014 and AMD 3500 in 2016, or slightly over US \$7 using 2016 conversion rates, for enrollment and completion of the survey. Respondents received

an additional secondary compensation—AMD 2000 in 2012 and 2014 and AMD 1600 in 2016—for each peer they recruited who enrolled and completed the survey.

The network size question is crucial to RDS studies as a proxy for a person's propensity to be included in the sample. Participants were asked how many individuals they know who meet the study eligibility requirements and then, as a follow-up, how many of them they have seen in the previous month. An individual's network size is considered to be the second, more restrictive, of these numbers. For example, the precise question for FSW in Vanadzor was "How many women do you know, whom also know you, are 18 years of age and older, are living in Yerevan, and have exchanged vaginal or anal sex for money or other reward? How many of them have you seen in the past month?"

Successive Sampling-Population Size Estimation and Visibility Imputation

Population size estimations were conducted using SS-PSE [25,26]. The approach approximates the RDS sampling mechanism via the successive sampling model of Gile [27] and uses the order of selection of the sample to provide information on the distribution of network sizes over the population members. SS-PSE uses a Bayesian framework, treating the population size N as unknown, but with a specified prior distribution. The SS-PSE framework allows for the incorporation of prior belief about the population size, which is often available via expert knowledge or PSEs from other sources, such as enumeration through mapping, network scale-up, multiplier, or capture-recapture methods [7]. The population unit sizes are treated as independent and identically distributed samples generated from a superpopulation model based on some unknown distribution. This setup is common in model-based sampling theory [28]; in it, the unit values of the finite population are conceived of as a random sample from an infinite

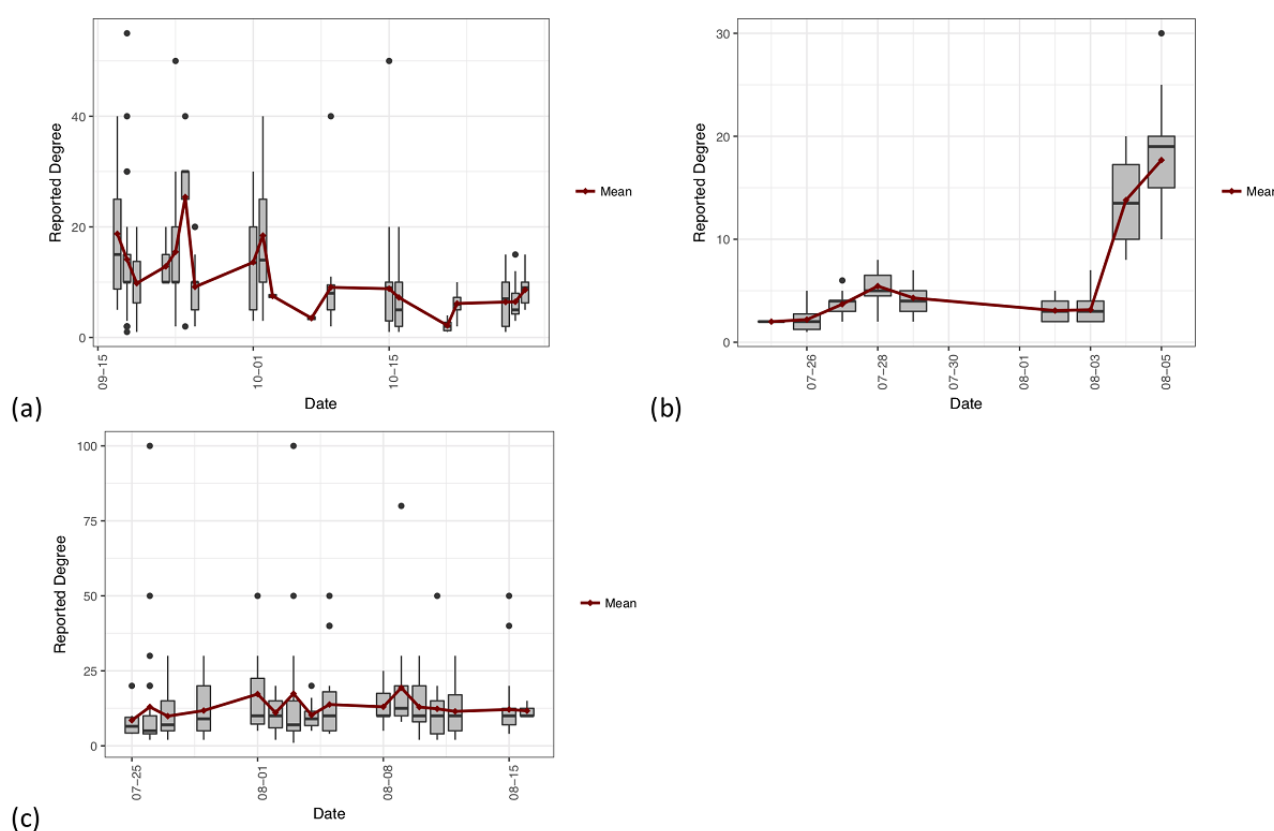
population or superpopulation rather than fixed, but are unknown. We observe a subset $n < N$ of members of the population in our sample, as well as the self-reported degree for each individual and order of participation (ie, enrollment date).

The successive sampling model assumes that individuals with a higher degree are more likely to be recruited earlier in the RDS process, since they are more connected and easily accessible in the social network. Thus, if there are fewer large-degree individuals in later waves than earlier waves, this suggests a depletion of members of the population and a large sample fraction; the population is likely not much larger than the sample. However, if the reported degrees stay roughly the same across recruitment waves, the sample size is likely a smaller portion of the population. If the reported degrees increase notably across waves, this may be an indication that the RDS recruitment process is not operating as expected and would merit caution when interpreting the results of various estimators. Figure 2 shows plots of enrollment date versus reported degree for selected populations. Panel (a) demonstrates a situation in which few large-degree individuals are observed in the later waves and the overall trend is slowly decreasing. Panel (b) shows a strongly increasing trend, which indicates

that the SS-PSE method may not perform well for this population. Panel (c) shows a relatively constant degree across waves, with some large-degree individuals still present in the later waves of the sample. These types of exploratory plots aid in understanding how RDS recruitment dynamics affect SS-PSE estimates and can alert us to possible violations of sampling assumptions.

The original SS-PSE method relies on self-reported network sizes. However, these values are subject to bias due to heaping or rounding and both intentional and unintentional misreporting; additionally, they may contain missing or impossibly low or high values [11]. We therefore use a modified version of SS-PSE that jointly models the visibility of each individual using a measurement error model [18]. Visibility is viewed as an adjusted or underlying degree that attempts to account for the aforementioned issues that arise from self-reports. We use a Conway-Maxwell-Poisson measurement error model that allows for the proportional inflation of the self-reported degree relative to the visibility and for relative error of the self-reported degrees around this inflated value. Computationally, this modification adds two additional components that need to be estimated during each step of the SS-PSE Markov chain Monte Carlo algorithm, but the outputs from the method are the same.

Figure 2. Plots of enrollment date versus self-reported network size for selected populations: (a) female sex workers (FSW), Yerevan 2014, showing a depletion in mean reported degree over the study period; (b) men who have sex with men (MSM), Vanadzor 2016, showing an increasing trend over the study period; and (c) FSW, Yerevan 2016, showing a constant trend over the study period. Note that the magnitude of trends is not comparable across plots due to different reported degree values for the different populations.



Imputed visibility SS-PSE is a Bayesian method, where information about unknown parameters is expressed through probability distributions over their possible values. Thus, the resulting estimates take the form of a distribution called the posterior distribution. We estimate the posterior distribution for the population size N , given our prior belief about the population size and observed data. The prior information used for each of the imputed visibility SS-PSE models of the 15 Armenian datasets was a median, obtained as the average of two medians for that population and city provided by local experts in 2016 through a consensus and extrapolation led by the second author (LGJ) [29]. Local experts included representatives from the National Center for AIDS Prevention, the Joint United Nations Programme on HIV/AIDS (UNAIDS), the Global Fund, and several governmental and nongovernmental organizations working directly with the populations. Local experts believe the population sizes of FSW, MSM, and PWID remained relatively constant over the period from 2012 to 2016, as reported by the Statistical Committee of the Republic of Armenia [17]; thus, we used the same prior medians for each of the three years for each population in Yerevan. To examine sensitivity to this choice of prior medians, we also fit each imputed visibility SS-PSE model using each of the two expert prior medians separately. Because population size distributions are skewed, we used the median of the posterior distribution as a point estimate and 90% credible intervals to express

uncertainty about the estimate. In addition to individual PSEs for each dataset, we also compared the estimates for the Yerevan populations across the 2012, 2014, and 2016 surveys. We assessed the trend of the estimates, considering their overall variability, using mirrored plots of the posterior distributions. Imputed visibility SS-PSE estimates were performed using the *posterior.size* function in the *sspse* package, version 0.8, for the R programming language (The R Foundation) [18].

Results

Population Size Estimates

We applied the imputed visibility SS-PSE method to 15 datasets of FSW, MSM, and PWID populations from Armenia. Table 1 reports the prior values and quantiles of the posterior distribution for population size from each of the populations. Reference values provided by local experts are shown as well, where the expert median is the value used as the prior median in the imputed visibility SS-PSE model. The expert low and high values are, respectively, the minimum of two expert values provided for the smallest that the population size could be and the maximum of two expert values provided for the largest that the population size could be. These numbers were not used in the estimation procedure, but are used to assess the model's goodness of fit and plausibility of the PSE.

Table 1. Prior values and quantiles of the posterior distribution for population size obtained from imputed visibility SS-PSE (successive sampling-population size estimation) estimates from 15 datasets of female sex workers (FSW), men who have sex with men (MSM), and people who inject drugs (PWID) populations in Armenia.

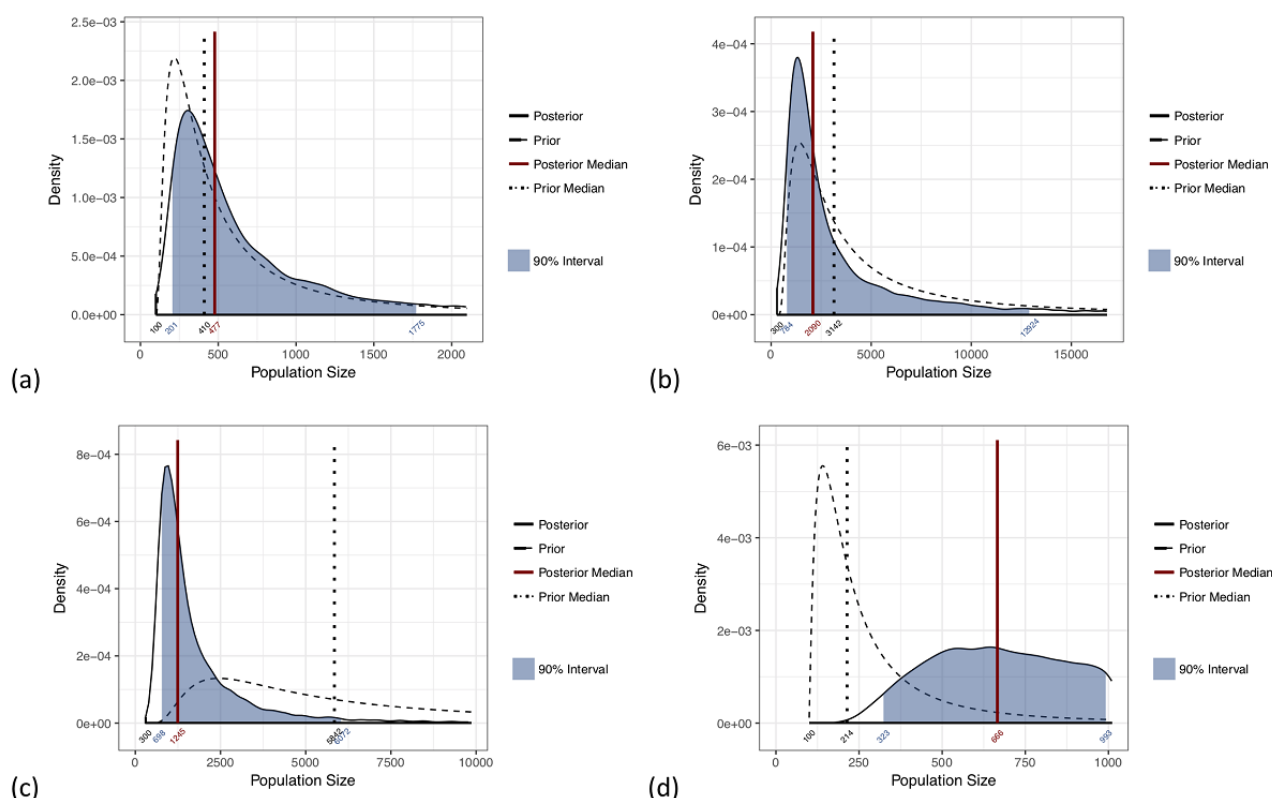
Population	Expert values, n			Posterior, n					
	Low ^a	Median (prior) ^b	High ^a	5%	25%	Median	75%	95%	Assessment ^c
FSW, Yerevan 2012	1500	3143	9900	1243	2235	3734	6975	16,599	Great
FSW, Yerevan 2014	1500	3143	9900	397	421	445	469	542	Bad
FSW, Yerevan 2016	1500	3143	9900	784	1340	2090	4169	12,924	Okay
FSW, Gyumri 2016	165	351	1089	126	152	171	217	340	Okay
FSW, Vanadzor 2016	115	239	759	133	164	205	280	551	Great
MSM, Yerevan 2012	2420	4202	6667	943	1550	2407	4335	14,296	Bad
MSM, Yerevan 2014	2420	4202	6667	836	1407	2299	4477	17,973	Bad
MSM, Yerevan 2016	2420	4202	6667	871	1121	1550	2264	4870	Bad
MSM, Gyumri 2016	176	306	485	127	176	249	409	974	Okay
MSM, Vanadzor 2016	123	214	339	323	506	666	836	993	Bad
PWID, Yerevan 2012	1667	5842	14,473	1196	2041	3236	5823	17,117	Good
PWID, Yerevan 2014	1667	5842	14,473	1196	2091	3435	3569	19,008	Good
PWID, Yerevan 2016	1667	5842	14,473	698	947	1245	2091	6072	Bad
PWID, Gyumri 2016	167	584	1446	197	354	596	1141	2960	Great
PWID, Vanadzor 2016	117	409	1013	201	318	477	812	1775	Great

^aThe prior low and high values are, respectively, the minimum of the expert prior lows and the maximum of the expert prior highs.

^bThe prior median is the average of the expert prior medians.

^cThe *Assessment* column describes how well the estimate aligned with expert knowledge, based on low and high values provided by experts that were not used in the size estimation model.

Figure 3. Posterior distributions for the size of selected populations: (a) people who inject drugs (PWID), Vanadzor 2016; (b) female sex workers (FSW), Yerevan 2016; (c) PWID, Yerevan 2016; and (d) men who have sex with men (MSM), Vanadzor 2016. The 90% credible interval is indicated by the shaded blue region and the posterior median by a vertical red line.



Several example posterior distributions are provided in [Figure 3](#) and the complete set for all 15 datasets is available in [Multimedia Appendix 1](#). The median of the posterior distribution provides a point estimate of the population size; the 90% credible interval demonstrates the uncertainty of the estimate and provides a range of likely values. The 90% credible interval is indicated on the posterior distribution plots as the shaded blue region and in [Table 1](#) as the 5% and 95% values.

Each PSE was assessed by comparing the posterior median with the low and high values provided by experts. A *Great* fit is one where the posterior median is between the maximum expert low and the minimum expert high values; a *Good* fit is one where the posterior median is between the average expert low and the average expert high values; an *Okay* fit is one where the posterior median is between the minimum expert low and the maximum expert high values; and a *Bad* fit is one where the posterior median is either smaller than the minimum expert low or larger than the maximum expert high value. The low and high prior values shown in [Table 1](#) represent the range used for an *Okay* fit, so if the posterior median falls between these two values, the fit will be *Okay* or better. Of the 15 populations considered, 4 (27%) size estimates were *Great*, 2 (13%) were *Good*, 3 (20%) were *Okay*, and 6 (40%) were *Bad*. Note that due to inconsistency of the expert prior medians provided for all MSM populations, *Great* fits were not possible for 5 out of the 15 (33%) datasets. A possible explanation for this is provided in the Discussion.

The posterior distributions for population size, like those shown in [Figure 3](#), can also be used to assess SS-PSE model fit. To

aid interpretation of the SS-PSE methodology, we show examples of several different-quality fits. Panel (a) PWID, Vanadzor 2016, and panel (b) FSW, Yerevan 2016, of [Figure 3](#) both demonstrate fits that look to be of good quality because the shape of the posterior distribution is similar to the prior distribution, with a long right tail and most of the mass of the distribution near smaller values. In panel (a), the posterior median is a little larger than the prior median, evidenced by the slight right shift of the posterior distribution relative to the prior distribution. This indicates that the RDS data provided evidence that the true population size was slightly larger than the prior belief specified. Conversely, in panel (b), the posterior median is a little smaller than the prior median and the posterior distribution is shifted slightly to the left relative to the prior distribution. This indicates that the RDS data provided evidence that the true population size was slightly smaller than the prior belief specified.

Panel (c) PWID, Yerevan 2016, provides an example of a case that is more difficult to interpret. Although the shape of the posterior distribution is acceptable and does not indicate problems with convergence of the SS-PSE method, it is clear that much of the mass from the posterior distribution falls below the prior distribution. This indicates that the PSE is much smaller than the prior median specified. Upon examining these data, we did not observe RDS assumption violations. A possible explanation is that a bottleneck in the underlying social network affected recruitment, making it difficult or impossible to sample from a portion of the population. This means that, in practice, the PSE is only for a subgroup within the overall PWID

population. When size estimates are much smaller than experts expect, this could be indicative of a disjoint network, bottleneck, or other reason why only a subset of the population was reachable in the sample. In this case, we advise study officials to return to the formative research study protocol to consider whether any of these scenarios were possible [11,30]. When the assessment of the SS-PSE model does not indicate convergence problems, but the estimates produced are very different from our prior beliefs, it is advisable to return to the study context and consider whether recruitment and study participation were impacted by any additional factors.

In contrast to panel (c), where the estimate was *Bad*, but the overall SS-PSE fit was acceptable, panel (d) MSM, Vanadzor 2016, provides an example of poor SS-PSE fit. The shape of the posterior distribution is much flatter than in the other examples, does not overlap with the prior distribution much, and has most of its mass on larger population sizes. The MSM, Vanadzor 2016, data show an increasing trend in network sizes over the time of the study period, as previously discussed in the Visibility Imputation section in the Methods and shown in Figure 2 (b). This is contrary to the typical RDS assumptions that high-degree individuals will be recruited earlier and that the depletion of high-degree individuals can be used to assess population size. Because the high-degree individuals were recruited toward the end of the study, the SS-PSE model estimates that the population size is actually quite a bit larger than the prior median provided. In cases such as this, where the distribution of network sizes throughout the recruitment chain does not meet the RDS assumptions, we recommend careful consideration of the data by experts to assess the RDS study. The SS-PSE results should only be used with extreme caution.

Overall, many of the point estimates tend to be lower than the expert prior median. This scenario may reflect the reality that RDS surveys may not be reaching the full hidden population, perhaps due to bottlenecks, clustering, or isolated individuals, resulting in a PSE only for the subpopulation that is reachable by RDS.

Comparison With Other Population Size Estimations

To place the estimates obtained using imputed visibility SS-PSE in context, we compare the posterior medians to PSEs obtained using service and unique object multiplier methods and wisdom of the crowds for the nine datasets in 2016; we also compare the posterior medians to SS-PSE without visibility imputation for all 15 datasets. The service multiplier method requires two overlapping data sources, including a count of nonduplicated clients accessing a service and a probability-based survey. For these estimations, the service data were unique counts of key populations who received an HIV test between January 1 and June 30, 2016. The PSE is this count divided by the proportion who reported having an HIV test in the probability-based survey (ie, the RDS surveys, also used for the SS-PSE models). Similarly, the unique object multiplier estimate is the number of unique objects distributed to the key population divided by the proportion who reported receiving that object in the probability-based survey. The unique object distributed was a

leather bracelet for all populations in 2016, given out one week prior to the start of the survey by outreach workers. Multiplier methods rely on several assumptions, including that no individual is counted more than once in each multiplier, that there is limited in-and-out-migration, that the two data sources are independent of each other, and that the probability-based survey is representative of the hidden population. Wisdom of the crowds assumes that, in aggregate, the responses of a sufficient number of key population members about the size of their population will provide a good estimate of the actual size of their population. Participants in the RDS survey were asked for their best guesstimate on the population size and the average was computed.

Table 2 compares the point estimates for the PSEs for the 15 datasets using object and service multipliers, wisdom of the crowds, SS-PSE without visibility imputation, and imputed visibility SS-PSE. The expert values are provided for reference as well. The SS-PSE estimates compare favorably to PSEs using object and service multiplier methods, which are commonly either much too small (eg, FSW, Yerevan 2016 and FSW, Gyumri 2016) or much too large (eg, PWID, Gyumri 2016 and PWID, Vanadzor 2016). Similarly, the wisdom of the crowds' estimates generally seem much too small (eg, 26 for PWID, Gyumri 2016, when the RDS sample size was 100) or much too large (eg, 19,342 for PWID, Yerevan 2016). Further, the SS-PSE models without using imputed visibility would not converge in 6 of the 15 datasets (40%) and produced poor estimates in many other cases; for example, the Yerevan PWID datasets. Imputed visibility SS-PSE makes size estimation possible in cases where they could not be previously calculated, both for older studies where needed questions were not correctly asked on the survey and for cases where SS-PSE models without visibility fail to converge. In cases where these other PSE methods are possible, imputed visibility SS-PSE still performs favorably.

Trend Analysis

The data considered included three rounds—2012, 2014, and 2016—of BBS for FSW, MSM, and PWID in Yerevan, Armenia. Imputed visibility SS-PSE models were fit for each year using the same prior median population size for each population, based on consultation with local experts. We compared the size estimates, descriptively, over these three years for each population. We present a visual inspection of trend in population size over time, as three years of data are not enough to do a time series analysis and a hypothesis test for equality depends on assumptions that may not be met by the RDS sampling process. Figure 4 shows the mirrored prior and posterior distributions for each year, with lines connecting the posterior median of each year. The prior distributions were the same for each year for a particular population; the placement of the posterior relative to the prior distribution indicates whether the estimate is being increased or decreased relative to the prior distribution based on the data. Because of sampling error, any time we draw a new random sample from the same population, we may get a slightly different estimate. This is a natural phenomenon in sampling and not a cause for concern.

Table 2. Comparison of imputed visibility SS-PSE (successive sampling-population size estimation) posterior medians with other population size estimation methods for the 15 Armenia datasets.

Population	Expert values (n)			Object multiplier	Service multiplier	WOC ^a (best mean)	SS-PSE median (no visibility)	SS-PSE median (visibility)
	Low	Median	High					
FSW ^b , Yerevan 2012	1500	3143	9900	N/A ^c	N/A	N/A	2041	3734
FSW, Yerevan 2014	1500	3143	9900	N/A	N/A	N/A	469	445
FSW, Yerevan 2016	1500	3143	9900	571	1283	1615	No fit ^d	2090
FSW, Gyumri 2016	165	351	1089	150	92	196	277	171
FSW, Vanadzor 2016	115	239	759	204	156	67	275	205
MSM ^e , Yerevan 2012	2420	4202	6667	N/A	N/A	N/A	No fit	2407
MSM, Yerevan 2014	2420	4202	6667	N/A	N/A	N/A	No fit	2299
MSM, Yerevan 2016	2420	4202	6667	749	8300	11,900	1121	1550
MSM, Gyumri 2016	176	306	485	3659	N/A	138	168	249
MSM, Vanadzor 2016	123	214	339	150	N/A	40	No fit	666
PWID ^f , Yerevan 2012	1667	5842	14,473	N/A	N/A	N/A	1245	3236
PWID, Yerevan 2014	1667	5842	14,473	N/A	N/A	N/A	1743	3435
PWID, Yerevan 2016	1667	5842	14,473	9000	N/A	19,342	997	1245
PWID, Gyumri 2016	167	584	1446	3000	6800	26	No fit	596
PWID, Vanadzor 2016	117	409	1013	3000	7000	198	No fit	477

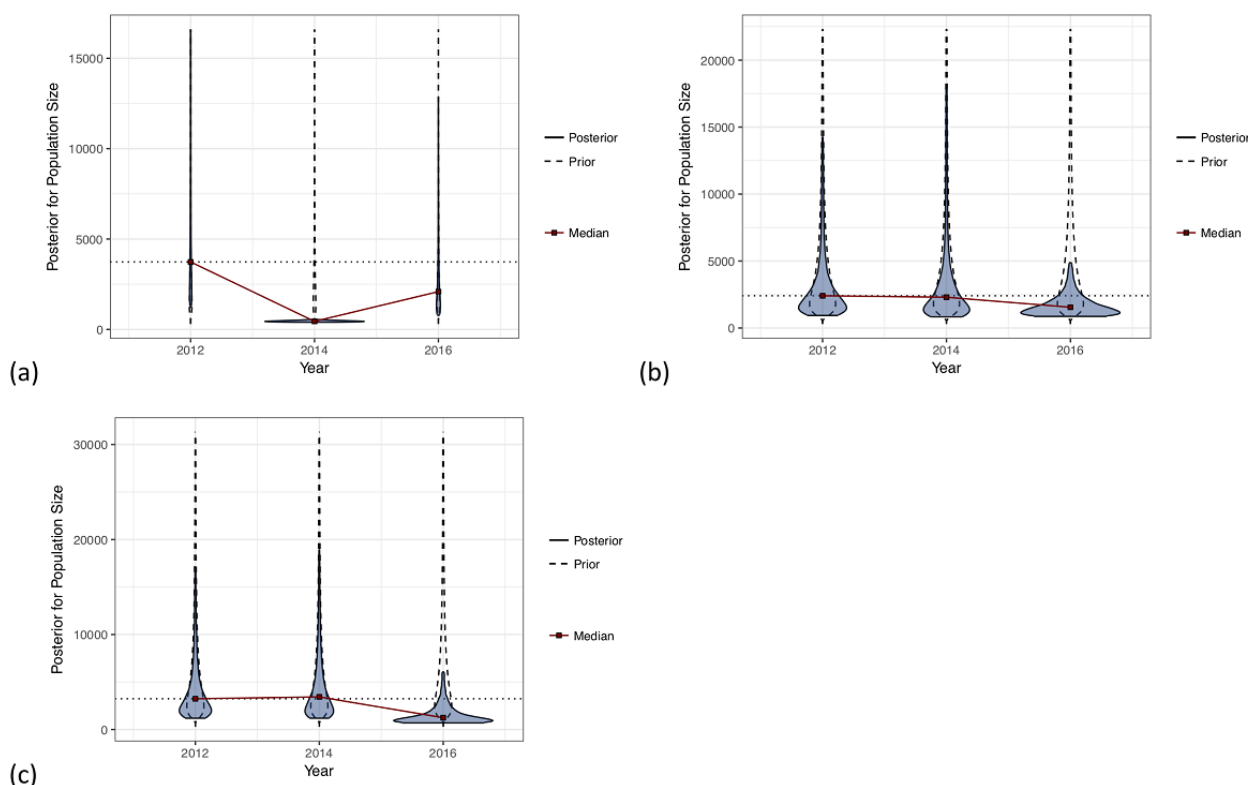
^aWOC: wisdom of the crowd estimates.^bFSW: female sex workers.^cN/A: not applicable. Information was not collected at the time the study was implemented that would enable calculation of these values.^dSS-PSEs without visibility imputation where the value is *No fit* indicate that the model would not converge.^eMSM: men who have sex with men.^fPWID: people who inject drugs.

Therefore, even if the size of, for example, the Yerevan MSM population remained exactly constant over the period from 2012 to 2016, we would expect to get different estimates each year due to sampling. We used the overall variability of the estimates, indicated by the 90% credible intervals, to assess how unusual any particular year's estimate was, and if it was actually indicative of a trend.

Panels (b) and (c) in [Figure 4](#) indicate that even though the yearly posterior medians are different, we should not interpret a strong trend from these data. It may be tempting to view the size of the MSM and PWID populations as decreasing. However, the change in median population size over time is small relative to the overall variability in the posterior distributions, so we would caution against a strong interpretation of trend.

Additionally, the appearance of trend in the Yerevan PWID populations relies heavily on the drop-off seen in 2016; the potential issues regarding the PWID posterior distribution from 2016 were previously mentioned in the discussion of [Figure 3](#). Since the estimates from 2012 and 2014 are relatively consistent, we would advise against drawing any conclusions that rely too heavily on the posterior distribution from 2016. Nevertheless, it is worth considering a potential trend when additional years of data are collected, such as the 2018 BBS in Yerevan. Panel (a) in [Figure 4](#) contains the 2014 FSW estimate from Yerevan, which, as noted previously, resulted in a *Bad* estimate, possibly due to a bottleneck in recruitment that resulted in only a subpopulation estimate. We therefore caution against reading too much into this apparent trend, as the 2012 and 2016 estimates are similar.

Figure 4. Trend analysis plots showing the prior and posterior distributions for 2012, 2014, and 2016 biobehavioral surveys in Yerevan of (a) female sex workers (FSW), (b) men who have sex with men (MSM), and (c) people who inject drugs (PWID). The median population size estimate is indicated by a red point for each year and the prior median for all three years is shown as a horizontal dotted line.



Discussion

Principal Findings

Imputed visibility SS-PSE provides an estimate for the size of a hidden population using data already routinely collected in an RDS survey. Unlike many other PSE methods, imputed visibility SS-PSE relies on only one data source and can therefore be performed retroactively, after an RDS study has already been conducted. Further, the visibility imputation modification allows for potentially erroneously self-reported network sizes to be modeled, making the method more robust to misreporting, missing values, and extreme values. However, given the difficulty measuring a hidden population and the potential for biases at many levels, including undetected bottlenecks in the network structure of the population, problems with the RDS sample, and errors in the prior size estimations, some estimates may not make sense. It is therefore always important to assess the quality of the PSE, rather than treating it as innately correct. Diagnostic plots, such as the plots of social network size by enrollment date, are useful tools to assess RDS and SS-PSE assumptions. The posterior distribution should also be examined to assess possible issues with model fit, which could be indicated by a flat distribution or one with a spike at large values of N .

When fitting the imputed visibility SS-PSE model, prior belief about the population size is specified. For these 15 datasets we used the prior median, as this was the information available. It is also possible to use the first and third quartiles or other distribution summary measures, based on available knowledge.

Prior values should be ascertained before fitting the model and not altered when an estimate does not make sense, in order to avoid introducing bias from the researcher. Instead, when the posterior distribution has the appropriate shape, but the estimate does not align with expert knowledge, it is advisable to engage additional stakeholders and examine the study in more detail. In this exercise, we found that the MSM populations in all three cities considered had inconsistencies with the expert prior values provided. The maximum low prior value was larger than the minimum high prior value, making a *Great* fit impossible. One possible explanation is the inclusion of transgender women in the MSM populations during these studies. Therefore, there could be differences between what experts perceive as the size of the population, based on who is an eligible member of that population, and what members of the population perceive. There could also be inconsistencies among different study participants, as some may include transgender peers in their accounting of personal network size, while others may not. Because of these difficulties, the transgender population was split apart from the MSM population for the 2018 BBS study.

To examine the sensitivity of the imputed visibility SS-PSE model fits to the choice of prior median, we fit each model with three different prior medians: the average of two expert values and each expert value individually. The average of the two expert values was the final prior median used for the models presented in the Results section. Using the other prior medians does not drastically change the PSE. Although the point estimates are slightly larger for the larger prior median and slightly smaller for the smaller prior median, the values are very similar given the overall variability of the distribution.

Superimposed posterior distributions for SS-PSE fits using these three prior medians for each of the 15 datasets are provided in [Multimedia Appendix 1](#).

Evaluating the results from the imputed visibility SS-PSE, as well as other PSEs used in conjunction with RDS (ie, unique object and service multipliers, wisdom of the crowds), is essential given that they are prone to biases, which may lead to unrealistic over- and underestimations. Many size estimation techniques can be used as part of each survey to triangulate and validate the most optimal size estimation [2,3,5,9,31]. Further validation of size estimations relies on expert input from many stakeholders, including governmental and nongovernmental organizations working with the population, persons directly involved with the sampling, and people with knowledge about statistics and epidemiology. These collaborative efforts are needed to explain biases and failures to meet assumptions in both the sampling and the population size methods.

Conclusions

The imputed visibility SS-PSE method of PSE can be used with existing RDS data sources to obtain reasonable estimates when benchmarked against prior expert knowledge. We demonstrate the performance of this method on 15 datasets of FSW, MSM, and PWID populations from three waves of BBS studies conducted using RDS from three cities in Armenia. This is the first assessment of the modification to the imputed visibility

SS-PSE methodology on such a large variety of data and the first to consider trend analysis for the same population over three time points. This is also the first presentation of how to interpret different outputs from SS-PSE in real data. These studies cover a variety of recruitment structures and sizes coming from nine different underlying social networks. The results provide examples of good model fits, where the RDS assumptions appear to be satisfied and the resulting posterior distributions are of the appropriate shape, and bad model fits, where the RDS assumptions appear to be violated in diagnostic plots or the posterior distributions depart greatly from expert opinions. We find that the imputed visibility SS-PSE model performs favorably compared to other PSE methods for these populations; these other methods have no basis on which to assess bias and often give impossibly large or small estimates or produce no estimate at all. Because SS-PSE does not rely on data from multiple surveys or census information, it is a valuable method of PSE. However, there are limitations to its use. If RDS assumptions are violated or there are issues with convergence in the model, results from SS-PSE should be interpreted with caution. To this end, we also provide guidance and suggested methods for goodness of fit to assess the SS-PSE methodology and the overall quality of the estimates. We recommend that SS-PSE be used in conjunction with other PSE techniques commonly used in RDS to ensure generation of the most accurate and acceptable size estimation.

Acknowledgments

We would like to thank everyone who participated in the surveys. In addition, we would like to thank Mark Handcock, University of California, Los Angeles, co-founder of the Hidden Population Methods Research Group and co-developer of the successive sampling population size estimation, for his consultation on this manuscript. We also want to thank Seda Abgaryan, Lilith Hovhannisyanyan, Ruben Hovhannisyanyan, and Tigran Hovsepyan of the National Center for AIDS Prevention, Yerevan, Armenia.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Diagnostic and sensitivity plots.

[[PDF File \(Adobe PDF File\), 2MB](#) - [publichealth_v5i1e12034_app1.pdf](#)]

References

1. Guidelines on Estimating the Size of Populations Most at Risk to HIV. Geneva, Switzerland: World Health Organization and UNAIDS; 2010. URL: http://www.unaids.org/sites/default/files/media_asset/2011_Estimating_Populations_en_0.pdf [accessed 2019-01-29] [WebCite Cache ID 75nhJmzAV]
2. Sabin K, Zhao J, Garcia Calleja JM, Sheng Y, Arias Garcia S, Reinisch A, et al. Availability and quality of size estimations of female sex workers, men who have sex with men, people who inject drugs and transgender women in low- and middle-income countries. PLoS One 2016;11(5):e0155150 [FREE Full text] [doi: [10.1371/journal.pone.0155150](https://doi.org/10.1371/journal.pone.0155150)] [Medline: [27163256](https://pubmed.ncbi.nlm.nih.gov/27163256/)]
3. Johnston LG, Prybylski D, Raymond HF, Mirzazadeh A, Manopaiboon C, McFarland W. Incorporating the service multiplier method in respondent-driven sampling surveys to estimate the size of hidden and hard-to-reach populations: Case studies from around the world. Sex Transm Dis 2013 Apr;40(4):304-310. [doi: [10.1097/OLQ.0b013e31827fd650](https://doi.org/10.1097/OLQ.0b013e31827fd650)] [Medline: [23486495](https://pubmed.ncbi.nlm.nih.gov/23486495/)]
4. Wu J, Crawford F, Raag M, Heimer R, Uusküla A. Using data from respondent-driven sampling studies to estimate the number of people who inject drugs: Application to the Kohtla-Järve region of Estonia. PLoS One 2017;12(11):e0185711 [FREE Full text] [doi: [10.1371/journal.pone.0185711](https://doi.org/10.1371/journal.pone.0185711)] [Medline: [29095866](https://pubmed.ncbi.nlm.nih.gov/29095866/)]

5. Johnston LG, Soe P, Aung MY, Ammassari S. Estimating the population size of males who inject drugs in Myanmar: Methods for obtaining township and national estimates. *AIDS Behav* 2019 Jan;23(1):295-301. [doi: [10.1007/s10461-018-2233-z](https://doi.org/10.1007/s10461-018-2233-z)] [Medline: [30046936](https://pubmed.ncbi.nlm.nih.gov/30046936/)]
6. Yu D, Calleja J, Zhao J, Reddy A, Seguy N, Technical Consultation on Lessons Learnt from Size Estimation among Key Populations in Asian Countries. Estimating the size of key populations at higher risk of HIV infection: A summary of experiences and lessons presented during a technical meeting on size estimation among key populations in Asian countries. *Western Pac Surveill Response J* 2014;5(3):43-49 [FREE Full text] [doi: [10.5365/WPSAR.2014.5.2.008](https://doi.org/10.5365/WPSAR.2014.5.2.008)] [Medline: [25320676](https://pubmed.ncbi.nlm.nih.gov/25320676/)]
7. Johnston LG, McLaughlin KR, El Rhilani H, Latifi A, Toufik A, Bennani A, et al. Estimating the size of hidden populations using respondent-driven sampling data: Case examples from Morocco. *Epidemiology* 2015 Nov;26(6):846-852 [FREE Full text] [doi: [10.1097/EDE.0000000000000362](https://doi.org/10.1097/EDE.0000000000000362)] [Medline: [26258908](https://pubmed.ncbi.nlm.nih.gov/26258908/)]
8. Paz-Bailey G, Jacobson J, Guardado M, Hernandez F, Nieto A, Estrada M, et al. How many men who have sex with men and female sex workers live in El Salvador? Using respondent-driven sampling and capture-recapture to estimate population sizes. *Sex Transm Infect* 2011 Jun;87(4):279-282. [doi: [10.1136/sti.2010.045633](https://doi.org/10.1136/sti.2010.045633)] [Medline: [21385892](https://pubmed.ncbi.nlm.nih.gov/21385892/)]
9. Wesson P, Qabazard R, Wilson E, McFarland W, Raymond H. Estimating the population size of transgender women in San Francisco using multiple methods, 2013. *Int J Transgend* 2018;19(1):107-112 [FREE Full text] [doi: [10.1080/15532739.2017.1376729](https://doi.org/10.1080/15532739.2017.1376729)]
10. Johnston LG. Introduction to HIV/AIDS and Sexually Transmitted Infection Surveillance. Geneva, Switzerland: World Health Organization; 2013. URL: http://applications.emro.who.int/dsaf/EMRPUB_2013_EN_1539.pdf [accessed 2019-01-29] [WebCite Cache ID 75niiB5Hd]
11. Gile KJ, Johnston LG, Salganik MJ. Diagnostics for respondent-driven sampling. *J R Stat Soc Ser A Stat Soc* 2015 Jan;178(1):241-269 [FREE Full text] [doi: [10.1111/rssa.12059](https://doi.org/10.1111/rssa.12059)] [Medline: [27226702](https://pubmed.ncbi.nlm.nih.gov/27226702/)]
12. Heckathorn D. Extensions of respondent-driven sampling: Analyzing continuous variables and controlling for differential recruitment. *Sociol Methodol* 2007 Dec;37(1):151-207. [doi: [10.1111/j.1467-9531.2007.00188.x](https://doi.org/10.1111/j.1467-9531.2007.00188.x)]
13. Heckathorn D. Respondent-driven sampling: A new approach to the study of hidden populations. *Soc Probl* 1997 May 01;44(2):174-199 [FREE Full text] [doi: [10.2307/3096941](https://doi.org/10.2307/3096941)]
14. Johnston LG, McLaughlin KR, Rouhani SA, Bartels SA. Measuring a hidden population: A novel technique to estimate the population size of women with sexual violence-related pregnancies in South Kivu Province, Democratic Republic of Congo. *J Epidemiol Glob Health* 2017 Dec;7(1):45-53 [FREE Full text] [doi: [10.1016/j.jegh.2016.08.003](https://doi.org/10.1016/j.jegh.2016.08.003)] [Medline: [27663900](https://pubmed.ncbi.nlm.nih.gov/27663900/)]
15. Wesson P, Handcock M, McFarland W, Raymond H. If you are not counted, you don't count: Estimating the number of African-American men who have sex with men in San Francisco using a novel Bayesian approach. *J Urban Health* 2015 Dec;92(6):1052-1064 [FREE Full text] [doi: [10.1007/s11524-015-9981-0](https://doi.org/10.1007/s11524-015-9981-0)] [Medline: [26392276](https://pubmed.ncbi.nlm.nih.gov/26392276/)]
16. Gile K, Handcock M. Respondent-driven sampling: An assessment of current methodology. *Sociol Methodol* 2010 Aug;40(1):285-327 [FREE Full text] [doi: [10.1111/j.1467-9531.2010.01223.x](https://doi.org/10.1111/j.1467-9531.2010.01223.x)] [Medline: [22969167](https://pubmed.ncbi.nlm.nih.gov/22969167/)]
17. Statistical Committee of the Republic of Armenia. 2016. Main statistical data URL: <https://www.armstat.am/en/> [accessed 2019-01-27] [WebCite Cache ID 75jkZISdd]
18. Handcock M, Fellows I, Gile K. Statistics UCLA. 2014. Hard-to-Reach Population Methods Research Group URL: http://wiki.stat.ucla.edu/hpmrg/index.php/Hard-to-Reach_Population_Methods_Research_Group [accessed 2019-01-30] [WebCite Cache ID 75ogk0RM9]
19. Bengtsson L, Thorson A. Global HIV surveillance among MSM: Is risk behavior seriously underestimated? *AIDS* 2010 Sep 24;24(15):2301-2303. [doi: [10.1097/QAD.0b013e32833d207d](https://doi.org/10.1097/QAD.0b013e32833d207d)] [Medline: [20634663](https://pubmed.ncbi.nlm.nih.gov/20634663/)]
20. Fenton KA, Johnson AM, McManus S, Erens B. Measuring sexual behaviour: Methodological challenges in survey research. *Sex Transm Infect* 2001 Apr;77(2):84-92 [FREE Full text] [Medline: [11287683](https://pubmed.ncbi.nlm.nih.gov/11287683/)]
21. Fisher RJ. Social desirability bias and the validity of indirect questioning. *J Consum Res* 1993 Sep;20(2):303-315 [FREE Full text] [doi: [10.1086/209351](https://doi.org/10.1086/209351)]
22. Bell DC, Belli-McQueen B, Haider A. Partner naming and forgetting: Recall of network members. *Soc Networks* 2007 May;29(2):279-299 [FREE Full text] [doi: [10.1016/j.socnet.2006.12.004](https://doi.org/10.1016/j.socnet.2006.12.004)] [Medline: [17940583](https://pubmed.ncbi.nlm.nih.gov/17940583/)]
23. Brewer DD. Forgetting in the recall-based elicitation of personal and social networks. *Soc Networks* 2000 May;22(1):29-43. [doi: [10.1016/S0378-8733\(99\)00017-9](https://doi.org/10.1016/S0378-8733(99)00017-9)]
24. Mills HL, Johnson S, Hickman M, Jones NS, Colijn C. Errors in reported degrees and respondent driven sampling: Implications for bias. *Drug Alcohol Depend* 2014 Sep 01;142:120-126 [FREE Full text] [doi: [10.1016/j.drugalcdep.2014.06.015](https://doi.org/10.1016/j.drugalcdep.2014.06.015)] [Medline: [24999062](https://pubmed.ncbi.nlm.nih.gov/24999062/)]
25. Handcock M, Gile K, Mar C. Estimating hidden population size using respondent-driven sampling data. *Electron J Stat* 2014;8(1):1491-1521 [FREE Full text] [doi: [10.1214/14-EJS923](https://doi.org/10.1214/14-EJS923)] [Medline: [26180577](https://pubmed.ncbi.nlm.nih.gov/26180577/)]
26. Handcock MS, Gile KJ, Mar CM. Estimating the size of populations at high risk for HIV using respondent-driven sampling data. *Biometrics* 2015 Mar;71(1):258-266 [FREE Full text] [doi: [10.1111/biom.12255](https://doi.org/10.1111/biom.12255)] [Medline: [25585794](https://pubmed.ncbi.nlm.nih.gov/25585794/)]
27. Gile KJ. Improved inference for respondent-driven sampling data with application to HIV prevalence estimation. *J Am Stat Assoc* 2011;106(493):135-146. [doi: [10.1198/jasa.2011.ap09475](https://doi.org/10.1198/jasa.2011.ap09475)]

28. Säärndal C, Swensson B, Wretman J. Model Assisted Survey Sampling. New York, NY: Springer-Verlag; 1992.
29. Integrated Biological-Behavioral Surveillance Survey Among People Who Inject Drugs, Female Sex Workers and Men Who Have Sex With Men in Yerevan, Gyumri, and Vanadzor, Armenia. Yerevan, Armenia: National Center for AIDS Prevention of Ministry of Health of Armenia; 2016. URL: http://www.arm aids.am/images/pdf/BBS_PWID_FSW_MSM_ARM_REPORT_FINAL_2016_eng.pdf [accessed 2019-01-27] [WebCite Cache ID 75jlefMgr]
30. Johnston LG, Whitehead S, Simic-Lawson M, Kendall C. Formative research to optimize respondent-driven sampling surveys among hard-to-reach populations in HIV behavioral and biological surveillance: Lessons learned from four case studies. *AIDS Care* 2010 Jun;22(6):784-792. [doi: [10.1080/09540120903373557](https://doi.org/10.1080/09540120903373557)] [Medline: [20467937](https://pubmed.ncbi.nlm.nih.gov/20467937/)]
31. Raymond HF, Bereknyei S, Berglas N, Hunter J, Ojeda N, McFarland W. Estimating population size, HIV prevalence and HIV incidence among men who have sex with men: A case example of synthesising multiple empirical data sources and methods in San Francisco. *Sex Transm Infect* 2013 Aug;89(5):383-387. [doi: [10.1136/sextrans-2012-050675](https://doi.org/10.1136/sextrans-2012-050675)] [Medline: [23620133](https://pubmed.ncbi.nlm.nih.gov/23620133/)]

Abbreviations

AMD: Armenian Dram
BBS: biobehavioral surveys
FSW: female sex workers
MSM: men who have sex with men
PSE: population size estimation
PWID: people who inject drugs
RDS: respondent-driven sampling
SS-PSE: successive sampling-population size estimation
UNAIDS: The Joint United Nations Programme on HIV/AIDS
WOC: wisdom of the crowd estimate

Edited by J Neal; submitted 24.08.18; peer-reviewed by A M, P Wesson; comments to author 13.10.18; revised version received 23.11.18; accepted 14.12.18; published 14.03.19.

Please cite as:

McLaughlin KR, Johnston LG, Gamble LJ, Grigoryan T, Papoyan A, Grigoryan S
Population Size Estimations Among Hidden Populations Using Respondent-Driven Sampling Surveys: Case Studies From Armenia
JMIR Public Health Surveill 2019;5(1):e12034
URL: <http://publichealth.jmir.org/2019/1/e12034/>
doi: [10.2196/12034](https://doi.org/10.2196/12034)
PMID: [30869650](https://pubmed.ncbi.nlm.nih.gov/30869650/)

©Katherine R McLaughlin, Lisa G Johnston, Laura J Gamble, Trdat Grigoryan, Arshak Papoyan, Samvel Grigoryan. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 14.03.2019. 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

Novel Approaches for Estimating Female Sex Worker Population Size in Conflict-Affected South Sudan

Alfred Geoffrey Okiria¹, MD, MPH; Alex Bolo², MD, MPH; Victoria Achut³, MD; Golda Ceasar Arkangelo³, BA; Acaga Taban Ismail Michael¹, BA; Joel Sua Katoro², MSc; Jennifer Wesson⁴, PhD; Steve Gutreuter⁵, PhD; Lee Hundley⁵, MPH; Avi Hakim⁵, MA, MPH

¹IntraHealth International, Juba, South Sudan

²United States Centres for Disease Control and Prevention, Juba, South Sudan

³South Sudan Ministry of Health, Juba, South Sudan

⁴IntraHealth International, Chapel Hill, NC, United States

⁵United States Centres for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:

Alfred Geoffrey Okiria, MD, MPH

IntraHealth International

Plot 193

Block 3-k South, Juba Nabari

Juba, 211

South Sudan

Phone: 221 404 374 4686

Email: aokiria@intrahealth.org

Abstract

Background: Limited data exist describing the population size of female sex workers (FSW) in South Sudan. A population size estimation exercise among FSW was undertaken in Juba and Nimule during the Eagle Survey.

Objective: The study aimed to estimate the number of FSW in Juba and Nimule to inform resource allocation and service provision for FSW.

Methods: We utilized service and unique object multipliers, and 3-source capture-recapture methods in conjunction with a respondent-driven sampling (RDS) survey to estimate the number of FSW in Juba and Nimule. For service multiplier, the number of FSW testing for HIV in 2015 (Juba) and 2016 (Nimule) was obtained from the LINKAGES program targeting FSW. Survey participants were asked whether they had been tested for HIV by LINKAGES during the relevant period. A total of 2 separate unique object distributions were conducted in Juba and Nimule. In Nimule, these were combined to produce a 3-source capture-recapture estimate. The exercise involved distribution of key chains and bangles to FSW, documentation of the number of those who received unique objects, and questions during RDS survey to assess whether participants received unique objects.

Results: In Juba, the service multiplier method yielded an estimate of 5800 (95% CI 4927-6673) FSW. The unique object estimate (key chain and RDS participation) yielded 5306 (95% CI 4673-5939). Another estimate using RDS participation and receipt of a bangle yielded a much lower estimate of 1863 (95% CI 1776-1951), as did a 2-source estimate of key chain and bangle (2120, 95% CI 2028-2211). A 3-source capture-recapture estimate could not be produced because aggregate rather than individual level data were collected during the third capture. The multiplier estimate using key chain and RDS participation was taken as the final population estimate for FSW in Juba, which constitutes more than 6% of the female population aged 15 to 64 years. In Nimule, the service multiplier method yielded an estimate of 9384 (95% CI 8511-10,257). The 2-source estimates for key chain and RDS yielded 6973 (95% CI 4759-9186); bangles and RDS yielded a higher estimate of 13,104 (95% CI 7101-19,106); key chains and bangles yielded a lower estimate of 1322 (95% CI 1223-1420). The 3-source capture-recapture method using Bayesian nonparametric latent-class model-based estimate yielded a population of 2694 (95% CI 1689-6945), and this was selected as the final estimate for Nimule, which constitutes nearly 40% of female population aged 15 to 64 years.

Conclusions: The service and unique object multiplier, and 3-source capture-recapture methods were successfully used to estimate the number of FSW in Nimule, whereas service and unique object multiplier methods were successfully used in Juba. These methods yielded higher than previously estimated FSW population sizes. These estimates will inform resource allocation and advocacy efforts to support services for FSW.

KEYWORDS

sex workers; population; methods; South Sudan

Introduction

Background

South Sudan, the newest country in the world, gained independence in 2011. This was followed by the return of refugees and foreigners from neighboring countries. The HIV prevalence in the neighboring countries is 6% in Uganda [1] and 5% in Kenya [2], and it is higher than 2.4% in South Sudan [3]. The relative stability allowed increased commerce and the apparent increase in the number of female sex workers (FSW) [4]. Little data exist on sex workers in South Sudan. Programs for FSW are guided by mapping conducted in 2011 and 2012 and formative assessments conducted in 2014. Despite a period of relative calm after independence, South Sudan experienced political crises in December 2013 and July 2016 that resulted in significant population movements [5].

Previous mappings and formative assessments indicated that Juba was home to the largest number of FSW in the country. The South Sudan AIDS Commission and World Health Organization mapped a study in 2012, which estimated 2511 (range: 2013-3008) FSW in Juba and 378 (range: 316-439) in Yambio [4,6]. Nimule, on the South Sudan-Uganda border, was estimated to have 400 sex workers [7].

Sex work is illegal in South Sudan and sex workers operate in a very stigmatizing environment with the constant threat of arrest, imprisonment, and sexual exploitation [6,8]. This context has led many sex workers to operate in a covert manner, in lodges, brothels, and homes. Street-based sex work is less common [4,6]. Few health services targeting FSW exist in South Sudan and until recently the South Sudan Ministry of Health (MOH) had limited data describing sex workers in the country to inform service provision.

Accurate estimation of the size of key populations using empirical methods that include data collection rather than conjecture provides important data for advocacy and resource allocation of HIV programs [9]. These estimates further facilitate intervention planning, monitoring and evaluation, and epidemic modeling. Reliable estimates of the size of key populations who are hidden, outlawed, stigmatized, and highly mobile are difficult to obtain [10]. There is no gold standard for size estimation, which sometimes prompts researchers to use multiple methods to identify a single best estimate [9,11,12]. Common size estimation methods include mapping and census, both of which provide underestimates of key populations, successive sampling, multiplier method, and capture-recapture [13-15]. The latter method was originally developed to count wildlife populations [16]. The method has evolved to include significant improvement in the accuracy and reliability of estimates, particularly through the use of 3 or more sources for capture-recapture, and the method may produce the best estimates available [9].

Objectives

The objective of this activity was to estimate the number of FSW in Juba and Nimule to inform resource allocation and service provision for FSW. This was done by the MOH together with the US Centers for Disease Control and Prevention (CDC) and IntraHealth International, who together implemented the first biobehavioral survey (BBS) and empirical size estimation activities, known as the Eagle Survey, among FSW in South Sudan.

Methods

Study Population

We included, in the survey and related size estimation activities, women and girls aged 15 years or more who received money, goods, or services in exchange for sex in the past 6 months; spoke English, Juba Arabic, or Kiswahili, and lived or worked in Juba or Nimule. Sexually exploited children and adult survivors of violence were referred to partner organizations experienced in providing counseling, health, social, and other protective services to these populations.

Procedures

The Eagle Survey utilized the service multiplier, unique object multiplier, and 3-source capture-recapture method to estimate population size of FSW. For the service multiplier, the number of FSW testing for HIV in Juba from January to December 2015 and from January to December 2016 in Nimule was obtained from the LINKAGES program that targets FSW. RDS survey participants were asked in an interview whether they had been tested for HIV by LINKAGES during the relevant period. All FSW getting services from LINKAGES receive unique identification numbers that are used each time they access services. This ensures that data from LINKAGES are unique counts reflecting the number of people tested rather than the number of tests conducted.

The second method used in both Juba and Nimule was the unique object multiplier method. This involved the distribution of key chains and bangles as unique objects to FSW. Paired together with the RDS survey, these 2 separate object distributions were used to produce a 3-source capture-recapture estimate.

Data collection began in Juba with the distribution of unique key chains to FSW in October 2015. After 5 weeks, data collection for the RDS survey began. Finally, in June 2016, unique bangles were distributed to FSW. In Nimule, unique key chains were distributed in June 2016 followed by a distribution of bangles 1 week later. The RDS survey was meant to begin 1 week after the bangle distribution but an eruption of violence days after the bangle distribution led to the postponement of the survey until January 2017. It was not possible to distribute new unique objects at that point.

In both Juba and Nimule, the number of FSW receiving a key chain during the distribution was the first data source. For Juba, the second data source was the RDS survey, which included questions to assess whether the participant had received a key chain. The third data source was the number of FSW who received a bangle after the RDS survey. Lessons learned in Juba resulted in a change in the sequence for Nimule with the unique object distributions occurring before the RDS survey.

Unique object distribution was conducted by FSW volunteers identified largely through LINKAGES. They were of different nationalities, operating in different parts of the city, and they were literate, respected, influential, and had the ability to explain the purpose of the object to potential recipients. In Juba, 54 FSW distributed key chains and 30 FSW distributed bangles, 6 of whom distributed both. In Nimule, 22 FSW distributed key chains and 14 FSW distributed bangles, 4 of whom distributed both. A 1-day volunteer training was conducted before each distribution covering the following: study background and objectives, eligibility criteria, safety and security, how to approach FSW, how to complete the distribution registration sheet and the unique object distribution form, the distribution process, and role plays. A second day was added for the key chain distribution training in Juba given the large number of volunteers. To facilitate recall and identification of volunteers, all volunteers wore a blue T-shirt bearing an eagle logo (the survey symbol). Volunteers were instructed to distribute objects to FSW in their neighborhoods.

In Juba, approximately 5 weeks before the start of the RDS survey in November 2015, the survey staff and volunteers identified and mapped hot spots with support from the LINKAGES project and with the aim of distributing at least 1300 key chains to FSW. Each of the FSW encountered received only 1 key chain, which she was instructed to keep because she may be asked about it in the near future by the survey staff. Volunteers used registration sheets to keep track of when, where, and how many objects were distributed and whether individuals had already received or rejected the object. Objects that were not distributed were returned to the study coordinator. Volunteers also told FSW about the Eagle Survey and let them know that they would not face stigma from the survey team. The volunteers received compensation for their time and transport to distribution locations.

Upon receipt of the object, the FSW was considered captured. Volunteers tracked the number of objects they distributed as this indicated the number of participants captured. During the RDS survey's eligibility screening process, all FSW were asked by the coupon manager whether they received the unique object, to show or describe it, and how they received it. Those who could not show or describe the object were asked to identify the correct object from a panel of 5 similar objects. The third recapture in Juba was a bangle distributed 3 months after the end of the RDS data collection because of procurement challenges. It utilized the same methods as the first unique object

distribution. During the bangle distribution, FSW were asked if they received a key chain, participated in the RDS survey, and received a bangle.

In Nimule, the order of activities was slightly modified. The 564 key chains were distributed in June 2016 followed by 546 bangles distributed 2 weeks later. Due to the conflict that started in July 2016, the RDS survey was delayed till January 2017. During the bangle distribution, participants were asked if they received a key chain or a bangle. The RDS survey again asked if the person received a key chain, bangle, or both. In both locations, all objects were distributed within 1 day. The number of key chains and bangles distributed was estimated to cover the entire FSW sample, 910 for Juba and 400 in Nimule with 40% extra key chains/bangles.

Analysis

The 2-source multipliers were calculated using the standard formula described elsewhere in conjunction with weighted estimates from the RDS survey or nonweighted estimates in the case of the key chain-bangle estimates [9]. The 3-source capture-recapture estimate was calculated using a Bayesian nonparametric latent-class capture-recapture model as implemented in the R package [17].

Ethical Approval

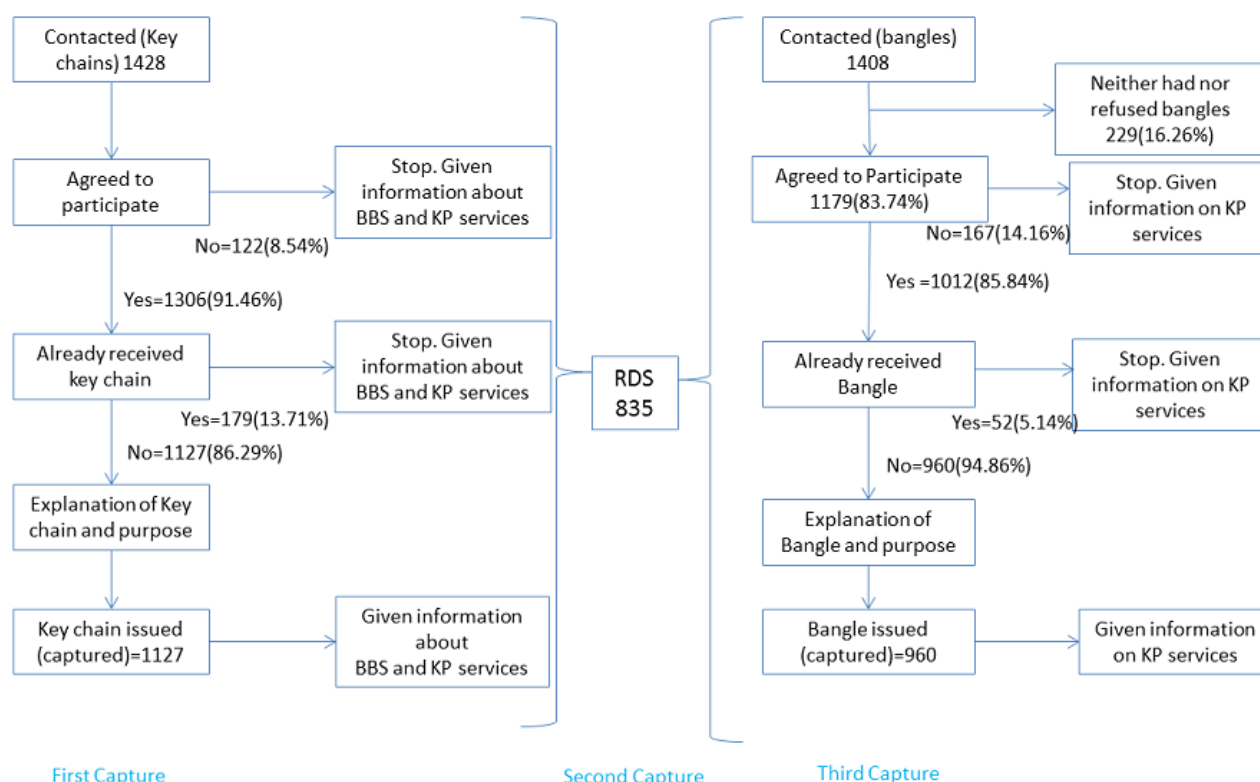
The study protocol was approved by the Government of South Sudan MOH Research Ethics Committee and CDC's Center for Global Health Associate Director for Science.

Results

Main Findings

In Juba, volunteers distributing key chains contacted 1428 FSW across 10 neighborhoods (Figure 1). Approximately 9 in 10, that is, 86.29% (1127/1306) of the FSW were not in possession of a key chain agreed to receive it. About 13.71% (179/1306) of the FSW contacted had already received a key chain. During the RDS survey, 179 FSW reported receiving a key chain. More neighborhoods where women sell sex were identified during the RDS survey; consequently, volunteers distributed bangles in 28 neighborhoods. Volunteers distributing bangles contacted 1179 FSW. Similar to the key chain distribution, 85.84% (1012/1179) of the FSW contacted had agreed to participate. Of those that participated, 94.86% (960/1012) of the FSW agreed to receive a bangle whereas 5.14% (52/1012) of the FSW had already received a bangle.

In Juba, LINKAGES service data indicated that 2204 FSW tested for HIV in 2015. The RDS survey included 835 FSW participants, 323 of whom tested for HIV at LINKAGES-Juba in 2015. Of the 9 seeds in the RDS survey, 3 were affiliated with LINKAGES. The longest chain had 17 waves. Among RDS participants, 314 received a key chain and 33 presented it during eligibility screening. A total of 498 RDS participants received a bangle.

Figure 1. Population size estimation process in Juba. BBS: biobehavioral survey; RDS: respondent-driven sampling; KP: key population.

The service multiplier method yielded an estimate of 5800 (95% CI 4927-6673) FSW. The unique object multiplier method of key chain and RDS participation yielded a similar result of 5306 (95% CI 4673-5939) but with tighter uncertainty bounds. Another unique object multiplier estimate using RDS participation and receipt of a bangle yielded a much lower estimate of 1863 (95% CI 1776-1951), as did an estimate produced by key chain and bangle (2120, 95% CI 2028-2211), as represented in Table 1. A 3-source capture-recapture estimate could not be produced because aggregate rather than individual level data were collected during the third capture. We should have asked a few more questions to be able to better disaggregate the data to know the precise number of FSW who were in each capture. The setup of Juba forms would not allow for this.

In Nimule key chain distribution, 22 volunteers contacted 788 FSW in 20 neighborhoods (Figure 2). Approximately 9 in 10, that is, 89.1% (702/788) of the FSW agreed to participate, of which 80.3% (564/702) accepted to receive a key chain. Nearly 1 in 5, that is, 19.7% (138/702) of the FSW had already received a key chain. Furthermore, 3 weeks after the key chain distribution, 14 volunteers contacted 770 FSW in 20 neighborhoods. Approximately 9 in 10, that is, 91.4% (704/770) of the FSW agreed to participate, of which 77.6% (546/704) accepted to receive the bangles. More than a fifth, that is, 22.4% (158/704) of the FSW had already received the bangles.

In Nimule, LINKAGES service data indicated that 2204 FSW tested for HIV in 2016. The RDS survey included 408 participants, out of which 31 tested for HIV at LINKAGES-Nimule in 2016. Of the 7 seeds in the RDS survey, 2 were affiliated with LINKAGES. The longest chain had 12 waves. Among RDS participants, 16 received a key chain, and 17 received a key chain and bangle. None received only a bangle. Of the 33 who received a key chain, 10 presented it during eligibility screening.

The service multiplier method yielded 9384 (95% CI 8511-10,257). The 2-source estimates for key chain and RDS yielded 6973 (95% CI 4759-9186); bangles and RDS yielded 13,104 (95% CI 7101-19,106), and key chains and bangles yielded 1322 (95% CI 1223-1420). The 3-source capture-recapture method using the Bayesian nonparametric latent-class model-based estimate yielded a population of 2694 (95% CI 1689-6945; see Table 1).

Female Population Estimates

On the basis of the National Bureau of Statistics population estimate of 2015 for females aged 15 to 64 years for Juba city, the service multiplier method estimates that 6.79% (5800/85386) of the female population are FSW, and based on 2-source capture-recapture, considering the key chain and RDS, 6.21% (5306/85386) are FSW. Using the females aged 15 to 64 years population estimates for 2017 for Nimule and taking the 3-source capture-recapture population estimates; we determined that 39.68% (2694/6790) are FSW.

Table 1. Population size estimates for female sex workers in Juba and Nimule using service multipliers and capture-recapture methods.

City or town and method	Population estimate	95% CI	Proportion of females aged 15 to 64 years ^a
Juba	— ^b	—	85,386
Service multiplier method ^c	5800	4927-6673	6.79
Capture-recapture method ^d	—	—	—
2-source (keychain and RDS ^e)	5306	4673-5939	6.21
2-source (RDS and bangle)	1863	1776-1951	2.18
2-source (keychain and bangle)	2120	2028-2211	2.48
Nimule	—	—	6790
Service multiplier method	9384	8511-10,257	138.20
Capture-recapture method	—	—	—
2-source (keychain and RDS)	6973	4759-9186	102.70
2-source (bangle and RDS)	13,104	7101-19,106	192.99
2-source (keychain and bangle)	1322	1223-1420	19.47
3-source (keychain, bangle, and RDS)	2694	1689-6945	39.68

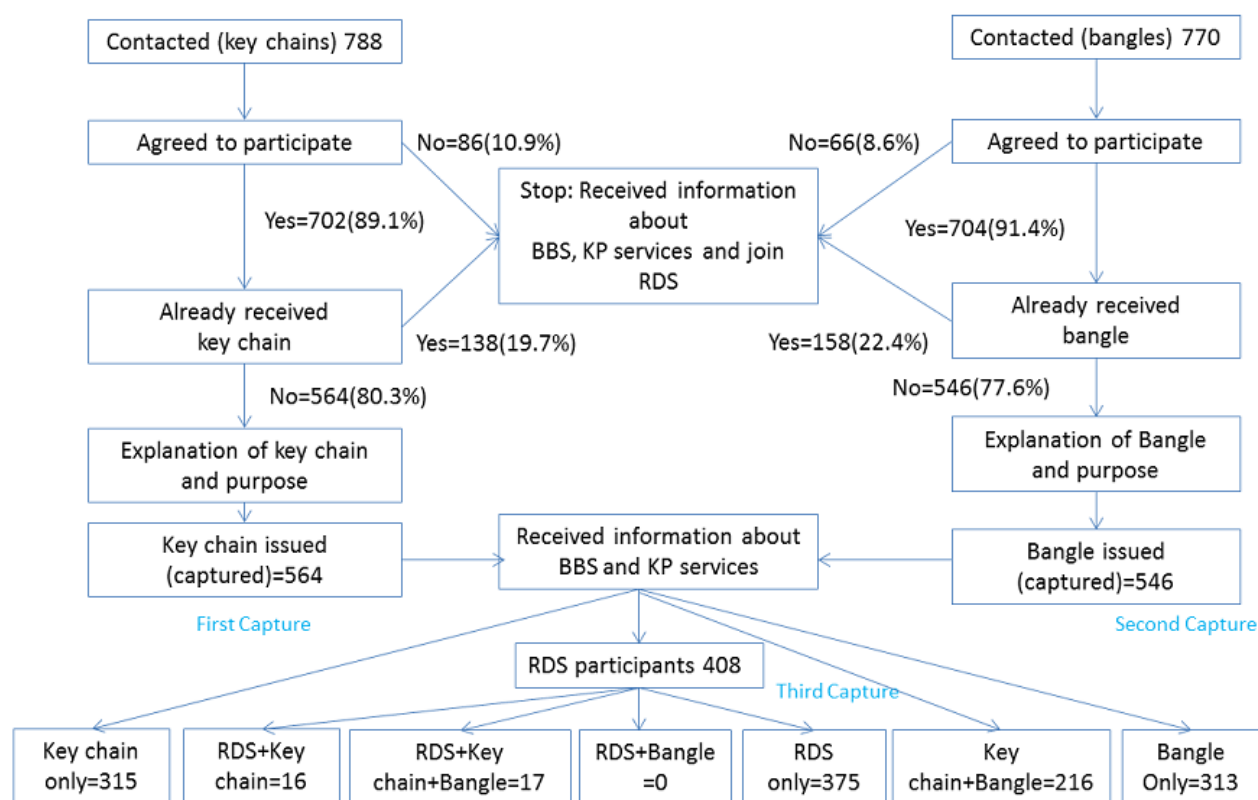
^aIndicates the National Bureau of Statistics population estimate of 2015 for females aged 15 to 64 years for Juba and estimates of 2017 for Nimule.

^bData captured under the separate source options.

^cService multiplier method: method used for population size estimation methods used in Juba and Nimule.

^dCapture-recapture method: method used for population size estimation in Juba and Nimule using unique objects that resulted into the different coerced combinations as indicated in the table.

^eRDS: respondent-driven sampling.

Figure 2. Population size estimation process in Nimule. BBS: biobehavioral survey; RDS: respondent-driven sampling; KP: key population.

Discussion

Principal Findings

In Juba, the service multiplier and the key chain unique object multiplier yielded similar population size estimates. These estimates were higher than previous estimates developed through mapping [4]. This may be because the mapping exercise defined FSW as those who received money for sex, excluding those that received goods or services for sex. The mapping was only conducted at hot spots (venues) and may have excluded hidden subpopulations operating in homes and compounds that were reached through the unique object distribution and RDS. The BBS widely defined FSW as those who exchange sex for money or goods or services, and through the peer referral, it was able to reach the hidden subpopulations. We noted divergent population size estimates in Nimule across all methods. This may result from a violation of the assumption of a closed population, stemming from displacement of FSW and other people as a result of the 2016 conflict.

The service multiplier method relied on the LINKAGES program data. The LINKAGES program uses a unique identification code for those who are reached with HIV testing and counseling; however, the difficulty in correct identification of FSW during outreach community-level testing may have resulted in some members of the general population being included in the counts, increasing the number of FSW testing and therefore the LINKAGES service beneficiaries. Our service count in Juba included all of 2015 to ease participant recall; however, the RDS survey began in November 2015. It is highly unlikely that people who tested negative in the survey in November to December 2015 would have gone for testing again at LINKAGES before the end of 2015. Survey participants living with HIV would have had no need for HIV testing after participating in the survey. Therefore, we do not feel that using the LINKAGES data through the end of 2015 risked excluding survey participants to produce a sizeable impact on the size estimate.

The proportion of FSW in the female population aged 15 to 64 in Juba is close to the estimates of 5% obtained in Kenya [10,18]. However, the absolute number for population estimates for FSW varies across the East Africa region and elsewhere [10,19,20]. In Nimule, the female population aged 15 to 64 years, who are FSW, was higher than that found in East Africa and elsewhere [10,19,20]. This is a result of high estimates of absolute numbers for FSW in this cross-border town. Ferguson and Morris estimated the number of FSW along the northern transport corridor of East Africa that links Mombasa port in Kenya and the rest of East African countries, and they found a higher proportion of FSW compared with general population women in the cross-border town of Busia (between Kenya and Uganda) compared with similar size towns along the transport corridor that are not on a border [21]. Our findings from the BBS also indicate that many of the sex workers who sell sex in Nimule reside or work across the border in the town of Elegu, Uganda.

Limitations

Capture-recapture has 4 conditions that need to be fulfilled to give reliable estimates [22]. First, encounter (capture) and recapture (second encounter) need to occur close to the capture visit (first encounter). The short time frame helps ensure that a very small number of sex workers move in or out of Juba and Nimule, meeting the condition of a closed population. However, this was not possible because of logistical challenges followed by operational challenges brought on by the conflict in the country. Second, all FSW should have the same probability of being captured. In order to address this, we mapped hot spots to facilitate volunteer selection. Unique object distribution occurred during the time FSW were most available. However, some FSW, particularly South Sudanese, operate in homes and likely had a lower probability of being captured. Third, we could not eliminate all dependency among samples; for instance, women captured during the first round may be more likely to be captured again if they recognize the volunteers or were recognized by volunteers. A positive dependency among samples, the most likely situation, will lead to an underestimation of the number of sex workers. The independence of samples is most important for multiplier and 2-source capture-recapture methods, but it can be relaxed for the 3-source method. The independence of service and unique object multiplier estimates was likely assured by using weighted estimates from the RDS survey. In addition, we used key chains and bangles that did not have significant monetary value but were unique and attractive so that they would be remembered by survey participants. Furthermore, volunteers wore distinct T-shirts to facilitate recognition and participant recall. Asking RDS participants to identify the key chain or bangle during the RDS survey helped ensure data quality and check that FSW were not biased to provide a certain response or guessing.

It was not possible to produce a 3-source capture-recapture estimate in Juba because we only collected aggregate data during the bangle distribution. Having finalized the RDS, we needed to have asked a few more questions during the distribution of the second unique object (bangle) to enable disaggregation of the data to precisely determine the number of FSW who were in each capture. The analysis method required individual level data at each capture, which were not collected. The 3-source would have given much more accurate results, though with wider confidence intervals compared with the 2-source.

The correct identification of sex workers at the hot spots is a key factor to the success of the object distributions. The use of FSW as volunteers was of utmost value [23]. Peers are not only in the best position to identify other FSW but they also build confidence and trust in potential participants. However, it is possible that some FSW were missed because of nonidentification, especially those operating out of their homes. There also could have been recall bias given that the interval between object distribution and the start of the RDS survey was quite long in Nimule, as was the bangle distribution in Juba. Using the same volunteers for the distribution of both objects may decrease the independence of samples; however, given the limited number of such volunteers in our study, we do not think our results were impacted by this.

Conclusions

By utilizing multiple size estimation methods that include people who are not easily counted in mapping activities, this study produced the robust estimates of the number of FSW in Juba and Nimule, South Sudan. Unsurprisingly, we found that the number of FSW in Juba and Nimule is higher than previously estimated. Despite the limitations brought about by the conflict in South Sudan, the 3-source capture-recapture method was successfully used to estimate the population of FSW in Nimule. It may also be the method best-suited for estimating the size of populations in conflict-affected settings and other environments with high mobility as the assumptions of independent samples are more relaxed and can be done without relying upon a BBS that can take months to plan and implement. These estimates will help inform resource allocation and advocacy efforts to support services for FSW in South Sudan.

Recommendations

The implementing partners that provide services to the FSW should improve on data quality during collection by ensuring correct identification of FSW and deduplication of the records to facilitate the use of the service multiplier method. For the unique object multiplier and capture-recapture methods, we recommend having different volunteers distribute objects for each distribution to increase the independence of samples.

When a BBS is used as part of a 3-source capture-recapture method, the 2 unique object distributions should occur before the BBS. This enables individual level data to be more accurately collected in a survey setting by data collectors who receive intensive training rather than by volunteers who may not be able to collect the required data.

Acknowledgments

This project has been supported by the President's Emergency Plan for AIDS Relief through the Centers for Disease Control and Prevention under the terms of cooperative agreement number 1U2GGH000678 with IntraHealth International.

Conflicts of Interest

None declared.

References

1. UNAIDS. Country factsheets: Uganda 2017 URL: <http://www.unaids.org/en/regionscountries/countries/uganda> [accessed 2018-12-04] [WebCite Cache ID 75w4GMI70]
2. UNAIDS. Country factsheets: Kenya 2017 URL: <http://www.unaids.org/en/regionscountries/countries/kenya> [accessed 2018-12-04] [WebCite Cache ID 75w4OjB8h]
3. UNAIDS. Overview: South Sudan URL: <http://www.unaids.org/en/regionscountries/countries/southsudan> [accessed 2018-12-10] [WebCite Cache ID 75w4WFCEk]
4. South Sudan AIDS Commission, World Health Organization. Mapping of female sex workers in South Sudan: A Geographical Mapping Approach. Juba: South Sudan AIDS Commission; Mar 10, 2012.
5. International Organization for Migration South Sudan. Population Displacements Continue to Rise in South Sudan URL: <https://southsudan.iom.int/media-and-reports/press-release/population-displacement-continues-rise-south-sudan> [accessed 2018-10-04] [WebCite Cache ID 75w5OyJaE]
6. Formative Assessment of Female Sex Workers in Juba and Yambio, South Sudan. South Sudan: Ministry of Health; 2014.
7. USAID. Nimule Female Sex Workers (FSW) Hot Spots Mapping Reports 2015. South Sudan: LINKAGES; Dec 15, 2015.
8. Groenendijk C, Veldwijk J. South Sudan Medical Journal. 2011. Sexual Exploitation and Abuse in Juba, South Sudan URL: http://www.southsudanmedicaljournal.com/assets/files/misc/CCC_Nov2011.pdf [accessed 2019-02-04] [WebCite Cache ID 75w5sA2GO]
9. Global HIV Strategic Information Working Group. Biobehavioral Survey Guidelines for Populations at Risk for HIV. Geneva: World Health Organization; 2017.
10. Vandepitte J, Lierla R, Dallabetta G, Crabbé F, Alary M, Buvé A. Estimates of the number of female sex workers in different regions of the world. Sex Transm Infect 2006 Jun;82 Suppl 3:iii18-iii25 [FREE Full text] [doi: [10.1136/sti.2006.020081](https://doi.org/10.1136/sti.2006.020081)] [Medline: [16735288](https://pubmed.ncbi.nlm.nih.gov/16735288/)]
11. Khalid FJ, Hamad FM, Othman AA, Khatib AM, Mohamed S, Ali AK, et al. Estimating the number of people who inject drugs, female sex workers, and men who have sex with men, Unguja Island, Zanzibar: results and synthesis of multiple methods. AIDS Behav 2014 Jan;18(Suppl 1):S25-S31. [doi: [10.1007/s10461-013-0517-x](https://doi.org/10.1007/s10461-013-0517-x)] [Medline: [23709254](https://pubmed.ncbi.nlm.nih.gov/23709254/)]
12. Quaye S, Fisher Raymond H, Atuahene K, Amenyah R, Aberle-Grasse J, McFarland W, et al. Critique and lessons learned from using multiple methods to estimate population size of men who have sex with men in Ghana. AIDS Behav 2015 Feb 24;19(S1):16-23. [doi: [10.1007/s10461-014-0943-4](https://doi.org/10.1007/s10461-014-0943-4)] [Medline: [25704987](https://pubmed.ncbi.nlm.nih.gov/25704987/)]
13. Vuylsteke B, Vandenhoudt H, Langat L, Semde G, Menten J, Odongo F, et al. Capture-recapture for estimating the size of the female sex worker population in three cities in Côte d'Ivoire and in Kisumu, western Kenya. Trop Med Int Health 2010 Dec;15(12):1537-1543 [FREE Full text] [doi: [10.1111/j.1365-3156.2010.02654.x](https://doi.org/10.1111/j.1365-3156.2010.02654.x)] [Medline: [21054693](https://pubmed.ncbi.nlm.nih.gov/21054693/)]

14. Johnston LG, McLaughlin KR, el Rhilani H, Latifi A, Toufik A, Bennani A, et al. Estimating the size of hidden populations using respondent-driven sampling data: case examples from Morocco. *Epidemiology* 2015 Nov;26(6):846-852 [FREE Full text] [doi: [10.1097/EDE.0000000000000362](https://doi.org/10.1097/EDE.0000000000000362)] [Medline: [26258908](https://pubmed.ncbi.nlm.nih.gov/26258908/)]
15. Holland CE, Kouanda S, Lougué M, Pitche VP, Schwartz S, Anato S, et al. Using population-size estimation and cross-sectional survey methods to evaluate HIV service coverage among key populations in Burkina Faso and Togo. *Public Health Rep* 2016 Dec;131(6):773-782 [FREE Full text] [doi: [10.1177/0033354916677237](https://doi.org/10.1177/0033354916677237)] [Medline: [28123223](https://pubmed.ncbi.nlm.nih.gov/28123223/)]
16. Biology Blog & Dictionary Online. Capture-recapture method URL: https://www.biology-online.org/dictionary/Capture-recapture_method [accessed 2018-07-20] [WebCite Cache ID 75w6MQL0H]
17. Manrique-Vallier D. Bayesian population size estimation using Dirichlet process mixtures. *Biom J* 2016;72(4):1246-1254 [FREE Full text] [doi: [10.1111/biom.12502](https://doi.org/10.1111/biom.12502)]
18. Odek WO, Githuka GN, Avery L, Njoroge PK, Kasonde L, Gorgens M, et al. Estimating the size of the female sex worker population in Kenya to inform HIV prevention programming. *PLoS One* 2014;9(3):e89180 [FREE Full text] [doi: [10.1371/journal.pone.0089180](https://doi.org/10.1371/journal.pone.0089180)] [Medline: [24595029](https://pubmed.ncbi.nlm.nih.gov/24595029/)]
19. Okal J, Geibel S, Muraguri N, Musyoki H, Tun W, Broz D, et al. Estimates of the size of key populations at risk for HIV infection: men who have sex with men, female sex workers and injecting drug users in Nairobi, Kenya. *Sex Transm Infect* 2013 Aug;89(5):366-371 [FREE Full text] [doi: [10.1136/sextrans-2013-051071](https://doi.org/10.1136/sextrans-2013-051071)] [Medline: [23761166](https://pubmed.ncbi.nlm.nih.gov/23761166/)]
20. Mutagoma M, Kayitesi C, Gwiza A, Ruton H, Koleros A, Gupta N, et al. Estimation of the size of the female sex worker population in Rwanda using three different methods. *Int J STD AIDS* 2015 Oct;26(11):810-814 [FREE Full text] [doi: [10.1177/0956462414555931](https://doi.org/10.1177/0956462414555931)] [Medline: [25336306](https://pubmed.ncbi.nlm.nih.gov/25336306/)]
21. Ferguson AG, Morris CN. Mapping transactional sex on the Northern Corridor highway in Kenya. *Health Place* 2007 Jun;13(2):504-519. [doi: [10.1016/j.healthplace.2006.05.009](https://doi.org/10.1016/j.healthplace.2006.05.009)] [Medline: [16815730](https://pubmed.ncbi.nlm.nih.gov/16815730/)]
22. World Health Organization. 2003. Estimating the size of populations at risk for HIV: issues and methods URL: <http://www.who.int/hiv/pub/surveillance/en/EstimatingSizePop.pdf> [accessed 2018-10-04] [WebCite Cache ID 75w9rtQVw]
23. Kruse N, Behets FM, Vaovola G, Burkhardt G, Barivelo T, Amida X, et al. Participatory mapping of sex trade and enumeration of sex workers using capture-recapture methodology in Diego-Suarez, Madagascar. *Sex Transm Dis* 2003 Aug;30(8):664-670. [doi: [10.1097/01.OLQ.0000079523.04451.82](https://doi.org/10.1097/01.OLQ.0000079523.04451.82)] [Medline: [12897692](https://pubmed.ncbi.nlm.nih.gov/12897692/)]

Abbreviations

BBS: biobehavioral survey
CDC: Centers for Disease Control and Prevention
FSW: female sex workers
MOH: Ministry of Health
RDS: respondent-driven sampling

Edited by J Neal; submitted 13.07.18; peer-reviewed by A Abdul-Quader, W Mcfarland; comments to author 23.10.18; revised version received 25.11.18; accepted 14.12.18; published 18.03.19.

Please cite as:

Okiria AG, Bolo A, Achut V, Arkangelo GC, Michael ATI, Katoro JS, Wesson J, Gutreuter S, Hundley L, Hakim A
Novel Approaches for Estimating Female Sex Worker Population Size in Conflict-Affected South Sudan
JMIR Public Health Surveill 2019;5(1):e11576
URL: <http://publichealth.jmir.org/2019/1/e11576/>
doi:[10.2196/11576](https://doi.org/10.2196/11576)
PMID:[30882356](https://pubmed.ncbi.nlm.nih.gov/30882356/)

©Alfred Geoffrey Okiria, Alex Bolo, Victoria Achut, Golda Ceasar Arkangelo, Acaga Taban Ismail Michael, Joel Sua Katoro, Jennifer Wesson, Steve Gutreuter, Lee Hundley, Avi Hakim. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 18.03.2019. 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

Kuantim mi tu (“Count me too”): Using Multiple Methods to Estimate the Number of Female Sex Workers, Men Who Have Sex With Men, and Transgender Women in Papua New Guinea in 2016 and 2017

Damian Weikum¹, MPH; Angela Kelly-Hanku^{2,3}, PhD; Parker Hou³; Martha Kupul³, BS; Angelyne Amos-Kuma³, MPH; Steven G Badman², MPH, JP; Nick Dala⁴, MBBS, MMed; Kelsey C Coy¹, MPH; John M Kaldor², PhD; Andrew J Vally², MBBS, MRCP, DTM&H, MSc, DLSHTM, PhD; Avi J Hakim¹, MA, MPH

¹US Centers for Disease Control and Prevention, Atlanta, GA, United States

²Kirby Institute, UNSW, Sydney, Australia

³Papua New Guinea Institute of Medical Research, Goroka, Papua New Guinea

⁴Papua New Guinea National Department of Health, Port Moresby, Papua New Guinea

Corresponding Author:

Damian Weikum, MPH

US Centers for Disease Control and Prevention

1600 Clifton Road, NE

Mail Stop E-30

Atlanta, GA, 30329

United States

Phone: 1 4047185688

Email: damian.weikum@gmail.com

Abstract

Background: Female sex workers (FSW), men who have sex with men (MSM), and transgender women (TGW) are at high risk of acquiring HIV in many settings, such as Papua New Guinea (PNG). An understanding of the approximate size of these populations can inform resource allocation for HIV services for FSW, MSM, and TGW.

Objective: An objective of this multi-site survey was to conduct updated population size estimations (PSE) of FSW and MSM/TGW.

Methods: Respondent-driven sampling (RDS) biobehavioral surveys of FSW and MSM/TGW were conducted in 3 major cities—(1) Port Moresby, (2) Lae, and (3) Mount Hagen—between June 2016 and December 2017. Eligibility criteria for FSW included: (1) ≥12 years of age, (2) born female, (3) could speak English or Tok Pisin (PNG Pidgin), and (4) had sold or exchanged sex with a man in the past six months. Eligibility for MSM/TGW included: (1) ≥12 years of age, (2) born male, (3) could speak English, or Tok Pisin, and (4) had engaged in oral or anal sex with another person born male in the past six months. PSE methods included unique object multiplier, service multiplier, and successive sampling-population size estimation (SS-PSE) using imputed visibility. Weighted data analyses were conducted using RDS-Analyst and Microsoft Excel.

Results: Sample sizes for FSW and MSM/TGW in Port Moresby, Lae, and Mount Hagen included: (1) 673 and 400, (2) 709 and 352, and (3) 709 and 111 respectively. Keychains were used for the unique object multiplier method and were distributed 1 week before the start of each RDS survey. HIV service testing data were only available in Port Moresby and Mount Hagen and SS-PSE estimates were calculated for all cities. Due to limited service provider data and uncertain prior size estimation knowledge, unique object multiplier weighted estimations were chosen for estimates. In Port Moresby, we estimate that there are 16,053 (95% CI 8232-23,874) FSW and 7487 (95% CI 3975-11,000) MSM/TGW, approximately 9.5% and 3.8% of the female and male populations respectively. In Lae, we estimate that there are 6105 (95% CI 4459-7752) FSW and 4669 (95% CI 3068-6271) MSM/TGW, approximately 14.4% and 10.1% of the female and male populations respectively. In Mount Hagen, we estimate that there are 2646 (95% CI 1655-3638) FSW and 1095 (95% CI 913-1151) MSM/TGW using service multiplier and successive sampling, respectively. This is approximately 17.1% and 6.3% of the female and male populations respectively.

Conclusions: As the HIV epidemic in PNG rapidly evolves among key populations, PSE should be repeated to produce current estimates for timely comparison and future trend analysis.

(*JMIR Public Health Surveill* 2019;5(1):e11285) doi:[10.2196/11285](https://doi.org/10.2196/11285)

KEYWORDS

Papua New Guinea; population size estimation; female sex workers; men who have sex with men; transgender women; key populations; respondent-driven sampling; unique object multiplier; service multiplier; successive sampling

Introduction

HIV disproportionately affects marginalized and stigmatized populations [1]. Female sex workers (FSW), men who have sex with men (MSM), and transgender women (TGW) are 3 key populations (KP) at greater risk for HIV [1-3]. Globally, FSW, MSM, and TGW are estimated to be 13.5 [2], 19.3 [3], and 48.8 [4] times more likely to be infected with HIV than the general population, respectively. This risk is accentuated by critical barriers to HIV-related prevention and treatment services, such as violence, criminalization, stigma, and discrimination [5-7].

Papua New Guinea (PNG) has the largest HIV epidemic in the Pacific region, with a national prevalence estimated at 0.9% [8]. Recent surveys of FSW and MSM/TGW estimated HIV prevalence among FSW and MSM/TGW at 14.9% and 8.5%, respectively, in the capital of Port Moresby. In Lae, the country's economic hub, it was estimated at 12.9% and 7.1%, respectively [9].

Population size estimates can inform resource targeting and program monitoring [1]. The only population size estimation of FSW and MSM in Port Moresby using empirical methods was conducted in 2006 and utilized the service multiplier in conjunction with a respondent-driven sampling (RDS) survey [10]. Social changes such as population growth and increased mobility have occurred since then [11,12]. Updated size estimates are needed in Port Moresby and other cities, such as Lae and Mount Hagen, require first-time estimates. We used unique object multiplier, service multiplier, and successive sampling-population size estimation methods, which are direct and empirical, to estimate the number of FSW and MSM/TGW in these 3 cities.

Methods

Community Consultation

Community consultation was undertaken within each city. Population members recognized that MSM and TGW are distinct populations but TGW are too few to achieve adequate sample size as an independent RDS sample. The 2 populations thus agreed to be combined into a single sample.

Recruitment

Seeds were purposively selected to be diverse with respect to age, sexual and gender identity, residence, region of origin, marital status, receipt of a unique object, and affiliation with a non-governmental or community-based organization.

Data Collection

Separate RDS biobehavioral surveys (BBS) of FSW and MSM/TGW were conducted in the 3 cities Port Moresby, Lae, and Mount Hagen between June 2016 and December 2017. Survey eligibility criteria for FSW included: (1) born female, (2) >12 years of age, (3) able to speak English or Tok Pisin, and (4) had sold or exchanged sex with a man in the past 6 months. Survey eligibility for MSM/TGW included: (1) born male, (2) >12 years of age, (3) able to speak English or Tok Pisin, and (4) had engaged in oral or anal sex in the past 6 months with another person born male.

Sample Size and Precision

We aimed to enroll 700 people into each BBS in each city. This assumed a design effect of 2 and was sufficiently powered to estimate an assumed HIV prevalence of 20% with an absolute precision of 10% [1]. The sample size was calculated so that the results of the present survey can be compared to anticipated follow-up studies ([Multimedia Appendix 1](#), [13-15]).

Unique Object Multiplier Method

Given the general unavailability of HIV testing or organization membership data in the survey cities, we primarily used the unique object multiplier (UOM) in conjunction with RDS surveys to estimate population size [16,17]. Volunteers, consisting of local KP peers and survey team members, distributed approximately 1000 keychains per population 1 week before survey rollout in each city. They were instructed to verify that each keychain recipient had not already received an object, that this person received only 1 object, and told the recipient to keep the object for the near future and not to give it to anyone else. Distributors were provided with 30 kina (US \$12) for distributing keychains, irrespective of the actual number distributed—this removed any incentive to report distributing more keychains than actuality as may occur if the compensation was given per keychain distributed. They also received 5 kina (US \$2) for transportation. To strengthen recall of receiving an object among BBS participants, keychain distributors wore distinctive hats that featured the survey logo “Kauntim mi tu.” During survey eligibility screening, the coupon manager asked participants if they received a unique object from a distributor. Those indicating that they had the unique object were asked to show it. If unable to do so, they were asked to select the keychain from among other keychains displayed by the coupon manager.

Service Multiplier Method

The service multiplier method was used only in Port Moresby and Mount Hagen [10,18,19]. Survey participants were asked during the face-to-face interview whether they had tested for

HIV at specific health facilities in 2015 (Port Moresby) or 2016 (Mount Hagen). Four HIV testing providers in Port Moresby and 1 in Mount Hagen were capable of providing key population-specific testing information. None were able to do so in Lae. Survey participants in each city were also asked whether they belonged to a KP organization. Responses to these questions were paired with data from HIV testing organizations and the KP community organization to develop multiple service multiplier estimates. In Lae and Mount Hagen, no KP organization could provide unique membership data.

Successive Sampling Method: Population Size Estimation

Using Respondent-Driven Sampling Analyst (RDS-A) version 0.62 [20,21] the successive sampling-population size estimation (SS-PSE) method was used to produce size estimates using routinely collected data in RDS surveys including: (1) self-reported network size, (2) number of participant's recruits enrolled in the survey, and (3) the date order of survey enrollment. We imputed visibility using these 3 routinely collected data items in order to smooth the network size distribution, reduce the effect of outliers, and minimize heaping of values (eg, around 5,10,15) [22]. Prior estimates were calculated using distribution of age and sex in each city, city general population sizes, proportions of FSW, MSM, and TGW in other countries, and previous knowledge of sex work in PNG.

Data Analysis

Data were analyzed using RDS-A with the Gile SS-PSE and Microsoft Excel 2016. Standard formulas for the UOM method, service multiplier method, and SS-PSE were used [1]. Both weighted and unweighted estimates were produced for the multiplier methods to compare results. The 95% CI were calculated around point estimates using RDS-A.

Protection of Minors

Participants <18 years were provided referrals as needed to organizations that offer counseling, protection, and other relevant services for victims of sexual exploitation and abuse.

Ethical Approval

This survey was approved by the PNG National Department of Health's Medical Research Advisory Committee, the Research Advisory Committee of the National AIDS Council Secretariat, the PNG Institute of Medical Research's Institutional Review Board, and the Human Research Ethics Committee at University of New South Wales Sydney, Australia. The activity was reviewed according to the Centers for Disease Control and Prevention (CDC) human research protection procedures and was determined to be research but the CDC was not engaged in research collection. Two peer-led organizations for KP (ie, Friends Frangipani and Kapul Champions) provided letters of endorsement.

Results

Sampling and Recruitment

Among the FSW, 673 were enrolled in Port Moresby, 709 in Lae, and 709 in Mount Hagen while for MSM/TGW there were 400 enrolled in Port Moresby, 352 in Lae, and 111 in Mount Hagen (Table 1). Similarly, more keychains were distributed to FSW in Port Moresby, Lae, and Mount Hagen (N=867, N=790, N=546, respectively) compared to MSM/TGW (N=598, N=777, N=152, respectively), suggesting that FSW are easier to reach, and they are likely to be better networked. The number of keychains to be distributed was determined with a calculator found in international guidelines [1]; we aimed to distribute more keychains than our sample size. The RDS data collection took approximately the same amount of time for each city and population (14-20 weeks).

Tables 2 and 3 present population size estimates for KP in each city. For FSW (Table 2), population size estimates in Port Moresby ranged from 3537-35,048 (2.1%-20.7% of the adult female population), in Lae from 4482-6105 (10.5%-14.4% of the adult female population), and in Mount Hagen from 2386-6315 (15.5%-40.9% of the adult female population). For MSM/TGW (Table 3), population size estimates in Port Moresby ranged from 501-18,644 (0.3%-9.6% of the adult male population), in Lae from 3455-4669 (7.5%-10.1% of the adult male population), and in Mount Hagen from 1095-3625 (6.3%-20.8% of the adult male population).

Table 1. Description of surveys conducted among female sex workers (FSW) and men who have sex with men (MSM) and transgender women (TGW) in Papua New Guinea in 2016 and 2017.

Target population in each city	Enrolled participants, (N)	Keychains distributed, (N)	Keychains not distributed, (N)	Participants receiving key-chains, (N)	Survey period	Data collection period, weeks
FSW						
Port Moresby	673	867	133	51	Jun 2016-Nov 2016	20
Lae	709	790	291	110	Jan 2017-May 2017	16
Mount Hagen	709	546	454	71	Sep 2017-Dec 2017	14
MSM/TGW						
Port Moresby	400	598	402	35	Jun 2016-Nov 2016	20
Lae	352	777	223	75	Jan 2017-May 2017	17
Mount Hagen	111	152	848	5	Sep 2017-Dec 2017	14

Port Moresby provided the most analytic possibilities, although there was variability in the number of unique individuals tested across these 4 organizations and only 1 KP organization could provide data on the number of members. As described above, in Lae, only the UOM and SS-PSE methods were used. In Mount Hagen, UOM, service multiplier (only for FSW), and SS-PSE estimates were derived for both KP. As many population estimates were developed in Port Moresby, meetings were held to discuss the estimates and the extent to which assumptions were met in order to identify a final estimate. In Lae and Mount Hagen, where fewer population estimates were developed, such meetings played a smaller role. Each of the 3 methods is direct and empirical, making them superior to other indirect methods such as census and enumeration [1]. These latter simplistic methods depend on counting everyone or a large number of people, which can be costly for census and often unfeasible for both with hidden groups like KP. Finally, weighted estimates

were selected over unweighted estimates in order to increase the likelihood of independence between the convenient nature of keychain distribution (the capture) and RDS recruitment (the recapture) by turning the RDS sample data into population-based data. Without weighting of estimates, key assumptions of multiplier methods are violated.

Population Size

Final weighted estimates chosen for each city were (1) 16,053 (95% CI 8232-23,874) FSW and 7487 (95% CI 3975-11,000) MSM/TGW in Port Moresby approximately 9.5% and 3.8% of the female and male populations, respectively, (2) 6105 (95% CI 4459-7752) FSW and 4669 (95% CI 3068-6271) MSM/TGW in Lae, approximately 14.4% and 10.1% of the female and male populations, respectively, and (3) 2646 (95% CI 1655-3638) FSW and 1095 (95% CI 913-1151) MSM/TGW in Mount Hagen approximately 17.1% and 6.3% of the female and male populations respectively (Tables 2 and 3).

Table 2. Population size estimates for female sex workers (FSW) in Port Moresby, Lae, and Mount Hagen using unique object multiplier (UOM), service multipliers organizations (ORG1, ORG2, ORG3, ORG4), and successive sampling-population size estimation (SS-PSE) in 2015 and 2016.

Population size estimation method	Multiplier number, N	Survey, %	Estimate, 95% CI	Female urban population size based on total, % ^a
Port Moresby, N=169,291^a				
UOM				
Unweighted	867	7.6	11,407 (8487-14,327)	6.7
Weighted	867	5.4	16,053 (8232-23,874)	9.5
ORG1				
KP membership				
Unweighted	3169	10.8	28,906 (22,764-35,048)	17.1
Weighted	3169	9.6	32,532 (30,324-34,743)	19.2
HIV testing				
Unweighted	908	16.6	5464 (4599-6329)	3.2
Weighted	908	14.4	6328 (4383-8273)	3.7
ORG2				
HIV testing				
Unweighted	77	1.2	6487 (2257-10,717)	3.8
Weighted	77	1.9	3907 (432-7383)	2.3
ORG3				
HIV testing				
Unweighted	208	0.6	35,048 (1134-68,962)	20.7
Weighted	208	1	18,773 (0-39,354)	11.1
ORG4				
HIV testing				
Unweighted	63	0.4	14,154 (0-29,750)	8.4
Weighted	63	0.4	14,078 (0-31,388)	8.3
SS-PSE	___ ^b	___ ^b	3537 (1062-6870)	2.1
Lae, N=42,532^a				
UOM				
Unweighted	790	15.5	5092 (4280-5903)	12
Weighted	790	13	6105 (4459-7752)	14.4
SS-PSE	___ ^b	___ ^b	4482 (1473-7388)	10.5
Mount Hagen, N=15,430^a				
UOM				
Unweighted	546	10	5452 (4330-6574)	35.3
Weighted	546	8.6	6315 (4668-7963)	40.9
ORG2				
HIV testing				
Unweighted	138	5.8	2386 (1792-2981)	15.5
Weighted	138	5.2	2646 (1655-3638)	17.1
SS-PSE	___ ^b	___ ^b	3843 (1303-7989)	24.9

^aValues provided by the 2011 census of Papua New Guinea.^bNot applicable.

Table 3. Population size estimates for men who have sex with men (MSM) and transgender women (TGW) in Port Moresby, Lae, and Mount Hagen using unique object multiplier (UOM), service multiplier organizations (ORG1, ORG2, ORG3, ORG4), and successive sampling-population size estimation (SS-PSE) in 2015 and 2016.

Population size estimation method	Multiplier number, N	Survey, %	Estimate, (95% CI)	Male urban population size based on total, % ^a
Port Moresby, N=194,834^a				
UOM				
Unweighted	598	8.8	6834 (4735-8932)	3.5
Weighted	598	8	7487 (3975-11,000)	3.8
ORG1				
KP membership				
Unweighted	792	5.8	13,773 (8388-19,158)	7.1
Weighted	792	4.2	18,644 (13,773-23,514)	9.6
HIV testing				
Unweighted	183	8	2288 (1597-2978)	1.2
Weighted	183	7.8	2380 (1218-3543)	1.2
ORG2				
HIV testing				
Unweighted	___b	___b	___b	___b
Weighted	___b	___b	___b	___b
ORG3				
HIV testing				
Unweighted	7	1.3	560 (299-821)	0.3
Weighted	7	1.5	501 (0-1175)	0.3
ORG4				
HIV testing				
Unweighted	8	0.8	1067 (116-2017)	0.5
Weighted	8	0.3	2185 (0-4941)	1.1
SS-PSE	___c	___c	3846 (3074-4200)	2
Lae, N=46,076^a				
UOM				
Unweighted	777	21.3	3647 (2951-4343)	7.9
Weighted	777	16.8	4669 (3068-6271)	10.1
SS-PSE	___c	___c	3455 (2752-3672)	7.5
Mount Hagen, N=17,400^a				
UOM				
Unweighted	152	4.5	3374 (532-6217)	19.4
Weighted	152	4.2	3625 (754-6497)	20.8
ORG2				
HIV testing				
Unweighted	___b	___b	___b	___b
Weighted	___b	___b	___b	___b
SS-PSE	___c	___c	1095 (913-1151)	6.3

^aValues provided by the 2011 census of Papua New Guinea.

^bNot available.

^cNot applicable.

Discussion

We employed several methods to develop population size estimates of FSW and MSM/TGW in PNG. We believe that the most robust population size estimates produced were with the UOM for both KP in Port Moresby and Lae, and in Mount Hagen, the service multiplier for FSW and SS-PSE for MSM/TGW. We present both unweighted and weighted estimates, the latter which adjusts for RDS recruitment. Individuals with smaller network sizes are up-weighted, whereby their responses are more valued as seen by the larger weighted estimates than unweighted estimates in our study.

Final population estimates were chosen through a series of rigorous meetings between experts in PNG, Australia, and the United States. Given the number and range of estimates produced, investigators reviewed each size estimation method, the extent to which their assumptions were met, and the resulting estimates. The investigators narrowed down the estimates to those that were most robust in each city. These estimates were then presented to key stakeholders including the National Department of Health, key population organizations, donors, and United Nation agencies who were tasked with agreeing on a single estimate. Details and results of these discussions follow below.

Several precautionary steps were taken to maximize the utility and accuracy of the unique object multiplier. We minimized the risk of participants falsely indicating that they had received a keychain when in fact they had not by asking participants to (1) show their keychain, or (2) identify the correct keychain from a group consisting of this keychain and other incorrect keychains. Distributing the unique object approximately one week before the start of the RDS survey helped to (1) limit in and out-migration of participants, (2) minimize the possibility of target population members giving their object away, and (3) reduce the possibility of someone receiving objects from multiple distributors.

It is worth noting that fewer keychains were distributed to MSM/TGW, which may be explained by their lower propensity toward gathering in public places due to stigma and discrimination and smaller network sizes, as compared to FSW. While keychain distribution is influenced partially by how well a distributor is trained, smaller network sizes pose a challenge because fewer people may be present in KP-friendly hotspots and there is greater reliance on tapping into an individual's limited social connections to distribute these keychains. We tried to address stigma/discrimination barriers in Mount Hagen prior to study initiation by building trust and rapport. For future work, we would consider using more volunteers who each distribute fewer objects, since individuals may know fewer people due to stigma.

We encountered no problems distributing keychains to either population and leftover keychains were returned to survey sites, suggesting that keychain distribution occurred with fidelity. In

Mount Hagen, UOM estimates indicated that FSW account for approximately 40% of the adult female population and MSM/TGW 20% of the adult male population. (Tables 2 and 3). These estimates were deemed unreasonably high, leading to their exclusion in favor of the service multiplier and SS-PSE estimates, respectively, in this city. All MSM/TGW estimates in Mount Hagen were hampered by the small RDS sample recruited.

While the data provided by HIV testing organizations in Port Moresby were meant to represent unique individuals, some service multiplier results there were excluded due to very wide CI values, while other population size estimates were too low to be plausible given our understanding of sex work in this city, and MSM/TGW globally [16]. Our results are higher than results from a previous survey in Port Moresby [10], which reported FSW and MSM service multiplier estimates to be 4212 (95% CI 3586-4839) and 2126 (1787-2468), respectively. We believe that our results are acceptable because those population estimates are almost a decade old, Port Moresby has undergone tremendous growth since then, and we included transgender women in our estimates.

Given the age of previous PNG estimates, we lacked a recent and more accurate prior value for use in producing an estimate with the SS-PSE method, suggesting a reason why the results from this method were relatively low. Likewise, the SS-PSE method estimates the number of people that fit our survey's time-frame eligibility criteria (ie, behaviors in the past 6 months) [23]. Therefore, individuals who sold sex, engaged in same-sex sexual behaviors, or TGW who had sex with men more than six months ago are not included in this estimate. As these people are still KP and face multiple vulnerabilities, we felt it important to select a larger and more plausible estimate, hence the selection of the UOM for Port Moresby and Lae. In both cities, the UOM produced population size estimates that were slightly larger than SS-PSE.

Our findings are limited in several ways, beginning with the self-reported nature of the interview data. Participants may have chosen to underreport or overreport HIV testing or organization membership due to stigma or fear, resulting in overestimation and underestimation, respectively. This social desirability bias could be mitigated in the future through the use of computer-assisted self-interviews [24-26].

The unique object multiplier relies on accurate reporting of keychains given out by distributors and received by KP. We cannot know whether all keychains that were said to be distributed were actually distributed correctly, that individuals received only one keychain each, or how many keychain recipients were ineligible for the survey. Additionally, though we distributed keychains immediately before survey initiation, data collection took up to four months to complete, so it is possible that KP members migrated out of the catchment area, causing an underestimation.

Few HIV testing providers in Port Moresby and Mount Hagen were able to provide data and those that were able tested only a small number of KP. HIV testing data were nonexistent in Lae, meaning the service multiplier method could not be used. There is also the concern of recall bias for participants in Port Moresby who were asked about services accessed in 2015. Furthermore, SS-PSE relies on prior estimates, which themselves may not be accurate in Port Moresby and we did not have in either Lae or Mount Hagen.

Our findings are important for describing the number of FSW and MSM/TGW in the survey cities. The direct and empirical size estimation methods used in this survey are superior to other indirect methods, such as census and enumeration. The population size estimates in this survey will inform efforts to improve resource targeting and monitoring of both existing and new services for these key populations. HIV testing providers in PNG should be encouraged to disaggregate their data by key population to facilitate the use of the service multiplier. This will also increase the utility of routinely collected data.

Lessons Learned

Much of this project's success, which was led by the national government of PNG with technical assistance from the United

States and Australia, should be credited to the Papua New Guinean "Kaumtim mi tu" survey team, which included KP members in strategic staff positions, and who nurtured trust among KP in each city. The staff was crucial in organizing keychain distribution and obtaining service provider data. In addition, the survey team was almost unchanged across all cities. This resulted in a consistency of operations and data collection, improved efficiency, and an overall increase in technical capacity to implement RDS surveys and population size estimation activities. We also used tablets to collect data electronically, which simplified data management and decreased chances of error. Furthermore, we found it valuable to start our project with the "easiest" site first, Port Moresby, because the capital city had the most visible KP, facilitating KP engagement and survey implementation. News of the survey's benefits to individuals and KP as a whole, as well as the friendliness and professionalism of the survey team, traveled to Lae and Mount Hagen motivating participation there. Given the lack of recent population size estimates, we found it indispensable to use several methods to estimate population size because each method has its strengths and weaknesses.

Multimedia Appendix 1

Study sample size and precision calculations.

[[PDF File \(Adobe PDF File\), 37KB - publichealth_v5i1e11285_app1.pdf](#)]

References

1. WHO, CDC. UNAIDS, FHI. Biobehavioural survey guidelines for populations at risk for HIV: World Health Organization; 2017:360-392.
2. Baral S, Beyrer C, Muessig K, Poteat T, Wirtz AL, Decker MR, et al. Burden of HIV among female sex workers in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet Infect Dis* 2012 Jul;12(7):538-549. [doi: [10.1016/S1473-3099\(12\)70066-X](#)] [Medline: [22424777](#)]
3. Baral S, Sifakis F, Cleghorn F, Beyrer C. Elevated risk for HIV infection among men who have sex with men in low- and middle-income countries 2000-2006: a systematic review. *PLoS Med* 2007 Dec;4(12):e339 [FREE Full text] [doi: [10.1371/journal.pmed.0040339](#)] [Medline: [18052602](#)]
4. Baral SD, Poteat T, Strömdahl S, Wirtz AL, Guadamuz TE, Beyrer C. Worldwide burden of HIV in transgender women: a systematic review and meta-analysis. *Lancet Infect Dis* 2013 Mar;13(3):214-222. [doi: [10.1016/S1473-3099\(12\)70315-8](#)] [Medline: [23260128](#)]
5. Baral S, Holland CE, Shannon K, Logie C, Semugoma P, Sithole B, et al. Enhancing benefits or increasing harms: community responses for HIV among men who have sex with men, transgender women, female sex workers, and people who inject drugs. *J Acquir Immune Defic Syndr* 2014 Aug 15;66 Suppl 3:S319-S328. [doi: [10.1097/QAI.0000000000000233](#)] [Medline: [25007203](#)]
6. Beyrer C. Global prevention of HIV infection for neglected populations: men who have sex with men. *Clin Infect Dis* 2010 May 15;50 Suppl 3:S108-S113. [doi: [10.1086/651481](#)] [Medline: [20397938](#)]
7. Baral S, Logie CH, Grosso A, Wirtz AL, Beyrer C. Modified social ecological model: a tool to guide the assessment of the risks and risk contexts of HIV epidemics. *BMC Public Health* 2013;13:482 [FREE Full text] [doi: [10.1186/1471-2458-13-482](#)] [Medline: [23679953](#)]
8. UNAIDS. Geneva. 2018. Papua New Guinea URL: <http://www.unaids.org/en/regionscountries/countries/papuanewguinea> [accessed 2018-12-21] [WebCite Cache ID 74pifdIC]
9. Kelly-Hanku A, Badman S, Willie B, Narakobi R, Amos-Kuma A, Pekon S, et al. Kelly-Hanku A, Badman S, Willie B, Narakobi R, Amos-Kuma A, Pekon S, Kupul M, Aeno H, Neo RB, Ase S, Nembari J, Hou P, Dala N, Weikum D, Vallely A, Hakim A. 90-90-90 and the HIV continuum of care: how well is Papua New Guinea doing amongst key populations? 9th IAS Conference on HIV Science (IAS 2017); Paris: International AIDS Society; 2017. 2017 Presented at: 9th IAS Conference on HIV Science; 2017; Paris.

10. Johnston LG, Prybylski D, Raymond HF, Mirzazadeh A, Manopaiboon C, McFarland W. Incorporating the service multiplier method in respondent-driven sampling surveys to estimate the size of hidden and hard-to-reach populations: case studies from around the world. *Sex Transm Dis* 2013 Apr;40(4):304-310. [doi: [10.1097/OLQ.0b013e31827fd650](https://doi.org/10.1097/OLQ.0b013e31827fd650)] [Medline: [23486495](https://pubmed.ncbi.nlm.nih.gov/23486495/)]
11. TheWorldBank. Papua New Guinea Country Profile URL: http://databank.worldbank.org/data/Views/Reports/ReportWidgetCustom.aspx?Report_Name=CountryProfile&Id=b450fd57&tbar=y&dd=y&inf=n&zm=n&country=PNG [accessed 2018-09-09] [WebCite Cache ID 6wJy6P2JJ]
12. The World Bank. Strategic Directions for Human Development in Papua New Guinea URL: <https://openknowledge.worldbank.org/bitstream/handle/10986/13543/40924.pdf?sequence=1&isAllowed=y> [accessed 2018-08-10] [WebCite Cache ID 6wJyvH0yS]
13. Kelly A, Kupul M, Man W, Nosi S, Lote N, Rawstorne P, et al. Askim na save (Ask and understand): People who sell and/or exchange sex in Port Moresby. Key Quantitative Findings. Sydney: Papua New Guinea Institute of Medical Research and the University of New South Wales; 2011:978.
14. Maibani G, Ryan C, Lote N, Edward G, Ralai A, Gare J, et al. Papua New Guinea-Australia Sexual Health Improvement Program (PASHIP). Report on the Baseline Survey. PNG IMR 2011.
15. Valley A, Page A, Dias S, Siba P, Lupiwa T, Law G, et al. The prevalence of sexually transmitted infections in Papua New Guinea: a systematic review and meta-analysis. *PLoS One* 2010 Dec 23;5(12):e15586 [FREE Full text] [doi: [10.1371/journal.pone.0015586](https://doi.org/10.1371/journal.pone.0015586)] [Medline: [21203468](https://pubmed.ncbi.nlm.nih.gov/21203468/)]
16. Quayle S, Fisher RH, Atuahene K, Amenyah R, Aberle-Grasse J, McFarland W, Ghana Men Study Group. Critique and lessons learned from using multiple methods to estimate population size of men who have sex with men in Ghana. *AIDS Behav* 2015 Feb;19 Suppl 1:S16-S23. [doi: [10.1007/s10461-014-0943-4](https://doi.org/10.1007/s10461-014-0943-4)] [Medline: [25704987](https://pubmed.ncbi.nlm.nih.gov/25704987/)]
17. Holland CE, Kouanda S, Lougué M, Pitche VP, Schwartz S, Anato S, et al. Using Population-Size Estimation and Cross-sectional Survey Methods to Evaluate HIV Service Coverage Among Key Populations in Burkina Faso and Togo. *Public Health Rep* 2016 Nov;131(6):773-782. [doi: [10.1177/0033354916677237](https://doi.org/10.1177/0033354916677237)] [Medline: [28123223](https://pubmed.ncbi.nlm.nih.gov/28123223/)]
18. Sharifi H, Karamouzian M, Baneshi MR, Shokoohi M, Haghdoust A, McFarland W, et al. Population size estimation of female sex workers in Iran: Synthesis of methods and results. *PLoS One* 2017;12(8):e0182755 [FREE Full text] [doi: [10.1371/journal.pone.0182755](https://doi.org/10.1371/journal.pone.0182755)] [Medline: [28796847](https://pubmed.ncbi.nlm.nih.gov/28796847/)]
19. Fearon E, Chabata ST, Thompson JA, Cowan FM, Hargreaves JR. Sample Size Calculations for Population Size Estimation Studies Using Multiplier Methods With Respondent-Driven Sampling Surveys. *JMIR Public Health Surveill* 2017 Sep 14;3(3):e59 [FREE Full text] [doi: [10.2196/publichealth.7909](https://doi.org/10.2196/publichealth.7909)] [Medline: [28912117](https://pubmed.ncbi.nlm.nih.gov/28912117/)]
20. McLaughlin K, Handcock M, Johnston L. UCLA College. 2015. Inference for the Visibility Distribution for Respondent-Driven Sampling URL: http://www.stat.ucla.edu/katherine.mclaughlin/JSMpaper_mclaughlin.pdf [WebCite Cache ID 6vrGceCFF]
21. Handcock M, Fellows I, Gile K. Statistics UCLA. RDS Analyst: Software for the Analysis of Respondent-Driven Sampling Data Version 0 URL: <http://hpmrg.org/> [accessed 2018-12-21] [WebCite Cache ID 74pjbwjny]
22. Johnston LG, McLaughlin KR, El RH, Latifi A, Toufik A, Bennani A, et al. Estimating the Size of Hidden Populations Using Respondent-driven Sampling Data: Case Examples from Morocco. *Epidemiology* 2015 Nov;26(6):846-852 [FREE Full text] [doi: [10.1097/EDE.0000000000000362](https://doi.org/10.1097/EDE.0000000000000362)] [Medline: [26258908](https://pubmed.ncbi.nlm.nih.gov/26258908/)]
23. Handcock MS, Gile KJ, Mar CM. Estimating hidden population size using Respondent-Driven Sampling data. *Electron J Stat* 2014;8(1):1491-1521 [FREE Full text] [doi: [10.1214/14-EJS923](https://doi.org/10.1214/14-EJS923)] [Medline: [26180577](https://pubmed.ncbi.nlm.nih.gov/26180577/)]
24. Kelly CA, Soler-Hampejsek E, Mensch BS, Hewett PC. Social desirability bias in sexual behavior reporting: evidence from an interview mode experiment in rural Malawi. *Int Perspect Sex Reprod Health* 2013 Mar;39(1):14-21 [FREE Full text] [doi: [10.1363/3901413](https://doi.org/10.1363/3901413)] [Medline: [23584464](https://pubmed.ncbi.nlm.nih.gov/23584464/)]
25. Adebajo S, Obianwu O, Eluwa G, Vu L, Oginni A, Tun W, et al. Comparison of audio computer assisted self-interview and face-to-face interview methods in eliciting HIV-related risks among men who have sex with men and men who inject drugs in Nigeria. *PLoS One* 2014;9(1):e81981 [FREE Full text] [doi: [10.1371/journal.pone.0081981](https://doi.org/10.1371/journal.pone.0081981)] [Medline: [24416134](https://pubmed.ncbi.nlm.nih.gov/24416134/)]
26. Ghanem KG, Hutton HE, Zenilman JM, Zimba R, Erbeling EJ. Audio computer assisted self interview and face to face interview modes in assessing response bias among STD clinic patients. *Sex Transm Infect* 2005 Oct;81(5):421-425 [FREE Full text] [doi: [10.1136/sti.2004.013193](https://doi.org/10.1136/sti.2004.013193)] [Medline: [16199744](https://pubmed.ncbi.nlm.nih.gov/16199744/)]

Abbreviations

BBS: biobehavioral surveys
CDC: Centers for Disease Control and Prevention
FSW: female sex workers
KP: key populations
MSM: men who have sex with men
ORG: organizations
PNG: Papua New Guinea

RDS: respondent-driven sampling

SS-PSE: successive sampling-population size estimation

TGW: transgender women

UOM: unique object multiplier

Edited by J Neal; submitted 13.06.18; peer-reviewed by L Johnston, A Gurunga; comments to author 06.08.18; revised version received 27.09.18; accepted 06.10.18; published 21.03.19.

Please cite as:

Weikum D, Kelly-Hanku A, Hou P, Kupul M, Amos-Kuma A, Badman SG, Dala N, Coy KC, Kaldor JM, Vallely AJ, Hakim AJ. Kuantim mi tu ("Count me too"): Using Multiple Methods to Estimate the Number of Female Sex Workers, Men Who Have Sex With Men, and Transgender Women in Papua New Guinea in 2016 and 2017

JMIR Public Health Surveill 2019;5(1):e11285

URL: <http://publichealth.jmir.org/2019/1/e11285/>

doi: [10.2196/11285](https://doi.org/10.2196/11285)

PMID: [30896432](https://pubmed.ncbi.nlm.nih.gov/30896432/)

©Damian Weikum, Angela Kelly-Hanku, Parker Hou, Martha Kupul, Angelyne Amos-Kuma, Steven G Badman, Nick Dala, Kelsey C Coy, John M Kaldor, Andrew J Vallely, Avi J Hakim. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 21.03.2019. 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

Period of Measurement in Time-Series Predictions of Disease Counts from 2007 to 2017 in Northern Nevada: Analytics Experiment

Amir Talaie-Khoei^{1,2*}, PhD; James M Wilson^{3*}, MD, PhD; Seyed-Farzan Kazemi⁴, PhD

¹Department of Information Systems, University of Nevada Reno, Reno, NV, United States

²School of Software, University of Technology Sydney, Sydney, Australia

³Nevada Medical Intelligence Center, School of Community Health Sciences and Department of Pediatrics, University of Nevada Reno, Reno, NV, United States

⁴Center for Research and Education in Advanced Transportation Engineering Systems, Rowan University, Glassboro, NJ, United States

* these authors contributed equally

Corresponding Author:

Amir Talaie-Khoei, PhD

Department of Information Systems

University of Nevada Reno

Ansari Business Building

1664 N Virginia Street, Room 314F

Reno, NV, 89557

United States

Phone: 1 7754407005

Email: atalaeikhoei@unr.edu

Abstract

Background: The literature in statistics presents methods by which autocorrelation can identify the best period of measurement to improve the performance of a time-series prediction. The period of measurement plays an important role in improving the performance of disease-count predictions. However, from the operational perspective in public health surveillance, there is a limitation to the length of the measurement period that can offer meaningful and valuable predictions.

Objective: This study aimed to establish a method that identifies the shortest period of measurement without significantly decreasing the prediction performance for time-series analysis of disease counts.

Methods: The data used in this evaluation include disease counts from 2007 to 2017 in northern Nevada. The disease counts for chlamydia, salmonella, respiratory syncytial virus, gonorrhea, viral meningitis, and influenza A were predicted.

Results: Our results showed that autocorrelation could not guarantee the best performance for prediction of disease counts. However, the proposed method with the change-point analysis suggests a period of measurement that is operationally acceptable and performance that is not significantly different from the best prediction.

Conclusions: The use of change-point analysis with autocorrelation provides the best and most practical period of measurement.

(*JMIR Public Health Surveill* 2019;5(1):e11357) doi:[10.2196/11357](https://doi.org/10.2196/11357)

KEYWORDS

autocorrelation; disease counts; prediction; public health surveillance; time-series analysis

Introduction

Overview

In a time-series prediction for a population, a measurement is the record of equally spaced disease counts over time. The length of these measurements, or equivalently, the interval between records, is the period of measurement [1]. Although time-series

predictions have been widely used in public health surveillance, a body of literature in statistics presents methods by which autocorrelation can detect the best periodicity. Periodicity detection refers to the detection of periodic patterns in a time-series database [2] and can improve the performance of time-series prediction [3-5]. Autocorrelation is a measure of the internal correlation within a time series [4] and a way of measuring and explaining internal association between

observations in a time series. The autocorrelation sequence of a periodic time series has the same cyclic characteristics as the time series itself. Thus, autocorrelation can help verify the presence of cycles and determine their durations [3]. Autocorrelation is often used to identify the best periodicity for time-series analysis [3-5]. This method selects the periodicity in which the autocorrelation is maximized, which would provide time series with better prediction performance [5]. The main aim of this study was to establish a period of measurement in which the autocorrelation is maximized and the periodicity is the interval that its prediction outperforms.

The selection of the period of measurement determines the interval for prediction. Thus, from the operational perspective, for the prediction to be meaningful and valuable in public health surveillance, there has to be a limit to its length. For example, when predicting influenza A cases for the next year, although the 8-week period of measurement may generate the best performance for the prediction, it also produces predicted values that are aggregated 8 weeks at a time. This period is too long to provide any value for practitioners. The 8-week period would cover most of the winter, which is expected to have more influenza A cases even if there is no prediction. For many diseases, particularly infectious diseases that are not disruptive to the healthcare infrastructure (eg, influenza), a 1-week prediction window is generally sufficient. However, for the healthcare infrastructure, a greater prediction window would be helpful to allow planning for potential changes, staffing, or resource allocation. Ultimately, identification of the optimal prediction window allows users to decide what is acceptable for their role in the community.

In response to the operational concern discussed above, this study aims to identify the shortest period of measurement without significantly decreasing the performance. Although autocorrelation provides the best period of measurement, this period may be too long to be practically acceptable. Therefore, we adopted a change-point analysis (CPA) to detect a shorter period of prediction with no change point in between in order to achieve similar performance with a shorter period. To this end, our method aims to apply CPA for autocorrelations of different periods of measurement. The objective is to identify the shortest period of measurement that has an autocorrelation value similar to the maximum value of autocorrelations.

Background and Significance

Public Health Surveillance

The initial target of public health surveillance was infectious diseases; however, with the recent advancements in analytics, data from surveillance systems are increasingly used to predict future trends in a wide range of noninfectious disease distributions. Data have been used for further resource planning and initiating warning systems [6,7]; for example, the Centers for Disease Control and Prevention organized a challenge to predict the 2013-2014 United States influenza season [8]. The ability to accurately forecast various diseases could facilitate key preparedness actions such as the development and use of medical countermeasures, communication strategies, and healthcare resource management [9]. To achieve this goal, different statistical methods have been used to forecast disease

counts; time-series prediction is a method often used in relevant literature [1,10-12], wherein the analysis predicts disease counts by modelling historical surveillance data [1,13]. However, the literature in this area recommends the use of a wide range of methods such as Autoregressive Integrated Moving Average (ARIMA) [14] and structural equation modelling [15].

Time-Series Prediction in Public Health Surveillance

Prior work in time-series prediction of public health surveillance has heavily relied on aberrancy-detection algorithms that are used to detect temporal changes in the data, which may be indicative of a disease outbreak [16]. The Centers for Disease Control and Prevention's Early Aberration Reporting Systems uses C algorithms. In terms of prediction capabilities, C1 only supports moving average with a 7-day window, whereas C2 and C3 offer moving average with a 7-day window and 2-day guard band. Similar to C1-C3, other algorithms [17-19] do not have long-term predictive features that allow public health authorities to achieve annual planning.

These algorithms are primarily designed on the basis of conventional hypothesis testing for the existence of disease outbreak. Aberrancy-detection algorithms only detect changes in static disease activity at a given time when the outbreak occurs and only notice the direction of changes in disease trends at a single time point [20]. However, when the prediction is for an annual disease count rather than a disease outbreak, ARIMA models and machine learning can address the limitation of aberrancy-detection algorithms [21].

ARIMA models are commonly used in public health surveillance [14] and are built on three basic ideas: (1) the present value of time-series is a linear function of its past values and random noise in the AR model [22], (2) the present value of time-series is a linear function of its present and past values of residuals in the moving average model [23], and (3) the AR moving average model [24] considers both the AR and moving average models as well as the historical values and residuals. The ARIMA model generally fits the time-series data based on a previous AR moving average model [24] and includes a differentiating process that effectively transforms nonstationary data required for the abovementioned models into stationary data used in ARIMA [14]. The ARIMA models have been widely used for time-series prediction in public health surveillance [13,25], including hemorrhagic fever with renal syndrome [26,27], dengue fever [28], tuberculosis [29], and mental health [30].

Although methods in conventional statistics are designed to assign most importance to immediate data, they work better with short-term predictions. In addition, these techniques are based on the notion that relationships among the constructs would continue in future, which may not be true [30]. A growing body of literature [31-35] addressed this issue through the use of machine-learning approaches such as Artificial Neural Networks (ANNs) for time-series prediction in public health surveillance. ANNs are inspired by the ways in which biological nervous systems such as the brain process information. It is composed of a large number of highly interconnected processing elements (similar to neurons) working in unison to recognize patterns in data. In addition, ANNs, like people, learn by example.

The ability of ANN to recognize patterns in data allows for better predictions and provides assistance for public health surveillance because it is able to self-organize and self-learn processes [36]. Public health surveillance uses ANN to forecast diseases distributions, whereas Guan et al (2004) used ANN to forecast incidents of hepatitis. Mehra et al (2016) also used ANN to predict the preplanting risk of *Stagonospora nodorum* blotch in winter wheat.

Since this study focuses on forecasting disease counts and limitations of aberrancy-detection algorithms to detect disease outbreak, we only discuss ARIMA and machine learning here.

Period of Measurement for Time-Series Prediction in Public Health Surveillance

Several studies have focused on predicting diseases for public health surveillance through the use of time-series methods such as ARIMA and machine learning. However, it is necessary to recognize that measurement periods play a significant role in the performance of time series, as time-series prediction methods may show different performances for the same population when predicting in different measurement periods [37-39]. For better surveillance of a disease, it is crucial to identify the period of measurement in which the time-series methods demonstrate the best performance for prediction in a particular population.

The performance indicators for time series, such as Q-score [40], can be used to identify the period of measurement that generates the best performance. However, they are computationally expensive to run across multiple time-series analysis for different periods of measurement and compare the performance using the indicator. Therefore, the literature in this field has suggested autocorrelation as one of the most commonly used algorithms to identify the best period of measurement in time series [5]. Autocorrelation refers to the correlation of a time series with its own past and future values [3]. The main objective of this method is to obtain an autocorrelation sequence of a periodic signal with the same cyclic characteristics as the signal itself, allowing autocorrelation to verify the presence of cycles and determine their durations [4]. Therefore, the overall goal is to determine the period of measurement that maximizes the autocorrelation to provide better performance prediction [5].

Although autocorrelation may suggest a periodicity mapped to a period of measurement that is operationally too long to be meaningful, the current study aims to use CPA in order to identify the shortest period of measurement with an autocorrelation value similar to the maximum autocorrelation value. Therefore, we do not expect to see a significant drop in the performance prediction.

Methods

Change-Point Analysis

CPA is exclusively designed to detect subtle changes and characterize changing trends in a time-series [20,41]. The literature has proposed several methods of CPA such as standard normal homogeneity, two-phase regressions with a common trend, and penalized likelihood criteria. In this study, we used

the pruned exact linear time (PELT) CPA method suggested by Killick et al (2012) [42]. This method is based on the CPA method of Jackson et al (2005) [43], but incorporates a pruning step that reduces the computational cost of the method and does not affect the exactness of the resulting segmentation. Although many CPA methods that can only detect the most significant change point, PELT can identify multiple change points. Therefore, owing to its computational performance, this study adopted the PELT method [44]. In addition, we used the R package for CPA [45], which implements PELT. In this algorithm, a change point is defined as the point that characterizes changing trends. As such, the value for the change point is significantly different from the point value immediately before the change point.

The PELT algorithm uses a common approach to detect change points through minimization of costs, which improves the computation performance of PELT. To find multiple change points, the PELT algorithm is first applied to the whole dataset and iteratively and independently to each partition until no further change points are detected. The main assumption of the PELT algorithm is that the numbers of change points increase linearly with the increase in the dataset; the change points are spread throughout the data and are not restricted to one portion of the data [44]. Since we used a small dataset in this study, this assumption is met.

Proposed Method

Our method sorts the autocorrelations based on their period of measurements, wherein the autocorrelation for the shortest period of measurement occupies the first place and the autocorrelation for the longest period of measurement occupies the last place. After conducting CPA using the PELT algorithm on autocorrelations, our method indicates the immediate ascending change point (ACP) before the highest autocorrelations. The autocorrelation of the ACP is the autocorrelation for the shortest period of measurement with similar performance as the highest autocorrelation. Since ACP indicates the closest ACP to the highest autocorrelations, there will be no ACP between the ACP and the highest autocorrelations. This would result in similar performance between the period of measurement associated with the ACP and the period of measurement for the highest autocorrelations. In addition, this will be the shortest period of measurement with similar performance as the highest autocorrelations, because we skip all periods of measurements between the ACP and the highest autocorrelations. As such, the ACP is the shortest period of measurement that has similar performance as the highest autocorrelations.

If the immediate change point before the highest autocorrelations is descending, there is no available period of measurement that is shorter than the highest autocorrelations and has similar performance as the highest autocorrelation. Therefore, the highest autocorrelations indicates the aimed period of measurement. If there is no change point before the highest autocorrelation, we consider the first point as the immediate change point prior to the highest autocorrelation. Figure 1 presents the evaluation of the proposed method.

Data Description

We used the notifiable disease case counts by epidemiological week from 2007 to 2017 in Washoe, Clark, and Carson Counties in Northern Nevada. The data included case counts for chlamydia, salmonella, respiratory syncytial virus (RSV), gonorrhea, viral meningitis, and influenza A. The data were deidentified and included patients of all age. For each disease, the dataset provides the number of reported cases in each epidemiological week. Therefore, for each week between 2007 and 2017, the dataset included all the reported cases of the abovementioned diseases in the three counties, separated according to the diseases.

Training and Test Datasets

The data were divided into training and test datasets in the ratio of 10:1. The scaling guidelines proposed by Guyon [46] were

adopted to identify the size of the training and test sets. The time-series analysis was trained using the dataset created from the data of 2007-2016 and tested on the data for 2017. The performance was subsequently reported.

The original datasets, mentioned in the Data Description section, are measured at 1-week periods. Therefore, the minimum period of measurement was 1 week. However, the study evaluated periods of measurement from 1 to 8 weeks. Depending on the period of measurement, the training and testing sets were aggregated into groups of 1-8 weeks. For example, when we look at the 3-week measurements, the 1-week measurements are aggregated into groups of three. This aggregation starts from week 1. Figure 2 presents the training and testing sets for period measurements.

Figure 1. Evaluation of the proposed method.

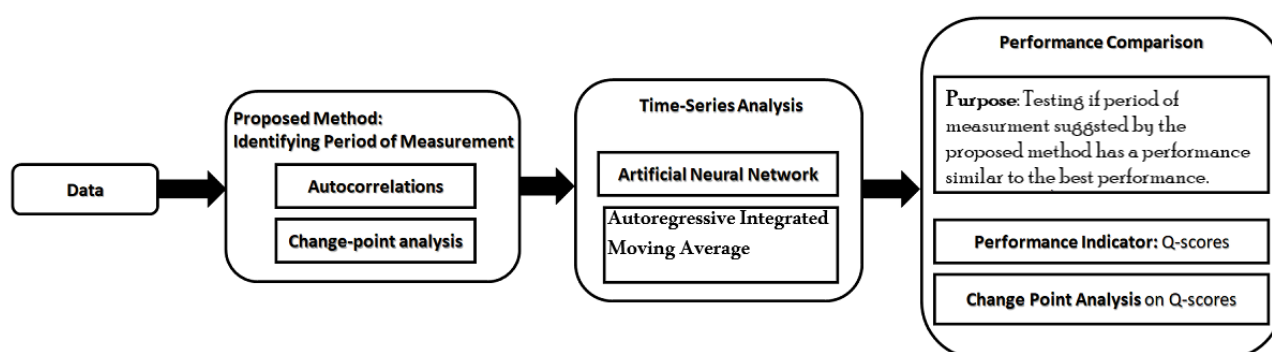
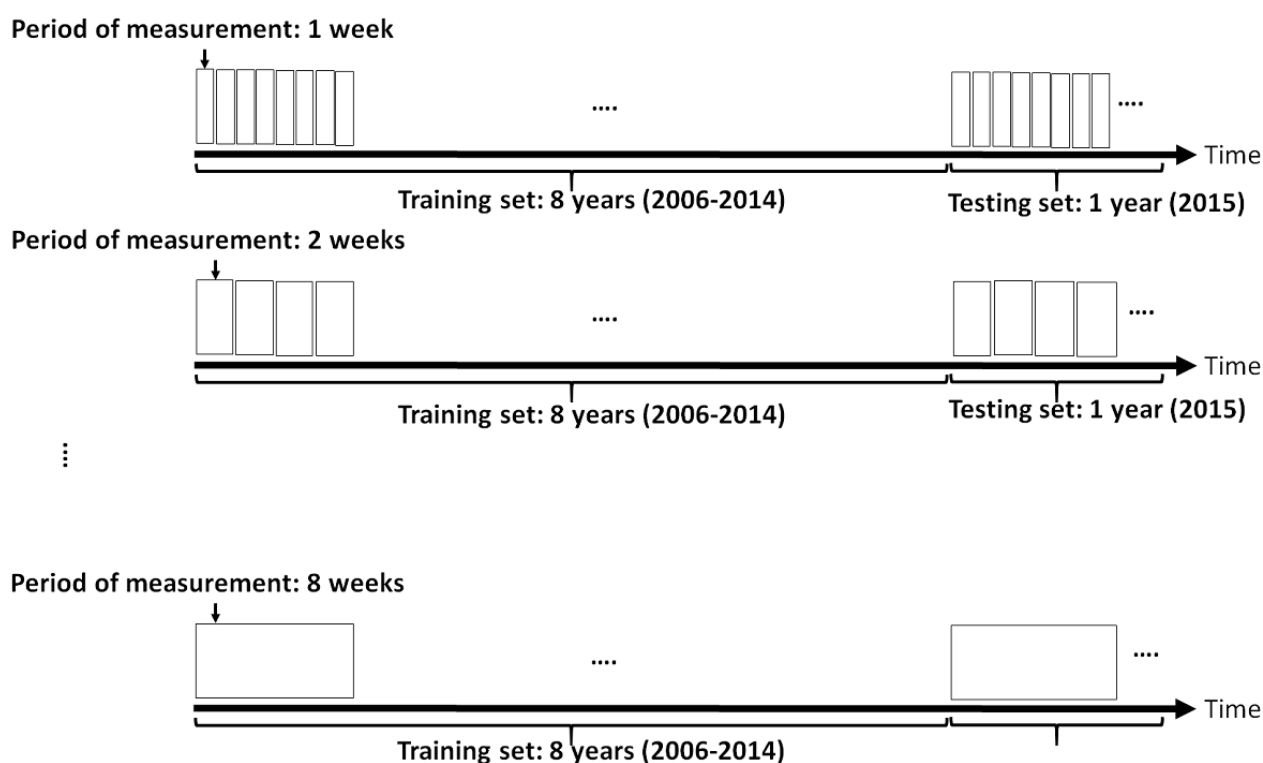


Figure 2. Training and testing sets.



Time-Series Analysis

In order to implement ARIMA, we used `auto.arima` from the R package of `forecast` [47]. Considering the growing body of literature on ANN for public health surveillance [36,48,49], we selected the ANN model for machine learning. Depending on the learning structure, there are many different types of ANNs. In this study, we adopted a feed-forward perceptron-based ANN [50] implemented by the R package CRAN: `nnet` (version, 7.3-5), as it was the most-suitable ANN for our data structure in the preliminary analysis. The parameters were model=multinomial log-linear models: maximum number of iterations=100, fitting=least squares, initial random weights=0.7, maximum allowable number of weights=1000, absolute stop fit criterion= 1.0^{e-4} , relative stop fit criterion= 1.0^{e-8} , size of single hidden layers=11, and weight decay=0.1. These parameters were run for each disease separately, and the predictor variable was time measured by the period of measurement. Figures 3-8 present the performance of ANN and ARIMA.

Performance Indicator: Q-Score

The performance of time-series analysis was measured using the Q-score indicator proposed by Ghil et al (2011) [40]. This indicator treats the data as continuous data, and therefore, the predicted value or observed value can be any positive number in the testing set. Formally, for each disease under the evaluation, we consider the prediction values of $P(t) \in [0, \infty)$ and the observation values of $O(t) \in [0, \infty)$ with integer time $1 \leq t \leq 52$ counting weeks within a year. The overall error of the prediction is quantified by the total squared discrepancy between the prediction values and observed values for the testing set (Figure 9).

To evaluate the performance of prediction, we compared the time-series analysis under evaluation with the unskilled prediction that predicts constant historic average count. This formula is defined in Figure 10.

Figure 3. Evaluation of the proposed method for chlamydia cases. ANN: Artificial Neural Network; ARIMA: Autoregressive Integrated Moving Average; AC: ascending change.

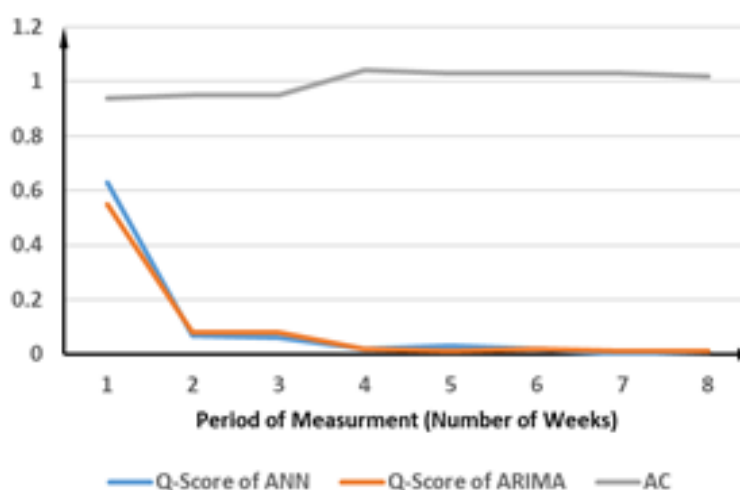
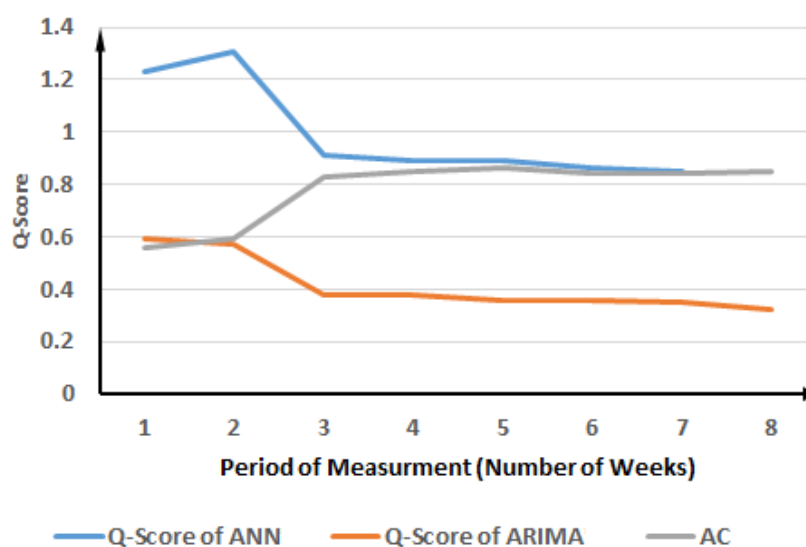


Figure 4. Evaluation of the proposed method for salmonella cases. ANN: Artificial Neural Network; ARIMA: Autoregressive Integrated Moving Average; AC: ascending change.



Finally, the Q-score was defined as the quadratic errors of prediction under evaluation and the unskilled prediction presenting as a constant average. Therefore, the Q-score was defined as presented in Figure 11.

The Q-score may take positive values. It takes Q-score=1 if the time-series prediction under evaluation generates similar results as the unskilled prediction, producing a constant average. A desired time-series analysis produces Q-score=1. Therefore, the aim was to minimize the Q-score.

The Q-score for each period of measurement was calculated for both ARIMA and ANN. Subsequently, a CPA was conducted to determine if the suggested period of measurement generated similar performance as the best performance prediction generating the smallest Q-score with ARIMA and ANN.

This provides a comparative indicator to show the extent to which a method improves unskilled random prediction, which fits our study requirements. The Q-score uses unskilled prediction as a basis and demonstrates how a method outperforms an unskilled prediction. Therefore, the Q-score is suitable for our purpose of comparing methods.

Figure 5. Evaluation of the proposed method for respiratory syncytial virus cases. ANN: Artificial Neural Network; ARIMA: Autoregressive Integrated Moving Average; AC: ascending change.

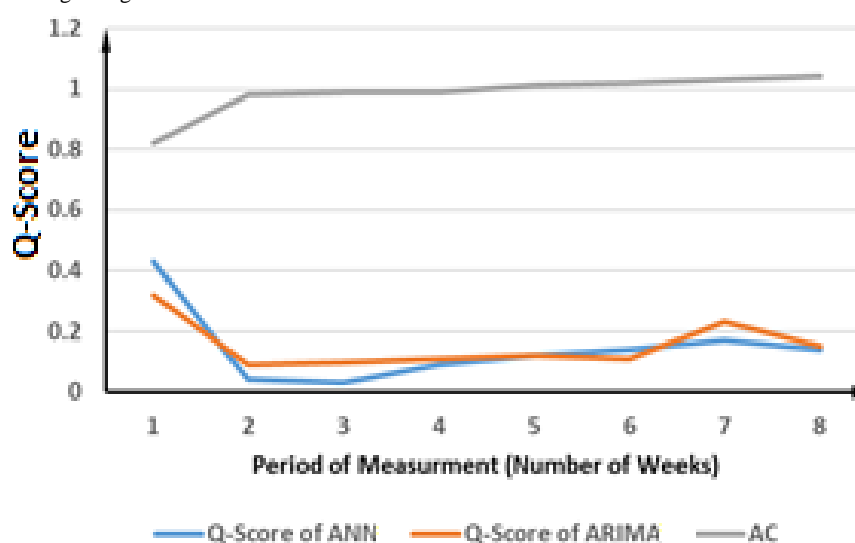


Figure 6. Evaluation of the proposed method for gonorrhea cases. ANN: Artificial Neural Network; ARIMA: Autoregressive Integrated Moving Average; AC: ascending change.

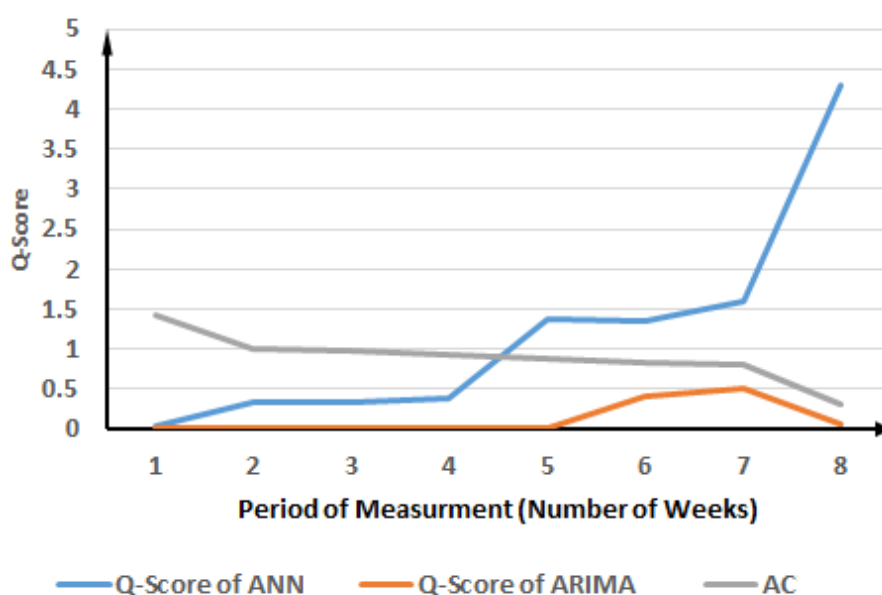


Figure 7. Evaluation of the proposed method for viral meningitis cases. ANN: Artificial Neural Network; ARIMA: Autoregressive Integrated Moving Average; AC: ascending change.

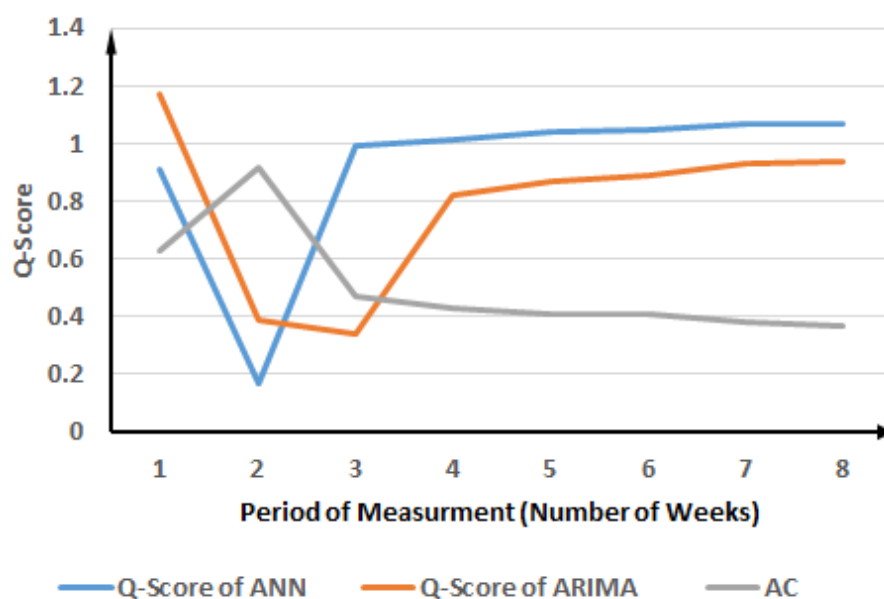


Figure 8. Evaluation of the proposed method for influenza A cases. ANN: Artificial Neural Network; ARIMA: Autoregressive Integrated Moving Average; AC: ascending change.

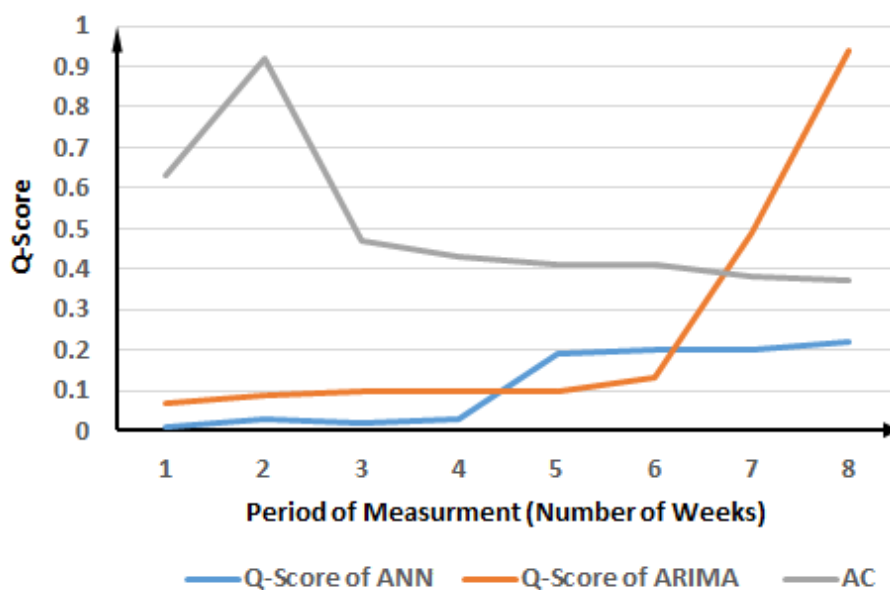


Figure 9. Prediction error.

$$R_{prediction} = \sum_t [P(t) - O(t)]^2$$

Figure 10. Historic average.

$$U(t) = \bar{P} = \frac{1}{n} \sum_{t=1}^{t=n} P(t)$$

Figure 11. Q-Score.

$$Q - Score = \frac{R_{Prediction}}{R_{Unskilled Prediction}} = \frac{\sum_{t=1}^{t=n} [P(t) - O(t)]^2}{\sum_{t=1}^{t=n} [U(t) - O(t)]^2}$$

Results

Figure 3 depicts the evaluation of the proposed method for chlamydia cases. The results show that the proposed method suggests a period of measurement of <3 weeks, which is operationally acceptable. Our result was validated against the performance of ANN and ARIMA, measured by the Q-score (**Table 1**).

Figure 3 and **Table 1** present the evaluation of the proposed method for chlamydia cases. The biggest AC is for the 4-week period of measurement. However, the immediate ACP is in the 2-week period of measurement. Therefore, the autocorrelations are similar in the 2- to 4-week periods of measurements. The proposed method suggests that the 2-week period of measurement yields a good performance, similar to the best performance. The best performance measured by Q-score occurs in 7-week period of measurement for ANN and the 5-week period of measurement for ARIMA. Although there is no ACP, the descending change point (DCP) is in the 2-week period of measurement. As such, performance of ANN and ARIMA remained similar for the 2-week period of measurement or longer. Although the 7-week period of measurement for ANN and 5-week period of measurement for ARIMA provided the best performance, and the smallest Q-scores, our results show that the 2-week period of measurement indicated by our proposed method showed similar performance.

Although our proposed method suggests the 3-week period of measurement for salmonella cases, the best performance occurs in the 8-week period of measurement for both ANN and ARIMA (**Figure 4** and **Table 2**). However, the results of CPs on Q-scores

shows that that 3-week period of measurement generates similar performance as the best Q-scores for ANN and ARIMA. The results for RSV (**Figure 5** and **Table 3**) and gonorrhea cases (**Figure 6** and **Table 4**) validate the proposed method.

Figure 7 and **Table 5** demonstrate an interesting example for viral meningitis. The 2-week period of measurement was suggested by the proposed method, whereas the highest and ACP for AC occurs in the 2-week period of measurement. For ANN, the best performance measured by the Q-score occurs for the 3-week period of measurement; however, the 2-week period of measurement shows a DCP for the Q-scores of ANN. Therefore, the 3-week period of measurement generates similar performance as the 2-week period of measurement suggested by the proposed method. For ARIMA, the best performance occurs in the 2-week period of measurement, which has the DCP as well. The proposed method was also validated for viral meningitis.

Influenza A has attracted a lot of attention from time-series analysis in public health. The biggest AC occurs in the 2-week period of measurement, but the best performance is in the 1-week period of measurement for both ANN and ARIMA. However, there is no change point until the 5-week period of measurement for ANN and the 7-week period of measurement for ARIMA when ACP occurs. Therefore, we can assume that in both ANN and ARIMA, the performance of the 1-week period of measurement with the best Q-score is similar to that of the 2-week period of measurement suggested by the proposed method, because of the biggest AC with the DCP in the 2-week period of measurement (**Figure 8** and **Table 6**). In addition, the proposed method improves the prediction of influenza A.

Table 1. Validation of the proposed method for chlamydia cases against the performance of Artificial Neural Networks and Autoregressive Integrated Moving Average, measured by the Q-score.

Period of measurement (week)	Q-score of the Artificial Neural Networks	Q-score of the Autoregressive Integrated Moving Average	Ascending change
1	0.63	0.55	0.94
2	0.07 (DCP ^a)	0.08 (DCP ^a)	0.95 (ACP ^b)
3	0.06	0.08	0.95
4	0.02	0.02	1.04 ^c
5	0.03	0.01 ^d	1.03
6	0.02	0.02	1.03
7	0 ^d	0.01	1.03
8	0.01	0.01	1.02

^aDCP: descending change point.

^bACP: ascending change point.

^cBiggest ascending change.

^dThe best performance measured by the Q-score for Artificial Neural Networks and Autoregressive Integrated Moving Average.

Table 2. Validation of the proposed method for salmonella cases against the performance of Artificial Neural Networks and Autoregressive Integrated Moving Average, measured by the Q-score.

Period of measurement (week)	Q-score of the Artificial Neural Networks	Q-score of the Autoregressive Integrated Moving Average	Ascending change
1	1.23	0.59	0.56
2	1.31	0.57	0.59
3	0.91 (DCP ^a)	0.38 (DCP ^a)	0.83 (ACP ^b)
4	0.89	0.38	0.85
5	0.89	0.36	0.86 ^c
6	0.86	0.36	0.84
7	0.85	0.35	0.84
8	0.82 ^d	0.32 ^d	0.85

^aDCP: descending change point.^bACP: ascending change point.^cBiggest ascending change.^dThe best performance measured by the Q-score for Artificial Neural Networks and Autoregressive Integrated Moving Average.**Table 3.** Validation of the proposed method for respiratory syncytial virus cases against the performance of Artificial Neural Networks and Autoregressive Integrated Moving Average, measured by the Q-score.

Period of measurement (week)	Q-score of the Artificial Neural Networks	Q-score of the Autoregressive Integrated Moving Average	Ascending change
1	0.43	0.32	0.82 ^a
2	0.04 (DCP ^b)	0.09 ^c (DCP ^b)	0.98 (ACP ^c)
3	0.03 ^d	0.1	0.99
4	0.09	0.11	0.99
5	0.12	0.12	1.01
6	0.14	0.11	1.02
7	0.17	0.23	1.03
8	0.14	0.15	1.04

^aBiggest ascending change.^bDCP: descending change point.^cACP: ascending change point.^dThe best performance measured by the Q-score for Artificial Neural Networks and Autoregressive Integrated Moving Average.

Table 4. Validation of the proposed method for gonorrhea cases against the performance of Artificial Neural Networks and Autoregressive Integrated Moving Average, measured by the Q-score.

Period of measurement (week)	Q-score of the Artificial Neural Networks	Q-score of the Autoregressive Integrated Moving Average	Ascending change
1	0.04 ^a	0.01 ^a	1.42 ^b
2	0.33	0.02	1.01
3	0.34	0.02	0.98
4	0.39	0.02	0.93
5	1.38	0.02	0.88
6	1.36	0.4	0.82
7	1.59	0.5	0.81
8	4.3	0.05	0.31

^aThe best performance measured by the Q-score for Artificial Neural Networks and Autoregressive Integrated Moving Average.

^bBiggest ascending change.

Table 5. Validation of the proposed method for viral meningitis cases against the performance of Artificial Neural Networks and Autoregressive Integrated Moving Average, measured by the Q-score.

Period of measurement (week)	Q-score of the Artificial Neural Networks	Q-Score of the Autoregressive Integrated Moving Average	Ascending change
1	0.91	1.17	0.63
2	0.17 ^a (DCP ^b)	0.39 (DCP ^b)	0.92 ^c (ACP ^d)
3	0.99	0.34 ^a	0.47
4	1.01	0.82	0.43
5	1.04	0.87	0.41
6	1.05	0.89	0.41
7	1.07	0.93	0.38
8	1.07	0.94	0.37

^aThe best performance measured by the Q-score for Artificial Neural Networks and Autoregressive Integrated Moving Average.

^bDCP: descending change point.

^cBiggest ascending change.

^dACP: ascending change point.

Table 6. Validation of the proposed method for influenza A cases against the performance of Artificial Neural Networks and Autoregressive Integrated Moving Average, measured by the Q-score.

Period of measurement (week)	Q-score of the Artificial Neural Networks	Q-score of the Autoregressive Integrated Moving Average	Ascending change
1	0.01 ^a	0.07 ^a	0.63
2	0.03	0.09	0.92 ^b (ACP ^c)
3	0.02	0.10	0.47
4	0.03	0.1	0.43
5	0.19 (ACP ^c)	0.1	0.41
6	0.2	0.13	0.41
7	0.2	0.49 (ACP ^c)	0.38
8	0.22	0.94	0.37

^aThe best performance measured by the Q-score for Artificial Neural Networks and Autoregressive Integrated Moving Average.

^bBiggest ascending change.

^cACP: ascending change point.

Discussion

Following the extensive use of time-series predictions in public health surveillance, autocorrelation is commonly used in statistics to identify the best period of measurement and improve the performance of predictions [3-5]. However, the forecast needs to address the operational perspective and offer meaningful and valuable predictions. Therefore, practitioners in public health surveillance may choose a shorter period of measurement wherein the forecast results may not be as accurate as those of analyses of longer periods of measurements. The literature in statistics shows how the best period of measurement suggested by autocorrelation can improve the performance of a time-series prediction [3-5]. In addition, our empirical results revealed that the most-outperforming period of measurement is not always the shortest one. However, the long periods of measurement that likely provide better prediction performance may not be useful to practitioners because they are too long. We have provided examples of such instances in the Introduction section of the manuscript.

This study proposed a method that runs CPA on autocorrelations and identifies the shortest period of measurement with a performance prediction similar to the best performance prediction. Our method was evaluated against ANN and ARIMA methods for a time-series analysis of disease counts in Clark, Carson, and Washoe Counties in Northern Nevada between 2007 and 2017, including case counts for chlamydia, salmonella, RSV, gonorrhea, viral meningitis, and influenza A.

Unfortunately, autocorrelation cannot guarantee the best performance for disease prediction. For example, for chlamydia, the greatest autocorrelation occurred in the 4-week period of measurement, the best performance of ANN was noted in the 7-week period, and the best performance of ARIMA was observed in the 5-week period of measurement. This was also the case for RSV, gonorrhea, viral meningitis, and influenza A. However, the proposed method adopting CPA suggests that the shortest period of measurement (to satisfy operational

perspective) ensures acceptable performance predictions similar to the best Q-scores.

The current study has two implications for academics. First, the study adds information on the importance of the period of measurement as a factor for providing better disease count forecasts. Second, it demonstrates the application of CPA in providing operationally focused autocorrelation for a more practical period of measurement that not only improves the prediction performance but also generates practical insights.

From a practical perspective, time-series prediction is an important tool for public health and clinical medicine to identify seasonal periods of changes in the relative risk for disease activity. Observed values that exceed predicted parameters do not necessarily reflect a “failed” prediction, but rather, a pattern of reported activity that was not observed in previous data. This is an important adjunct to other methodologies for aberration detection, such as the aforementioned Early Aberration Reporting System. Predictions offer value to the unaware practitioner by offering a “most likely” hypothesis for expected disease activity, which may carry implications for proactive education and disease-control policies.

Although the current study evaluated the proposed method for a variety of diseases, the data were limited to Northern Nevada. Therefore, expanding the datasets and re-evaluating the method with a wider range of diseases from various geographical locations and larger sample sizes would provide a better understanding of the performance prediction of this method. In addition, the proposed method was evaluated against only ARIMA and ANN. This limitation can be addressed in future studies by applying more time-series prediction methods. Although this method uses autocorrelation, Fourier Transforms have been used in the literature to identify the period of measurement [51]. Thus, further research can compare the performance of AC and Fourier Transforms adopted in the method proposed in this study. In addition, the use ARIMA as a predictive model despite its difficulties with periodic prediction has limited the evaluation of our study. However,

the purpose of this study was to compare ANN and ARIMA for their applicability with the method proposed.

Although we chose ARIMA and ANN to demonstrate the performance of the suggested method, researchers in this field are encouraged to use other conventional or machine-learning algorithms to evaluate the performance of this method in future.

The study has potential from a mathematical perspective, since the different time series generated by autocorrelation are mathematical manipulations of the original time series. For example, they could be modelled as reindexed discrete-time stochastic processes. This would open an avenue of future research to mathematically study the behavior of these time series.

The period of measurement plays an important role in the performance of time-series analysis for disease counts. The literature in statistics has been using autocorrelation to identify the outperforming period of measurement. However, in predicting disease counts, long periods may not provide sufficient values for public health and surveillance practitioners. Therefore, we used CPA to find the shortest period of

measurement, which has similar performance as the period identified by AC.

In conclusion, through the adoption of autocorrelation and CPA, we propose a novel method for identifying the period of measurement, which can improve the performance of time-series predictions for disease counts. Our method implements a practical perspective through which we aim to determine the shortest period of measurement that achieves a better prediction performance. This finding makes the method practically applicable in the field when longer periods of prediction, even with better performance, are not operationally valuable to public health professionals. Our method was evaluated against ANN and ARIMA analyses for disease counts of chlamydia, salmonella, RSV, gonorrhea, viral meningitis, and influenza A between 2007 and 2017 in Northern Nevada. Future work should focus on enhancing the evaluation of the method by using more diverse datasets as well as assessing the use of Fourier Transforms instead of AC. Moreover, we encourage researchers to use a wide range of machine learning and alternative CPA methods to improve the suggested approach.

Acknowledgments

The authors gratefully acknowledge Dr Randall Todd and Dr Lei Chen of the Washoe County Health District for sharing the data used in this study.

Conflicts of Interest

None declared.

References

1. Zeger SL, Irizarry R, Peng R. On time series analysis of public health and biomedical data. *Annu Rev Public Health* 2006;27:57-79. [doi: [10.1146/annurev.publhealth.26.021304.144517](https://doi.org/10.1146/annurev.publhealth.26.021304.144517)] [Medline: [16533109](https://pubmed.ncbi.nlm.nih.gov/16533109/)]
2. Hatkar, Shital P. International Journal of Computer Applications Technology and Research. 2014. Analysis of Various Periodicity Detection Algorithms in Time Series Data with Design of New Algorithm URL: <http://ijcat.com/archives/volume3/issue4/ijcatr03041008.pdf> [accessed 2018-12-21] [WebCite Cache ID 74qJ93vG8]
3. Dormann C, McPherson J, Araújo M, Bivand R, Bolliger J, Carl G, et al. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* 2007 Sep 27;30(5):609-628. [doi: [10.1111/j.2007.0906-7590.05171.x](https://doi.org/10.1111/j.2007.0906-7590.05171.x)]
4. McQuillan A, Aigrain S, Mazeh T. Measuring the rotation period distribution of field M dwarfs with Kepler. *Monthly Notices of the Royal Astronomical Society* 2013;432:16. [doi: [10.1093/mnras/stt536](https://doi.org/10.1093/mnras/stt536)]
5. Otunba R, Lin J. KSI Research. 2014. APT: Approximate Period Detection in Time Series URL: https://ksiresearch.org.ipage.com/seke/seke14paper/seke14paper_9.pdf [accessed 2018-12-21] [WebCite Cache ID 74qGQn2hA]
6. Groseclose SL, Buckeridge D. Public Health Surveillance Systems: Recent Advances in Their Use and Evaluation. *Annu Rev Public Health* 2017 Dec 20;38:57-79. [doi: [10.1146/annurev-publhealth-031816-044348](https://doi.org/10.1146/annurev-publhealth-031816-044348)] [Medline: [27992726](https://pubmed.ncbi.nlm.nih.gov/27992726/)]
7. Kandula S, Hsu D, Shaman J. Subregional Nowcasts of Seasonal Influenza Using Search Trends. *J Med Internet Res* 2017 Dec 06;19(11):e370 [FREE Full text] [doi: [10.2196/jmir.7486](https://doi.org/10.2196/jmir.7486)] [Medline: [29109069](https://pubmed.ncbi.nlm.nih.gov/29109069/)]
8. Biggerstaff M, Alper D, Dredze M, Fox S, Fung IC, Hickmann KS, Influenza Forecasting Contest Working Group. Results from the centers for disease control and prevention's predict the 2013-2014 Influenza Season Challenge. *BMC Infect Dis* 2016 Dec 22;16:357 [FREE Full text] [doi: [10.1186/s12879-016-1669-x](https://doi.org/10.1186/s12879-016-1669-x)] [Medline: [27449080](https://pubmed.ncbi.nlm.nih.gov/27449080/)]
9. Chretien J, George D, Shaman J, Chitale RA, McKenzie FE. Influenza forecasting in human populations: a scoping review. *PLoS One* 2014;9(4):e94130 [FREE Full text] [doi: [10.1371/journal.pone.0094130](https://doi.org/10.1371/journal.pone.0094130)] [Medline: [24714027](https://pubmed.ncbi.nlm.nih.gov/24714027/)]
10. Reis BY, Mandl KD. Time series modeling for syndromic surveillance. *BMC Med Inform Decis Mak* 2003 Jan 23;3:2 [FREE Full text] [Medline: [12542838](https://pubmed.ncbi.nlm.nih.gov/12542838/)]
11. Rounds J, Charles-Smith L, Corley C. Soda Pop: A Time-Series Clustering, Alarming and Disease Forecasting Application. *OJPHI* 2017 May 02;9(1):A. [doi: [10.5210/ojphi.v9i1.7582](https://doi.org/10.5210/ojphi.v9i1.7582)]

12. Zhang X, Zhang T, Young AA, Li X. Applications and comparisons of four time series models in epidemiological surveillance data. *PLoS One* 2014;9(2):e88075 [FREE Full text] [doi: [10.1371/journal.pone.0088075](https://doi.org/10.1371/journal.pone.0088075)] [Medline: [24505382](https://pubmed.ncbi.nlm.nih.gov/24505382/)]
13. Gopakumar S, Tran T, Luo W, Phung D, Venkatesh S. Forecasting Daily Patient Outflow From a Ward Having No Real-Time Clinical Data. *JMIR Med Inform* 2016 Jul 21;4(3):e25 [FREE Full text] [doi: [10.2196/medinform.5650](https://doi.org/10.2196/medinform.5650)] [Medline: [27444059](https://pubmed.ncbi.nlm.nih.gov/27444059/)]
14. Zhang G. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 2003 Jan;50:159-175. [doi: [10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)]
15. Abdullah N, Kelly JT, Graham SC, Birch J, Gonçalves-Carneiro D, Mitchell T, et al. Structure-Guided Identification of a Nonhuman Morbillivirus with Zoonotic Potential. *J Virol* 2018 Dec 01;92(23):1248 [FREE Full text] [doi: [10.1128/JVI.01248-18](https://doi.org/10.1128/JVI.01248-18)] [Medline: [30232185](https://pubmed.ncbi.nlm.nih.gov/30232185/)]
16. Murphy SP, Burkom H. Recombinant temporal aberration detection algorithms for enhanced biosurveillance. *J Am Med Inform Assoc* 2008;15(1):77-86 [FREE Full text] [doi: [10.1197/jamia.M2587](https://doi.org/10.1197/jamia.M2587)] [Medline: [17947614](https://pubmed.ncbi.nlm.nih.gov/17947614/)]
17. Reis B, Pagano M, Mandl K. Using temporal context to improve biosurveillance. *Proc Natl Acad Sci U S A* 2003 Feb 18;100(4):1961-1965 [FREE Full text] [doi: [10.1073/pnas.0335026100](https://doi.org/10.1073/pnas.0335026100)] [Medline: [12574522](https://pubmed.ncbi.nlm.nih.gov/12574522/)]
18. Wallenstein S, Naus J. CDC - MMWR. 2004. Scan statistics for temporal surveillance for biologic terrorism URL: <https://www.cdc.gov/MMWR/preview/mmwrhtml/su5301a17.htm> [accessed 2018-12-21] [WebCite Cache ID 74qHP5ZXP]
19. Brillman BJ, Burr T, Forslund D, Joyce E, Picard R, Umland E. Modeling emergency department visit patterns for infectious disease complaints: results and application to disease surveillance. *BMC Med Inform Decis Mak* 2005 Mar 02;5:4 [FREE Full text] [doi: [10.1186/1472-6947-5-4](https://doi.org/10.1186/1472-6947-5-4)] [Medline: [15743535](https://pubmed.ncbi.nlm.nih.gov/15743535/)]
20. Kass-Hout TA, Xu Z, McMurray P, Park S, Buckeridge DL, Brownstein JS, et al. Application of change point analysis to daily influenza-like illness emergency department visits. *J Am Med Inform Assoc* 2012;19(6):1075-1081 [FREE Full text] [doi: [10.1136/amiajnl-2011-000793](https://doi.org/10.1136/amiajnl-2011-000793)] [Medline: [22759619](https://pubmed.ncbi.nlm.nih.gov/22759619/)]
21. Mooney SJ, Pejaver V. Big Data in Public Health: Terminology, Machine Learning, and Privacy. *Annu Rev Public Health* 2018 Apr 01;39:95-112. [doi: [10.1146/annurev-publhealth-040617-014208](https://doi.org/10.1146/annurev-publhealth-040617-014208)] [Medline: [29261408](https://pubmed.ncbi.nlm.nih.gov/29261408/)]
22. Akaike H. Fitting autoregressive models for prediction. *Ann Inst Stat Math* 1969 Dec;21(1):243-247. [doi: [10.1007/BF02532251](https://doi.org/10.1007/BF02532251)]
23. Haining RP. The Moving Average Model for Spatial Interaction. *Transactions of the Institute of British Geographers* 1978;3(2):202. [doi: [10.2307/622202](https://doi.org/10.2307/622202)]
24. Rojas I, Valenzuela O, Rojas F, Guillen A, Herrera L, Pomares H, et al. Soft-computing techniques and ARMA model for time series prediction. *Neurocomputing* 2008 Jan;71(4-6):519-537. [doi: [10.1016/j.neucom.2007.07.018](https://doi.org/10.1016/j.neucom.2007.07.018)]
25. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, Mostashari F, et al. Implementing syndromic surveillance: a practical guide informed by the early experience. *J Am Med Inform Assoc* 2004;11(2):141-150 [FREE Full text] [doi: [10.1197/jamia.M1356](https://doi.org/10.1197/jamia.M1356)] [Medline: [14633933](https://pubmed.ncbi.nlm.nih.gov/14633933/)]
26. Li Q, Guo NN, Han ZY, Zhang YB, Qi SX, Xu YG, et al. Application of an autoregressive integrated moving average model for predicting the incidence of hemorrhagic fever with renal syndrome. *Am J Trop Med Hyg* 2012 Aug;87(2):364-370 [FREE Full text] [doi: [10.4269/ajtmh.2012.11-0472](https://doi.org/10.4269/ajtmh.2012.11-0472)] [Medline: [22855772](https://pubmed.ncbi.nlm.nih.gov/22855772/)]
27. Liu Q, Liu X, Jiang B, Yang W. Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. *BMC Infect Dis* 2011 Aug 15;11:218 [FREE Full text] [doi: [10.1186/1471-2334-11-218](https://doi.org/10.1186/1471-2334-11-218)] [Medline: [21838933](https://pubmed.ncbi.nlm.nih.gov/21838933/)]
28. Wongkoon S, Jaroensutasinee M, Jaroensutasinee K. Development of temporal modeling for prediction of dengue infection in northeastern Thailand. *Asian Pac J Trop Med* 2012 Mar;5(3):249-252 [FREE Full text] [doi: [10.1016/S1995-7645\(12\)60034-0](https://doi.org/10.1016/S1995-7645(12)60034-0)] [Medline: [22305794](https://pubmed.ncbi.nlm.nih.gov/22305794/)]
29. Ríos M, García JM, Sánchez JA, Pérez D. A statistical analysis of the seasonality in pulmonary tuberculosis. *Eur J Epidemiol* 2000 May;16(5):483-488. [Medline: [10997837](https://pubmed.ncbi.nlm.nih.gov/10997837/)]
30. McClellan C, Ali MM, Mutter R, Kroutil L, Landwehr J. Using social media to monitor mental health discussions - evidence from Twitter. *J Am Med Inform Assoc* 2017 May 01;24(3):496-502. [doi: [10.1093/jamia/ocw133](https://doi.org/10.1093/jamia/ocw133)] [Medline: [27707822](https://pubmed.ncbi.nlm.nih.gov/27707822/)]
31. Fisichella M, Stewart A, Denecke K. Unsupervised public health event detection for epidemic intelligence. 2010 Presented at: Proceedings of the 19th ACM International Conference on Information and Knowledge Management; 2010; Boston.
32. Gomide J, Veloso A, Meira JW. Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. 2011 Presented at: Proceedings of the 3rd International Web Science Conference; 2011; Arizona.
33. Maenner MJ, Yeargin-Allsopp M, Van Naarden Braun K, Christensen DL, Schieve LA. Development of a Machine Learning Algorithm for the Surveillance of Autism Spectrum Disorder. *PLoS One* 2016;11(12):e0168224 [FREE Full text] [doi: [10.1371/journal.pone.0168224](https://doi.org/10.1371/journal.pone.0168224)] [Medline: [28002438](https://pubmed.ncbi.nlm.nih.gov/28002438/)]
34. Wang J, McMichael A, Meng B, Becker NG, Han W, Glass K, et al. Spatial dynamics of an epidemic of severe acute respiratory syndrome in an urban area. *Bull World Health Organ* 2006 Dec;84(12):965-968 [FREE Full text] [Medline: [17242832](https://pubmed.ncbi.nlm.nih.gov/17242832/)]
35. Zhou SM, Fernandez-Gutierrez F, Kennedy J, Cooksey R, Atkinson M, Denaxas S, UK Biobank Follow-up Outcomes Group, et al. Defining Disease Phenotypes in Primary Care Electronic Health Records by a Machine Learning Approach: A Case Study in Identifying Rheumatoid Arthritis. *PLoS One* 2016;11(5):e0154515 [FREE Full text] [doi: [10.1371/journal.pone.0154515](https://doi.org/10.1371/journal.pone.0154515)] [Medline: [27135409](https://pubmed.ncbi.nlm.nih.gov/27135409/)]

36. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw* 2015 Jan;61:85-117. [doi: [10.1016/j.neunet.2014.09.003](https://doi.org/10.1016/j.neunet.2014.09.003)] [Medline: [25462637](https://pubmed.ncbi.nlm.nih.gov/25462637/)]
37. Shafia MA, Sadjadi SJ, Jamili A, Tavakkoli-Moghaddam R, Pourseyed-Aghae M. The periodicity and robustness in a single-track train scheduling problem. *Applied Soft Computing* 2012 Jan;12(1):440-452. [doi: [10.1016/j.asoc.2011.08.026](https://doi.org/10.1016/j.asoc.2011.08.026)]
38. Vlachos M, Yu P, Castelli V. On periodicity detection and structural periodic similarity. 2005 Presented at: Proceedings of the SIAM International Conference on Data Mining; 2005; San Francisco.
39. Lu FS, Hou S, Baltrusaitis K, Shah M, Leskovec J, Sosis R, et al. Accurate Influenza Monitoring and Forecasting Using Novel Internet Data Streams: A Case Study in the Boston Metropolis. *JMIR Public Health Surveill* 2018 Jan 09;4(1):e4 [FREE Full text] [doi: [10.2196/publichealth.8950](https://doi.org/10.2196/publichealth.8950)] [Medline: [29317382](https://pubmed.ncbi.nlm.nih.gov/29317382/)]
40. Ghil M, Yiu P, Hallegatte S, Malamud BD, Naveau P, Soloviev A, et al. Extreme events: dynamics, statistics and prediction. *Nonlin. Processes Geophys* 2011 May 18;18(3):295-350. [doi: [10.5194/npg-18-295-2011](https://doi.org/10.5194/npg-18-295-2011)]
41. Kim HJ, Fay M, Feuer E, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med* 2000 Feb 15;19(3):335-351. [Medline: [10649300](https://pubmed.ncbi.nlm.nih.gov/10649300/)]
42. Killick R, Fearnhead P, Eckley IA. Optimal Detection of Changepoints With a Linear Computational Cost. *Journal of the American Statistical Association* 2012 Oct 17;107(500):1590-1598. [doi: [10.1080/01621459.2012.737745](https://doi.org/10.1080/01621459.2012.737745)]
43. Jackson B, Scargle J, Barnes D, Arabhi S, Alt A, Gioumoussis P, et al. An algorithm for optimal partitioning of data on an interval. *IEEE Signal Process. Lett* 2005 Feb;12(2):105-108. [doi: [10.1109/LSP.2001.838216](https://doi.org/10.1109/LSP.2001.838216)]
44. Dorcas Wambui G. The Power of the Pruned Exact Linear Time (PELT) Test in Multiple Changepoint Detection. *AJTAS* 2015;4(6):581. [doi: [10.11648/j.ajtas.20150406.30](https://doi.org/10.11648/j.ajtas.20150406.30)]
45. Killick R, Eckley I. An R Package for Changepoint Analysis. *J. Stat. Soft* 2014;58(3):19. [doi: [10.18637/jss.v058.i03](https://doi.org/10.18637/jss.v058.i03)]
46. Guyon I. AT & T Bell Laboratories Report. US: AT&T; 1997.
47. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. URL: <http://www.r-project.org/> [accessed 2018-12-21] [WebCite Cache ID 74qEKVUVo]
48. Guan P. Forecasting model for the incidence of hepatitis A based on artificial neural network. *WJG* 2004;10(24):3579. [doi: [10.3748/wjg.v10.i24.3579](https://doi.org/10.3748/wjg.v10.i24.3579)]
49. Mehra LK, Cowger C, Gross K, Ojiambo PS. Predicting Pre-planting Risk of Stagonospora nodorum blotch in Winter Wheat Using Machine Learning Models. *Front Plant Sci* 2016;7:390 [FREE Full text] [doi: [10.3389/fpls.2016.00390](https://doi.org/10.3389/fpls.2016.00390)] [Medline: [27064542](https://pubmed.ncbi.nlm.nih.gov/27064542/)]
50. Ncibi K, Sadraoui T, Faycel M. A Multilayer Perceptron Artificial Neural Networks Based a Preprocessing and Hybrid Optimization Task for Data Mining and Classification. *International Journal of Econometrics and Financial Management* 2017;5:21. [doi: [10.12691/ijefm-5-1-3](https://doi.org/10.12691/ijefm-5-1-3)]
51. Marshall AR, Verdun FR. Fourier Transforms in NMR, Optical and Mass Spectrometry: A User's Handbook. New York: Elsevier Science; 1989.

Abbreviations

ACP: ascending change point
ANN: Artificial Neural Network
ARIMA: Autoregressive Integrated Moving Average
CPA: change-point analysis
DCP: descending change point
PELT: pruned exact linear time
RSV: respiratory syncytial virus

Edited by G Eysenbach; submitted 20.06.18; peer-reviewed by D Smiths, J Daniel, F Lu; comments to author 06.10.18; revised version received 23.10.18; accepted 30.10.18; published 15.01.19.

Please cite as:

Talaie-Khoei A, Wilson JM, Kazemi SF

Period of Measurement in Time-Series Predictions of Disease Counts from 2007 to 2017 in Northern Nevada: Analytics Experiment *JMIR Public Health Surveill* 2019;5(1):e11357

URL: <http://publichealth.jmir.org/2019/1/e11357/>

doi: [10.2196/11357](https://doi.org/10.2196/11357)

PMID: [30664479](https://pubmed.ncbi.nlm.nih.gov/30664479/)

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.

Viewpoint

Innovative Approaches to Improve Public Health Practice in the Eastern Mediterranean Region: Findings From the Sixth Eastern Mediterranean Public Health Network Regional Conference

Bashiruddin Noormal¹, PhD; Elmuez Eltayeb², PhD; Mohannad Al Nsour³, PhD; Ezzeddine Mohsni³, PhD; Yousef Khader⁴, PhD; Mark Salter⁵, PhD; Scott McNabb⁶, PhD; Dionisio Herrera Guibert⁷, PhD; Salman Rawaf⁸, PhD; Amrisha Baidjoe⁸, PhD; Aamer Ikram⁹, PhD; Christophe Longuet¹⁰, PhD; Abdulwahed Al Serouri¹¹, PhD; Faris Lami¹², PhD; Asmae Khattabi¹³, PhD; Sami AlMudarra¹⁴, PhD; Ibrahim Iblan¹⁵, MPH; Sahar Samy¹⁶, MPH; Nissaf Bouafif ép Ben Alaya¹⁷, PhD; Qahtan Al-Salihi¹⁸, PhD

¹Ministry of Public Health Afghanistan, Afghanistan, Afghanistan

²Federal Ministry of Health, Khartoum, Sudan

³Global Health Development, Amman, Jordan

⁴Jordan University of Science and Technology, Irbid, Jordan

⁵Public Health England, London, United Kingdom

⁶Emory University, Rollins School of Public Health, North Carolina, NC, United States

⁷Training Programs in Epidemiology and Public Health Interventions Network, Decatur, GA, United States

⁸Imperial College London, London, United Kingdom

⁹National Institute of Health, Islamabad, Pakistan

¹⁰Connecting Organizations for Regional Disease Surveillance, Lyon, France

¹¹Yemen Field Epidemiology Training Program, Sanaa, Yemen

¹²Iraq Field Epidemiology Training Program, Baghdad, Iraq

¹³Ecole Nationale de Santé Publique, Rabat, Morocco

¹⁴Saudi Field Epidemiology Training Program, Riyadh, Saudi Arabia

¹⁵Field Epidemiology Training Program, Amman, Jordan

¹⁶Ministry of Health and Population, Cairo, Egypt

¹⁷National Observatory of New and Emerging Diseases, Tunis, Tunisia

¹⁸Iraq Ministry of Health, Baghdad, Iraq

Corresponding Author:

Yousef Khader, PhD

Jordan University of Science and Technology

Al-Ramtha-Amman Street

Irbid, 22110

Jordan

Phone: 962 0796802040

Email: yskhader@just.edu.jo

Abstract

Public health professionals in the Eastern Mediterranean region (EMR) have limited access to continuing education, including workshops and conferences in public health. Held under the theme *Innovative Approaches: Adapting to the Current EMR Context*, the Eastern Mediterranean Public Health Network (EMPHNET) organized and conducted the *Sixth EMPHNET Regional Conference* from March 26 to 29, 2018. This paper summarizes the key activities including workshops, roundtable discussions, oral and poster presentations, keynote speeches, and side meetings. Before the opening, 5 preconference workshops were held: “Field Epidemiology Training Program (FETP) Accreditation,” “Innovative Public Health Surveillance,” “Human and Animal Brucellosis,” “Rapid Response Teams,” and “Polio Transition and Routine Immunization.” The conference hosted 6 roundtable discussions: “Consolidation of the FETP Network,” “One Health to Achieve Global Health Security,” “Polio Eradication Efforts and Transition Planning for Measles Elimination,” “Mobile Data Collection and Other Innovative Tools to Enhance Decision Making,” “Confronting *Candida auris*: An Emerging Multidrug-resistant Global Pathogen,” and “Functioning and Sustainable Country

Public Health Emergency Response Operation Framework.” One of the conference’s key objectives was to provide a space for FETP residents, graduates, and public health professionals to showcase achievements. A total of 421 abstracts were submitted and after professional review, 34.9% (147/421) were accepted (111 for oral presentations and 36 for poster presentations) and published by Iproceeding. The conference met the primary objectives of showcasing the public health accomplishments and contributions of the EMR, encouraging the exchange of ideas and coordination among stakeholders, and engaging cross-sectoral workforce in producing recommendations for approaching regional and global health concerns. Moreover, the conference presented a unique opportunity for FETPs and other public health professionals from the Mediterranean region to present their significant scientific work and also facilitated networking among professionals. EMPHNET strives to continue to present similar exchange opportunities for public health professionals in the region.

(*JMIR Public Health Surveill* 2019;5(1):e11382) doi:[10.2196/11382](https://doi.org/10.2196/11382)

KEYWORDS

workshops; public health; one health; capacity building

Background

The Eastern Mediterranean region (EMR) has complex and diverse health challenges especially because of the political instability, conflicts and forced displacement, frequent outbreaks of infectious diseases, increasing burden of noncommunicable disease, and poverty [1-3]. This demands public health professionals with a diverse set of skills including medical, behavioral, social, and environmental sciences. Continuing education for public health professionals is also essential to deliver high-quality public health services. However, public health professionals in the EMR have limited access to continuing education [4].

Workshops and conferences provide a critical infusion of public health capital. The Eastern Mediterranean Public Health Network (EMPHNET) holds workshops and conferences, and also supports the Field Epidemiology Training Programs (FETPs) in the EMR (ie, Afghanistan, Egypt, Iraq, Jordan, Morocco, Pakistan, Sudan, Tunisia, Saudi Arabia, and Yemen) [5].

EMPHNET organized and conducted the *Sixth EMPHNET Regional Conference* from March 26 to 29, 2018. It created the space for public health professionals to present accomplishments, share experiences, and provide networking opportunities. The target audiences were public health professionals and policy makers in EMR and regional and international organizations. This manuscript summarizes the key activities of the *Sixth EMPHNET Regional Conference* including workshops, roundtable discussions, oral and poster presentations, keynote speeches, and side meetings.

The Conference Theme and Objectives

The central theme of the Sixth EMPHNET Regional Conference was addressing a variety of innovative approaches to public health and how these approaches can be adapted to the current EMR context. The conference’s main objectives were to create an opportunity for public health professionals from the region to present their accomplishments to a wide range of audience, discuss innovative approaches and solutions in counteracting public health challenges and problems in the EMR countries, and engage members of the public health community in a

dialogue that focuses on reducing the impact of public health problems in the region.

The conference program offered a platform to showcase the scientifically grounded work of FETP residents and graduates, as well as other public health professionals. The abstracts presented at the conference covered a wide range of topics including outbreak investigations for respiratory diseases, outbreak investigations for vaccine preventable diseases, surveillance systems, zoonotic and vector-borne diseases, noncommunicable diseases (NCDs), mental health, child health, hepatitis, and HIV. Presenters included teams of public health experts and FETP graduates and residents from different countries across the EMR including Afghanistan, Algeria, Bahrain, Egypt, Ethiopia, Iraq, Iran, Jordan, Lebanon, Morocco, Oman, Pakistan, Palestine, Qatar, Saudi Arabia, Sudan, Tunisia, United Arab Emirates, Yemen, and other countries.

The preconference workshops were facilitated by experts from the respective fields to highlight public health concepts relevant to the EMR, whereas the roundtable sessions hosted a wide range of expert panelists. Roundtable sessions tackled issues related to the high burden of NCDs, communicable disease outbreaks, emerging and reemerging infections, weak surveillance systems, public health threats in mass gatherings, risks to biosecurity, and public health emergencies.

Oral and Poster Presentations

One of the conference’s key objectives was to provide a space for FETP residents, graduates, and public health professionals to showcase achievements. A total of 421 abstracts were submitted and after being reviewed, 34.9% (147/421) were accepted (111 for oral presentations and 36 for poster presentations; [Table 1](#)) and published by Iproceeding [6]. [Multimedia Appendix 1](#) shows the abstract book. Presentations were given by FETP residents and graduates, as well as other public health professionals from 13 countries (ie, Pakistan, Iraq, Egypt, Yemen, Morocco, Jordan, Saudi Arabia, Lebanon, Bangladesh, Palestine, Tunisia, Oman, and Sudan). Abstracts covered vaccine-preventable diseases, public health service evaluation, food- and waterborne diseases, vector-borne diseases, maternal, child, and reproductive health, respiratory diseases, zoonotic diseases, NCDs, and sexually transmitted infections.

Table 1. The distribution of abstracts and posters that were accepted and presented at the Sixth Eastern Mediterranean Public Health Network Regional Conference according to the abstract track and country of origin.

Distribution of abstracts and posters	Statistics (n)
Track	
Vaccine preventable diseases	26
Surveillance system evaluation	23
Food- and waterborne diseases	17
Vector-borne diseases	11
Maternal, child, and reproductive health	11
Respiratory diseases	15
Zoonotic diseases	9
Noncommunicable diseases	13
Sexually transmitted infection	4
Other	18
Country	
Pakistan	52
Iraq	16
Egypt	13
Yemen	13
Morocco	16
Jordan	9
Afghanistan	4
Sudan	1
Oman	1
Palestine	2
Tunisia	2
Bangladesh	3
Lebanon	4
Saudi Arabia	7
Other	4

Preconference Workshops

A total of 5 preconference workshops were organized: FETP Accreditation, Innovative Public Health Surveillance, Human and Animal Brucellosis, Rapid Response Teams, and Polio Transition and Routine Immunization.

Field Epidemiology Training Program Accreditation Workshop

FETP accreditation by Training Programs in Epidemiology and Public Health Interventions Network (TEPHINET) [7] is a process aimed at improving and maintaining the quality of FETP training and graduates and their involvement and effectiveness in supporting the country's public health priorities. The accreditation process references a set of minimum quality standards, provides momentum for continuous program quality improvement, and facilitates the identification of program needs.

The purpose of the workshop was to share with program directors and coordinators and FETP management staff both the process of becoming accredited and the ways that programs can be improved even if not seeking accreditation status. Specifically, the objectives of the preconference workshop were to review the purpose and benefits of FETP accreditation, describe in detail the eligibility requirements, indicators, and standards for FETP accreditation, and guide the program directors and coordinators to get their programs accredited.

The main recommendation of this workshop was to sustain the FETPs by institutionalization and integration of curricular at academia with the active contribution of countries to guarantee continuous funding and central program coordination and networking. Moreover, the FETP directors and coordinators and management staff in the EMR countries should work to get their FETPs accredited.

Innovative Public Health Surveillance

The emerging field of new technologies and innovative tools has led to an evolution in public health surveillance and epidemic intelligence gathering [8]. Informal or “nontraditional” public health surveillance should enable quicker recognition of outbreaks. In response, TEPHINET—in joint venture with the Skoll Global Threats Fund, HealthMap, and ProMED-mail—developed EpiCore (a system that aims to complement existing surveillance methods and speed up the process of finding, reporting, and verifying public health events) [9]. The regional contributions made through the utilization of the EpiCore platform can create an impact on global health development that is significantly different than the one produced by traditional surveillance.

The objectives of this workshop were to discuss types, functions, challenges, and opportunities of traditional and nontraditional public health surveillance, recognize the importance and use of nontraditional public health surveillance in early detection and response to disease outbreak, and to demonstrate the functions and significance of EpiCore in improving public health surveillance. The workshop highlighted innovative public health surveillance tools to detect and respond to public health events in a more timely manner and raised awareness of EpiCore. All participants agreed on the following specific recommendations: Identify and prioritize various applications, platforms, and data collection devices through carefully listed selection criteria, develop a mobile app for EpiCore to allow easier access to the EpiCore platform and in turn easier and faster reporting, and share the concept of EpiCore with other relevant groups including schools of public health.

Human and Animal Brucellosis

Brucellosis is known to be widespread in Middle Eastern and North African (MENA) countries, yet many of these countries do not have well-established surveillance systems or laboratory capacity to confirm brucellosis in humans or animals. Therefore, true disease burden and circulating species causing illness in humans and animals are not known.

This workshop aimed at improving brucellosis diagnosis and disease management and related practices. Specifically, the main objectives of the workshop were to explore the importance of strengthening the capacity of the public and veterinary health professionals in surveillance, diagnosis, and control of *Brucella* disease in animals and humans; emphasize EMPHNET experience in brucellosis surveillance, diagnosis, and control in Jordan and Iraq; present the laboratory diagnostic methods for brucellosis, advanced technology, and biotyping; discuss brucellosis control strategy in MENA region; and propose contextualized solutions for improving zoonotic diseases management and related best practices with focusing on awareness.

This workshop discussed the impacts of zoonotic diseases and explored collaborative opportunities to work on human and animal health related to brucellosis. It focused on detecting and preventing brucellosis in rural areas where borders are porous and neighboring countries should collaborate. The participants agreed on the following recommendations:

- Strengthen the surveillance and diagnostic capacities of the EMR countries including veterinary public health surveillance and response systems.
- Implement public awareness programs for the public, health workers, home industries, and farmers, by producing printing materials (eg, brochures, posters, and leaflets).
- Develop a national strategic action plan for the control of the major zoonotic diseases in the region. Evidence-based policy framework should be designed to strengthen information systems within each country in the EMR, and methods should be developed for estimating the total economic burden of zoonotic diseases in each country.
- Establish a regional mechanism that can serve the region and overcome the challenges and delayed responses witnessed in some countries such as Yemen.

Rapid Response Teams

The EMR is one of the most vulnerable regions in the world. Furthermore, 10 EMR countries have suffered one or more disasters or emergencies just in the last couple of years, ranging from communicable disease outbreaks to complex emergencies. In 2017, more than 76 million people were directly or indirectly affected by political conflicts, environmental threats, or natural disasters [10]. Public health suffers from prolonged responses and the system itself is under immense pressure. EMPHNET and Global Health Development (GHD) has significant experience in the arena of rapid response including training, coordination, and deployment within and outside the region. EMPHNET has trained an important number of rapid response teams from 7 countries from 2012 onward.

This workshop was conducted to update knowledge and understanding of participants on rapid response, update rapid response procedures and protocols as well as structure and management, and expose participants to real-time information on disasters and complex emergencies at the regional level. Most of the participants agreed on the importance of working under a framework for emergency management. There was an agreement among participants on the importance and need for a regional rapid response team mechanism that can serve the region and overcome the challenges and delayed responses witnessed in some countries such as Yemen. The discussion also highlighted the importance of sustained rapid response team training programs at the national and subnational levels. The participants agreed on and recommended the following:

- Adequate public health emergency response operations plan for all the countries of the region. While developing and revising their public health emergency response plans, countries have to consider functionality, sustainability, and flexibility among the key characteristics of the end product and operate accordingly.
- Secure strong synergies and interactions between the country’s public health emergency operation center, rapid response teams, and the available health care workforce.
- FETPs should be considered (among other opportunities such as academia) as an important asset while building the country’s public health emergency response capacities as well as a mechanism to ensure sustainability of the country’s public health emergency response system.

- Countries have to cost their public health emergency response plans and fund them from their government's basic budget lines to secure functionality and sustainability.

Polio Transition and Routine Immunization

The progress toward the objectives of the eastern Mediterranean regional vaccine action plan from 2016 to 2020, except polio eradication, has been very slow and some of them are estimated to be off-track. The main reason is that a number of countries are facing acute or protracted emergencies.

The objectives of the workshop were to brief participants on the polio transition planning and postcertification strategy, share the country's experiences with progress in developing transitional plans (Somalia and Sudan), discuss the remaining nonendemic countries' plans to develop a transitional plan (Iraq, Syria, and Yemen), focus more on the country's efforts, achievements, and constraints using polio assets to improve routine immunization (all countries), and discuss the best mechanisms to synergize and support priority countries in maintaining polio-free status.

This workshop offered an opportunity for senior immunization staff to interact with senior officers from the global polio eradication initiative partners (World Health Organization Eastern Mediterranean Regional Office, United Nations International Children's Emergency Fund Middle East and North Africa [UNICEF MENA], Centers for Disease Control and Prevention [CDC] Atlanta, and Bill and Melinda Gates

Foundation) and focus on the country's major debated concerns. The workshop recommended the following:

- All countries in the EMR should have a polio transition plan, whether these countries were identified, either globally or regionally, as priority countries or not.
- Increasing expanded program on immunization (EPI) coverage is crucial for sustaining gains for polio eradication.
- Countries should document the experience of polio eradication, its best practices, and how it is being used for improving EPI and other health programs.
- In endemic and high priority countries, transition should be initiated in subnational levels, which have been polio-free for a long time or have relatively stronger health systems than others.
- Orientation on transition and provision of technical support if needed for the staff in nonpriority countries should be taken into consideration.

Roundtable Discussions

The Conference hosted 6 roundtable discussions: Consolidation of the FETP Network, One Health to Achieve Global Health Security, Polio Eradication Efforts and Transition Planning for Measles Elimination, Mobile Data Collection and Other Innovative Tools to Enhance Decision Making, Confronting *Candida auris*: An Emerging Multidrug-resistant Global Pathogen, and Functioning and Sustainable Country Public Health Emergency Response Operation Framework ([Textbox 1](#)).

Textbox 1. Roundtable discussions hosted by the Sixth Eastern Mediterranean Public Health Network Regional Conference.

1. Polio Eradication Efforts and Transition Planning for Measles Elimination:
 - Moderator: Dr Wael Hayejneh, Dean of Faculty of Medicine at Jordan University of Science and Technology
 - Presenter: Dr Dastagir Nazary, National Expanded Program on Immunization manager, Ministry of Health, Afghanistan
 - Panellists:
 - Mr Christopher Maher, Manager, Polio Eradication and Emergency Response, World Health Organization, Eastern Mediterranean Region (EMR)
 - Dr Jay Wenger, Manager, Polio Eradication Program, Bill and Melinda Gates Foundation
 - Dr Noha Farag, Deputy Team Lead for Science and focal point for research with Centers for Disease Control and Prevention's (CDC) Eastern Mediterranean Team in the Polio Eradication Branch
 - Dr Sittana Ahmed, Polio C4D, UNICEF (United Nations International Children's Emergency Fund) Regional office for MENA (Middle East and North Africa)
2. Toward the Consolidation of Field Epidemiology Training Program (FETP) Network in the EMR:
 - Moderator: Dr Mohamed Chahed, Director for the Center of Excellence of Applied Epidemiology, GHD/EMPHNET (Global Health Development/ Eastern Mediterranean Public Health Network)
 - Presenter: Dr Mohannad Al-Nsour, Executive Director, GHD/EMPHNET
 - Panellists:
 - Dr Mohamed Chahed, Director for the Center of Excellence of Applied Epidemiology, GHD/EMPHNET
 - Dr Kip Baggett, Chief of the US CDC's Workforce and Institute Development Branch
 - Dr Dionisio Herrera, Director of the Training Programs in Epidemiology and Public Health Interventions Network
 - Dr Christophe Longuet, Executive Director of Connecting Organizations for Regional Disease Surveillance
 - Dr El Fatih El Semani, Professor at Ahfad University for Women
 - Dr Amrish Baidjoe, President of the European Programme for Intervention Epidemiology Training Alumni Network, currently an Operational Research Associate and Coordinator for outbreak analyses at Imperial College within the R Epidemics Consortium
 - Dana Shalabi, Communication Specialist, GHD/EMPHNET
3. Mobile Data Collection and Other Innovative Tools to Enhance Decision Making:
 - Moderators:
 - Dr Soloman Chandrasegarar, Immunization Specialist, UNICEF MENA Regional Office
 - Scott JN McNabb, Research Professor, Emory University, Rollins School of Public Health
 - Dr Mirwais Amiri, Operational Research Specialist, GHD/EMPHNET
 - Presenters:
 - Dr Moazzem Hossain, Chief of Health and Nutrition in UNICEF Iraq
 - Dr Farida Al Hosani, Manager of Communicable Diseases Department United Arab Emirates (UAE)
 - Panellists:
 - Dr Amr Youssef Ali Kotb, Ethics Committee and FETP Mentor, Egypt
 - Dr Scott Mc Nabb, Research Professor, Emory University, Rollins School of Public Health
4. Confronting *Candida auris*: An Emerging Multidrug-Resistant Global Pathogen:
 - Moderator:
 - Dr Tarek Al-Sanouri
 - Presenters:
 - Snigdha Vallabhaneni, CDC Mycotic Diseases Branch
 - Adnan Alatoom, Consultant physician, Head of Clinical Microbiology, Cleveland Clinic Abu Dhabi, UAE

- Dr Rana Jawad Asghar, Resident Advisor of the Field Epidemiology and Laboratory Training Program (FELTP) in Pakistan
- Dr Mirza Amir Baig, FELTP Pakistan

5. Roundtable on building Functioning and Sustainable Country Public Health Emergency Response Operation Framework:

- Moderator:
 - Dr Ezzeddine Mohsni, Advisor Global Health Security, GHD/EMPHNET
- Presenters:
 - Mahmoud Kaed Yousef, Crisis Management Unit Director, MoH Jordan
 - Dr Elmuez Eltayeb Elnaiem, General Director, Primary Health care General Directorate, Federal Ministry of Health, Sudan
 - Dr Abdel Wahed Al Serour, professor of Community Health, Faculty of Medicine and Health Sciences, Sana'a University, Yemen; and technical advisor of Yemen (FETP)
- Panellists:
 - Sheik Dr Mohammed Bin Hamad Bin J. Al-Thani, Director of Public Health Department. Ministry of Public Health – Qatar; Associate Professor of Clinical Health care Policy and Research at Weill Cornell Medical College - Qatar
 - Dr Mark Salter - Consultant for Global Health
 - Tasha Stehling-Ariza, Epidemiologist with the Global Rapid Response Team at the US CDC

6. One Health to Achieve Global Health Security:

- Moderator:
 - Scott JN McNabb, Research Professor, Emory University, Rollins School of Public Health
- Presenters:
 - Professor Mahmudur Rahman, Former Director Institute of Epidemiology, Disease Control and Research and National Influenza Center, Bangladesh
 - Professor Ahmad Al-Majali, Professor of Infectious Diseases and Epidemiology at the Faculty of Veterinary Medicine, Jordan University for Science and Technology
 - Dr Zahida Fatima, Deputy Director/ Senior Scientific Officer, Pakistan Agricultural Research Council
 - Dr Ekhlas Qasem Hailat, Senior Disease Control Specialist/GHD/EMPHNET

Consolidation of the Field Epidemiology Training Program Network

For the past 10 years, EMPHNET has been providing its support to existing FETPs and supporting the establishment of new FETPs in the region (6 new FETPs). The network is growing over time and the number of FETP graduates is now more than 500, which represents an invaluable core capacity for the countries' health systems. EMPHNET has already undertaken some activities and initiatives to interconnect FETP graduates (via website, social media, and the ambassador initiative). EMPHNET would like to strengthen the consolidation of the network to value field experiences and support the public health systems.

The main objective of this roundtable was to discuss how to secure the emergence of field epidemiology in the EMR based on the lessons learned through the panelist experiences. The roundtable discussion focused on how to reinforce a productive and useful network through the region as a strategic orientation. The panel emphasized the willingness of EMPHNET to mountain and support an FETP network and the big need for

the FETPs in the region to be strongly linked and supported by EMPHNET. The participants recommended to develop an interactive communication with graduates through "EpiShares" to consolidate the EMR FETP network.

One Health to Achieve Global Health Security

One Health is based on the concept that the health of humans, animals, and environment is interconnected [11]. The most effective way to reduce human health threats is to interoperate all health disciplines. Protecting human health by combating outbreaks and preventing international spread remain a universal national priority. Addressing emerging and reemerging diseases (most being zoonotic) as a global health security issue will promote the prevention and rapid detection of novel biological threats, as well as assist in contextualized solutions for the response to these disease threats.

This roundtable aimed to present an overview of priority zoonotic global health security concerns and the importance of the interface of human, animal, and environmental public health sciences; describe regional zoonotic disease prevention, detection, and response using brucellosis as a model; and

contextualize solutions to improve zoonotic disease management through best practices.

The panelists presented an overview of priority zoonotic global health security concerns and suggested different ways to strengthen collaboration between human and veterinary health and laboratory systems to prevent, detect, and respond to emerging and reemerging global health security threats and enhance information exchange. Facilitators agreed that zoonotic diseases represent significant challenges to human health and identified some gaps in both the veterinary and medical sectors, such as insufficient information to build a strategy; lack of training programs for veterinarians, technicians, clinicians, farmers, and laboratory workers; lack of diagnostic capacity that can ensure early detection of zoonotic pathogens; surveillance gaps for important diseases; limited capacity in field epidemiology and rapid field investigations; delays in reporting insufficient preparedness to control epidemics; and lack of monitoring and evaluation and weak coordination between entities such as ministries of agriculture, health, and environment.

The roundtable discussed the way forward to enable, enhance, and empower One Health and participants agreed on the below steps to enhance and empower the One Health approach:

- A national strategic action plan must be launched for the control of major zoonotic diseases.
- An evidence-based policy framework should be designed to strengthen information systems within each country.
- Methods should be developed for estimating the total economic burden of zoonotic diseases in each country.
- Strengthening of veterinary public health surveillance and response systems.
- Adoption of internet-based information technologies to improve disease reporting, facilitate emergency communications, and dissemination of information.
- Implementing effective preventive and therapeutic interventions.
- Implementing awareness programs in collaboration with media partners.
- Strengthening international collaboration and communication.
- Strengthening collaboration and coordination between different entities such as ministries of health, agriculture, and environment.
- Securing political commitment and adequate resources to address underlying socioeconomic factors.

Polio Eradication Efforts and Transition Planning for Measles Elimination

Stopping polio transmission has been possible through improvement in implementing supplementary immunization activities (SIAs), supported by important innovative tools such as the precampaign readiness assessment, strong precampaign monitoring, and drastic postcampaign evaluation. Unlike polio SIAs, most priority countries have not implemented strong measles campaigns. Accordingly, countries continue to experience frequent measles outbreaks.

This roundtable aimed to review the package of assets, innovative tools, mechanisms, and procedures that allowed countries to make the change and achieve high-level performance in polio SIAs and to discuss some “low hanging fruit” solutions that countries and partners can support to allow smooth transition of polio assets in terms of measles immunization campaigns.

Participants agreed that polio SIAs benefit from a high government ownership and commitment translated into a strong accountability framework at all administrative levels of the country, an excellent coordination between implementing and supporting partners, and from a sustained high-quality acute flaccid paralysis surveillance. Moreover, they agreed that implementation of measles elimination programs suffers from low political commitment, lack of accountability, limited number of staff at national levels, and lack of dedicated staff in provinces. There was an agreement among the majority that the challenges of measles campaigns can be addressed through better campaign preparation, coordination, monitoring, and evaluation, as well as ownership and accountability, and that polio infrastructure should be used wherever and whenever possible (without jeopardizing the polio eradication efforts) to support required measles elimination vaccination campaigns that are much less frequent than polio SIAs but cost a lot of money.

The roundtable recommendations included the following:

- The country’s decision makers and stakeholders, as well as national, regional, and global partners, have to multiply efforts to support the implementation of high-performance measles vaccination campaigns.
- All priority countries that succeeded to stop polio transmission should start planning for and using, wherever and whenever relevant, polio assets, tools, and mechanisms to tackle public health priorities, with focus on routine immunization and measles elimination, while securing essential polio functions to sustain polio-free status.
- Routine immunization stakeholders and partners should not hide behind the structural differences between polio and measles to start transitioning polio SIAs assets and mechanisms to improve measles campaigns and surveillance performance.
- Social mobilization mechanisms and tools, such as barriers analysis, community empowerment, communication through front-line mothers, and others, used for polio, should be expanded to incorporate routine immunization as well as measles elimination.
- Where polio and EPI management teams are not integrated, countries and partners should encourage establishing of regular communication, information sharing, and coordination mechanisms to secure optimal utilization of resources, avoid duplication, and synergize efforts and messages for the benefits of both programs.

Mobile Data Collection and Other Innovative Tools to Enhance Decision Making

Mobile data collection (MDC) and other innovative tools can effectively improve the decision-making process by providing on-demand reports and information in the shortest time and

across all levels of the health system. Use of MDC is important in many areas such as minimizing cost of data collection and data entry, minimizing human errors, and eliminating quality issues related to data capturing, management, processing, and utilization [12,13]. MDC tools can improve current health systems by providing accurate results and information for public health practice while supporting the decision-making processes at various levels.

The main objectives of this roundtable were to share and highlight experiences from selected efforts in the region on how MDC or other innovative tools improved data quality and timeliness and discuss the most important challenges or lessons learned. The roundtable raised awareness of the importance and use of MDC in enhancing decision making and focused on programmatic implications and other innovative tools. One of the concerns was related to data security and sensitivity of the use of technology in conflict areas. Therefore, it is recommended to identify and prioritize them through carefully listed selection criteria, researching MDC platform options and considering data needs.

Confronting *Candida auris*: An Emerging Multidrug-Resistant Global Pathogen

Candida auris is an emerging serious threat and causes severe illness in hospitalized patients [14]. *Candida auris* is an emerging global pathogen that poses the problem of multidrug resistance.

This roundtable aimed to raise awareness and increase knowledge about *Candida auris* and discuss concerns and issues related to it as an emerging global pathogen, and discuss strategies to prevent and control infections caused by it. All agreed that accurate identification is important to estimate the prevalence of this under-reported pathogen in different geographic areas. Some isolates are resistant to antifungal drugs which is a concern for getting the attention of the health care community. All agreed that implementation of stringent infection prevention and control for cases along with regular audits for compliance should be undertaken. Participants concluded that there was a need to stay alert and vigilant in monitoring the epidemiology of *Candida auris* globally through strengthening public health surveillance and laboratory capacities to detect and identify the organism.

Functioning and Sustainable Country Public Health Emergency Response Operation Framework

Current turbulences and conflicts in the region affected tremendously at least half of the EMR countries, which conjugated with the increased number of threats from infectious diseases as well as other public health events [15]. Public health emergency response plan is a key component of overall emergency preparedness and response.

While developing the public health emergency response operation framework, countries should take measures that the final “product” is functional, sustainable, and flexible (dynamic and adaptable across time and circumstances). Unfortunately, this has not always been the case in the EMR, where the public health emergency response operation framework is fragmented and not built on available opportunities that can ensure

sustainability with almost no links with the FETP, academia, and even national public health institutes.

This roundtable highlighted the importance for countries in the EMR to consider a comprehensive approach while developing their public health emergency response operation capacities. In addition, this roundtable helped to understand the importance for countries to consider strong links between their rapid response teams, public health emergency operation centers, and health care workforce development opportunities, particularly FETP graduates, while building their emergency response operations system.

Moreover, it highlighted the importance for countries to regularly check the functionality and performance of their public health emergency response operation framework, in particular after each outbreak or with periodic simulations and tabletop exercises, and undertake the required corrections and revisions. Panellists highlighted the fact that the prerequisites for a good preparedness include adequate knowledge about all potential hazards, perfect risk analysis, resource mapping and prioritization process, and develop capacities to prevent occurrence of public health events, reduce impact when they occur, and recover from their impact.

Other Activities

Launch of EpiShares

The conference launched *EpiShares*, a networking platform powered by GHD and EMPHNET. Developed with the aim to increase opportunities for the exchange of knowledge among public health professionals, the platform was developed by GHD and EMPHNET’s team to ensure a mechanism for sharing information and experience and to be a space that can attract public health experts, FETP residents, FETP graduates, or any community of practice. It comes with a host of features including a social media networking platform, the capacity for people with mutual interests to form groups, the capacity for members to start blogs, and a feature listing members within FETP directories or rosters of experts.

Launch of Alumni Association

The conference also launched an alumni association aimed at bringing FETP alumni together from the region to share their rich experiences. The process of launching was discussed in a meeting well-participated by the interested FETP graduates. The agenda included a presentation by the President of the European Programme for Intervention Epidemiology Training (EPIET) Alumni Network, Dr Amrish Baidjoe, who gave an overview of the work done by EPIET and the challenges faced when initiating its launch. The presentation was deemed beneficial, as it allowed prospective FETP alumni members from the EMR to benefit from past experiences gained by their European peers. In his presentation, Dr Amrish stressed that the success of an alumni association was based on several key factors, the most important of which is having a passionate group of core committee members and a membership base willing to see this initiative succeed.

Conclusions

The conference met the primary objectives of showcasing the public health accomplishments and contributions of the EMR, encouraging the exchange of ideas and coordination among stakeholders, and engaging cross-sectoral workforce in producing recommendations for approaching regional and global health concerns. Moreover, the conference presented a unique opportunity for FETPs and other public health professionals

from the Mediterranean region to present their prominent work and network with other international professionals.

At the heart of action on global health security is a commitment to protecting the health of each community and bridging initiatives of all geographical and political regions. Hence, EMPHNET will continue to present similar exchange opportunities for public health professionals in the region. Conference participants and expert panellists provided many insights for moving onward.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The abstracts book.

[[PDF File \(Adobe PDF File\), 4MB - publichealth_v5i1e11382_app1.pdf](#)]

References

- Kennedy J. Why have the majority of recent polio cases occurred in countries affected by Islamist militancy? A historical comparative analysis of the political determinants of polio in Nigeria, Somalia, Pakistan, Afghanistan and Syria. *Med Confl Surviv* 2016 Mar;32(4):295-316. [doi: [10.1080/13623699.2017.1294338](#)] [Medline: [28266226](#)]
- Rahim HF, Sibai A, Khader Y, Hwalla N, Fadhil I, Alsiyabi H, et al. Non-communicable diseases in the Arab world. *Lancet* 2014 Jan 25;383(9914):356-367. [doi: [10.1016/S0140-6736\(13\)62383-1](#)] [Medline: [24452044](#)]
- WHO Regional Office for the Eastern Mediterranean. 2015 Nov. Refugees and internally displaced persons in the Eastern Mediterranean region: a health perspective URL: http://www.emro.who.int/images/stories/eha/documents/migrants_refugees_position_paper.pdf?ua=1 [accessed 2019-01-24] [WebCite Cache ID 75feEgBLR]
- White F. The imperative of public health education: a global perspective. *Med Princ Pract* 2013;22(6):515-529 [FREE Full text] [doi: [10.1159/000354198](#)] [Medline: [23969636](#)]
- The Eastern Mediterranean Public Health Network. URL: http://emphnet.net/?page_id=4334 [accessed 2019-01-24] [WebCite Cache ID 75feN6hXv]
- Iproceedings: iproc. EMPHNET 6th Regional Conference 2018: "Innovative approaches: Adapting to the current EMR context" URL: <https://www.iproc.org/collection/view/650> [accessed 2019-01-24] [WebCite Cache ID 75feVPe13]
- Training Programs in Epidemiology and Public Health Interventions Network. TEPHINET Accreditation of Field Epidemiology Training Programs URL: <http://www.tephinet.org/accreditation> [accessed 2019-01-24] [WebCite Cache ID 75febhl2]
- Groseclose SL, Buckeridge DL. Public health surveillance systems: recent advances in their use and evaluation. *Annu Rev Public Health* 2017 Dec 20;38:57-79. [doi: [10.1146/annurev-publhealth-031816-044348](#)] [Medline: [27992726](#)]
- Haddad Z, Madoff L, Cohn E, Olsen J, Crawley A, Brownstein J, et al. The EpiCore project: using innovative surveillance methods to verify outbreaks of emerging infectious diseases. *Int J Infect Dis* 2016 Apr;45:19. [doi: [10.1016/j.ijid.2016.02.076](#)]
- ReliefWeb. WHO EMRO | Emergencies in the Eastern Mediterranean Region in 2017: the year in review URL: <https://reliefweb.int/report/world/who-emro-emergencies-eastern-mediterranean-region-2017-year-review> [accessed 2019-01-24] [WebCite Cache ID 75fejjGZX]
- One Health Initiative. One Health Initiative will unite human and veterinary medicine URL: <http://onehealthinitiative.com/> [accessed 2019-01-27] [WebCite Cache ID 75kGhcVuD]
- Mukhi SN, Chester TL, Klaver-Kibria JD, Nowicki DL, Whitlock ML, Mahmud SM, et al. Online Journal of Public Health Informatics. 2011 Jun 22. Innovative technology for web-based data management during an outbreak URL: <https://journals.uic.edu/ojs/index.php/ojphi/article/view/3514> [accessed 2019-01-27] [WebCite Cache ID 75kGsmbDK]
- Mukhi S, Dhiravani K, Micholson B, Yan L, Hatchard J, Mubareka S, et al. An innovative mobile data collection technology for public health in a field setting. *Online J Public Health Inform* 2018;10(2):e202 [FREE Full text] [doi: [10.5210/ojphi.v10i2.9114](#)] [Medline: [30349620](#)]
- Tsay S, Welsh RM, Adams EH, Chow NA, Gade L, Berkow EL, et al. Notes from the field: ongoing transmission of *Candida auris* in health care facilities-United States, June 2016-May 2017. *MMWR Morb Mortal Wkly Rep* 2017;66(19):514-515. [doi: [10.15585/mmwr.mm6619a7](#)] [Medline: [28520710](#)]
- WHO Regional Office for the Eastern Mediterranean. Countries in crisis URL: <http://www.emro.who.int/eha/countries-in-crisis/index.html> [accessed 2019-01-27] [WebCite Cache ID 75kHaVotk]

Abbreviations

CDC: Centers for Disease Control and Prevention

EMR: Eastern Mediterranean region

EMPHNET: Eastern Mediterranean Public Health Network

EPI: Expanded Program on Immunization

EPIET: European Programme for Intervention Epidemiology Training

FELTP: Field Epidemiology and Laboratory Training Program

FETP: Field Epidemiology Training Program

GHD: Global Health Development

MDC: mobile data collection

MENA: Middle Eastern and North African

NCDs: noncommunicable diseases

SIAs: supplementary immunization activities

TEPHINET: Training Programs in Epidemiology and Public Health Interventions Network

UNICEF: United Nations International Children's Emergency Fund

Edited by G Eysenbach; submitted 02.07.18; peer-reviewed by Y Acharya, P Amuna; comments to author 15.11.18; revised version received 20.11.18; accepted 09.12.18; published 07.03.19.

Please cite as:

Noormal B, Eltayeb E, Al Nsour M, Mohsni E, Khader Y, Salter M, McNabb S, Herrera Guibert D, Rawaf S, Baidjoe A, Ikram A, Longuet C, Al Serouri A, Lami F, Khattabi A, AlMudarra S, Iblan I, Samy S, Bouafif ép Ben Alaya N, Al-Salihi Q

Innovative Approaches to Improve Public Health Practice in the Eastern Mediterranean Region: Findings From the Sixth Eastern Mediterranean Public Health Network Regional Conference

JMIR Public Health Surveill 2019;5(1):e11382

URL: <http://publichealth.jmir.org/2019/1/e11382/>

doi: [10.2196/11382](https://doi.org/10.2196/11382)

PMID: [30843869](https://pubmed.ncbi.nlm.nih.gov/30843869/)

©Bashiruddin Noormal, Elmuez Eltayeb, Mohannad Al Nsour, Ezzeddine Mohsni, Yousef Khader, Mark Salter, Scott McNabb, Dionisio Herrera Guibert, Salman Rawaf, Amrish Baidjoe, Aamer Ikram, Christophe Longuet, Abdulwahed Al Serouri, Faris Lami, Asmae Khattabi, Sami AlMudarra, Ibrahim Iblan, Sahar Samy, Nissaf Bouafif ép Ben Alaya, Qahtan Al-Salihi. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 07.03.2019. 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

Relationship Between Media Coverage and Measles-Mumps-Rubella (MMR) Vaccination Uptake in Denmark: Retrospective Study

Niels Dalum Hansen¹, PhD; Kåre Mølbak¹, MD, DMSc; Ingemar Johansson Cox^{2,3}, PhD; Christina Lioma², PhD

¹Statens Serum Institut, Copenhagen, Denmark

²Department of Computer Science, University of Copenhagen, Copenhagen, Denmark

³Department of Computer Science, University College London, London, United Kingdom

Corresponding Author:

Niels Dalum Hansen, PhD

Statens Serum Institut

Artillerivej 5

Copenhagen, 2300

Denmark

Phone: 45 32683268

Email: niels.dalum.hansen@gmail.com

Abstract

Background: Understanding the influence of media coverage upon vaccination activity is valuable when designing outreach campaigns to increase vaccination uptake.

Objective: To study the relationship between media coverage and vaccination activity of the measles-mumps-rubella (MMR) vaccine in Denmark.

Methods: We retrieved data on media coverage (1622 articles), vaccination activity (2 million individual registrations), and incidence of measles for the period 1997-2014. All 1622 news media articles were annotated as being provaccination, antivaccination, or neutral. Seasonal and serial dependencies were removed from the data, after which cross-correlations were analyzed to determine the relationship between the different signals.

Results: Most (65%) of the anti-vaccination media coverage was observed in the period 1997-2004, immediately before and following the 1998 publication of the falsely claimed link between autism and the MMR vaccine. There was a statistically significant positive correlation between the first MMR vaccine (targeting children aged 15 months) and provaccination media coverage ($r=.49$, $P=.004$) in the period 1998-2004. In this period the first MMR vaccine and neutral media coverage also correlated ($r=.45$, $P=.003$). However, looking at the whole period, 1997-2014, we found no significant correlations between vaccination activity and media coverage.

Conclusions: Following the falsely claimed link between autism and the MMR vaccine, provaccination and neutral media coverage correlated with vaccination activity. This correlation was only observed during a period of controversy which indicates that the population is more susceptible to media influence when presented with diverging opinions. Additionally, our findings suggest that the influence of media is stronger on parents when they are deciding on the first vaccine of their children, than on the subsequent vaccine because correlations were only found for the first MMR vaccine.

(*JMIR Public Health Surveill* 2019;5(1):e9544) doi:[10.2196/publichealth.9544](https://doi.org/10.2196/publichealth.9544)

KEYWORDS

online news media; vaccination uptake; media influence on vaccination uptake; MMR; autism

Introduction

Reaching all children with two doses of a measles vaccine is an important aim of all national immunization programs. However, many countries have difficulties achieving the

declared aim of measles elimination [1]. Achieving and maintaining measles elimination is possible through a two-dose vaccination program with vaccination coverage of at least 95% for both doses [2]. According to World Health Organization statistics for 2016, only 41/160 (25.6%) countries have a coverage of 95% for the second measles-mumps-rubella (MMR)

vaccine [3]. A 2009 assessment of measles elimination in Europe attributed differences in measles incidence in European countries to the varying degrees of success of the national immunization programs [1] (ie, lower coverage equals higher incidence). Several factors influence the success of an immunization program, including accessibility and availability of vaccination clinics, knowledge about vaccination-preventable diseases, and vaccination cost [4]. In this study, we aim to investigate the relationship between media coverage, the incidence of measles, and the vaccination uptake for the MMR vaccine in Denmark. For our analysis, we use historical data on vaccination and media activity over an 18-year period (January 1, 1997, to December 31, 2014).

The safety of the MMR vaccine became an important topic after 1998 when Wakefield [5] falsely claimed a link between the MMR vaccine and autism. This reduced the public confidence in the vaccine and resulted in a drop in vaccination uptake from above 90% to 79% in England [6,7]. The uptake of the MMR vaccine has also in Denmark been vulnerable to negative media attention. In 1993, the safety of the vaccine was questioned in a nationwide TV program, resulting in record low vaccination coverage [8]. This vulnerability of a vaccination program to public distrust is not limited to the MMR vaccine. Recently the fear of adverse reactions to the human papillomavirus vaccine caused a significant decline in vaccination uptake in Denmark [9]. Hypothesizing that there is a link between media coverage and changes in vaccination coverage is not new [10-13]. However, no study has examined the relationship over an extended period. Understanding the relationship between media coverage and vaccination uptake may underpin the design of public health communication strategies and the development of new surveillance strategies.

In this paper, we take advantage of an 18-year long time series to analyze the correlation between MMR vaccinations, the incidence of measles and media coverage in Denmark. Additionally, we look at the effect of provaccination versus antivaccination media coverage.

Methods

Register Data: Vaccination and Measles Incidence

The MMR vaccination program was introduced in Denmark on January 1, 1987 [14]. The vaccination program consists of 2

vaccinations: 1 targeted at 15-month-old children (MMR-1), and another targeted at 12-year-old children (MMR-2). Since April 1, 2008, the MMR-2 vaccination schedule has changed to target 4-year-old children [15]. Every time a general practitioner vaccinates a child, the date and civil registry number (CRN) of the child are recorded in order for the doctor to receive a reimbursement [16]. These reports are saved in the childhood vaccination database, an immunization information system containing reports from 1997 onwards [16]. Using the CRN, we looked up the birthday of the vaccinated person and calculated the child's age when receiving the vaccine. We separated the registered MMR vaccines into groups based on the recommended vaccination schedule of 15 months, 4 years or 12 years. Each registered vaccine was assigned to the group where the age of the child at vaccination was closest to the target age of the group. We excluded data on the 4-year-old children because they were not represented in the full study period (this corresponds to 374,867 vaccinations). Table 1 shows a summary of the number of registered vaccinations.

We defined vaccination activity as 100 times the number of children vaccinated in a given month divided by the number of eligible children (ie, for MMR-1 the number of children turning 15 months that month). We controlled for yearly and seasonal variations in vaccination activity by dividing by the birth cohort size.

The reported dates of vaccination contain some errors, mainly when doctors report the date of the reimbursement claim instead of the vaccination date. We, therefore, aggregated data on a monthly basis. The top plot in Figure 1 shows the monthly vaccination activity for the 2 vaccines.

To evaluate to what extent MMR vaccination numbers and media coverage about MMR were correlated with the number of measles cases, we also retrieved information about the number of reported measles cases during the study period. Measles is a notifiable disease, and each case is reported to Statens Serum Institut. We aggregated data on a monthly basis, which is shown in Figure 1 (bottom plot). Table 1 shows the total number of reported measles cases in the study period.

Table 1. Summary of the study data. Measles-mumps-rubella (MMR) vaccinations were grouped into MMR-1 (15-month-old children) and MMR-2 (12-year-old children).

Variable	Value
MMR-1	
Number of vaccinations	1,098,389
Age (years) at vaccination, mean (SD) ^a	1.7 (1.4)
MMR-2	
Number of vaccinations	1,108,205
Age (years) at vaccination, mean (SD) ^a	12.3 (1.8)
Reported measles cases	
Number of cases	334
Media coverage^b	
All media, N	1622
National media, n (%)	390 (24.0)
Analysis of all media content	
Relevant to MMR ^c , N	681
Provaccination, n (%)	430 (63.1)
Neutral, n (%)	500 (73.4)
Antivaccination, n (%)	72 (10.6)

^aThe dataset for MMR-1 (targeted 15-month-old children) and MMR-2 (targeted 12-year-old children) only contains age at vaccination in years. This should be considered when interpreting the mean age (SD).

^bMedia coverage is quantified as the number of news items containing MMR related keywords over the 18-year study period.

^cThis row denotes the number of news items that have been labeled as either provaccination, neutral or antivaccination. A news item can get more than one label; hence, the numbers do not sum to 681.

Figure 1. A plot of monthly vaccination activity, media coverage, and measles incidence.

Web-Mined Data: Media Coverage of Measles-Mumps-Rubella

To determine media coverage of MMR, we used the Infomedia archive [17], an online Danish news archive. The archive covers

9 major Danish newspapers, as well as a variety of other news sources. The number of sources indexed is continuously expanding as local newspapers, magazines, news agencies, web media, radio news, and TV news are added to the archive [17].

Radio news and TV news are included in the archive as written summaries.

To measure media coverage related to the MMR vaccine, we constructed a query to retrieve relevant news items from the Infomedia archive (this is standard practice when mining health information from the web [18,19]). The query was designed to have high sensitivity, in other words, most relevant news items should be retrieved. The high sensitivity came with a loss of specificity, since all articles that merely mentioned the MMR vaccine would be retrieved. The query, which we will refer to as the MMR-query, was:

((“mæslinger” OR “mæslinge” OR “fåresyge” OR “røde hunde” OR “mfr”) AND “vaccine”) OR “mæslingevaccine” OR “fåresygevaccine” OR “røde hunde-vaccine OR mfr-vaccine

where “mæslinger” is the Danish word for measles, “fåresyge” means mumps, “røde hunde” means rubella, and “mfr” is the Danish abbreviation for MMR. This query retrieved all news items mentioning “mæslinger” or “mæslinge” (plural or singular) or “fåresyge” or “røde hunde” together with “vaccine,” or news items where either one of the compound phrases (as shown in the second line of the MMR-query) was present. We did not add search terms regarding vaccination, on the assumption that relevant news items will also mention the vaccine. After retrieval, we counted the number of news items returned for this query, for each month of our study period. This type of analysis, which is based on frequency counts, is inspired by computational epidemiology approaches that use web search frequencies to predict health events (eg, influenza-like illness [20], vaccination coverage [21], or antimicrobial drug consumption [22]).

The Infomedia archive has expanded throughout the 18-year study period. In 1997 the archive indexed news items from 20 sources, while in 2014 this number was 1389. As the number of news sources increased, the number of news items added to the archive each month also increased. To accommodate for this change in archive size, we applied 2 sets of frequency counts: (1) 8 major nationwide newspapers that were present in the full duration of the study and (2) all news sources in the archive. We refer to approach (1) as national media and (2) as all media. The middle plot in Figure 1 shows the monthly number of news items retrieved using the MMR-query for each approach, and Table 1 shows the total number of retrieved news items for the 18-year period.

Textbox 1. Annotation criteria for the news items.

Provaccination
<ul style="list-style-type: none"> News items expressing positive views about the vaccine Encouraging people to get vaccinated
Antivaccination
<ul style="list-style-type: none"> News items expressing negative views about the vaccine Discouraging people to get vaccinated
Neutral
<ul style="list-style-type: none"> Neutral information about the vaccine (eg, reports on the number of people vaccinated per year or diseases covered by the vaccine)

Annotation of News Items

The MMR-query was designed with high sensitivity and low specificity. All retrieved news items were subsequently annotated as being relevant to MMR vaccination or irrelevant to improve the specificity. In addition, relevant news items were labeled as having either provaccination, antivaccination, and neutral stance towards the vaccine. The 3 labels are defined in Textbox 1.

Relevant news items were categorized into 1 or more of the three categories. For example, an article with an interview of an antivaccination group accompanied by comments from a doctor explaining the medical reasons and benefits of getting vaccinated would be categorized as both pro and antivaccination. News items whose main focus was not the MMR vaccine (eg, vaccines for pets, annual accounts of vaccination producers, charities for developing countries) would be viewed as irrelevant and would not be categorized.

Data Analysis

The data described above is a time-series (ie, it consists of MMR/media signals that have timestamps). We analyzed this data as follows. First, we removed any seasonality to avoid general seasonal trends biasing the results. Second, we quantified the relationship between the MMR and media (or measles) signals.

Adjusting for Seasonal Correlations

In the analysis, we are not interested in effects due to seasonality. For example, reduced vaccination activity during Christmas. Any seasonality or serial dependencies in the signals were therefore removed by fitting an autoregressive model to the signal and subsequently using the residual of the fitted model. An autoregressive model is defined in Figure 2 where where x is a time series, t is a time point, p is the number of autoregressive terms, the α is the model coefficient, and ϵ_t is the residual at time t .

To quantify seasonality and serial dependencies we calculated the autocorrelation and partial autocorrelation [23] for all signals. Autocorrelation refers to calculating the Pearson correlation (Pearson r) between the signal and a lagged version of itself. The Pearson correlation for 2 time series, x and y , with mean μ and length n is defined in Figure 3.

Figure 2. Autoregressive model for x at time t with p autoregressive terms.

$$X_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i} + \epsilon_t$$

Figure 3. The Pearson r for time series x and y , with mean μ and length n .

$$(X, Y) = \frac{\sum_{i=1}^n (X_i - \mu_X)(Y_i - \mu_Y)}{\sqrt{\sum_{i=1}^n (X_i - \mu_X)^2} \sqrt{\sum_{i=1}^n (Y_i - \mu_Y)^2}}$$

The partial autocorrelation can be used to determine the value of p in Figure 2 because as the partial autocorrelation approaches zero, the value of additional autoregressive terms is reduced. The partial autocorrelation consists of calculating the correlation between the signal x , and a version of itself with a lag of k (ie, x_k), while at the same time controlling for the autocorrelation of the $k-1$ previous lags [23]. The partial autocorrelation at lag k can be calculated by fitting an autoregressive model, as defined in Figure 1, with k autoregressive terms. The value of the k^{th} coefficient (ie, α_k , corresponds to the partial autocorrelation at lag k).

Quantifying the Relationship Between Signals

To quantify the relationship between 2 signals we estimated the cross-correlation. The cross-correlation consists of calculating the Pearson correlation (Figure 3) between 2 signals using different lags. We applied lags between -12 and $+12$ (ie, up to one year before and after). A cross-correlation of 1 means perfect positive correlation, while a correlation of -1 corresponds to perfect negative correlation. To measure the significance of the correlations we treated a series of n cross-correlations as random variables from a student t distribution with degrees of freedom $n-1$. We only reported results for the lags where the correlation was significant (ie, $P < .01$).

Quantifying the Quality of the News Item Annotation

The first author of this paper (NDH) annotated all news items. A random subset of 200 news items was annotated by a second annotator (the second annotator has no medical or computer science background and works as a legal advisor) to assess the reliability of the annotation. The interannotator agreement is quantified by calculating the Cohen kappa coefficient (κ) [24], which measures interannotator reliability while taking into account chance agreement. The coefficient ranges from -1 to 1, with a common interpretation for κ being that <0 is poor agreement, 0 to .20 is slight, .21 to .40 is fair, .41 to .60 is moderate, .61 to .80 is substantial and .81 to 1.00 almost perfect [25]. Because the 3 categories are not mutually exclusive (ie, a news item can be categorized as both neutral and antivaccination) a kappa coefficient is calculated for each label.

Software

The Python packages StatsModels (version 0.8.0) and SciPy (version 0.19.0) were used for calculating autocorrelation, partial autocorrelation, and cross-correlation.

Results

Modeling Expected Variations

Figure 1 (top plot) shows the monthly vaccination activity for MMR-1 and MMR-2. There was a marked periodicity of the number of MMR-2 vaccinations and a visible change in vaccination pattern around 2009. From 1997-2008, inclusive, a reminder letter was sent at the beginning of the year to all children turning 12 that year. The letter was sent at the beginning of each year, and we assume that this was responsible for the annual peak around March. Since May 2014 a reminder letter was sent at age 2, 6.5, and 14 if a child was missing a vaccination [26]. The letters were sent every month, and the effect would, therefore, be spread evenly throughout the year. Figure 4 shows the autocorrelation of vaccination activity, national media (390 articles), all media (1622 articles), and measles incidence. Based on this plot we see that the vaccination activity has a peak at 0, 12, and 24 months. This shows that vaccination activity repeats an annual pattern. For MMR-2 this annual autocorrelation was more pronounced than for MMR-1, likely caused by the pronounced yearly peaks from 1997-2008. For all media, we observed consistent high autocorrelation due to a steady increase in media coverage throughout the study period in the number of retrieved news items. Since this increase was not observed for the national media, the increase is likely explained by the increasing number of media sources in the Infomedia archive, as opposed to generally increased media attention towards the MMR vaccine.

Figure 5 shows the partial autocorrelation for the vaccination activity, media coverage, and measles incidence. The partial autocorrelation reflects the number of autoregressive terms (ie, p in Figure 2) in the autoregressive models used to control for seasonality and serial dependencies. The partial autocorrelation for MMR-1 and MMR-2 quickly drops after the first lag and subsequently peaks again at a 12 months lag. For all media, partial autocorrelation remains close to .2 until a 7-month lag after which it fluctuates around zero.

Based on the observation above we applied an autoregressive model with 12 terms, corresponding to the peak in partial autocorrelation for the vaccination activity. We fit the model both to the vaccination activity and to the media coverage time series. The residual (ie, ϵ_t from Figure 2) will be used in the remaining analysis since this part of the signal is not accounted for by seasonality or serial dependencies.

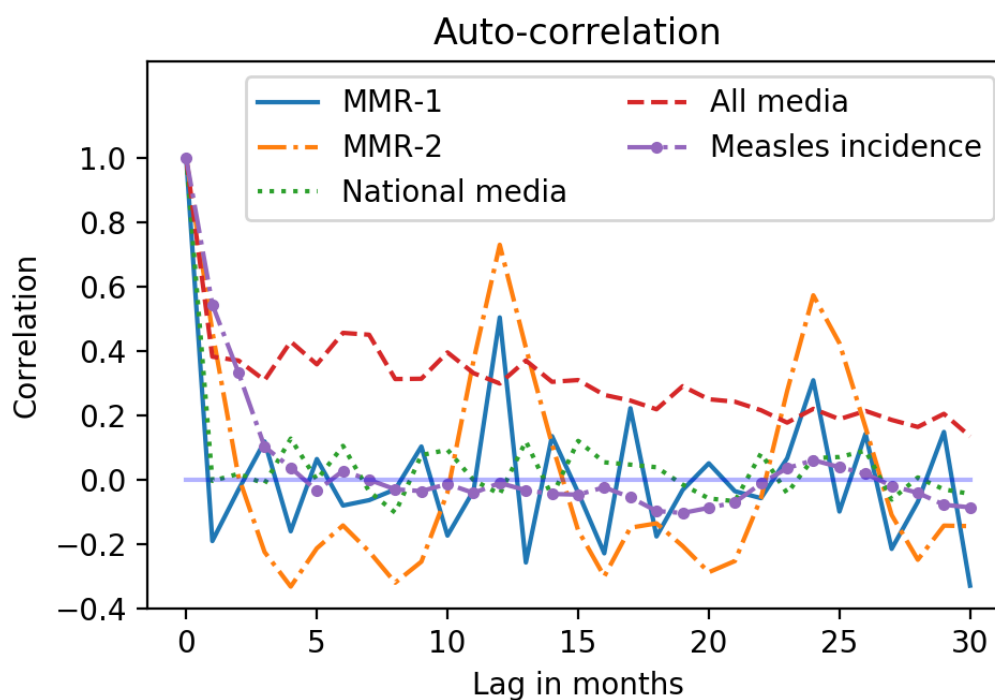
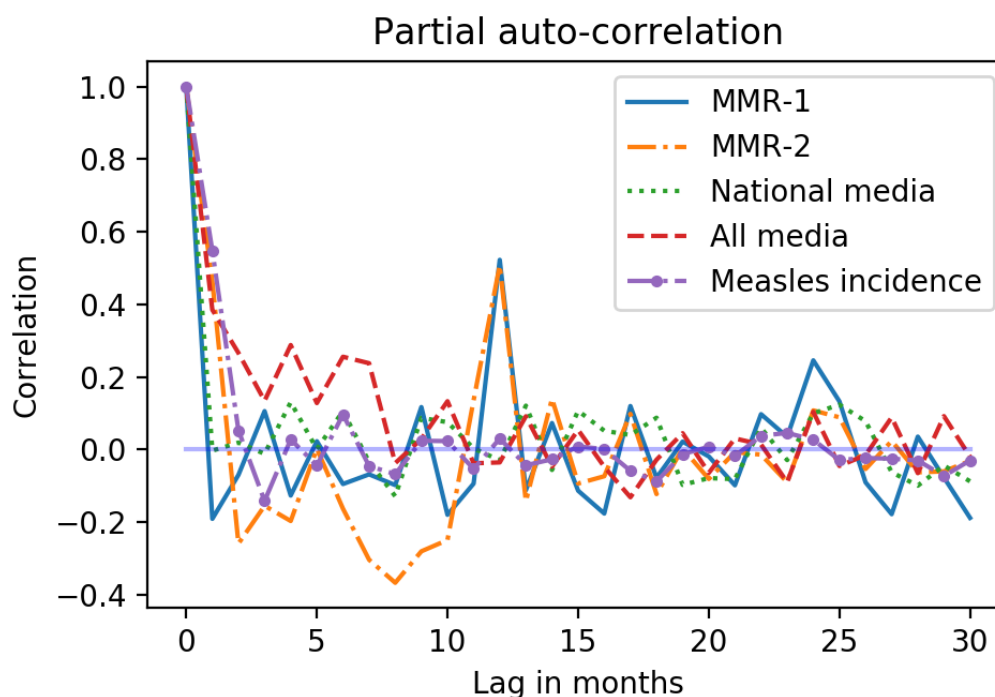
Figure 4. Autocorrelation for MMR-1, MMR-2, national media coverage, all media coverage, and measles incidence. (MMR: measles-mumps-rubella).**Figure 5.** Partial autocorrelation for MMR-1, MMR-2, national media coverage, all media coverage, and measles incidence. (MMR: measles-mumps-rubella).

Figure 5 shows that when an autoregressive model with 12 terms was fitted to the all media signal, then seasonal dependencies are removed. To assess to what extent this was also the case for a general upwards or downwards trend, we fitted a linear model with only a trend term and intercept to the all media signal and the residual of the autoregressive model. For the original signal, the trend was 0.0672 with $P < .001$, while for the residual the

trend was 0.0072 with $P = .32$. In other words, with a 0.0672 monthly increase in the number of news items over 18 years, we would expect to see 14.5 additional news items in the last month of the period compared to the first. While for the residual this increase is only 1.6 news items over an 18-year period. Because controlling for seasonal dependencies also removed the bias from a general upwards trend in the media coverage,

we will, for brevity, in the remainder of the article focus on the results from all media, and disregard the results using national media.

Annotation of News Items

Table 1 shows the number of news items in each category. The results clearly show low specificity of the retrieval method, with only 42.0% (681/1622) of the news items being relevant to the MMR vaccine. Figure 6 shows the distribution of each category during the study period. The peaks in provaccination and neutral information in 2002, 2006, and 2011 correspond to measles outbreaks. The majority of antivaccination news items occurred in the period 1997-2004. This coincided with the retracted study by Wakefield et al [5] published in 1998 that linked autism to the MMR vaccine. The antivaccination news items were primarily about the now falsified link between autism and MMR, but also about Danish court cases on allegations of adverse reactions to the MMR vaccine.

The first author annotated the complete dataset of 1622 news items. A random subset of 200 news items has been annotated by a second annotator, and the interannotator agreement was evaluated using the Cohen kappa coefficient to assess the quality of the annotation. For provaccination, the Cohen kappa coefficient is .54, for neutral it is .35, and for antivaccination, the score is .56. In other words, there is general agreement on provaccination and antivaccination, while less so for neutral.

Relationship Between Media Coverage and Vaccination Activity

For the whole period 1997-2014, we found no significant correlations between vaccination activity and media coverage. This was the case both when we calculated the cross-correlation between MMR-1 vaccination activity and media coverage, and MMR-2 vaccination activity and media coverage, and similarly when using the annotated media data.

Most of the negative media coverage 65.3% (47/72) occurred in the period 1997-2004 (Figure 6). To assess if parents in this period were more susceptible to media influence than in the following period, we separated the dataset into 2: 1998-2004 and 2005-2014 (1997 is omitted because of the 12 months autoregressive models used to control for the seasonal changes and serial dependencies). Stratifying the data on these 2 periods, we found that for the period 1998-2004 there was a small but significant correlation at lag 0 between MMR-1 and all media ($r=.32$, $P=.009$). When using the annotated data, we saw that for the period 1998-2004 there was a statistically significant correlation between provaccination media and MMR-1 vaccination activity ($r=.49$, $P=.004$) and a statistically significant correlation between neutral media and MMR-1 vaccination activity ($r=.45$, $P=.003$). For MMR-2 we observed no significant correlation. Figures 7 and 8 show the cross-correlation at different lags for the 2 periods and 2 vaccines.

Figure 6. Vaccination attitude (stance) in media. For readability, we plotted a 12 months rolling mean. The rolling mean is calculated based on the number of articles published in a window of 6 months before and after a given data point.

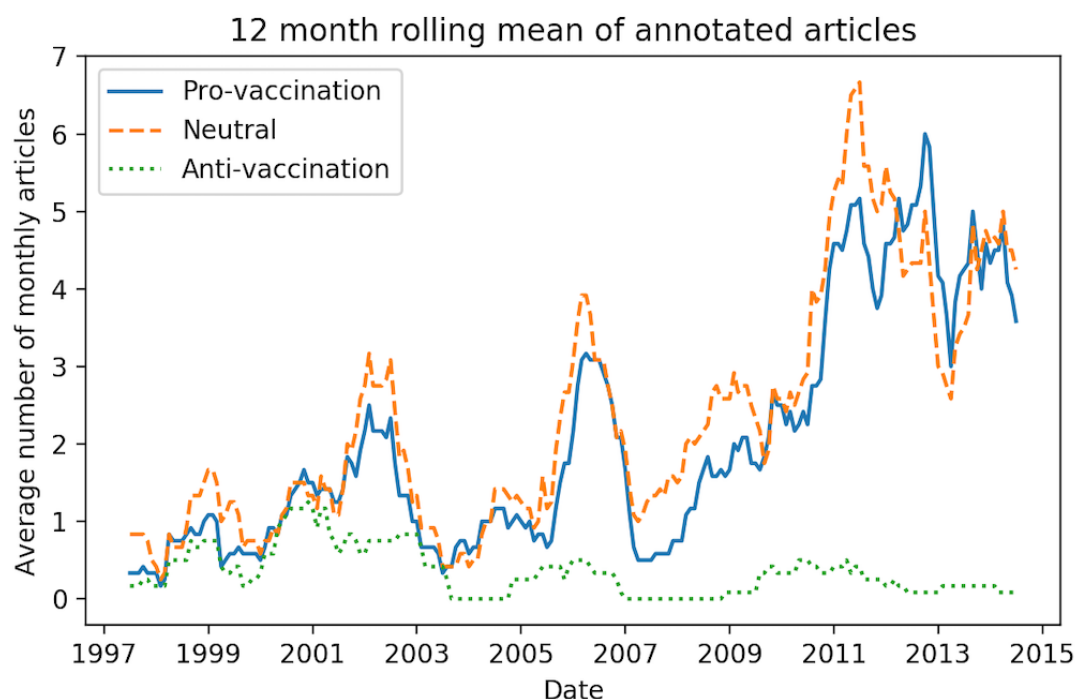
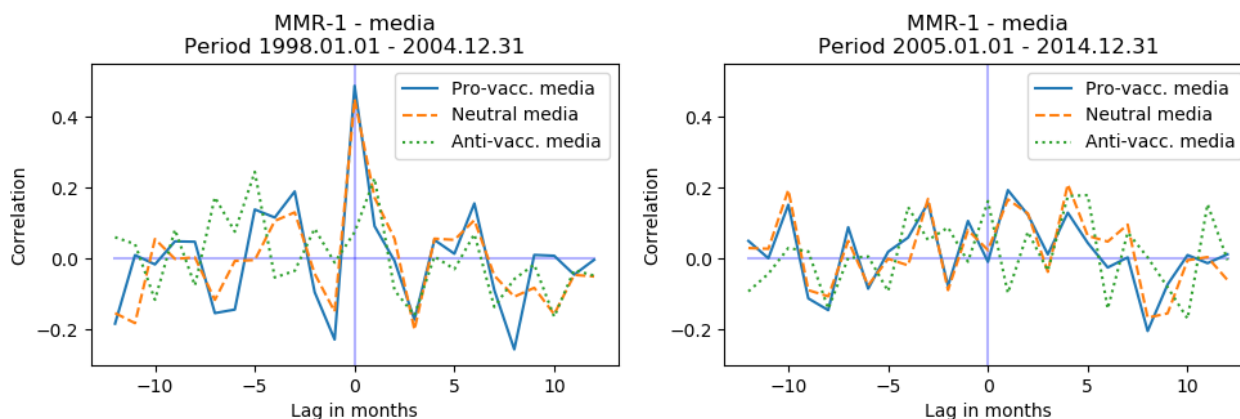
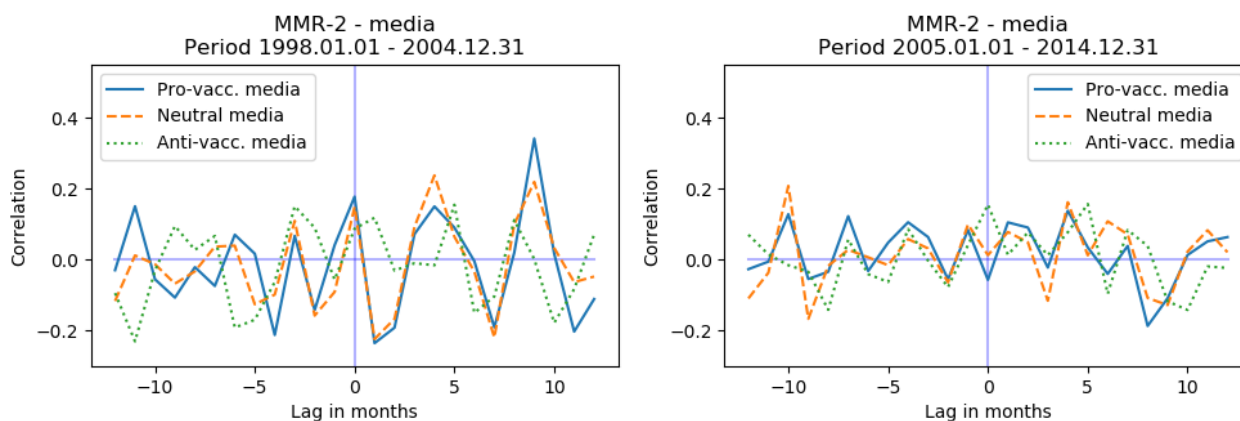


Figure 7. Cross-correlation for vaccination activity of MMR-1 and annotated media data for the 2 periods. (MMR: measles-mumps-rubella).**Figure 8.** Cross-correlation for vaccination activity of MMR-2 and annotated media data for the 2 periods. (MMR: measles-mumps-rubella).

Relationship With Measles Incidence

A possible confounder could be media coverage of measles outbreaks. To quantify this, we have analyzed the cross-correlations between vaccination activity and measles incidences, and between media coverage and measles incidence. We observe that the correlation between measles incidence and MMR-1 ($r=.31$, $P=.005$) was statistically significant at shift 1, meaning that an increase in measles incidence was followed the next month by an increase in MMR-1 vaccinations. For the media data, we found a statistically significant correlation at lag 0 between provaccination media and measles incidence ($r=.38$, $P=.007$). Though not statistically significant, the correlation between neutral media and measles incidence was also relatively high ($r=.35$). We observed no statistical correlations for MMR-2.

Discussion

Principal Results

Our study covered the period 1997-2014 and investigated the relationship between written media coverage and vaccination activity for the MMR vaccine in Denmark. Treating the whole period as 1 time series revealed no relationship between media and vaccination activity. However, the majority of antivaccination media coverage occurred in the beginning of the period (1998-2004). This represents a period where fear of

adverse reactions to the vaccine was high, and the public discourse was tainted by the work of Wakefield [5] and others on the link between autism, as well as other diseases, and the MMR vaccination. During this period there is a statistically significant positive correlation between both provaccination media and vaccination activity for MMR-1 ($r=.49$, $P=.004$), and between neutral media coverage and vaccination activity for MMR-1 ($r=.45$, $P=.003$). In the period 2005-2014 we found no significant correlations. The observed correlations were small, indicating only a limited relationship between media coverage and vaccination activity. Additionally, we only observed the relationships for the first MMR vaccine, targeted the 15-month-old children. This could indicate that parents are more susceptible to media influence when deciding on the first vaccine.

Analysis of the media coverage shows that peaks in provaccination and neutral media coverage often coincided with measles outbreaks. To quantify to what extent measles incidence is a confounder, we calculated the cross-correlation between media coverage and measles incidence. For provaccination media, there was a significant positive correlation of $r=.38$. This shows that there is a temporal relationship, but also that measles incidence does not fully explain the variations in the media coverage.

Strengths and Limitations

The long study period of 18 years strengthens the research because the dynamics between media coverage and vaccination uptake could be studied both in a period with debate and in one without. The Danish vaccination register [16] ensures very reliable vaccination data on a per person level, which allows us to investigate timely changes in the vaccination activity. This is not possible with vaccination uptake data accumulated for each birth cohort.

There are some limitations to the study design. First, not all Danish media have been included, and information about news on radio and television are only present from May 2009 [17]. Additionally, social media have not been analyzed at all. However, we know from other studies on the relationship between social media and news media during a measles outbreak in the Netherlands, that the correlation between social media and news media is very high [27].

Another limitation is the annotation approach, specifically the threshold for when to rate a news item as relevant. Based on the subset of 200 news items annotated by a second annotator, it is evident that the threshold is unclear. This means that conclusions based on the absolute number of news items within a specific category can be questioned. For our analysis, this is not a problem, since we are using the Pearson correlation, which only considers changes relative to the mean. In other words, the information about the absolute number of news items is not used in the analysis.

Finally, it should be noted that we cannot make any statements on causality based on our results. Additionally, the general vaccination activity throughout the period is relatively stable, indicating a priori that external events only have a limited effect on the vaccination activity.

Comparison With Prior Work

There has been previous work on analyzing the effect of media coverage on public behavior. In the 1970s during the US presidential elections, McCombs and Shaw [28] observed a correlation between people's news consumption and their political opinions, which they defined as an agenda-setting effect. The agenda-setting effect depends on the issue at hand. If the issue affects people directly (eg, raising gas prices) the effect will be minimal; however, for more abstract issues (eg, trade deficits or balancing the national budget) the effect will be strong [29]. In our analysis of the media coverage, we saw that measles outbreaks are one of the strong drivers of provaccination and neutral media content, while antivaccination content is driven by fears of adverse reactions. We observe a significant correlation between media coverage and vaccination activity in the period with the most focus on adverse reactions. Analyzing the observation within the context of agenda-setting effects, one explanation could be that the risk of adverse reactions is hard to grasp, and the debate is often filled with discussions of abstract concepts such as relative risks of vaccination versus infection. We hypothesize that because the fear of adverse reactions is hard to relate to everyday life, people are more affected by the media when the discourse is dominated by safety concern, as we saw in the period 1998-2004. Another

explanation for only observing a relationship in the period with a focus on safety maybe because it was a period where opposing views on vaccinations were expressed in the media. Similar observations have been made with respect to political debates [29], where a correlation between media coverage and people's opinions was observed for countries where the politicians did not agree, but no correlation was observed if the politicians agreed. In other words, when the media come to a consensus, their impact vanishes.

Related work more directly comparable to ours shows similar results. The effect of media coverage on vaccination uptake has been studied with respect to the influenza vaccine [12], HPV vaccine [10], and MMR vaccine [11,13]. Smith et al [11] focused on selective MMR nonrecipients, meaning children who received all recommended vaccinations except the MMR vaccine, and concentrated on media related to Wakefield et al [5] and its now discredited link between the MMR vaccine and autism. They concluded that there was a limited influence of mainstream media on MMR vaccinations in the United States. This fits with our results, where we also observe a limited effect. In a study by Mason and Donnelly [13] they compare vaccination uptake in different areas of Wales for the period 1997-1998. They observed a lower vaccination uptake in areas where a series of anti-MMR vaccine articles had been published. Ma et al [30] concluded that media coverage together with recommendations from physicians was associated with increased influenza vaccination coverage in young children. Finally, Kelly et al [10] looked at the relationship between media exposure and knowledge about the HPV vaccine. They found that people exposed to health-related media had more knowledge about HPV than people with less exposure. These results indicate that, to some extent, there is an agenda-setting effect from the media on people's vaccination behavior.

Future Work

Vaccination programs are an essential part of most countries' public health programs, and maintaining a sufficient vaccination coverage is high priority. With disinformation being used as a part of cyberwarfare [31], and the easy spreading of fake news [32,33] surveillance of traditional media and social media is an essential task for public health authorities. Digital media has made the publishing of information easy by both qualified and unqualified persons. The resulting variety of publication outlets of various authority make detailed surveillance an increasingly time-consuming task. One solution to this problem could be automation of the surveillance task. In our study, the crude retrieval method based on only a query showed very low specificity, only 42% (681/1622) of retrieved news items were judged relevant. Manual labeling was required to improve the specificity. This could be work we need to automate. Traditional sentiment detection will likely not suffice, since articles do not necessarily express negative views about the vaccine, but could, for example, emphasize benefits of "natural" immunization (ie, getting infected by measles). A related approach, namely stance detection [34], aims at automatically determining the stance expressed in ideological debates. Such approaches could potentially be used for detecting changes in attitudes expressed in the continuous stream of published media.

An important consideration when continuing the work on media monitoring is to assess to what extent the cost associated with the monitoring corresponds to the potential gain. Could changes in media coverage have been an early indicator of the reduced public trust in the MMR vaccine? And would this signal be strong enough to launch a proactive information campaign, potentially reducing vaccination distrust and the associated costs? The fact that media contains a potential for improved public health communication was illustrated in a study by Bahri et al [35], who showed that active monitoring of the HPV media debate and creation of derived questions could support proactive communication and preparedness. They estimated that the work corresponded to 49% of a full-time position. Extrapolating this to a full vaccination program corresponds to several full-time positions. This raises the question of whether new research within natural language processing, information retrieval, and

machine learning could be used to automate this process and make it accessible at a low cost?

Conclusion

This paper assesses the overall effect of media coverage on the rate of the MMR vaccination in Denmark during the period 1997-2014. The study shows that while for the whole period 1997-2014 there is no correlation between vaccination uptake and media coverage, there is a significant positive correlation in the period 1998-2004 between provaccination and neutral media coverage and vaccination activity for the first MMR vaccine. The period 1998-2004 was characterized by having both provaccination and antivaccination views expressed in the media. The results indicate 2 things: (1) the influence of media is stronger on parents when they are deciding on the first vaccine and (2) the effect of media coverage is stronger when it presents opposing viewpoints.

Acknowledgments

The study was in part supported by a grant from Innovation Fund Denmark grant 4135-00047B and IBM Denmark. NDH did the work while being a PhD student at the University of Copenhagen and IBM Denmark.

Conflicts of Interest

None declared.

References

1. Muscat M, Bang H, Wohlfahrt J, Glismann S, Mølbak K, EUVAC.NET Group. Measles in Europe: an epidemiological assessment. *Lancet* 2009 Jan 31;373(9661):383-389. [doi: [10.1016/S0140-6736\(08\)61849-8](https://doi.org/10.1016/S0140-6736(08)61849-8)] [Medline: [19131097](https://pubmed.ncbi.nlm.nih.gov/19131097/)]
2. WHO Europe. Eliminating Measles and Rubella and Preventing Congenital Rubella Infection. Denmark: World Health Organization; 2005.
3. WHO. Measles, 2nd dose (MCV2) Immunization coverage estimates by country. 2017 Jul 05. Global Health Observatory data repository URL: <http://apps.who.int/gho/data/node.main.MCV2n> [accessed 2018-11-20] [WebCite Cache ID 744t33PhP]
4. Falagas ME, Zarkadoulia E. Factors associated with suboptimal compliance to vaccinations in children in developed countries: a systematic review. *Curr Med Res Opin* 2008 Jun;24(6):1719-1741. [doi: [10.1185/03007990802085692](https://doi.org/10.1185/03007990802085692)] [Medline: [18474148](https://pubmed.ncbi.nlm.nih.gov/18474148/)]
5. Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M, et al. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet* 1998 Feb 28;351(9103):637-641. [Medline: [9500320](https://pubmed.ncbi.nlm.nih.gov/9500320/)]
6. Elliman D, Bedford H. MMR: where are we now? *Arch Dis Child* 2007 Dec;92(12):1055-1057 [FREE Full text] [doi: [10.1136/adc.2006.103531](https://doi.org/10.1136/adc.2006.103531)] [Medline: [17626143](https://pubmed.ncbi.nlm.nih.gov/17626143/)]
7. Ford JA, Mahgoub H, Shankar AG. Vaccine acceptance: the UK perspective. *Hum Vaccin Immunother* 2013 Dec;9(12):2658-2660 [FREE Full text] [doi: [10.4161/hv.26411](https://doi.org/10.4161/hv.26411)] [Medline: [24025731](https://pubmed.ncbi.nlm.nih.gov/24025731/)]
8. Begg N, Ramsay M, White J, Bozoky Z. Media dents confidence in MMR vaccine. *BMJ* 1998 Feb 14;316(7130):561-561. [doi: [10.1136/bmj.316.7130.561](https://doi.org/10.1136/bmj.316.7130.561)]
9. Mølbak K, Hansen ND, Valentiner-Branth P. Pre-Vaccination Care-Seeking in Females Reporting Severe Adverse Reactions to HPV Vaccine. A Registry Based Case-Control Study. *PLoS One* 2016;11(9):e0162520 [FREE Full text] [doi: [10.1371/journal.pone.0162520](https://doi.org/10.1371/journal.pone.0162520)] [Medline: [27611088](https://pubmed.ncbi.nlm.nih.gov/27611088/)]
10. Kelly BJ, Leader AE, Mittermaier DJ, Hornik RC, Cappella JN. The HPV vaccine and the media: how has the topic been covered and what are the effects on knowledge about the virus and cervical cancer? *Patient Educ Couns* 2009 Nov;77(2):308-313 [FREE Full text] [doi: [10.1016/j.pec.2009.03.018](https://doi.org/10.1016/j.pec.2009.03.018)] [Medline: [19395221](https://pubmed.ncbi.nlm.nih.gov/19395221/)]
11. Smith MJ, Ellenberg SS, Bell LM, Rubin DM. Media coverage of the measles-mumps-rubella vaccine and autism controversy and its relationship to MMR immunization rates in the United States. *Pediatrics* 2008 Apr;121(4):e836-e843. [doi: [10.1542/peds.2007-1760](https://doi.org/10.1542/peds.2007-1760)] [Medline: [18381512](https://pubmed.ncbi.nlm.nih.gov/18381512/)]
12. Yoo B, Holland ML, Bhattacharya J, Phelps CE, Szilagyi PG. Effects of mass media coverage on timing and annual receipt of influenza vaccination among Medicare elderly. *Health Serv Res* 2010 Oct;45(5 Pt 1):1287-1309 [FREE Full text] [doi: [10.1111/j.1475-6773.2010.01127.x](https://doi.org/10.1111/j.1475-6773.2010.01127.x)] [Medline: [20579128](https://pubmed.ncbi.nlm.nih.gov/20579128/)]
13. Mason BW, Donnelly PD. Impact of a local newspaper campaign on the uptake of the measles mumps and rubella vaccine. *J Epidemiol Community Health* 2000 Jun;54(6):473-474 [FREE Full text] [Medline: [10818125](https://pubmed.ncbi.nlm.nih.gov/10818125/)]

14. Andersen PH, Valentiner-Branth P, Knudsen LK. MFR-vaccination til børn og voksne. EPI-NYT 2015 Mar 11.
15. Glismann S, Christiansen AH. MFR-2 vaccination fremrykkes til 4 års alderen. EPI-NYT 2008 Feb 27.
16. Grove Krause T, Jakobsen S, Haarh M, Mølbak K. The Danish vaccination register. Euro Surveill 2012 Apr 26;17(17) [FREE Full text] [Medline: 22551494]
17. Infomedia. URL: <https://infomedia.dk/> [accessed 2018-11-27] [WebCite Cache ID 74EvovC1L]
18. Dragusin R, Petcu P, Lioma C, Larsen B, Jørgensen HL, Cox IJ, et al. Specialized tools are needed when searching the web for rare disease diagnoses. Rare Dis 2013;1:e25001 [FREE Full text] [doi: 10.4161/rdis.25001] [Medline: 25002998]
19. Dragusin R, Petcu P, Lioma C, Larsen B, Jørgensen H, Winther O. Rare Disease Diagnosis as an Information Retrieval Task. In: Advances in Information Retrieval Theory.: Giambattista Amati, Fabio Crestani; 2011 Presented at: Conference on the Theory of Information Retrieval; September 12-14, 2011; Bertinoro, Italy p. 356-359 URL: https://link.springer.com/chapter/10.1007/978-3-642-23318-0_38 [doi: 10.1007/978-3-642-23318-0_38]
20. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
21. Hansen ND, Mølbak K, Cox IJ, Lioma C. Time-Series Adaptive Estimation of Vaccination Uptake Using Web Search Queries. In: Proceedings of the 26th International Conference on World Wide Web Companion. 2017 Presented at: International Conference on World Wide Web; April 03 - 07, 2017; Perth, Australia p. 773-774. [doi: 10.1145/3041021.3054251]
22. Hansen N, Cox I, Mølbak K, Lioma C. Predicting antimicrobial drug consumption using web search data. In: Proceedings of the 2018 International Conference on Digital Health. 2018 Apr 26 Presented at: International Conference on Digital Health; April 23 - 26, 2018; Lyon, France. [doi: 10.1145/3194658.3194667]
23. Chatfield C. The Analysis of Time Series An Introduction (Texts in Statistical Science Series). UK: Taylor & Francis, Inc; 2003.
24. Cohen J. A coefficient of agreement for nominal scales. Educational and psychological measurement 1960;20(1):37-46.
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
26. Suppli CH, Rasmussen M, Valentiner-Branth P, Mølbak K, Krause TG. Written reminders increase vaccine coverage in Danish children - evaluation of a nationwide intervention using The Danish Vaccination Register, 2014 to 2015. Euro Surveill 2017 Apr 27;22(17) [FREE Full text] [doi: 10.2807/1560-7917.ES.2017.22.17.30522] [Medline: 28488995]
27. Mollema L, Harmsen IA, Broekhuizen E, Clijnk R, De Melker H, Paulussen T, et al. Disease detection or public opinion reflection? Content analysis of tweets, other social media, and online newspapers during the measles outbreak in The Netherlands in 2013. J Med Internet Res 2015 May 26;17(5):e128 [FREE Full text] [doi: 10.2196/jmir.3863] [Medline: 26013683]
28. McCombs M, Shaw D. The Agenda-Setting Function of Mass Media. The Public Opinion Quarterly 1972;36(2):176-187 [FREE Full text]
29. McCombs M, Reynolds A. How the news shapes our civic agenda. In: Media Effects: Advances in Theory and Research. UK: Taylor & Francis, Inc; 2008.
30. Ma KK, Schaffner W, Colmenares C, Howser J, Jones J, Poehling KA. Influenza vaccinations of young children increased with media coverage in 2003. Pediatrics 2006 Feb;117(2):e157-e163. [doi: 10.1542/peds.2005-1079] [Medline: 16452325]
31. Force EEST. 2018. EU vs disinformationabout/ Internet URL: <https://euvsdisinfo.eu/about/> [accessed 2018-05-02] [WebCite Cache ID 6z7CWaUyx]
32. Allcott H, Gentzkow M. Social Media and Fake News in the 2016 Election. Journal of Economic Perspectives 2017 May;31(2):211-236. [doi: 10.1257/jep.31.2.211]
33. Lioma C, Simonsen J, Larsen B. Evaluation Measures for Relevance and Credibility in Ranked Lists. In: Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval. 2017 Oct 04 Presented at: Conference on Theory of Information Retrieval; October 01 - 04, 2017; Amsterdam, The Netherlands p. 91-98. [doi: 10.1145/3121050.3121072]
34. Hasan KS, Ng V. Stance Classification of Ideological Debates: Data, Models, Features, and Constraints. In: Proceedings of the International Joint Conference on Natural Language Processing. 2013 Oct 18 Presented at: International Joint Conference on Natural Language Processing; 14-18 October 2013; Nagoya, Japan p. 1348-1356.
35. Bahri P, Fogd J, Morales D, Kurz X, ADVANCE consortium. Application of real-time global media monitoring and 'derived questions' for enhancing communication by regulatory bodies: the case of human papillomavirus vaccines. BMC Med 2017 Dec 02;15(1):91 [FREE Full text] [doi: 10.1186/s12916-017-0850-4] [Medline: 28464900]

Abbreviations

CRN: civil registry number

MMR: measles-mumps-rubella

Edited by G Eysenbach; submitted 01.12.17; peer-reviewed by B Ghose, B Worm, M Fadda; comments to author 16.03.18; revised version received 12.08.18; accepted 24.09.18; published 23.01.19.

Please cite as:

Hansen ND, Mølbak K, Cox JJ, Lioma C

Relationship Between Media Coverage and Measles-Mumps-Rubella (MMR) Vaccination Uptake in Denmark: Retrospective Study
JMIR Public Health Surveill 2019;5(1):e9544

URL: <https://publichealth.jmir.org/2019/1/e9544/>

doi: [10.2196/publichealth.9544](https://doi.org/10.2196/publichealth.9544)

PMID: [30672743](https://pubmed.ncbi.nlm.nih.gov/30672743/)

©Niels Dalum Hansen, Kåre Mølbak, Ingemar Johansson Cox, Christina Lioma. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 23.01.2019. 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

Characterizing Trends in Human Papillomavirus Vaccine Discourse on Reddit (2007-2015): An Observational Study

Yuki Lama¹, MPH; Dian Hu², BSc; Amelia Jamison³, MAA, MPH; Sandra Crouse Quinn^{1,3}, PhD; David A Broniatowski², PhD

¹Department of Family Science, School of Public Health, University of Maryland, College Park, MD, United States

²George Washington University, Department of Engineering Management and Systems Engineering, Washington, DC, United States

³Maryland Center for Health Equity, School of Public Health, University of Maryland, College Park, MD, United States

Corresponding Author:

Yuki Lama, MPH

Department of Family Science

School of Public Health

University of Maryland

4200 Valley Drive, Bldg 225

College Park, MD, 20742

United States

Phone: 1 301 405 7574

Fax: 1 301 314 9161

Email: lamay@terpmail.umd.edu

Abstract

Background: Despite the introduction of the human papillomavirus (HPV) vaccination as a preventive measure in 2006 for cervical and other cancers, uptake rates remain suboptimal, resulting in preventable cancer mortality. Social media, widely used for information seeking, can influence users' knowledge and attitudes regarding HPV vaccination. Little is known regarding attitudes related to HPV vaccination on Reddit (a popular news aggregation site and online community), particularly related to cancer risk and sexual activity. Examining HPV vaccine-related messages on Reddit may provide insight into how HPV discussions are characterized on forums online and influence decision making related to vaccination.

Objective: We observed how the HPV vaccine is characterized on Reddit over time and by user gender. Specifically, this study aimed to determine (1) if Reddit messages are more related to cancer risks or sexual behavior and (2) what other HPV vaccine-related discussion topics appear on Reddit.

Methods: We gathered all public Reddit comments from January 2007 to September 2015. We manually annotated 400 messages to generate keywords and identify salient themes. We then measured the similarity between each comment and lists of keywords associated with sexual behavior and cancer risk using Latent Semantic Analysis (LSA). Next, we used Latent Dirichlet Allocation (LDA) to characterize remaining topics within the Reddit data.

Results: We analyzed 22,729 messages containing the strings *hpv* or *human papillomavirus* and *vaccin*. LSA findings show that HPV vaccine discussions are significantly more related to cancer compared with sexual behavior from 2008 to 2015 ($P < .001$). We did not find a significant difference between genders in discussions of cancer and sexual activity ($P > .05$). LDA analyses demonstrated that although topics related to cancer risk and sexual activity were both frequently discussed (16.1% and 14.5% of word tokens, respectively), the majority of online discussions featured other topics. The most frequently discussed topic was politics associated with the vaccine (17.2%). Other topics included HPV disease and/or immunity (13.5%), the HPV vaccine schedule (11.5%), HPV vaccine side effects (9.7%), hyperlinks to outside sources (9.1%), and the risks and benefit of HPV vaccination (8.5%).

Conclusions: Reddit discourse on HPV vaccine encompasses a broad range of topics among men and women, with HPV political debates and cancer risk making up the plurality of the discussion. Our findings demonstrated that women and men both discussed HPV, highlighting that Reddit users do not perceive HPV as an issue that only pertains to women. Given the increasing use of social media as a source of health information, these results can inform the development of targeted online health communication strategies to promote HPV vaccination to young adult users of Reddit. Analyzing online discussions on Reddit can inform health communication efforts by identifying relevant, important HPV-related topics among online communities.

KEYWORDS

papillomavirus infections; prevention & control; cancer prevention; cervical cancer; HPV; vaccination; papillomavirus vaccines; immunology; administration & dosage; social media; health communication; infodemiology

Introduction

Background

Long-lasting infections with high-risk human papillomaviruses (HPVs) can cause cancer in parts of the body where HPV infects cells, such as in the cervix, oropharynx, anus, rectum, penis, vagina, and vulva. Virtually all cervical cancers are caused by HPV. HPV vaccination acceptance is a critical public health issue as low completion rates of the vaccine series place adolescents and young adults at risk for HPV-associated infection and cancers. The multiple factors that influence vaccine acceptance have been examined on traditional mass media and social media, including the understudied platform, Reddit. This study is the first to examine HPV discourse on Reddit, which provides insight about attitudes of HPV vaccination among a unique population that may not be readily accessible in traditional data surveillance methods [1]. Our findings build upon the literature examining attitudes toward HPV vaccination online and in the media. Our goals were to assess whether discussions on Reddit reflected the competing narratives (cancer risk vs sexual promiscuity) described in the broader discourse [2-5] and, secondly, to document and describe the full range of HPV vaccine-related topics discussed on Reddit.

Human Papillomavirus Vaccination Rates and Factors Affecting Uptake

The Advisory Committee on Immunization Practices (ACIP) at the Centers for Disease Control and Prevention (CDC) recommends routine HPV vaccination for all girls and boys aged 11 and 12 years, and catch-up immunization for teens and young adults up to age 26 years, to prevent HPV infection and HPV-associated cancers [6]. The vaccine series is safe and widely effective when taken before the onset of sexual activity and potential HPV exposure. However, national vaccination rates remain low, with only 49.5% of girls and 37.5% of boys completing the recommended series [6]. Increasing vaccine uptake is a preeminent national public health concern as demonstrated by the Healthy People 2020 objective to increase HPV vaccination series completion for adolescents aged from 13 to 15 years to 80% [7].

Factors contributing to lagging vaccination rates include parental concerns about vaccine safety [8], low confidence in adolescent vaccination [9], lack of provider recommendation [10], general lack of knowledge about the vaccine [11], and concerns about adolescent sexual behavior [12,13]. In particular, as HPV is sexually transmitted, some parents fear that vaccination will give license for adolescents to engage in early or risky sexual activity [2] (this has been refuted by prior research [2,14]). Owing to the stigma associated with sexually transmitted infections (STIs), parents and providers have faced challenges addressing the sexual nature of HPV [3].

This stigma is exacerbated by the fact that the initial recommendation of HPV vaccination was for girls only. Consequently, marketing of the HPV vaccine as primarily for cervical cancer prevention in 2006 has led to continued framing of HPV as a women's health issue despite the fact that HPV affects both men and women [15,16]. A strong "gender bias" in vaccine uptake has been observed, as parents of male teens may feel their child is at low risk and the vaccine is unnecessary [17]. This enduring, gendered narrative has contributed to low vaccine uptake and opportunity loss for cancer prevention among men, who are increasingly affected by HPV-related anal, penile, and oropharyngeal cancers [18].

Traditional and Social Media

The media are influential in relaying HPV-related cancer risk communication and contributing to fears of adolescent sexual behavior, including promiscuity concerns. Prior content analyses of US and Canadian newspapers have found that HPV-related articles mentioning sexual promiscuity concerns ranged from 30% to 60% [4,5,19], detracting from a focus on more pertinent details of vaccination (eg, reducing HPV-associated cancers and vaccination schedule and benefits). The media's framing of the HPV vaccine as competing narratives of adolescent sexual behaviors and cancer risk may have implications for influencing public knowledge, awareness, and beliefs on HPV vaccination.

Although HPV vaccination has been studied on other media, social media may have a particularly strong impact on public perception of HPV vaccination risks. In total, 69% of adults use some form of social media [20], and research has shown that an increasing number of people seek health information online [21,22]. Furthermore, the open nature of social media provides a critical opportunity to investigate personal attitudes and beliefs regarding HPV vaccination. Thus, an understanding of the online discourse regarding HPV vaccination can aid public health communicators in their efforts to address public misconceptions.

Reddit is a popular online forum that allows users to post, share, and rank content through a voting system. Reddit has 542 million monthly visitors, nominating it as the fifth most visited website in the United States [23,24]. Although HPV has been studied on social media platforms such as Twitter, YouTube, and Pinterest [25-27], to our knowledge, no published study has examined Reddit content related to HPV vaccination. Reddit data capture attitudes and trends that are not readily accessible in traditional data surveillance methods, as highlighted in previous studies that used Reddit data [28,29]. Moreover, Reddit's features allow users to post anonymously, allowing users to disclose personal behaviors more readily [28,29].

Reddit's relative lack of restrictions on posting requirements and its significant number of publicly available posts related to HPV vaccination make it a critical site to assess HPV discourse among its user base. Reddit is typically associated with a male adolescent culture as 67% of users are men and 59% of users

are between the ages of 18 and 29 years [30], providing an opportunity to communicate with a group that is eligible for vaccination, at risk of HPV-associated cancers, and otherwise hard to reach. In a survey of the Reddit community [31], 83.5% of participants identified themselves as aged between 18 and 34 years. Of these, 80.4% identified themselves as male. Thus, an analysis of Reddit provides insight into the male perspective of HPV vaccination on social media, an understudied area in the previous literature where women have been the focus.

Study Objectives

This study was the first to utilize Reddit data to gain a better understanding of online public discussions of the HPV vaccine, including determining whether online discussions mirror previous research and media coverage. We observed how the HPV vaccine is characterized on Reddit over time and by user gender. Furthermore, to the extent that those two topics may not be representative of the total HPV-related discourse on Reddit, we conducted an additional topic analysis. Our research questions (RQs) are as follows:

RQ1: Are Reddit messages more related to cancer risks or to sexual behavior?)

RQ2: What other topics characterize the discussion on HPV vaccination on Reddit?)

Studying cancer risks, sexual activity, and other emerging topics from Reddit content would provide greater insight into extant online discussions about HPV vaccine, which can inform concerted, tailored health communication efforts to promote HPV vaccination and improve uptake rates.

Methods

Data Source

We used Reddit as our main data source. Unlike microblogging social media platforms such as Twitter, Reddit does not limit message length, allowing users to express their opinions in depth. Communities can be organized by different topics through “subreddits,” which are forums dedicated to a specific topic such as news (/r/news) or HPV (/r/HPV). Reddit also has a strong community-based moderation culture; repetitive messages or unrelated messages will often be removed by moderators of each subreddit. Although it is possible that messages in our sample are due to malicious actors, such as bots and trolls [32], Reddit’s moderation culture also helps to mitigate these concerns. Therefore, Reddit comments are a promising data source for understanding how a segment of the online community perceives the risks of HPV and the risks and benefits of the HPV vaccine. The study was determined to be research that is exempt from the Institutional Review Board at the George Washington University (IRB-180804).

Data Collection

We downloaded all recent Reddit comments from January 1, 2007, to September 30, 2015, using a platform which allows researchers to collect and share complete Reddit datasets for research purposes [33]. At the time of analysis in 2016, the most up-to-date data were culled, resulting in our use of data from 2007 to 2015. From the total set of all messages, we identified

a subset of HPV-related messages by filtering all comments containing the text “hpv” or “human papillomavirus”. From the set of hpv-related messages, we further identified a subset of vaccine-related messages by filtering messages for the string “vaccin.”

Data Analysis

Our first analysis was designed to answer RQ1.

Our aim was to determine whether HPV-related messages discussed sexual behavior or cancer risk more often. First, 2 annotators (YL and AJ) manually annotated 50 messages and selected an initial set of keywords associated with both topic areas. The same annotators then annotated another 350 messages, in rounds of 50 each, adjusting the keywords and refining coding techniques. We agreed upon a final set of keywords as follows: a single keyword, “cancer,” was sufficient to identify messages pertaining to cancer. Messages pertaining to sexual behavior included keywords derived from the “purity” category in the Moral Foundation Dictionary [34]: “piety, pious, purity, pure, clean, sterile, sacred, chaste, saint, innocent, unclean, slut, whore, dirty, impiety, impious, profane, promiscuity, promiscuous, adulter, unchaste, sexual, sex, intercourse, coitus, lovemaking, and premarital.” Using these specific keywords, annotators initially began annotating for promiscuity; however, owing to a limited number of messages, this topic was expanded to include all sexual activity related to HPV (including promiscuity). Through this iterative process, we created a codebook to identify messages as either pertaining to cancer risk, sexual activity, or other (Table 1). There was moderate agreement on messages indexing sexual behaviors, Cohen kappa=0.62 (95% CI 0.54 to 0.69), and high agreement on messages indexing cancer risk, kappa=0.81 (95% CI 0.74 to 0.88).

Next, we employed Latent Semantic Analysis (LSA) [35], a commonly used technique in natural language processing, to measure the semantic similarity between each Reddit message and the keyword lists. Specifically, we rendered the set of all HPV-related Reddit messages into a term-document matrix and transformed this corpus into a “latent semantic space.” When preparing the corpus for the LSA analysis, we removed all but the 500 most frequent unigrams. We next applied term frequency-inverse document frequency weighting to the corpus. Finally, consistent with the past literature, we retained dimensions 2 to 101 when computing the Singular Value Decomposition underlying LSA. We used the underlying LSA space to measure the cosine similarity between each Reddit post and each keyword list and calculated the average similarity score per month (see Multimedia Appendix 1).

In addition, we developed a classifier to assess the gender of the user posting each message. This tool helped us understand the overall difference between male and female user discussions related to the HPV vaccine. The classifier works by identifying gender-indicating keywords and expressions included in Reddit messages. For instance, statements such as “As a man...” or “I am a straight woman...” are used as indicators of a user’s gender (Multimedia Appendix 2). In some instances, there was not enough information to identify a user’s gender with confidence, so those users were classified as having an unknown gender.

Table 1. Qualitative codebook and sample messages for Research Question 1.

Domain	Description	Sample messages ^a
Cancer risk	Messages are related to HPV ^b and relationship to cancer and HPV vaccine. This includes concerns related to cancer and awareness of cancer risk. Comments can be any related type of cancer: cervical, throat, penile, and anal.	“it's not just genital warts, the hpv vaccine can prevent cervical cancer.”; “the vaccines don't prevent all causes of cervical cancer, they protect against some strains of hpv which can cause cervical cancer. also, they only started giving that vaccine to women over 25 recently.”
General sexual behavior	Messages are related to HPV and sexual activity including discussions of age of sexual debut, risky sexual behavior, multiple sex partners, and associations with promiscuity. Messages need an explicit connection to HPV. Can include screening behavior if it is related to sexual activity.	“I've only had one sexual partner in my history and we were both virgins, so i'm certain i'm clear. not that hpv is a death sentence or anything, but i'd rather not have to deal with it . ”

^aMessages have been lightly edited to protect user anonymity.

^bHPV: human papillomavirus.

We manually annotated over 100 messages to test the accuracy of the classifier. A total of 2 annotators (DH and DB) read messages identified by the classifier and determined that 43 of 50 messages were correctly classified as “male” and 45 of 50 messages were correctly classified as “female.” This indicates an overall accuracy of 88% (see [Multimedia Appendix 3](#)).

We also developed a classifier to assess the age of users into 7 age categories (10 to 18, 18 to 24, 25 to 34, 35 to 44, 45 to 54, 55 to 64, and 65 and older) consistent with Reddit surveys [31]. When reading the gender-indicating expressions (eg, “I am 27 years old”), we observed the linguistic features related to age-indicating expressions were sometimes complicated and open to misclassification. For example, some expressions discussing time or gestational age were coded as ages, resulting in the classifier achieving 67% accuracy. Nevertheless, the current accuracy is higher than the random baseline of 7 categories (14.28%; [Multimedia Appendix 4](#)). This classifier, which is being developed, cannot be used to reach a definite conclusion of the age groups. Instead, it can be applied to report Reddit users' demographic information for the scope of this study.

Our second analysis addressed RQ2.

To further explore the topics discussed on Reddit, we used Latent Dirichlet Allocation (LDA) [36], a Bayesian topic model, to automatically segment Reddit messages into 10 probabilistic “topics” (using unigram features; hyperparameter values were $\alpha=.1$ and $\eta=0.01$). LDA assigns each word token to one of these topics, allowing one to summarize the content of each Reddit post. After examining models with 20 and 50 topics, we determined that a 10-topic model best captured the variety of discussion subjects with minimal redundancy.

For each month, we calculated how many word tokens belonged to each topic. A total of 2 annotators (YL and AJ) read the sample messages and top word tokens of each topic to determine the theme of each LDA topic. The proportional topic weight of each year was calculated by the following equation:

proportional topic weight of Topic X=(The frequency of all words that belongs to Topic X)/(The total frequency of all words in given years)

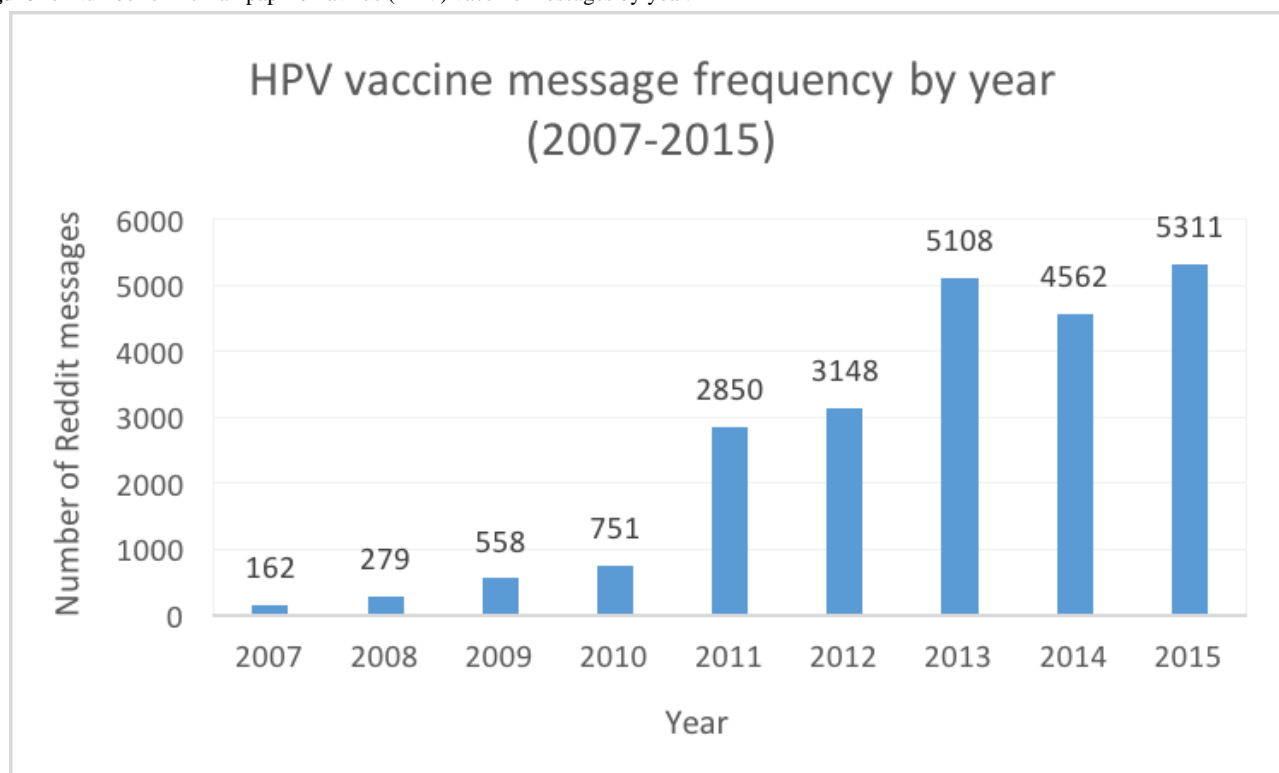
The detailed results and interpretation are reported below.

Results

Research Question 1: Cancer Risk and Sexual Activity

Our final dataset included 22,729 messages filtered to include “hpv” or “human papillomavirus” and “vaccin”. Mitigating concerns about behavior by malicious actors, such as bots, we found that no two messages in our dataset are exactly the same, a common indicator of bot behavior [37]. The number of messages varied by year and increased over time, from 162 messages in 2007 to a peak of 5311 messages in 2015 ([Figure 1](#)). Among those posting HPV vaccine messages (n=12,952), 53.0% (6869/12,952) were male, 8.8% (1143/12,952) were female, and 38.1% (4940/12,952) were of unknown gender ([Table 2](#)). Nearly 17% of users were aged from 25 to 34 years, followed by 11.2% of users aged 18 to 24 years, although 60.0% of users did not specify their age ([Table 2](#)).

In manual annotation, it was clear that messages describing cancer risk were more identifiable than messages discussing HPV-related sexual behavior. Cancer risk messages tended to focus on specific themes, including the relationship between HPV infection and cancer incidence (especially related to specific HPV strains), and on HPV vaccination as a resource to prevent future cancer incidence. Messages about sexual behaviors were more varied. Initially, we attempted to code for instances strictly related to HPV vaccine and sexual promiscuity but found very few instances (<5 in the first 100 messages) that we were unable to do so. The widespread use of sarcasm and irony on Reddit also made it difficult to know the meaning of some messages out of context. Many of the messages we encountered incorporated the moral foundation keywords, but only to refute the idea that HPV vaccine promotes promiscuity. To capture the complexity of this topic, we broadened our coding scheme to include any discussion of sexual activity and HPV vaccination and were able to identify a wider range of subjects including safe sex practices, personal stories, and debates over sexual risk.

Figure 1. Number of Human papillomavirus (HPV) vaccine messages by year.**Table 2.** Age and gender distribution of the Reddit sample data.

Age group (years)	Female users	Male users	Unspecified gender	Age subtotal
10 to 18	128	339	57	524
18 to 24	293	862	291	1446
25 to 34	376	1489	297	2162
35 to 44	103	549	65	717
45 to 54	42	195	26	263
55 to 64	7	37	2	46
65 and above	7	13	0	20
Unspecified age	187	3385	4202	7774
Gender subtotal	1143	6869	4940	12,952

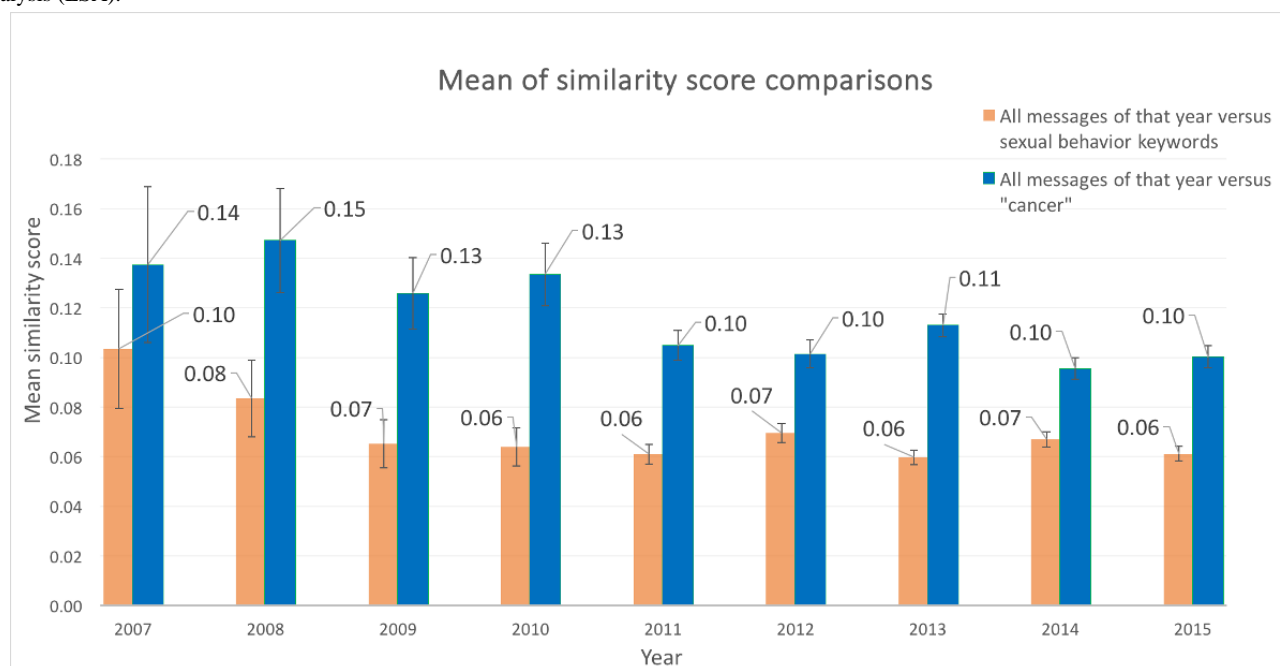
The LSA analysis showed that cancer is significantly more discussed than sexual activity (Table 3). Discussions about sexuality and cancer did not typically co-occur: The cosine similarity score between the cancer risk and sexual activity keyword lists in the semantic space was -0.01 , indicating that the 2 lists were essentially orthogonal. Discussions about HPV vaccine were more similar to the keyword “cancer” than to the sexual purity keywords, particularly after 2007 (Figure 2). However, the average monthly Reddit message discussed neither cancer nor sexuality (cosine similarity values were 0.11 and 0.08, respectively). These low similarity scores indicate that

other topics must exist in the corpus that are neither about cancer nor sexual behavior. Furthermore, we observed a decrease in average similarity to both cancer and sexuality over time, indicating that other topics of discourse may have emerged within these forums.

In general, we did not notice a significant difference between genders in discussions of cancer and sexual activity (all P values were greater than .05, except in 2012; note that we would expect one P value to be less than .05 owing to chance alone; Figure 2).

Table 3. *t* test of similarity of messages between human papillomavirus (HPV) cancer risk and sexual activity on Reddit by Latent Semantic Analysis (LSA).

Year	Message count	Sexual similarity, mean	Cancer similarity, mean	Yearly <i>t</i> test score
2007	162	0.10	0.14	-1.70
2008	279	0.08	0.15	-4.83 ^a
2009	558	0.07	0.13	-6.84 ^a
2010	751	0.06	0.13	-9.26 ^a
2011	2850	0.06	0.11	-11.97 ^a
2012	3148	0.07	0.10	-9.35 ^a
2013	5108	0.06	0.11	-19.20 ^a
2014	4562	0.07	0.10	-10.25 ^a
2015	5311	0.06	0.10	-14.64 ^a

^a*P*<.001.**Figure 2.** Comparison of message similarity between human papillomavirus (HPV) cancer risk and sexual activity on Reddit by Latent Semantic Analysis (LSA).

Research Question 2: Selected Topic Analysis

In our close reading of messages for manual annotation, we recognized that although messages included keywords related to cancer risk or sexual activity, there were additional topics that extended beyond the scope of either topic (Table 4). Some of these conversations were narrowly focused on HPV vaccination, for instance, debates over whether vaccinating the entire population was a cost-effective prevention method or discussions of vaccine efficacy related to particular strains of HPV. Other conversations were focused on sexual activity and health, in general. We observed a subset of conversations that detailed personal experiences with STIs and described the various steps individuals had taken to avoid all types of STIs (eg, condom use, HPV vaccination, and regular STI testing).

Another major topic was related to circumcision, as some users felt that primary prevention measures such as HPV vaccination reduced the need to circumcise (as circumcision is often cited as a method to reduce STI infection). Finally, another subset of conversations was focused on broader political and philosophical discussions that sometimes used HPV vaccination as a talking point. For instance, political discussions of the role of government sometimes questioned vaccine mandates as infringing on parental rights as part of a larger justification for libertarianism. Alternatively, users would discuss sexual education, access to contraceptives, and HPV vaccination as talking points in discussions of abortion rights (Table 4). Although these conversations were not specifically focused on HPV vaccination, we believed they reflect the broader discourse on the topic and were an important area of study.

Table 4. Summary of human papillomavirus (HPV) vaccine topics generated by the Latent Dirichlet Allocation (LDA) approach.

Topic name	Topic keywords	Description
General vaccine debate (Topic 0)	peopl, immun, diseases, get, like, flu, one, hpv, caus, children	Broadly concerned with vaccine decision making; Vaccine safety; Risks of side effects; Importance of vaccination; Personal stories
HPV ^a -related cancer and genital warts (Topic 1)	hpv, strain, cancer, caus, wart, protect, type, infect, genit, genit-wart	HPV strains related to cancer and genital warts; Strains covered by vaccine
Sharing government links (Topic 2)	http, www, http-www, gov, cdc, cdc-gov, nih, www-cdc, nih-gov, nlm	Vaccine concerns corroborated by links to CDC ^b and PubMed research articles
HPV cancer prevention (Topic 3)	cancer, cervic, cervic-cancer, hpv, get, women, prevent, caus, hpv-vaccin, men	Gender-related vaccine and HPV concerns; cervical cancer; other HPV related cancers; recommendations for men and women
HPV-specific vaccine debate (Topic 4)	effect, gardasil, studi, report, gt, hpv, death, year, side, hpv-vaccin	Adverse events linked to HPV vaccination; Side effect risk; Vaccine efficacy; Debating facts (fewer personal stories than topic 0).
HPV vaccine political debate (Topic 5)	peopl, hpv, make, gt, would, like, hpv-vac-cin, think, one thing	Parental vaccine rights; Reach of federal power; Abortion debate; Sexual education policy; General political debate; Individual politicians; Corruption
Sharing general links on HPV (Topic 6)	http, com, www, http-www, org, amp, en, hpv, comment, wikipedia	HPV concerns with links to news stories; Links to other Reddit threads; Generally sharing single links rather than a list of links (Topic 2)
HPV vaccination schedule (Topic 7)	get, hpv, hpv-vaccin, year, got, doctor, go, shot, age, girl	Age of vaccination; Preventive screenings; Medical advice; Personal stories; Deciding whether to vaccinate
Circumcision debate (Topic 8)	circumcis, men, gt, risk, women, male, hiv, infect, hpv, benefit	Risks and benefits of circumcision; Likelihood of contracting STI ^c ; Cancer prevention
Sexual behaviors including HPV (Topic 9)	get, hpv, sex, test, condom, peopl, like, know, partner, std	Discussions of sexual behaviors; STI prevention; Vaccination; Condom use; Relationship norms

^aHPV: human papillomavirus.

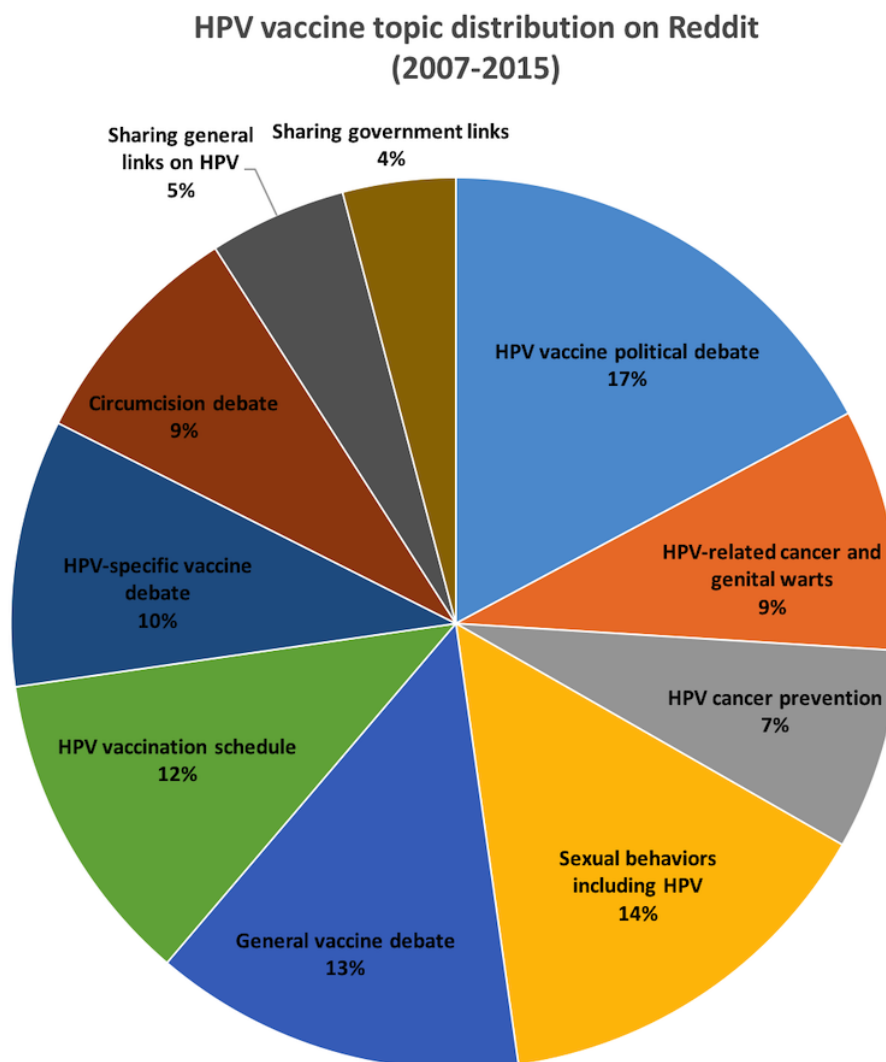
^bCDC: Centers for Disease Control and Prevention.

^cSTI: sexually transmitted infection.

In addition to the diversity in topics, it appears users were turning to Reddit for a variety of reasons: some were posing questions and were looking for answers; others seemed to want to debate; some just wanted to discuss topical issues in the news; and still others seemed to want to make jokes. Messages ranged from one-word responses to posts a few thousand words long. Some were based on personal experiences and anecdotes, and others were based on facts and statistics, often linking to cited research. The understanding gained from reading HPV vaccine-related messages led us to our second research question, as we realized we needed to explore the full range of topics that were being discussed on Reddit.

Expanding upon the LSA findings, the LDA topics (Table 4) generated a more comprehensive list of all HPV vaccine-related topics of interest. The most widely discussed topic was HPV-related political debate (17.2%), wherein users posted messages on political topics including vaccine policy, abortion rights, and sex education policy. Cancer was the second most

widely discussed (16.1%), after combining the two closely related topics of HPV strains causing cancer and genital warts and HPV cancer prevention. This was followed by sexual activity and preventive behaviors (14.5%), which focused on behaviors such as using condoms, receiving HPV vaccination, and regular testing and screenings for STIs. The fourth most popular topic was HPV disease and immunity (13.5%), which included broad concerns with diseases and infections associated with HPV. The fifth was the vaccine schedule related to HPV (11.5%), which included messages on age of vaccination initiation and requesting information on vaccination from physicians. The sixth, vaccine side effects and risks (9.7%) included messages on adverse events linked to HPV vaccination. The seventh topic was related to messages that included use of government or research (eg, CDC and PubMed) papers and general websites (eg, Wikipedia and news sites) to corroborate user opinions (9.1%). The final topic discussed circumcision, particularly the risks and benefits as related to STI and cancer prevention (8.5%) (Figure 3).

Figure 3. Human papillomavirus (HPV) vaccine topic distribution on Reddit.

The political debate surrounding HPV vaccines peaked in 2011 (24%) and decreased in 2013 (14%), when the number of messages plateaued ([Multimedia Appendix 5](#)). General HPV vaccine messages peaked in 2008 (20%), whereas general vaccine messages sharply decreased during 2008 (8%) and increased in 2009 (22%). Discussion of sexual behavior sharply increased in 2010 (18%) and 2014 (17%). Circumcision debate sharply increased in 2012 (11%; [Multimedia Appendix 5](#)).

Across the LDA topics, notable spikes were apparent in certain time periods (ie, 2008, 2009, 2011, 2013, and 2014). We observed peaks for certain years across LDA topics ([Multimedia Appendix 5](#)). Specifically, we observed an increase in the discussion related to HPV-associated cancers and genital warts in 2008 (10%; [Multimedia Appendix 5](#)). Furthermore, increases were observed across HPV cancer (11% and 0.092) and sexual activity discussions (20% and 18%, respectively) in 2010 ([Multimedia Appendix 5](#)). Discussions of cancer increased in 2013 (8% and 10%, respectively) and 2014 (17%).

Discussion

Principal Findings

Our findings emphasize increased discussion of HPV-related cancer risks on Reddit, in contrast with the mainstream media's focus on sexual behavior linked to HPV vaccine. Of the total set of HPV vaccine-related Reddit messages, 16.1% were focused on cancer risk topics and 14.5% were focused on HPV related to sexual activity. However, even when combined, these two topics represented less than one-third of the total HPV vaccine-related messages on Reddit. LSA findings did not show significant gender differences in discussions of HPV-related sexual behaviors and cancer risks, though discussions of both topics decreased over time. We identified additional topics that provide insight into the broader HPV discourse on Reddit, which included HPV disease and immunity, vaccine schedule, side effects, risks and benefits of vaccination, and the circumcision debate ([Figure 3](#)), providing further public health surveillance of nuanced HPV attitudes on Reddit.

Our LSA findings demonstrated that, on Reddit, cancer is significantly more discussed than sexual activity, consistent with the majority of prior research [[14,38,39](#)]. Although HPV

cancer risk prevention and HPV-related sexuality are presented as competing narratives in the broader media discourse, these two topics alone do not reflect the focus of the majority of HPV-related conversations on Reddit. Indeed, the political debate regarding vaccination was the most discussed topic on Reddit. The HPV vaccine has been the subject of political discussion since its introduction in 2006 [40]. Messages spiked in 2008 for the HPV vaccine debate, perhaps associated with media attention surrounding state consideration for school vaccination mandates for middle-school age girls [41]. The following year also saw spikes in messages for general vaccine debate, perhaps associated with the introduction of a second HPV vaccine, Cervarix, in 2009. That same year, Gardasil was approved for boys aged 9 to 26 years to reduce genital warts; however, the ACIP did not recommend routine vaccination among males until 2011 [40]. We observed continued increases in discussion through 2011 and 2012, most likely due to CDC HPV vaccine recommendations for boys in fall 2011 coinciding with the 2012 presidential campaign, which incited debates around vaccination [40]. Increases in cancer discussions during 2013 and sexual activity in 2014 do not coincide with major policy changes or announcements and thus warrant additional research on factors that contributed to these trends.

Our analyses did not support the idea of HPV as primarily a woman's health issue on Reddit, highlighting the differences between narratives on Reddit and general public discourse. This discrepancy is surprising, given the initial recommendation and marketing of the vaccine toward girls to prevent cervical cancer [15] and given the majority of Reddit users are adolescent males. Furthermore, until recently, media coverage of HPV vaccine has largely been focused on women, particularly regarding promiscuity [42] even with expanded vaccine recommendations to include males. Our LDA findings also challenge the notion that HPV is perceived as an issue affecting women exclusively, with sizable discussion on Reddit on HPV prevention for men related to circumcision. The circumcision debate indicates broader concerns of men concerned with STI and cancer prevention related to HPV. These results further emphasize the difference between discussions on Reddit and media representations on the HPV vaccine.

The implications of inconsistencies between HPV vaccination discussion on Reddit compared with media outlets are significant. Continued media coverage of controversies may detract from focus on benefits of vaccination, which may translate to lower rates of vaccine uptake and increased incidence of cancer [43]. Public health agencies can utilize traditional and digital media to share targeted and tailored health messages promoting HPV vaccination to the broader population. Our findings demonstrated a wide spectrum of viewpoints, ranging from evidence-based posts to subjective personal experience posts with varying accuracy. Users reported questions about HPV prevention strategies (eg, vaccination vs using condoms), best practices for regular screenings (eg, Pap tests), and the benefits of vaccination while considering potential side effects. By highlighting the topics most salient to Reddit users, public health communication efforts can be targeted to suit the needs of this online community. In addition, the range and extent of messages from users seeking information

highlights a gap in accessible credible information online. These messages on Reddit can pose a challenge for a user to navigate and make an informed decision regarding vaccination.

Critical to vaccine uptake is the receipt of health care provider (HCP) recommendation [42], which may increase awareness of the benefits of HPV vaccine as well as cancer risks of nonvaccination. Despite recommendations being a key driver in uptake, physicians are more likely to recommend other vaccines (ie, tetanus and diphtheria) compared with HPV vaccine [44]. Lacking HCP recommendations or accurate knowledge, users may turn to online forums such as Reddit seeking advice, which may be filled with inaccurate information. In addition, without provider and public health engagement in these forums, misconceptions surrounding HPV vaccination may proliferate, fueling vaccine hesitancy.

Our analysis of Reddit posts can inform health communication efforts tailored to users' queries. Topics such as HPV vaccination as cancer prevention can be promoted whereas other topics, such as fear of vaccine side effects, can be combated with accurate messaging. Our findings suggested an increasing interest in HPV vaccination on Reddit over time, indicating a need for public health communication through social media platforms such as Reddit. Public health officials can develop health communication strategies that engage users, such as answering questions through Ask Me Anything events on Reddit or enlisting experts or well-known public figures to promote vaccination. Providers can share accurate HPV information and engage with patients regarding their questions and concerns regarding HPV vaccination, particularly through national campaigns and events. Massey et al found that HCPs were less likely to utilize hashtags and engage in Twitter chat events, resulting in decreased online presence and missed opportunities to productively engage with users [45].

Given the increasing number of people using social media and seeking health information online, public health practitioners and agencies need to take advantage of these opportunities to connect with users online to advocate for accurate, timely information. This may help to combat misconceptions related to HPV vaccination behavior. For example, in posts related to the vaccine schedule, users questioned the need for age cutoffs for HPV vaccines, a question which can be addressed with clear messaging about eligibility, especially given the recent Food and Drug Administration approval for expanded use of individuals aged from 27 to 45 years [46]. The continued failure to address these communication needs may contribute to increased vaccine hesitancy, further decreasing vaccination rates and, subsequently, increasing HPV-related morbidities. Although the increased effort of targeted engagement with users online will require more resources (ie, dedicated time) from public health agencies, the opportunities to reach and impact HPV vaccination greatly outweigh the challenges.

Limitations

The messages are drawn from Reddit users and do not necessarily represent the range of attitudes and beliefs across the broader population. Consistent with Barthel et al, our dataset revealed that the majority of users who posted HPV vaccine messages were males (53%) compared with females (8.8%)

[30]. However, the fact that Reddit users skew younger in age, with 59% users between the ages of 18 and 29 years, is in line with target audiences for vaccination, particularly at the catch-up stage (aged 18 to 26 years) [30]. Our findings do not examine other demographic variables besides binary gender (ie, male or female) and only examined messages from 2007 to 2015. During this period, we noticed spikes in conversation that did not coincide with major policy recommendations or noteworthy events, warranting a further examination of underlying causes of these upticks in discussion.

Future studies would benefit by examining other demographic variables across further time points. Although our data limited us to examine only gender and age, future research can be expanded to parse messages by user characteristics such as age and geographical area to examine complex HPV attitudes with another level of granularity. This is especially important given the disparities in HPV vaccine uptake in rural areas compared with urban areas. Furthermore, sentiment analysis may provide insight regarding differences in message content (positive, neutral, and negative) across users. Topic analysis can also be applied across social media to examine any differences in discussion across different platforms.

Conclusions

This study provided critical insight about HPV vaccination discourse on Reddit over time. In particular, we found Reddit users were discussing a wider variety of topics, beyond cancer risks and sexual activity, in contrast to prevailing media focus. Reddit can be a surveillance tool to examine emerging trends among users. In response, health communication stakeholders can utilize this platform to address concerns of Web-based communities to the extent that they are representative of the broader public, instead of concerns of the news cycle. Finally, our results indicate that HPV is discussed among both men and women on Reddit, and it is heartening that HPV is not perceived as an issue that only pertains to women. These findings may inform the development of strategies to address HPV vaccine information, while dispelling misinformation and misconceptions, to increase vaccination and promote sexual health online. As young adults increasingly use Web-based resources to seek health information, public health communication efforts are needed to leverage this opportunity to deliver timely, accurate health information to internet communities.

Acknowledgments

This research was supported by the National Institute of General Medical Sciences, National Institutes of Health (NIH; award 5R01GM114771). The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Latent Semantic Analysis (LSA) similarity results.

[[XLSX File \(Microsoft Excel File\), 51KB - publichealth_v5i1e12480_app1.xlsx](#)]

Multimedia Appendix 2

Gender classifier accuracy assessment.

[[XLSX File \(Microsoft Excel File\), 34KB - publichealth_v5i1e12480_app2.xlsx](#)]

Multimedia Appendix 3

Development strategy of the gender classifier.

[[DOCX File, 58KB - publichealth_v5i1e12480_app3.docx](#)]

Multimedia Appendix 4

Annotation to assess accuracy age classifier.

[[XLSX File \(Microsoft Excel File\), 20KB - publichealth_v5i1e12480_app4.xlsx](#)]

Multimedia Appendix 5

Latent Dirichlet allocation (LDA) results.

[[XLSX File \(Microsoft Excel File\), 82KB - publichealth_v5i1e12480_app5.xlsx](#)]

References

1. Dredze M, Broniatowski DA, Smith MC, Hilyard KM. Understanding vaccine refusal: why we need social media now. *Am J Prev Med* 2016 Apr;50(4):550-552 [FREE Full text] [doi: [10.1016/j.amepre.2015.10.002](https://doi.org/10.1016/j.amepre.2015.10.002)] [Medline: [26655067](https://pubmed.ncbi.nlm.nih.gov/26655067/)]
2. Forster A, Wardle J, Stephenson J, Waller J. Passport to promiscuity or lifesaver: press coverage of HPV vaccination and risky sexual behavior. *J Health Commun* 2010 Mar;15(2):205-217. [doi: [10.1080/10810730903528066](https://doi.org/10.1080/10810730903528066)] [Medline: [20390987](https://pubmed.ncbi.nlm.nih.gov/20390987/)]
3. Friedman AL, Sheppard H. Exploring the knowledge, attitudes, beliefs, and communication preferences of the general public regarding HPV: findings from CDC focus group research and implications for practice. *Health Educ Behav* 2007 Jun;34(3):471-485. [doi: [10.1177/1090198106292022](https://doi.org/10.1177/1090198106292022)] [Medline: [17000622](https://pubmed.ncbi.nlm.nih.gov/17000622/)]
4. Quintero Johnson J, Sionean C, Scott A. Exploring the presentation of news information about the HPV vaccine: a content analysis of a representative sample of U.S. newspaper articles. *Health Commun* 2011 Sep;26(6):491-501. [doi: [10.1080/10410236.2011.556080](https://doi.org/10.1080/10410236.2011.556080)] [Medline: [21469005](https://pubmed.ncbi.nlm.nih.gov/21469005/)]
5. Casciotti DM, Smith KC, Tsui A, Klassen AC. Discussions of adolescent sexuality in news media coverage of the HPV vaccine. *J Adolesc* 2014 Feb;37(2):133-143 [FREE Full text] [doi: [10.1016/j.adolescence.2013.11.004](https://doi.org/10.1016/j.adolescence.2013.11.004)] [Medline: [24439619](https://pubmed.ncbi.nlm.nih.gov/24439619/)]
6. Walker T, Elam-Evans L, Singleton J, Yankey D, Markowitz LE, Fredua B, et al. National, regional, state, and selected local area vaccination coverage among adolescents aged 13-17 years-United States, 2016. *MMWR Morb Mortal Wkly Rep* 2017 Aug 25;66(33):874-882 [FREE Full text] [doi: [10.15585/mmwr.mm6633a2](https://doi.org/10.15585/mmwr.mm6633a2)] [Medline: [28837546](https://pubmed.ncbi.nlm.nih.gov/28837546/)]
7. Office of Disease Prevention and Health Promotion. Immunization and Infectious Diseases URL: <https://www.healthypeople.gov/2020/topics-objectives/topic/immunization-and-infectious-diseases/objectives?topicId=23> [accessed 2019-02-11] [WebCite Cache ID 767B4Bloc]
8. Darden PM, Thompson DM, Roberts JR, Hale JJ, Pope C, Naifeh M, et al. Reasons for not vaccinating adolescents: National Immunization Survey of Teens, 2008-2010. *Pediatrics* 2013 Apr;131(4):645-651. [doi: [10.1542/peds.2012-2384](https://doi.org/10.1542/peds.2012-2384)] [Medline: [23509163](https://pubmed.ncbi.nlm.nih.gov/23509163/)]
9. Gilkey M, McRee AL. Provider communication about HPV vaccination: a systematic review. *Hum Vaccin Immunother* 2016 Dec 02;12(6):1454-1468 [FREE Full text] [doi: [10.1080/21645515.2015.1129090](https://doi.org/10.1080/21645515.2015.1129090)] [Medline: [26838681](https://pubmed.ncbi.nlm.nih.gov/26838681/)]
10. Wong KY, Do YK. Are there socioeconomic disparities in women having discussions on human papillomavirus vaccine with health care providers? *BMC Womens Health* 2012 Oct 04;12:33 [FREE Full text] [doi: [10.1186/1472-6874-12-33](https://doi.org/10.1186/1472-6874-12-33)] [Medline: [23033931](https://pubmed.ncbi.nlm.nih.gov/23033931/)]
11. Katz I, Bogart L, Fu C, Liu Y, Cox J, Samuels R, et al. Barriers to HPV immunization among blacks and latinos: a qualitative analysis of caregivers, adolescents, and providers. *BMC Public Health* 2016 Dec 25;16(1):874 [FREE Full text] [doi: [10.1186/s12889-016-3529-4](https://doi.org/10.1186/s12889-016-3529-4)] [Medline: [27558506](https://pubmed.ncbi.nlm.nih.gov/27558506/)]
12. Rosenberg K, Mehcatie E. HPV vaccination and sexual behaviors in adolescents and young women. *Am J Nurs* 2018 Dec;118(4):69. [doi: [10.1097/01.NAJ.0000532080.57125.48](https://doi.org/10.1097/01.NAJ.0000532080.57125.48)] [Medline: [29596260](https://pubmed.ncbi.nlm.nih.gov/29596260/)]
13. Reiter PL, Brewer NT, Gottlieb SL, McRee AL, Smith JS. Parents' health beliefs and HPV vaccination of their adolescent daughters. *Soc Sci Med* 2009 Aug;69(3):475-480. [doi: [10.1016/j.socscimed.2009.05.024](https://doi.org/10.1016/j.socscimed.2009.05.024)] [Medline: [19540642](https://pubmed.ncbi.nlm.nih.gov/19540642/)]
14. Jena AB, Goldman DP, Seabury SA. Incidence of sexually transmitted infections after human papillomavirus vaccination among adolescent females. *JAMA Intern Med* 2015 Apr;175(4):617-623 [FREE Full text] [doi: [10.1001/jamainternmed.2014.7886](https://doi.org/10.1001/jamainternmed.2014.7886)] [Medline: [25664968](https://pubmed.ncbi.nlm.nih.gov/25664968/)]
15. Daley EA, Vamos CA, Zimet GD, Rosberger Z, Thompson EL, Merrell L. The feminization of HPV: reversing gender biases in US human papillomavirus vaccine policy. *Am J Public Health* 2016 Jun;106(6):983-984. [doi: [10.2105/AJPH.2016.303122](https://doi.org/10.2105/AJPH.2016.303122)] [Medline: [27153015](https://pubmed.ncbi.nlm.nih.gov/27153015/)]
16. Wailoo K, Livingston J, Epstein S, Aronowitz R. Three shots at prevention: The HPV vaccine and the politics of medicine's simple solutions: Johns Hopkins University Press; 2010:320 p.
17. Johnson KL, Lin MY, Cabral H, Kazis LE, Katz IT. Variation in human papillomavirus vaccine uptake and acceptability between female and male adolescents and their caregivers. *J Community Health* 2017 Jun;42(3):522-532 [FREE Full text] [doi: [10.1007/s10900-016-0284-5](https://doi.org/10.1007/s10900-016-0284-5)] [Medline: [27778139](https://pubmed.ncbi.nlm.nih.gov/27778139/)]
18. Mourad M, Jetmore T, Jategaonkar A, Moubayed S, Moshier E, Urken M. Epidemiological trends of head and neck cancer in the United States: a SEER population study. *J Oral Maxillofac Surg* 2017 Dec;75(12):2562-2572 [FREE Full text] [doi: [10.1016/j.joms.2017.05.008](https://doi.org/10.1016/j.joms.2017.05.008)] [Medline: [28618252](https://pubmed.ncbi.nlm.nih.gov/28618252/)]
19. Abdelmutti N, Hoffman-Goetz L. Risk messages about HPV, cervical cancer, and the HPV vaccine gardasil in North American news magazines. *J Cancer Educ* 2010 Sep;25(3):451-456. [doi: [10.1007/s13187-010-0087-9](https://doi.org/10.1007/s13187-010-0087-9)] [Medline: [20232189](https://pubmed.ncbi.nlm.nih.gov/20232189/)]
20. Smith A, Anderson M. Pew Internet. Washington, DC; 2018 Mar 08. Social media use in 2018 URL: <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/> [accessed 2019-02-11] [WebCite Cache ID 767CkiQXd]
21. Betsch C, Brewer N, Brocard P, Davies P, Gaissmaier W, Haase N, et al. Opportunities and challenges of Web 2.0 for vaccination decisions. *Vaccine* 2012 May 28;30(25):3727-3733. [doi: [10.1016/j.vaccine.2012.02.025](https://doi.org/10.1016/j.vaccine.2012.02.025)] [Medline: [22365840](https://pubmed.ncbi.nlm.nih.gov/22365840/)]
22. McRee AL, Reiter PL, Brewer NT. Parents' internet use for information about HPV vaccine. *Vaccine* 2012 May 28;30(25):3757-3762 [FREE Full text] [doi: [10.1016/j.vaccine.2011.11.113](https://doi.org/10.1016/j.vaccine.2011.11.113)] [Medline: [22172505](https://pubmed.ncbi.nlm.nih.gov/22172505/)]
23. Reddit. 2018. Traffic Statistics URL: <https://www.alexa.com/siteinfo/reddit.com> [accessed 2019-02-11] [WebCite Cache ID 767Cvg3Ce]
24. Reddit. 2017. URL: <https://www.redditinc.com/> [accessed 2019-02-17] [WebCite Cache ID 76GaNbjQN]

25. Briones R, Nan X, Madden K, Waks L. When vaccines go viral: an analysis of HPV vaccine coverage on YouTube. *Health Commun* 2012;27(5):478-485. [doi: [10.1080/10410236.2011.610258](https://doi.org/10.1080/10410236.2011.610258)] [Medline: [22029723](https://pubmed.ncbi.nlm.nih.gov/22029723/)]
26. Guidry JP, Carlyle K, Messner M, Jin Y. On pins and needles: how vaccines are portrayed on Pinterest. *Vaccine* 2015 Sep 22;33(39):5051-5056. [doi: [10.1016/j.vaccine.2015.08.064](https://doi.org/10.1016/j.vaccine.2015.08.064)] [Medline: [26319742](https://pubmed.ncbi.nlm.nih.gov/26319742/)]
27. Shapiro G, Surian D, Dunn A, Perry R, Kelaher M. Comparing human papillomavirus vaccine concerns on Twitter: a cross-sectional study of users in Australia, Canada and the UK. *BMJ Open* 2017 Oct 05;7(10):e016869 [FREE Full text] [doi: [10.1136/bmjopen-2017-016869](https://doi.org/10.1136/bmjopen-2017-016869)] [Medline: [28982821](https://pubmed.ncbi.nlm.nih.gov/28982821/)]
28. 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 Dec 01;188:364-369. [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/)]
29. Sowles SJ, McLeary M, Optican A, Cahn E, Krauss MJ, Fitzsimmons-Craft EE, et al. A content analysis of an online pro-eating disorder community on Reddit. *Body Image* 2018 Mar;24:137-144. [doi: [10.1016/j.bodyim.2018.01.001](https://doi.org/10.1016/j.bodyim.2018.01.001)] [Medline: [29414146](https://pubmed.ncbi.nlm.nih.gov/29414146/)]
30. Barthel M, Stocking G, Holcomb J, Mitchell A. Pew Research Center. Washington, DC; 2016. Reddit news users more likely to be male, young and digital in their news preferences URL: <http://www.journalism.org/2016/02/25/reddit-news-users-more-likely-to-be-male-young-and-digital-in-their-news-preferences/> [accessed 2019-02-11] [WebCite Cache ID 767DCmEaT]
31. Reddit. 2011. Who in the World is reddit? Results are in... URL: <https://redditblog.com/2011/09/12/who-in-the-world-is-reddit-results-are-in/> [accessed 2019-02-17] [WebCite Cache ID 76GakchqO]
32. Broniatowski DA, Jamison AM, Qi S, AlKulaib L, Chen T, Benton A, et al. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *Am J Public Health* 2018 Oct;108(10):1378-1384. [doi: [10.2105/AJPH.2018.304567](https://doi.org/10.2105/AJPH.2018.304567)] [Medline: [30138075](https://pubmed.ncbi.nlm.nih.gov/30138075/)]
33. Pushshift. Learn about big data and social media ingest and analysis URL: <https://pushshift.io/using-bigquery-with-reddit-data/> [accessed 2019-02-18] [WebCite Cache ID 76HFiVv0l]
34. Graham J, Haidt J, Nosek BA. Liberals and conservatives rely on different sets of moral foundations. *J Pers Soc Psychol* 2009 May;96(5):1029-1046. [doi: [10.1037/a0015141](https://doi.org/10.1037/a0015141)] [Medline: [19379034](https://pubmed.ncbi.nlm.nih.gov/19379034/)]
35. Deerwester S, Dumais ST, Furnas GW, Landauer TK, Harshman R. Indexing by latent semantic analysis. *J Am Soc Inf Sci* 1990;41(6):391-407. [doi: [10.1002/\(SICI\)1097-4571\(199009\)41:6%3C391::AID-ASII%3E3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-4571(199009)41:6%3C391::AID-ASII%3E3.0.CO;2-9)]
36. Blei D, Ng A, Jordan M. Latent Dirichlet Allocation. *J Mach Learn Res* 2003;3:993-1022 [FREE Full text]
37. Gilbert E. Widespread Underprovision on Reddit. In: Proceedings of the 2013 conference on Computer supported cooperative work.: Association for Computing Machinery; 2013 Presented at: Conference on computer-supported cooperative work and social computing; February 23-27, 2013; San Antonio, Texas p. 803-808.
38. Bednarczyk RA, Davis R, Ault K, Orenstein W, Omer SB. Sexual activity-related outcomes after human papillomavirus vaccination of 11- to 12-year-olds. *Pediatrics* 2012 Nov;130(5):798-805. [doi: [10.1542/peds.2012-1516](https://doi.org/10.1542/peds.2012-1516)] [Medline: [23071201](https://pubmed.ncbi.nlm.nih.gov/23071201/)]
39. Mullins TL, Rosenthal SL, Zimet GD, Ding L, Morrow C, Huang B, et al. Human papillomavirus vaccine-related risk perceptions do not predict sexual initiation among young women over 30 months following vaccination. *J Adolesc Health* 2018 Feb;62(2):164-169 [FREE Full text] [doi: [10.1016/j.jadohealth.2017.09.008](https://doi.org/10.1016/j.jadohealth.2017.09.008)] [Medline: [29198772](https://pubmed.ncbi.nlm.nih.gov/29198772/)]
40. Krakow M, Rogers B. Collateral damage and critical turning points: public health implications of HPV vaccine news coverage for boys and men in 2011. *Health Commun* 2016 Dec;31(9):1081-1088. [doi: [10.1080/10410236.2015.1038773](https://doi.org/10.1080/10410236.2015.1038773)] [Medline: [26799666](https://pubmed.ncbi.nlm.nih.gov/26799666/)]
41. Ueda Y, Enomoto T, Sekine M, Egawa-Takata T, Morimoto A, Kimura T. Japan's failure to vaccinate girls against human papillomavirus. *Am J Obstet Gynecol* 2015 Mar;212(3):405-406. [doi: [10.1016/j.ajog.2014.11.037](https://doi.org/10.1016/j.ajog.2014.11.037)] [Medline: [25434842](https://pubmed.ncbi.nlm.nih.gov/25434842/)]
42. Burdette AM, Webb NS, Hill TD, Jokinen-Gordon H. Race-specific trends in HPV vaccinations and provider recommendations: persistent disparities or social progress? *Public Health* 2017 Jan;142:167-176. [doi: [10.1016/j.puhe.2016.07.009](https://doi.org/10.1016/j.puhe.2016.07.009)] [Medline: [27592005](https://pubmed.ncbi.nlm.nih.gov/27592005/)]
43. Hilton S, Hunt K, Langan M, Bedford H, Petticrew M. Newsprint media representations of the introduction of the HPV vaccination programme for cervical cancer prevention in the UK (2005-2008). *Soc Sci Med* 2010 Mar;70(6):942-950 [FREE Full text] [doi: [10.1016/j.socscimed.2009.11.027](https://doi.org/10.1016/j.socscimed.2009.11.027)] [Medline: [20064682](https://pubmed.ncbi.nlm.nih.gov/20064682/)]
44. Gilkey MB, Magnus BE, Reiter PL, McRee AL, Dempsey AF, Brewer NT. The Vaccination Confidence Scale: a brief measure of parents' vaccination beliefs. *Vaccine* 2014 Oct 29;32(47):6259-6265 [FREE Full text] [doi: [10.1016/j.vaccine.2014.09.007](https://doi.org/10.1016/j.vaccine.2014.09.007)] [Medline: [25258098](https://pubmed.ncbi.nlm.nih.gov/25258098/)]
45. Massey PM, Budenz A, Leader A, Fisher K, Klassen AC, Yom-Tov E. What drives health professionals to tweet about #HPV vaccine? Identifying strategies for effective communication. *Prev Chronic Dis* 2018 Dec 22;15:E26 [FREE Full text] [doi: [10.5888/pcd15.170320](https://doi.org/10.5888/pcd15.170320)] [Medline: [29470166](https://pubmed.ncbi.nlm.nih.gov/29470166/)]
46. US Food and Drug Administration. FDA approves expanded use of Gardasil 9 to include individuals 27 through 45 years old. 2018. URL: <https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm622715.htm> [accessed 2019-02-11] [WebCite Cache ID 767EXxXhK]

Abbreviations

ACIP: Advisory Committee on Immunization Practices

CDC: Centers for Disease Control and Prevention

HCP: health care provider

HPV: human papillomavirus

LDA: Latent Dirichlet Allocation

LSA: Latent Semantic Analysis

NIH: National Institutes of Health

STI: sexually transmitted infection

Edited by T Sanchez; submitted 12.10.18; peer-reviewed by N O'Kane, C Zhang; comments to author 12.11.18; revised version received 08.01.19; accepted 18.01.19; published 18.03.19.

Please cite as:

Lama Y, Hu D, Jamison A, Quinn SC, Broniatowski DA

Characterizing Trends in Human Papillomavirus Vaccine Discourse on Reddit (2007-2015): An Observational Study

JMIR Public Health Surveill 2019;5(1):e12480

URL: <http://publichealth.jmir.org/2019/1/e12480/>

doi: [10.2196/12480](https://doi.org/10.2196/12480)

PMID:

©Yuki Lama, Dian Hu, Amelia Jamison, Sandra Crouse Quinn, David A Broniatowski. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 18.03.2019. 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 Syndrome-Based Surveillance System for Infectious Diseases Among Asylum Seekers in Austrian Reception Centers, 2015-2018: Analysis of Reported Data

Ziad El-Khatib^{1,2}, PhD; Karin Taus¹, MSc; Lukas Richter¹, MSc; Franz Allerberger¹, MD; Daniela Schmid¹, MD

¹Department of Surveillance and Infectious Disease Epidemiology, Institute of Medical Microbiology and Hygiene, Austrian Agency for Health and Food Safety, Vienna, Austria

²Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:

Daniela Schmid, MD

Department of Surveillance and Infectious Disease Epidemiology

Institute of Medical Microbiology and Hygiene

Austrian Agency for Health and Food Safety

Währingerstraße 25a

Vienna, 1096

Austria

Phone: 43 50 555 37304

Email: daniela.schmid@ages.at

Abstract

Background: Austria has been among the main European countries hosting incoming asylum seekers since 2015. Consequently, there was an urgent need to predict any public health threats associated with the arriving asylum seekers. The Department of Surveillance and Infectious Disease Epidemiology at the Austrian Agency for Health and Food Safety (AGES) was mandated to implement a national syndrome-based surveillance system in the 7 reception centers by the Austrian Ministry of Interior and Ministry of Health.

Objective: We aimed to analyze the occurrence and spread of infectious diseases among asylum seekers using data reported by reception centers through the syndrome-based surveillance system from September 2015 through February 2018.

Methods: We deployed a daily data collection system for 13 syndromes: rash with fever; rash without fever; acute upper respiratory tract infection; acute lower respiratory tract infection; meningitis or encephalitis; fever and bleeding; nonbloody gastroenteritis or watery diarrhea; bloody diarrhea; acute jaundice; skin, soft tissue, or bone abnormalities; acute flaccid paralysis; high fever with no other signs; and unexplained death. General practitioners, the first professionals to consult for health problems at reception centers in Austria, sent the tally sheets on identified syndromes daily to the AGES.

Results: We identified a total of 2914 cases, presenting 8 of the 13 syndromes. A total of 405 signals were triggered, and 6.4% (26/405) of them generated alerts. Suspected acute upper respiratory tract infection (1470/2914, 50.45% of cases), rash without fever (1174/2914, 40.29% of cases), suspected acute lower respiratory tract infection (159/2914, 5.46% of cases), watery diarrhea (73/2914, 2.51% of cases), and skin, soft tissue, or bone abnormalities (32/2914, 1.10% of cases) were the top 5 syndromes.

Conclusions: The cooperation of the AGES with reception center health care staff, supported by the 2 involved ministries, was shown to be useful for syndromic surveillance of infectious diseases among asylum seekers. None of the identified alerts escalated to an outbreak.

(*JMIR Public Health Surveill* 2019;5(1):e11465) doi:[10.2196/11465](https://doi.org/10.2196/11465)

KEYWORDS

Austria; refugee health; asylum seekers; syndrome surveillance system; mass health monitoring; refugees; population surveillance; public health surveillance; epidemiological monitoring

Introduction

Background

Since 2015, over 1.5 million refugees and asylum seekers have reached Europe [1], where Austria, Sweden, and Hungary were the top 3 countries per capita in hosting this vulnerable population [1]. The majority of these asylum seekers came from Syria (49%), Afghanistan (21%), and Iraq (9%) [2,3].

The asylum seekers entered Austria largely through the Balkan corridor. The countries they travel through include, in the order, are Turkey, Greece, the Former Yugoslav Republic of Macedonia, Serbia, Croatia, Slovenia, and finally Austria [2,4], during a period ranging between 1 and 7 months [5]. Among the top 5 countries of origin of the asylum seekers who entered Austria in 2015-2016, over half of those were from Afghanistan and Syria, followed by Iraq, Iran, and Pakistan [6]. On behalf of the Republic of Austria, since 2012, a medical organization, called in German *Organisation für Regie und Spezialaufträge* (ORS Service AG, Zurich, Switzerland), has looked after all asylum seekers who are in federal care, including the operation of 3 permanent reception centers (RCs) for asylum seekers. Health services, including physical examination at arrival and need-based health care, are provided in the RCs by a physician and a nurse.

A refugee population is considered as a vulnerable group due to their living conditions during migration and their psychological stress, combined with limited access to sanitation, health services, and nutritious food [4]. This places asylum seekers, especially children and women, at a higher risk of acquiring communicable diseases during their migration in transit and in host countries [7]. A systematic evaluation of health conditions of migrants at the RCs will inform health care service providers [8]. In July 2015, following a recommendation of the Austrian Agency for Health and Food Safety (AGES), the Austrian Ministry of Health together with the Ministry of Interior decided to implement a syndrome-based surveillance system (SbSS) in all of the 7 national RCs for asylum seekers. The AGES implemented this SbSS in 3 phases, starting in August 2015. Phase 1 was preparatory, which included the selection of the RCs and collection of information on the number of beds in the RCs, applied hygiene measures, and procedure of the health status assessment of asylum seekers at their arrival at the center. We assessed the risk for epidemic-prone infectious diseases existing in the country of migrant origin, prevailing in transit countries, present in the host country, and favored by the immunization status, as recommended by the European Centre for Disease Prevention and Control (ECDC) [9]. Finally, we prepared a protocol on the SbSS. Phase 2 was piloting the SbSS in 2 of the 7 RCs in September 2015. These 2 were located in 2 of the 9 Austrian provinces (Lower Austria and Tyrol). Phase 3 included the implementation of the SbSS. In November 2015, the ECDC published a report based on expert opinion on the public health needs of asylum seekers at the southern and southeastern borders of the European Union (EU) [4]. According to this document, the public health measures to be considered

included (1) disease screening, (2) vaccination, (3) access to health care, free of charge, general hygiene measures, and prevention of overcrowding in RCs, and (4) implementation of an SbSS [10].

Overall, SbSS is a reliable method for public health surveillance that has been used in different countries and contexts. It provides timely information and supports suitable public health responses [11]. According to EU member state experiences [4], the following syndromes were suggested: (1) respiratory tract disease, (2) suspected pulmonary tuberculosis, (3) bloody diarrhea, (4) watery diarrhea, (5) fever and rash, (6) meningitis or encephalitis, or encephalopathy or delirium, (7) lymphadenitis with fever, (8) botulism-like illness, (9) sepsis or unexplained shock, (10) hemorrhagic illness, (11) acute jaundice, (12) parasitic skin infection, and (13) unexplained death [4,10,12].

Objective

The main objective of the Austrian SbSS was to enhance early detection of single cases or outbreaks of infectious diseases that require an assessment to initiate and guide appropriate public health measures in order to prevent spread of these diseases among the migrant population and the host population. To our knowledge, the use of the SbSS method among asylum seekers in the EU has not been thoroughly documented during the current refugee crisis. In this paper, we present and describe the results of using a national SbSS for asylum seekers in Austrian RCs.

Methods

Description of the Syndromic Surveillance System

Population and Syndromes Under Surveillance

The population under surveillance comprised asylum seekers hosted by the 7 Austrian RCs, which were located in 7 of the 9 Austrian provinces (Table 1). The SbSS was rolled out in these 7 centers one after another, 6 between September and December 2015, and the last 1 in August 2016. Table 1 gives the RCs by province and dates of implementation of the SbSS. The longest period of surveillance was in the permanent RC, which is located in the province of Lower Austria (RC1 east) for a duration up to 2.5 years.

The surveillance included 13 syndromes (Table 2), which we selected according to the infectious disease risks assessed specifically for asylum seekers in Austria and to the ECDC's assessment for risks related to asylum seekers in Europe [13-15]. The syndromes were rash with fever; rash without fever; suspected acute upper respiratory tract infection (URTI); suspected acute lower respiratory tract infection (LRTI); meningitis or encephalitis; fever and bleeding; non-bloody gastroenteritis or watery diarrhea; bloody diarrhea; acute jaundice; skin, soft tissue, or bone abnormalities; acute flaccid paralysis; high fever with no other signs; and unexplained death. Table 2 presents the 13 syndromes by name and gives the definition, underlying target diseases, and relevant public health actions.

Table 1. Reception centers in Austria by duration of syndrome-based surveillance in place and number of beds.

Province (geographical location inside Austria)	Reception center	Surveillance start date	Duration of surveillance (days)	Beds (N)
Lower Austria	RC1 (east)	September 2015	890	1800
Tyrol	RC2 (west)	September 2015	873	200
Upper Austria	RC3 (north)	October 2015	852	180
Carinthia	RC4 (south)	December 2015	804	200
Vienna	RC5 (east)	December 2015	797	150
Styria	RC6 (south)	December 2015	787	150
Salzburg	RC7 (west)	August 2016	538	160

Case Finding and Case Reporting

In Austrian RCs, health care staff, including general practitioners and nurses, are the first professionals to consult for health problems of asylum seekers. The health care staff assessed the presence of syndromes, as defined in Table 2, during the medical entry examination (active case finding) and any other consultation required by asylum seekers during their stay at the RC (passive case finding). The RC's SbSS focal person accordingly filled in a tally sheet, which listed the 13 syndromes with their clinical definition by age group (0-4, 5-14, 15-44, 45-64, ≥65 years) and was asked to send it on a daily basis to the SbSS project manager at the AGES by fax, email, or an online data entry form.

Quality Indicators Used for Syndromic Surveillance System Evaluation

We defined completeness and timeliness as a completed tally sheet (including "zero" reporting) sent within 24 hours of detection to the AGES.

Data Analysis

We adapted and modified the model for signal and alert generation as described by Napoli et al and applied in refugee RCs in Italy [16]. We determined the number of observed (reported) daily cases for each syndrome (syndrome-specific ODCs) and RC, including zero reporting. By calculating the 1-sided moving average of the syndrome-specific ODCs of the

previous 14 days with the formula shown in Figure 1, we determined the expected daily cases (EDCs). We created a threshold for the expected daily cases (ETH) by adding 2 times the standard deviation to the EDCs (ie, $ETH_t = EDCs_t + 2 \times SD_t$). We measured the ODCs daily against the ETH for each syndrome. A signal was generated when the number of syndrome-specific ODCs exceeded the ETH for the syndrome, and an alert was defined as signals occurring over 2 consecutive days. We analyzed the data and generated outputs (syndrome-specific number of signals and alerts) automatically on a daily, weekly, and monthly basis using R [17]. We shared the weekly reports with the Federal Ministry of Health and the monthly reports with the contact person at each RC. In case the RCs did not send the daily report, the manager of the SbSS called the contact person at the respective RC to follow up on the status of the report. In case of a signal, the SbSS manager at the AGES immediately called the respective RC contact person to clarify the underlying diseases and to verify the cluster of cases. In case of an alert, the AGES immediately informed the Ministry of Health.

Ethics and Informed Consent

The SbSS was conducted as part of the health care services offered by the Austrian Ministry of Interior, in collaboration with the Ministry of Health and the Department of Surveillance and Infectious Disease Epidemiology at the AGES. As data were collected in an anonymous fashion, there was no need for ethical clearance.

Table 2. List of syndromes surveilled and their definition, target diseases, and public health action^a.

Name of syndrome	Definition of syndrome	Target disease	Public health action or measures
Rash with fever	Temperature $\geq 38.0^{\circ}\text{C}$ and generalized rash of any nature	Measles, rubella, varicella, small-pox, louse-borne diseases (relapsing fever due to <i>Borrelia recurrentis</i> , trench fever due to <i>Bartonella quintana</i> , epidemic typhus due to <i>Rickettsia prowazekii</i>).	Outbreak confirmation and investigation, contact tracing, isolation, vaccination
Rash without fever	N/A ^b	Scabies.	N/A
Suspected acute upper respiratory tract infection	Fever, cough, sore throat, runny nose	Pharyngitis, tonsillitis caused by adenovirus, rhinovirus, respiratory syncytial virus, influenza, parainfluenza.	N/A
Suspected acute lower respiratory tract infection with fever	Temperature $\geq 38.0^{\circ}\text{C}$ and at least one of the following signs or symptoms: breathing difficulties; chest rales or increased respiratory rate	Tracheitis, bronchitis, pneumonia, bronchopneumonia, or bronchiolitis, including those caused by, for example, adenovirus, streptococci, pneumococci, <i>Mycoplasma</i> , <i>Legionella</i> .	Outbreak investigation in case of clustering of cases
Meningitis or encephalitis	Temperature $\geq 38.0^{\circ}\text{C}$ and at least one of the following signs or symptoms: severe, persistent headache; neck stiffness; altered consciousness; altered mental status; confusion; delirium; or disorientation	Bacterial, viral, fungal, or other infectious meningitis or encephalitis. This could be caused by meningococci, <i>Haemophilus influenzae</i> , pneumococci, <i>Listeria</i> , <i>Leptospira</i> , <i>Mycobacterium tuberculosis</i> , <i>Treponema pallidum</i> , enteroviruses, poliovirus, measles virus, mumps virus, rubella virus, influenza virus, West Nile virus, other arboviruses.	Outbreak confirmation and investigation, contact tracing, isolation
Fever and bleeding	Temperature $\geq 38.0^{\circ}\text{C}$ and at least one of the following signs or symptoms: petechial rash with any purpuric areas; hemorrhagic exanthema; hematuria; conjunctival hemorrhage; gingival bleeding; epistaxis; bloody diarrhea; unexplained bleeding from other sites; or clinical suspicion of a viral hemorrhagic illness	Hemorrhagic fevers due to infectious disease agents. These could include yellow fever, dengue, or Crimean-Congo hemorrhagic fever and other arboviral diseases, Ebola.	Contact tracing, isolation
Non-bloody watery diarrhea	≥ 3 watery stools per day, nausea, vomiting	Gastroenteritis caused by norovirus, rotavirus, bacterial toxins. <i>Campylobacter</i> , <i>Salmonella</i> , <i>Escherichia coli</i> , <i>Yersinia</i> , <i>Vibrio cholerae</i> .	Outbreak investigation for source and vehicle, and control in case of clustering of cases
Bloody diarrhea	Red blood in the stool	Amoebic dysentery, <i>Shigella</i> , <i>Campylobacter</i> , verotoxin-producing <i>E coli</i> .	Outbreak investigation for source and vehicle, and control in case of clustering of cases
Acute jaundice	Acute onset of jaundice and at least one of the following signs or symptoms: temperature $\geq 38.0^{\circ}\text{C}$; malaise or hepatomegaly	Acute viral hepatitis A and E; other hepatitis.	Outbreak investigation for source and vehicle
Skin, soft tissue, or bone abnormalities	Skin or soft tissue lesions, ulceration, inflammation	Cutaneous diphtheria, cutaneous tuberculosis, cutaneous leishmaniasis, bacterial wound infection.	N/A
Acute flaccid paralysis	Person under the age of 15 years with acute flaccid nonsymmetrical paralysis	Acute flaccid paralysis, or paralytic or poliomyelitis.	Contact tracing, immunization
High fever with no other signs	High fever up to 40°C or more, persisting, intermittent, long-lasting	Typhoid fever; malaria, or visceral leishmaniasis.	N/A
Unexplained death	N/A	N/A	N/A

^aSources: European Centre for Disease Prevention and Control [4,10].^bN/A: not applicable.

Figure 1. Formula for the expected daily cases (EDCs) at day t defined as the 1-sided moving average of the syndrome-specific observed daily cases (ODCs) of the previous 14 days.

$$EDC_t = \frac{\sum_{i=1}^{14} ODC_{t-i}}{14}$$

Results

Reports of Cases, Signals, and Alerts

During the period of September 2015 through February 2018, a total of 2914 cases, showing 8 of the 13 syndromes surveilled, were reported. The majority of patients were aged between 15 and 44 years (2179/2913, 74.80%), followed by children aged between 5 and 14 years (369/2913, 12.67%) and under 5 years (311/2913, 10.68%; [Table 3](#)).

The top 5 syndromes by number of cases were (1) suspected URTI (1470/2914, 50.45% of cases), (2) rash without fever (entirely due to scabies; 1174/2914, 40.29% of cases), (3) suspected acute LRTI with fever (159/2914, 5.46% of cases), (4) watery diarrhea (73/2914, 2.51% of cases), and (5) skin, soft tissue, or bone abnormalities (32/2914, 1.10% of cases; [Table 4](#)). [Figure 2](#) shows a plot of the data output for the example of the number of ODCs of scabies.

Among the total 405 triggered signals, the predominating signal-causing syndromes were suspected URTI (151/405, 37.3%), rash without fever (142/405, 35.1%), and suspected LRTI (67/405, 16.5%). Of the 405 signals, 26 (6.4%) triggered an alert, which were due to suspected acute URTI (10/26, 39%),

rash without fever (10/26, 39%), suspected acute LRTI (4/26, 15%), and skin, soft tissue, or bone abnormalities (2/26, 8%; see [Table 5](#)). [Figure 3](#) shows a plot of the data output for the example of the number of signals and alerts for scabies. [Figure 4](#) shows the monthly number of signals and alerts for all syndromes during September 2015 through February 2018.

Laboratory Diagnostics of the Case-Patients

Of the 1470 case-patients with URTI, 21 of 30 patients (70%) with influenza-like illness tested positive for influenza during the 2015-2016 influenza season. Among asylum seekers with acute LRTI with fever and cough, 3 cases were diagnosed as pulmonary tuberculosis. Of the 73 cases of watery diarrhea, 3 cases were culture-confirmed shigellosis. As the others had short-term diarrhea, the causative pathogen was not identified. One case of malaria was identified among those with high fever with no other signs.

Quality Indicators

Among the RCs, only RC1 and RC2, with permanent health care staff, sent tally sheets on syndromes daily within 24 hours. The remaining RCs sent the sheets whenever the health care service was delivered at the site.

Table 3. Number of cases of each syndrome in total and by age group recorded by the syndrome-based surveillance system, Austria, September 2015-February 2018.

Name of syndrome	All ages ^a	Age groups (years)				
		0-4	5-14	15-44	45-64	≥65
Suspected acute URTI ^b , n (%)	1469 (50.43)	235 (75.6)	291 (78.9)	911 (41.81)	31 (60)	1 (50)
Rash without fever, n (%)	1174 (40.30)	20 (6.4)	55 (14.9)	1095 (50.25)	4 (8)	0 (0)
Suspected acute LRTI ^c with fever, n (%)	159 (5.46)	16 (5.1)	12 (3.3)	115 (5.28)	16 (31)	0 (0)
Watery diarrhea, n (%)	73 (2.51)	34 (10.9)	11 (3.0)	27 (1.24)	1 (2)	0 (0)
Skin, soft tissue, or bone abnormalities, n (%)	32 (1.10)	1 (0.03)	0 (0)	30 (1.38)	0 (0)	1 (50)
Rash with fever, n (%)	4 (0.14)	4 (0.13)	0 (0)	0 (0)	0 (0)	0 (0)
Meningitis or encephalitis, n (%)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
High fever (≥39°C) with no other signs, n (%)	1 (0.03)	0 (0)	0 (0)	1 (0.05)	0 (0)	0 (0)
Unexplained death, n (%)	1 (0.03)	1 (0.3)	0 (0)	0 (0)	0 (0)	0 (0)
Total (N)	2913	311	369	2179	52	2

^aOne case-patient had missing information on age.

^bURTI: upper respiratory tract infection.

^cLRTI: lower respiratory tract infection.

Table 4. Number of cases of each syndrome, in total and stratified by reception center (RC; RC1-RC7), recorded by the syndrome-based surveillance system, Austria, September 2015-February 2018.

Name of syndrome	All RCs	RC (by geographical location in Austria)						
		RC1 (east)	RC2 (west)	RC3 (north)	RC4 (south)	RC5 (east)	RC6 (south)	RC7 (west)
Start date	—	September 2015	September 2015	October 2015	December 2015	December 2015	December 2015	April 2016
Suspected acute URTI ^a , n (%)	1470 (50.45)	722 (38.16)	127 (70.6)	250 (71.4)	240 (83.0)	28 (39)	64 (73)	39 (74)
Rash without fever, n (%)	1174 (40.29)	1031 (54.49)	38 (21.1)	55 (15.7)	38 (13.2)	6 (8)	0 (0)	6 (11)
Suspected acute LRTI ^b with fever, n (%)	159 (5.46)	84 (4.44)	5 (2.8)	12 (3.4)	4 (1.4)	26 (36)	22 (25)	6 (11)
Watery diarrhea, n (%)	73 (2.51)	29 (1.53)	10 (5.6)	30 (8.6)	2 (0.7)	8 (11)	2 (2)	2 (4)
Skin, soft tissue, or bone abnormalities, n (%)	32 (1.10)	24 (1.27)	0 (0)	0 (0)	5 (1.7)	3 (4)	0 (0)	0 (0)
Skin rash with fever, n (%)	4 (0.14)	0 (0)	0 (0)	3 (0.9)	0 (0)	1 (1)	0 (0)	0 (0)
Meningitis or encephalitis, n (%)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
High fever with no other signs, n (%)	1 (0.03)	1 (0.05)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Unexplained death, n (%)	1 (0.03)	1 (0.05)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Total (N)	2914	1892	180	350	289	72	88	53

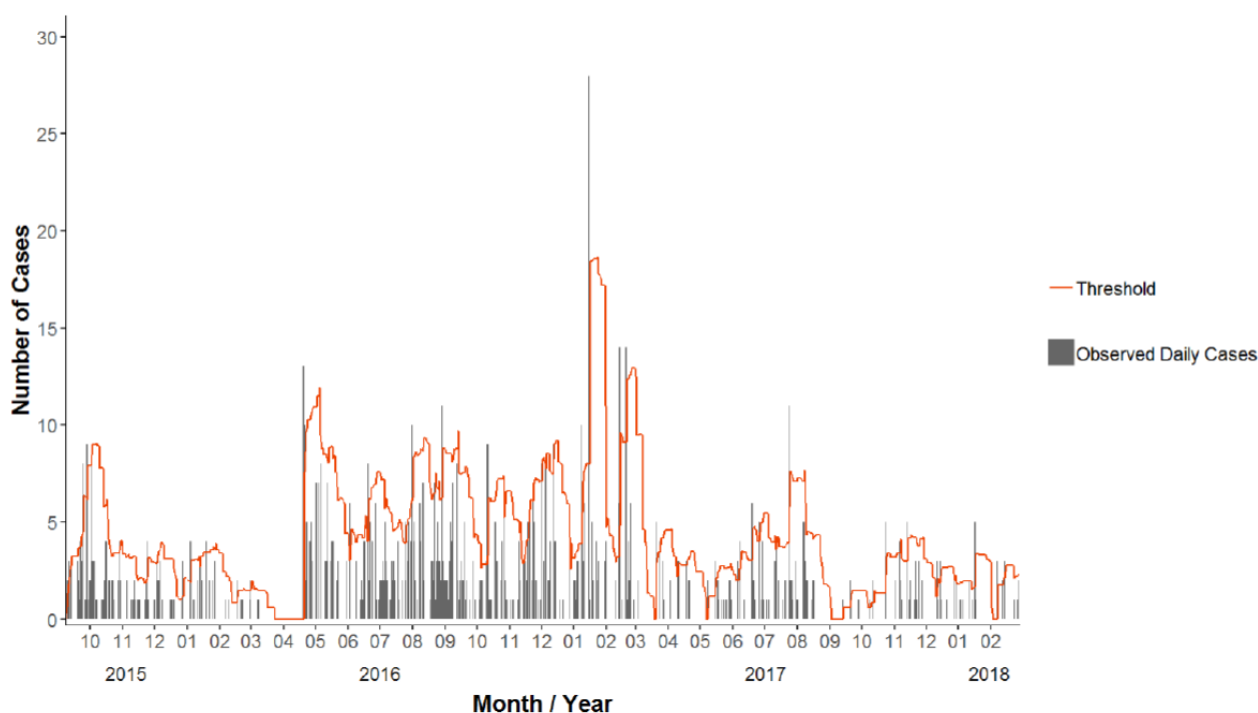
^aURTI: upper respiratory tract infection.^bLRTI: lower respiratory tract infection.**Figure 2.** Observed daily cases of rash without fever, illustrated by gray bars, and threshold of the expected daily cases, shown by the orange line (n=151), 2015-2018, Austria.

Table 5. Number of signals and alerts for each syndrome recorded by the syndrome-based surveillance system, Austria, September 2015-February 2018.

Name of syndrome	Signals	Alerts
Suspected acute URTI ^a , n (%)	151 (37.3)	10 (39)
Rash without fever, n (%)	142 (35.1)	10 (39)
Suspected acute LRTI ^b with fever, n (%)	67 (16.5)	4 (15)
Watery diarrhea, n (%)	27 (6.7)	0 (0)
Skin, soft tissue, or bone abnormalities, n (%)	14 (3.5)	2 (8)
Rash with fever, n (%)	2 (0.5)	0 (0)
High fever ($\geq 39^{\circ}\text{C}$) with no other signs, n (%)	1 (0.3)	0 (0)
Unexplained death, n (%)	1 (0.3)	0 (0)
Total (N)	405	26

^aURTI: upper respiratory tract infection.

^bLRTI: lower respiratory tract infection.

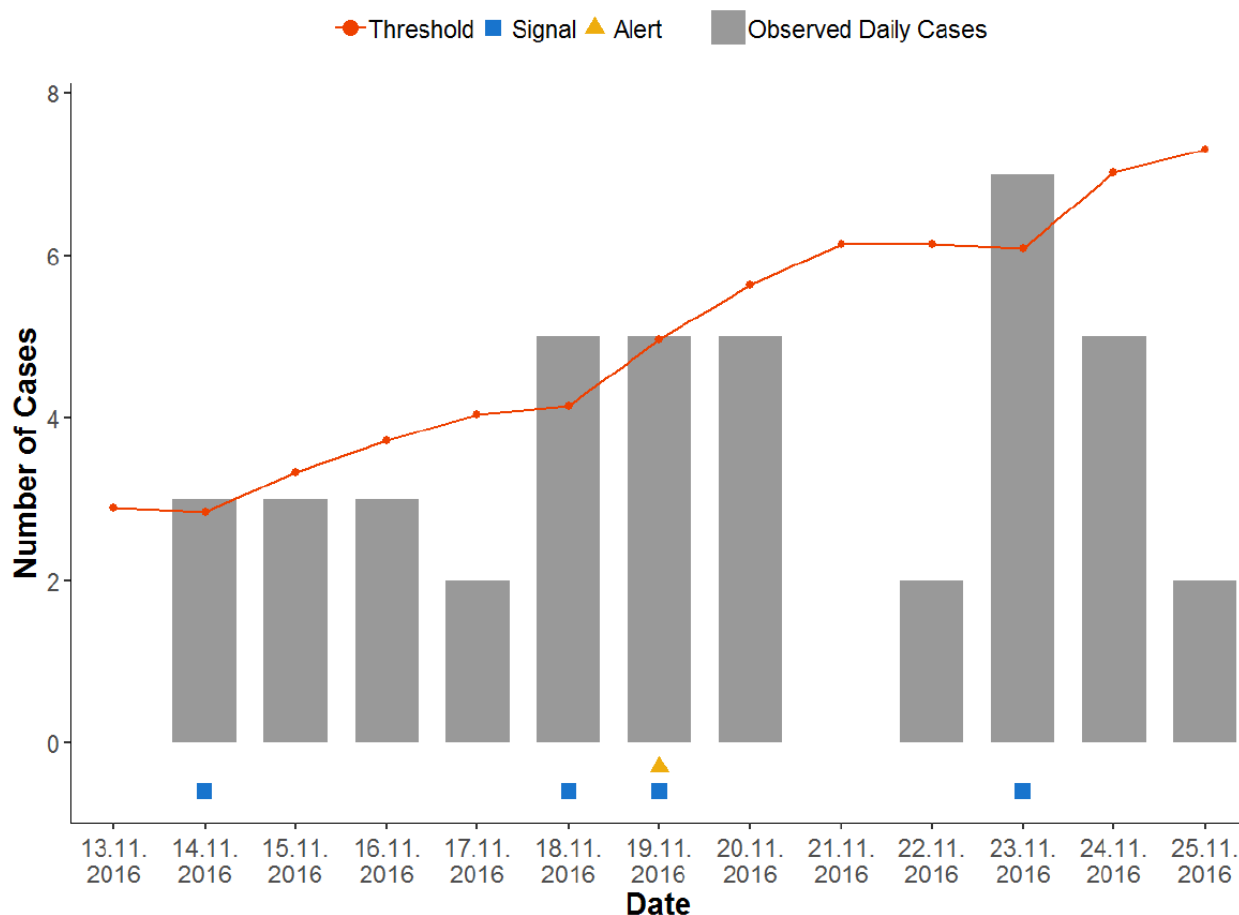
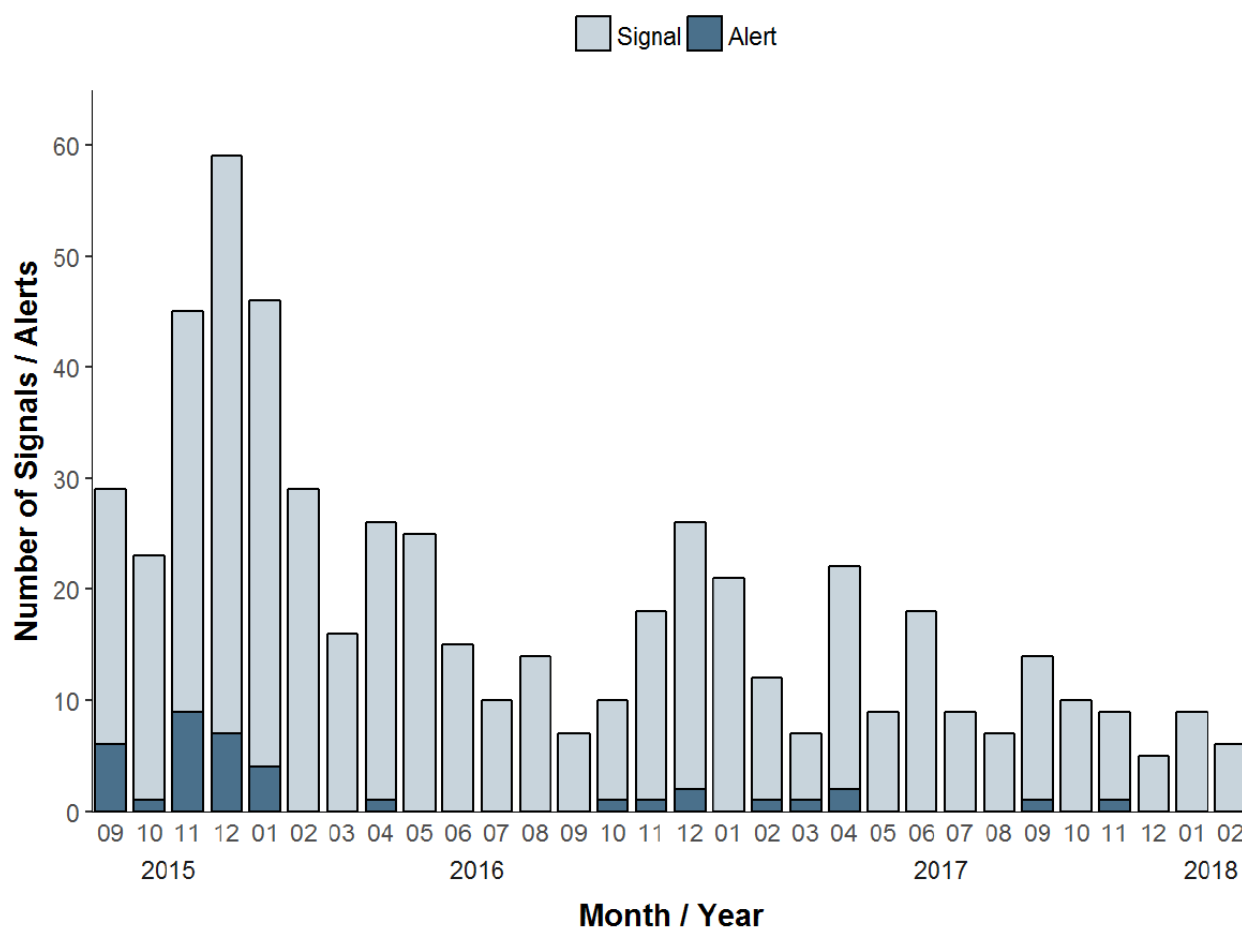
Figure 3. Selected 2 weeks of surveillance, during November 13-25, 2016, illustrating the observed daily cases of rash without fever by bars, threshold of the expected daily cases by orange line, the signals by blue squares (n=4; at November 14, 18, 19, and 23, 2016), and the alerts by a yellow triangle (n=1; at November 19, 2016).

Figure 4. Numbers of signals and alerts generated by the syndrome-based surveillance system for all syndromes, per month, September 2015-February 2018, Austria.



Discussion

Principal Findings

This was the first attempt to use an SbSS in Austria during an unusual influx of asylum-seeking refugees, facing harsh travelling conditions, since World War II. During the total surveillance period from September 2015 to February 2018, more than 400 statistical signals were triggered, caused by 5 predominating syndromes: suspected acute URTI, rash without fever, suspected acute LRTI, watery diarrhea, and skin, soft tissue, or bone abnormalities. The small number of alerts (n=26) indicated sufficient control of infectious disease spread at the 7 Austrian RCs. The signals triggered appropriate and prompt public health actions by the RCs' health care staff and responsible public health authorities: investigations of alerts; exclusion of people suspected to be contagious from crowded activities, such as the RC refectory; contact tracing; scabies-related hygiene measures; intensified cleaning and disinfection in case of environmental contamination; and referral to specialized, secondary health care for adequate treatment. We identified suspected acute URTI as the predominant syndrome among the asylum seekers, which was similar to the experience with refugees arriving at the Greek-Turkish border [18] and among 51 RCs in Greece [19]. Our findings support previous experiences that asylum seekers, compared with the host population, mainly acquire infections with agents present

among the host population, even though they are at higher risk due to many infection-favoring factors.

Interestingly, the second most frequently reported syndrome was rash without fever (scabies-like), for which over half of the signals (1032/1892, 54.49%) were reported in RC1. This was relatively higher than those reported among refugees in other countries, such as Greece [19]. Over half of the bed capacity of all RCs was provided by RC1 (1800/2840, 63.38%), which is considered as a large RC. According to the ECDC, exposure to crowded shelters increases the risk of the spread of lice, fleas, and mites [4]. Watery diarrhea accounted for less than 3% of reported syndromes among the RC residents, which is likely due to strict national food safety and control according to the Austrian Food Safety and Consumer Protection Act. Surprisingly, there was no norovirus outbreak in any of the RCs. A cluster of 21 cases of shigellosis occurred between July and November 2015 among refugees, mainly affecting transit centers in Austria. However, 3 cases of this cluster were identified in 3 of the 7 RCs, which indicates the sensitivity of the Austrian SbSS. In addition, completeness and timeliness, which were measured as system quality indicators, were high. Surprisingly, in RC5, acute URTI accounted for the smallest proportion across all RCs (39% vs 49% among the remaining RCs), whereas the proportion of LRTI was the highest in RC5. The opening of RC5 in December 2015 in the crowded urban setting of Vienna

may explain the high rate of tracheobronchitis; 1 case also turned out to be pulmonary tuberculosis.

Syndrome-Based Surveillance System Implementations in the European Union

Systematic data collection on asylum seekers' health is limited in the EU [20]; therefore, there is a need to enhance national health surveillance systems [21], including being innovative in health surveillance methods to monitor the health of asylum seekers [22-24]. SbSS requires a near real-time automated data collection and analysis system [25]. Therefore, the Austrian SbSS was established in the form of a collaboration, and with a strong commitment, among different national authorities [13], including the Ministry of Interior, the RC health care staff, and the surveillance experts at the AGES. All signals and alerts were addressed on a daily basis through direct contact between the AGES and the focal RC SbSS. Sharing daily reports on syndromes with each RC and communicating signals and alerts immediately to both the focal RC SbSS and the Ministry of Health made appropriate action possible. We have observed that during the rollout of SbSS, the medical staff in nearby hospitals became more aware of the potential risk of communicable diseases among refugees. For example, in 1 of the RCs, located in Salzburg, a case of louse-borne relapsing fever was identified prior to the implementation of the SbSS (DS, unpublished data, 2016).

SbSS has been deployed in other high-income countries to supplement national public health surveillance systems. This includes during the 2006 heat wave in France [25], during the Olympic Games in each of Greece [26] and Italy [27], in Virginia, USA [28] during a national youth camp, in Sweden [29] using a Web-based query system for influenza, and in the United Kingdom [30] for the national telephone health helpline. In Austria, the AGES has introduced SbSS as a supplement to routine tuberculosis screening in selected migrant groups and

to the routine national surveillance of 65 notifiable infectious diseases.

Strengths and Limitations

The AGES had to face some limitations in operating the SbSS. The daily number of asylum seekers registered at the RCs was not reported in a consistent fashion. Thus, we could not calculate the incidence of syndromes among the asylum-seeking population. However, this is considered as a common limitation [16]. Of the 7 RCs, 5 were established on an ad hoc basis as temporary asylum seeker-hosting centers. In these ad hoc RCs, the availability of health care-providing staff was not consistent on a daily basis. This may have caused underdetection and underreporting of syndrome cases, particularly when asylum seekers were immediately transferred to the hospital. We analyzed the reported data in this study without comparison with a reference standard SbSS [30]; however, we need to acknowledge that, in general, SbSSs have a low specificity [31-33].

Strengths of the Austrian SbSS were, first, that data were collected in near real time [31] and that the largest RCs reported on a daily basis, which made the public health response time-efficient. Second, the SbSS had high sensitivity and practicability due to the use of easily ascertainable clinical signs without requiring laboratory testing [32,33]. Third, the Austrian SbSS could be deployed and implemented rapidly in the emergency situation of the most recent refugee crisis [10,22,31]. The SbSS in Austrian RCs proved to be highly eligible for identifying infectious diseases and detecting clusters among the asylum-seeking population.

Conclusion

This was the first time that an SbSS was used in Austria for an increased number of incoming refugees seeking asylum. The SbSS was reliable at identifying and controlling the spread of infectious diseases among the asylum-seeking population from September 2015 onward.

Acknowledgments

We would like to acknowledge and thank (1) all the health care staff at the reception centers and ORS Service, the emergency health services (144-Notruf Niederösterreich) and the Austrian Red Cross for reporting to the SbSS; (2) the Ministry of Health, the local and regional staff of the Ministry of Interior, and the local and regional public health authorities for supporting the SbSS activities; and (3) Dr Birgit Ngedly for supporting and helping in initiating, piloting, training, and implementing the SbSS at the RCs.

Authors' Contributions

KT, DS, LR, and FA developed the initial SbSS surveillance protocol. KT managed the daily operations of the SbSS activities, under the supervision of DS. ZEK, LR, FA, and DS wrote, revised, and approved the manuscript.

Conflicts of Interest

None declared.

References

1. Murray D. UNHCR applauds Austria and Germany as refugees march across Hungary. Geneva, Switzerland: United Nations High Commissioner for Refugees (UNHCR); 2015. URL: <https://www.unhcr.org/news/latest/2015/9/55eae4116/unhcr-applauds-austria-germany-refugees-march-across-hungary.html>

2. Arsenijević J, Schillberg E, Ponthieu A, Malvisi L, Ahmed WAE, Argenziano S, et al. A crisis of protection and safe passage: violence experienced by migrants/refugees travelling along the Western Balkan corridor to Northern Europe. *Confl Health* 2017;11:6 [FREE Full text] [doi: [10.1186/s13031-017-0107-z](https://doi.org/10.1186/s13031-017-0107-z)] [Medline: [28416965](#)]
3. Médecins Sans Frontières. One year after the EU-Turkey deal: migrants and asylum seekers are paying the price with their health. Geneva, Switzerland: MSF; 2017 Mar 14. URL: <https://www.msf.org/one-year-after-eu-turkey-deal-migrants-and-asylum-seekers-are-paying-price-their-health> [accessed 2018-07-02] [WebCite Cache ID 70bi4Nc6M]
4. European Centre for Disease Prevention and Control. Communicable disease risks associated with the movement of refugees in Europe during the winter season. Stockholm, Sweden: ECDC; 2015 Nov 10. URL: <https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/refugee-migrant-health-in-european-winter-rapid-risk-assessment.pdf> [accessed 2019-01-24] [WebCite Cache ID 75feNzf59]
5. Mockenhaupt FP, Barbre KA, Jensenius M, Larsen CS, Barnett ED, Stauffer W, et al. Profile of illness in Syrian refugees: a GeoSentinel analysis, 2013 to 2015. *Euro Surveill* 2016;21(10):30160 [FREE Full text] [doi: [10.2807/1560-7917.ES.2016.21.10.30160](https://doi.org/10.2807/1560-7917.ES.2016.21.10.30160)] [Medline: [26987893](#)]
6. Statistics for the asylum seekers in Austria. Vienna, Austria: Austrian Federal Ministry of the Interior; 2018. URL: <https://www.bmi.gv.at/301/Statistiken/> [accessed 2018-07-02] [WebCite Cache ID 70biaWwry]
7. World Health Organization, Regional Office for Europe. Migration and health: key issues. Copenhagen, Denmark: WHO; 2018. URL: <http://www.euro.who.int/en/health-topics/health-determinants/migration-and-health/migrant-health-in-the-european-region/migration-and-health-key-issues> [accessed 2018-07-02] [WebCite Cache ID 70bjCBPby]
8. Blitz BK, d'Angelo A, Kofman E, Montagna N. Health challenges in refugee reception: dateline Europe 2016. *Int J Environ Res Public Health* 2017 Dec 30;14(12) [FREE Full text] [doi: [10.3390/ijerph14121484](https://doi.org/10.3390/ijerph14121484)] [Medline: [29189766](#)]
9. European Centre for Disease Prevention and Control. Infectious diseases of specific relevance to newly-arrived migrants in the EU/EEA. Stockholm, Sweden: ECDC; 2015 Nov 19. URL: <https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/Infectious-diseases-of-specific-relevance-to-newly-arrived-migrants-in-EU-EEA.pdf> [accessed 2018-07-02] [WebCite Cache ID 70bjM2Kbh]
10. European Centre for Disease Prevention and Control. Expert opinion on the public health needs of irregular migrants, refugees or asylum seekers across the EU's southern and south-eastern borders. Stockholm, Sweden: ECDC; 2015 Oct. URL: <https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/Expert-opinion-irregular-migrants-public-health-needs-Sept-2015.pdf> [accessed 2019-01-24] [WebCite Cache ID 75feTfXK]
11. Jossieran L, Caillère N, Brun-Ney D, Rottner J, Filleul L, Brucker G, et al. Syndromic surveillance and heat wave morbidity: a pilot study based on emergency departments in France. *BMC Med Inform Decis Mak* 2009 Feb 20;9:14 [FREE Full text] [doi: [10.1186/1472-6947-9-14](https://doi.org/10.1186/1472-6947-9-14)] [Medline: [19232122](#)]
12. European Centre for Disease Prevention and Control. Risk of importation and spread of malaria and other vector-borne diseases associated with the arrival of migrants to the EU. Stockholm, Sweden: ECDC; 2015 Oct 21. URL: <https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/risk-malaria-vector-borne-diseases-associated-with-migrants-october-2015.pdf> [accessed 2019-01-24] [WebCite Cache ID 75fenhgpx]
13. Napoli C, Riccardo F, Declich S, Dente MG, Pompa MG, Rizzo C, National Working Group. An early warning system based on syndromic surveillance to detect potential health emergencies among migrants: results of a two-year experience in Italy. *Int J Environ Res Public Health* 2014 Aug 20;11(8):8529-8541 [FREE Full text] [doi: [10.3390/ijerph110808529](https://doi.org/10.3390/ijerph110808529)] [Medline: [25140999](#)]
14. Weinberg M, Waterman S, Lucas CA, Falcon VC, Morales PK, Lopez LA, et al. The U.S.-Mexico Border Infectious Disease Surveillance project: establishing bi-national border surveillance. *Emerg Infect Dis* 2003 Jan;9(1):97-102 [FREE Full text] [doi: [10.3201/eid0901.020047](https://doi.org/10.3201/eid0901.020047)] [Medline: [12533288](#)]
15. Harcourt SE, Fletcher J, Loveridge P, Bains A, Morbey R, Yeates A, et al. Developing a new syndromic surveillance system for the London 2012 Olympic and Paralympic Games. *Epidemiol Infect* 2012 Dec;140(12):2152-2156. [doi: [10.1017/S0950268812001781](https://doi.org/10.1017/S0950268812001781)] [Medline: [22892324](#)]
16. European Centre for Disease Prevention and Control. Handbook on implementing syndromic surveillance in migrant reception/detention centres and other refugee settings. Stockholm, Sweden: ECDC; 2016 Oct. URL: <https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/syndromic-surveillance-migrant-centres-handbook.pdf>
17. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.
18. Eonomopoulou A, Pavli A, Stasinopoulou P, Giannopoulos LA, Tsiodras S. Migrant screening: lessons learned from the migrant holding level at the Greek-Turkish borders. *J Infect Public Health* 2017;10(2):177-184 [FREE Full text] [doi: [10.1016/j.jiph.2016.04.012](https://doi.org/10.1016/j.jiph.2016.04.012)] [Medline: [27161593](#)]
19. Pavli A, Maltezos H. Health problems of newly arrived migrants and refugees in Europe. *J Travel Med* 2017 Dec 01;24(4). [doi: [10.1093/jtm/tax016](https://doi.org/10.1093/jtm/tax016)] [Medline: [28426115](#)]

20. Semenza JC, Carrillo-Santisteve P, Zeller H, Sandgren A, van der Werf MJ, Severi E, et al. Public health needs of migrants, refugees and asylum seekers in Europe, 2015: infectious disease aspects. *Eur J Public Health* 2016 Dec;26(3):372-373. [doi: [10.1093/eurpub/ckw023](https://doi.org/10.1093/eurpub/ckw023)] [Medline: [27053729](#)]
21. Fleischauer AT, Gaines J. Enhancing surveillance for mass gatherings: the role of syndromic surveillance. *Public Health Rep* 2017;132(1_suppl):95S-98S. [doi: [10.1177/0033354917706343](https://doi.org/10.1177/0033354917706343)] [Medline: [28692398](#)]
22. Memish ZA, Zumla A, McCloskey B, Heymann D, Al Rabeeah AA, Barbeschi M, et al. Mass gatherings medicine: international cooperation and progress. *Lancet* 2014 Jun 14;383(9934):2030-2032. [doi: [10.1016/S0140-6736\(14\)60225-7](https://doi.org/10.1016/S0140-6736(14)60225-7)] [Medline: [24857704](#)]
23. Tam JS, Barbeschi M, Shapovalova N, Briand S, Memish ZA, Kieny M. Research agenda for mass gatherings: a call to action. *Lancet Infect Dis* 2012 Mar;12(3):231-239. [doi: [10.1016/S1473-3099\(11\)70353-X](https://doi.org/10.1016/S1473-3099(11)70353-X)] [Medline: [22252148](#)]
24. World Health Organization. Communicable disease alert and response for mass gatherings: key considerations. Geneva, Switzerland: WHO; 2008 Jun. URL: https://www.who.int/csr/Mass_gatherings2.pdf [accessed 2019-01-28] [WebCite Cache ID 75IZYARX1]
25. Flamand C, Larrieu S, Couvy F, Jouvès B, Josseran L, Filleul L. Validation of a syndromic surveillance system using a general practitioner house calls network, Bordeaux, France. *Euro Surveill* 2008 Jun 19;13(25) [FREE Full text] [Medline: [18761939](#)]
26. Dafni UG, Tsiodras S, Panagiotakos D, Gkolfinopoulou K, Kouvatsas G, Tsourti Z, et al. Algorithm for statistical detection of peaks--syndromic surveillance system for the Athens 2004 Olympic Games. *MMWR Suppl* 2004 Sep 24;53:86-94 [FREE Full text] [Medline: [15714635](#)]
27. Epidemiological Consultation Team, Demicheli V, Raso R, Tiberti D, Barale A, Ferrara L, et al. Results from the integrated surveillance system for the 2006 Winter Olympic and Paralympic Games in Italy. *Euro Surveill* 2006 Aug 17;11(8):E060817.5. [Medline: [16966785](#)]
28. Centers for Disease Control and Prevention (CDC). Surveillance for early detection of disease outbreaks at an outdoor mass gathering--Virginia, 2005. *MMWR Morb Mortal Wkly Rep* 2006 Jan 27;55(3):71-74 [FREE Full text] [Medline: [16437057](#)]
29. Hulth A, Rydevik G, Linde A. Web queries as a source for syndromic surveillance. *PLoS One* 2009;4(2):e4378 [FREE Full text] [doi: [10.1371/journal.pone.0004378](https://doi.org/10.1371/journal.pone.0004378)] [Medline: [19197389](#)]
30. Smith GE, Cooper DL, Loveridge P, Chinemana F, Gerard E, Verlander N. A national syndromic surveillance system for England and Wales using calls to a telephone helpline. *Euro Surveill* 2006;11(12):220-224. [Medline: [17370968](#)]
31. Abat C, Chaudet H, Rolain J, Colson P, Raoult D. Traditional and syndromic surveillance of infectious diseases and pathogens. *Int J Infect Dis* 2016 Jul;48:22-28 [FREE Full text] [doi: [10.1016/j.ijid.2016.04.021](https://doi.org/10.1016/j.ijid.2016.04.021)] [Medline: [27143522](#)]
32. Josseran L, Fouillet A, Caillère N, Brun-Ney D, Ille D, Brucker G, et al. Assessment of a syndromic surveillance system based on morbidity data: results from the Oscour network during a heat wave. *PLoS One* 2010 Aug 09;5(8):e11984 [FREE Full text] [doi: [10.1371/journal.pone.0011984](https://doi.org/10.1371/journal.pone.0011984)] [Medline: [20711252](#)]
33. Kleinman KP, Abrams AM. Assessing surveillance using sensitivity, specificity and timeliness. *Stat Methods Med Res* 2006 Oct;15(5):445-464. [doi: [10.1177/0962280206071641](https://doi.org/10.1177/0962280206071641)] [Medline: [17089948](#)]

Abbreviations

AGES: Austrian Agency for Health and Food Safety
ECDC: European Centre for Disease Prevention and Control
EDC: expected daily case
ETH: expected threshold
EU: European Union
LRTI: lower respiratory tract infection
ODC: observed daily case
RC: reception center
SbSS: syndrome-based surveillance system
URTI: upper respiratory tract infection

Edited by Y Khader, M Smolinski; submitted 02.07.18; peer-reviewed by F Lami, K Shatnawi; comments to author 25.10.18; revised version received 08.11.18; accepted 03.12.18; published 27.02.19.

Please cite as:

El-Khatib Z, Taus K, Richter L, Allerberger F, Schmid D

A Syndrome-Based Surveillance System for Infectious Diseases Among Asylum Seekers in Austrian Reception Centers, 2015-2018: Analysis of Reported Data

JMIR Public Health Surveill 2019;5(1):e11465

URL: <http://publichealth.jmir.org/2019/1/e11465/>

doi: [10.2196/11465](https://doi.org/10.2196/11465)

PMID: [30810535](https://pubmed.ncbi.nlm.nih.gov/30810535/)

©Ziad El-Khatib, Karin Taus, Lukas Richter, Franz Allerberger, Daniela Schmid. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 27.02.2019. 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

Evaluating the Usefulness of Translation Technologies for Emergency Response Communication: A Scenario-Based Study

Anne M Turner^{1,2}, MD, MPH, MLIS; Yong K Choi², MPH, PhD; Kristin Dew³, MS; Ming-Tse Tsai², MD, MS; Alyssa L Bosold¹, MPH; Shuyang Wu², MS; Donahue Smith², MHIIM; Hendrika Meischke¹, MPH, PhD

¹School of Public Health, Department of Health Services, University of Washington, Seattle, WA, United States

²School of Medicine, Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, United States

³College of Engineering, Department of Human Centered Design and Engineering, University of Washington, Seattle, WA, United States

Corresponding Author:

Anne M Turner, MD, MPH, MLIS

School of Public Health

Department of Health Services

University of Washington

1959 NE Pacific Street

Seattle, WA,

United States

Phone: 1 206 221 3615

Fax: 1 206 616 9415

Email: amturner@uw.edu

Abstract

Background: In the United States, language barriers pose challenges to communication in emergency response and impact emergency care delivery and quality for individuals who are limited English proficient (LEP). There is a growing interest among Emergency Medical Services (EMS) personnel in using automated translation tools to improve communications with LEP individuals in the field. However, little is known about whether automated translation software can be used successfully in EMS settings to improve communication with LEP individuals.

Objective: The objective of this work is to use scenario-based methods with EMS providers and nonnative English-speaking users who identified themselves as LEP (henceforth referred to as LEP participants) to evaluate the potential of two automated translation technologies in improving emergency communication.

Methods: We developed mock emergency scenarios and enacted them in simulation sessions with EMS personnel and Spanish-speaking and Chinese-speaking (Mandarin) LEP participants using two automated language translation tools: an EMS domain-specific fixed-sentence translation tool (QuickSpeak) and a statistical machine translation tool (Google Translate). At the end of the sessions, we gathered feedback from both groups through a postsession questionnaire. EMS participants also completed the System Usability Scale (SUS).

Results: We conducted a total of 5 group sessions (3 Chinese and 2 Spanish) with 12 Chinese-speaking LEP participants, 14 Spanish-speaking LEP participants, and 17 EMS personnel. Overall, communications between EMS and LEP participants remained limited, even with the use of the two translation tools. QuickSpeak had higher mean SUS scores than Google Translate (65.3 vs 48.4; $P=.04$). Although both tools were deemed less than satisfactory, LEP participants showed preference toward the domain-specific system with fixed questions (QuickSpeak) over the free-text translation tool (Google Translate) in terms of understanding the EMS personnel's questions (Chinese 11/12, 92% vs 3/12, 25%; Spanish 12/14, 86% vs 4/14, 29%). While both EMS and LEP participants appreciated the flexibility of the free-text tool, multiple translation errors and difficulty responding to questions limited its usefulness.

Conclusions: Technologies are emerging that have the potential to assist with language translation in emergency response; however, improvements in accuracy and usability are needed before these technologies can be used safely in the field.

(*JMIR Public Health Surveill* 2019;5(1):e11171) doi:[10.2196/11171](https://doi.org/10.2196/11171)

KEYWORDS

Chinese; Emergency Medical Services; emergency response; language barriers; language translation; public health informatics; Spanish; limited English proficient; translation technologies

Introduction

The United States is linguistically diverse, with over 350 spoken languages [1]. In 2016, approximately 63.2 million US residents spoke a language other than English [2], and approximately 40% of these individuals (25.4 million people) are considered limited English proficient (LEP) [3]. LEP is defined as having a primary language that is not English and limited ability to read, speak, write, or understand English [4].

With growth of the foreign-born population in the United States, the number of LEP individuals is also growing [5]. From 1990 to 2010, the number of LEP individuals in the United States increased by 80%, meaning that in 2010, about 25.2 million or 9% of the US population over the age of 5 years was considered LEP [5].

Health care providers in many parts of the country likely experience challenges with language translation on a frequent basis. In hospital settings, language barriers have contributed to disparities in care for LEP individuals, including longer hospital stays, greater risk of hospital-acquired infections, and increased likelihood of readmission after discharge [6]. In worst-case scenarios, LEP individuals are misdiagnosed and experience serious consequences from improper or delayed treatment [7]. In the emergency response setting, lack of clear communication between LEP individuals and Emergency Medical Services (EMS) personnel can interfere with prompt and accurate dispatching of aid [8]. Language barriers were listed as the second most common reason for delay in care delivery among EMS providers in Minnesota [9]. While the use of interpreters and telephone language lines are recommended in emergency situations involving LEP individuals, time constraints and perceived delays in connecting with interpreters present barriers to their use.

With the advent of new technologies, options for communication with LEP individuals are expanding. A variety of automated translation tools have been developed to assist with translation and interpretation between individuals that have language incongruence. In one study, EMS personnel report using digital applications on their personal mobile devices, such as Google Translate (a freely available Web-based system developed for general translation use), to attempt to communicate with their patients [10]. Many fire departments are also using electronic tools in the field, such as tablets [11]. Access to tablets in the field has opened the door to the use of other translation software. For example, some EMS departments are considering the use of “QuickSpeak,” a tablet-based translation app and one of the few translation tools designed specifically for use in emergency response.

Although digital communication devices could be promising, these tools have not been systematically evaluated for use in the field by EMS, and there is little to no evidence regarding the usefulness of these newer strategies in facilitating communication between LEP individuals and EMS providers.

Our prior work and review of the literature has revealed that in clinical or public health settings, most automated translation systems are not accurate enough to be safely used [12-19].

This study takes place in King County, Washington, where in a recent survey, 78% (96/123) of 911 dispatchers reported that communication difficulties with LEP individuals affect the medical care these callers receive [20]. In Washington, Spanish and Chinese (Mandarin) are the two most commonly spoken non-English languages. Among non-English speakers in the United States, Chinese and Spanish-speaking individuals are also some of the most likely to have limited English proficiency [21]. Many EMS agencies in the King County area are considering the use of automated language tools to improve communication with LEP individuals but are concerned about the safety of these tools in the field. The purpose of this study was to gather evidence on how QuickSpeak and Google Translate (which were both being considered for use by EMS personnel in King County) performed in emergency situations where clear communication is critical for rapid identification, treatment, and transport of patients. Specifically, we tested how QuickSpeak and Google Translate performed in mock emergency response settings requiring prompt EMS response and translation from English to Spanish and Chinese (Mandarin).

Methods

Participants

For our study, we focused on the two most common languages spoken by LEP individuals in the King County area: Spanish and Chinese (Mandarin). We recruited Spanish- and Chinese-speaking individuals whose native language was not English and who self-identified as LEP (henceforth referred to as LEP participants) from local community organizations in King County, Washington. Bilingual research team members collaborated with community organization staff from programs serving LEP individuals in King County to recruit participants. Additionally, we used convenience sampling, through research team members’ personal contacts, to enhance recruitment. To be eligible for the study, LEP participants had to be 18 years or older, speak at least some English but identify themselves as having difficulty communicating in English, and prefer to receive medical care in their native language (Spanish or Chinese) [22]. Given the challenges of recruiting and collaborating with LEP individuals, we sought to minimize the burden of participant screening procedures and did not add a quantitative language assessment instrument to the screening. The University of Washington Institutional Review Board approved all study protocols and materials.

We recruited EMS personnel, including on-duty fire fighters and emergency medical technicians, from local fire departments in King County through convenience sampling. The research team contacted battalion chiefs at local fire departments located in close proximity to communities where there are a large number of LEP residents and asked for permission to recruit

firefighters and conduct a simulation session at their station. We based the number of sessions on when data saturation was reached and additional responses were not forthcoming [23].

Tools

The King County Vulnerable Populations Strategic Initiative [24] team identified two translation tools, QuickSpeak and Google Translate, which EMS personnel were piloting for use (QuickSpeak) or were using on rare occasions (Google Translate). We investigated the potential use of these tools for improving communication between EMS and LEP individuals through simulation sessions involving emergency scenarios.

QuickSpeak

QuickSpeak is an EMS domain-specific translation software that provides EMS personnel access to internally validated, verbal translations of a set of standard English questions asked by first responders. QuickSpeak is one of the few translation tools designed specifically for emergency response. The EMS personnel can select written questions using a touchscreen, and the software provides recorded translations in the requested language. All questions are posed in a yes or no response format. Questions and answers are not recorded or archived. At the time of this study, QuickSpeak could respond in 7 languages: Spanish, Italian, French, German, Finnish, Chinese (Mandarin), and Vietnamese. Figures 1 and 2 present screenshots of QuickSpeak.

Figure 1. Screenshot of QuickSpeak translation software. (Source: www.esosolutions.com).

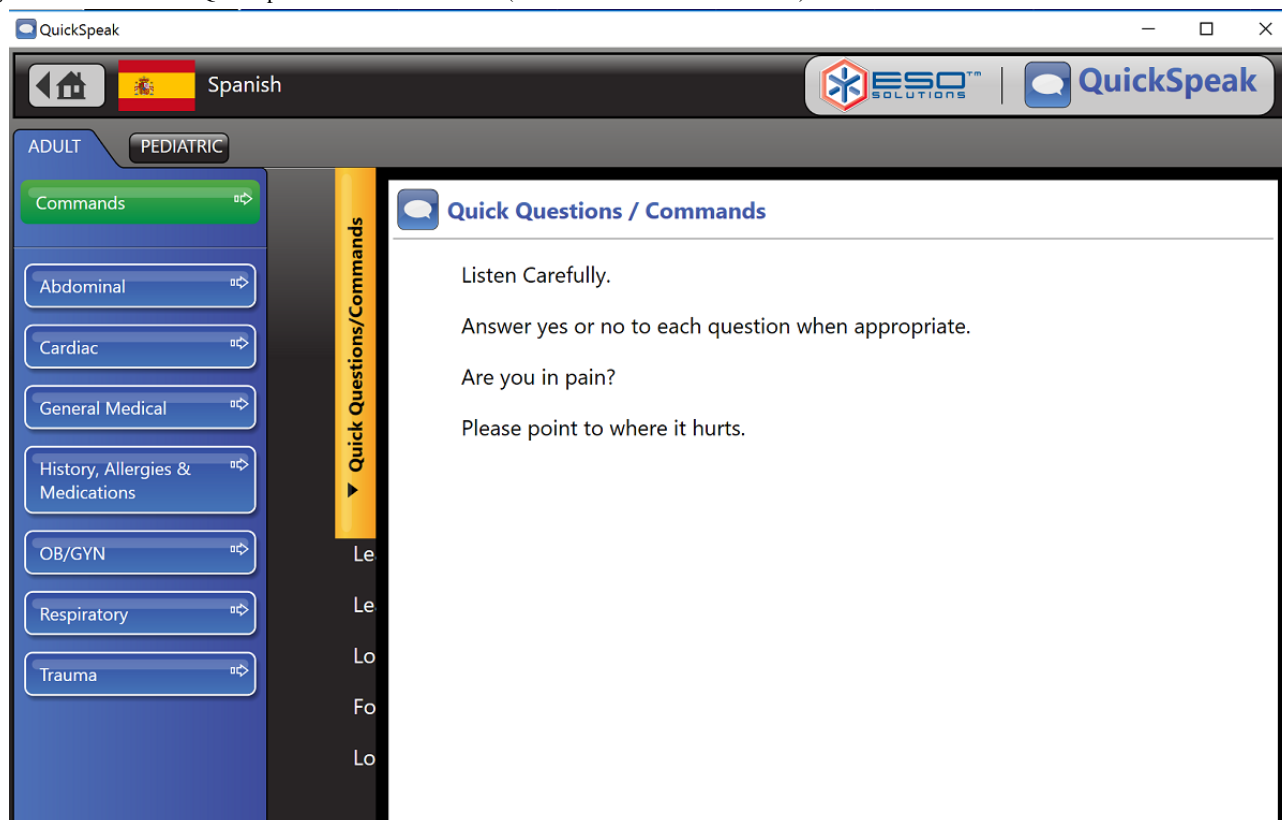
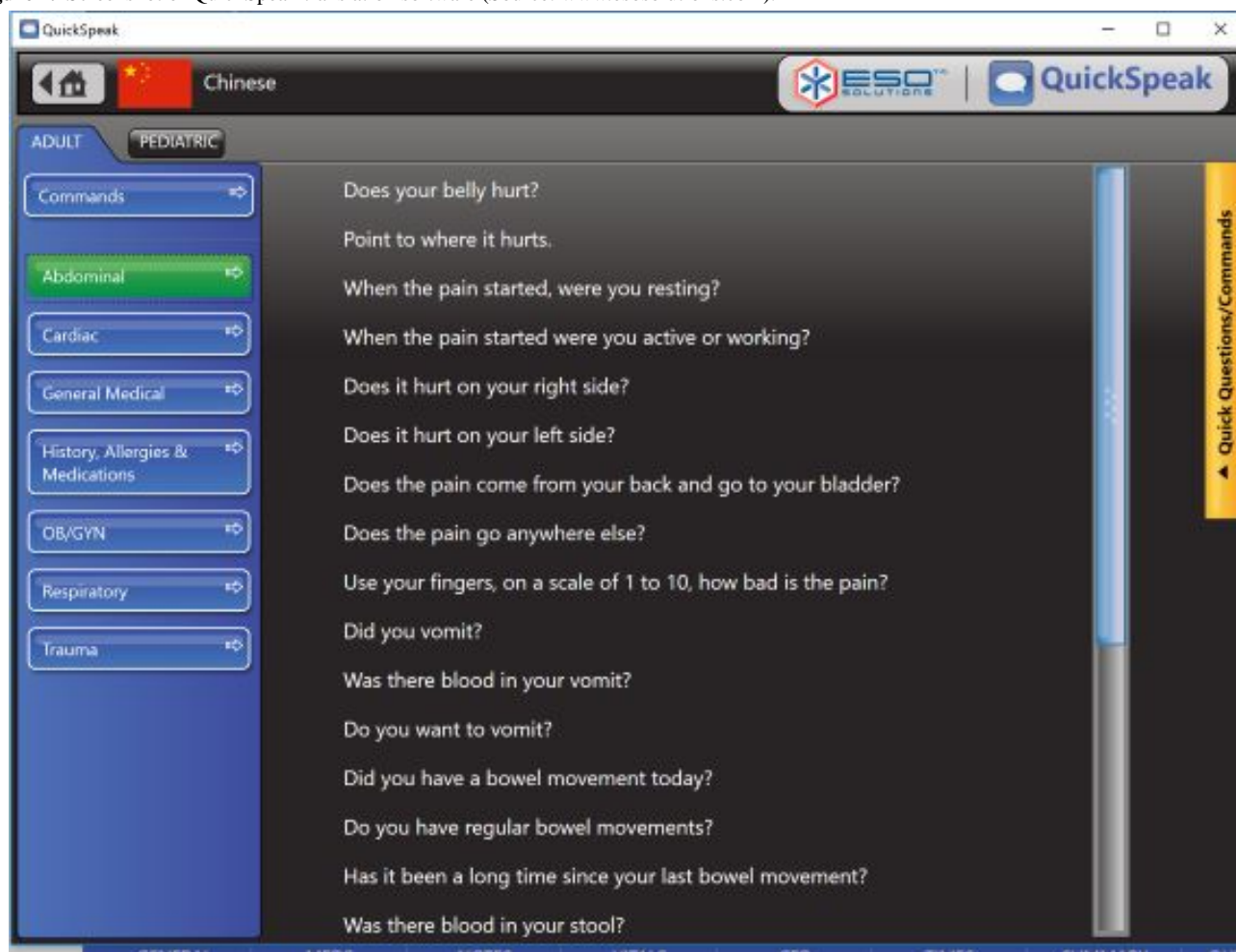
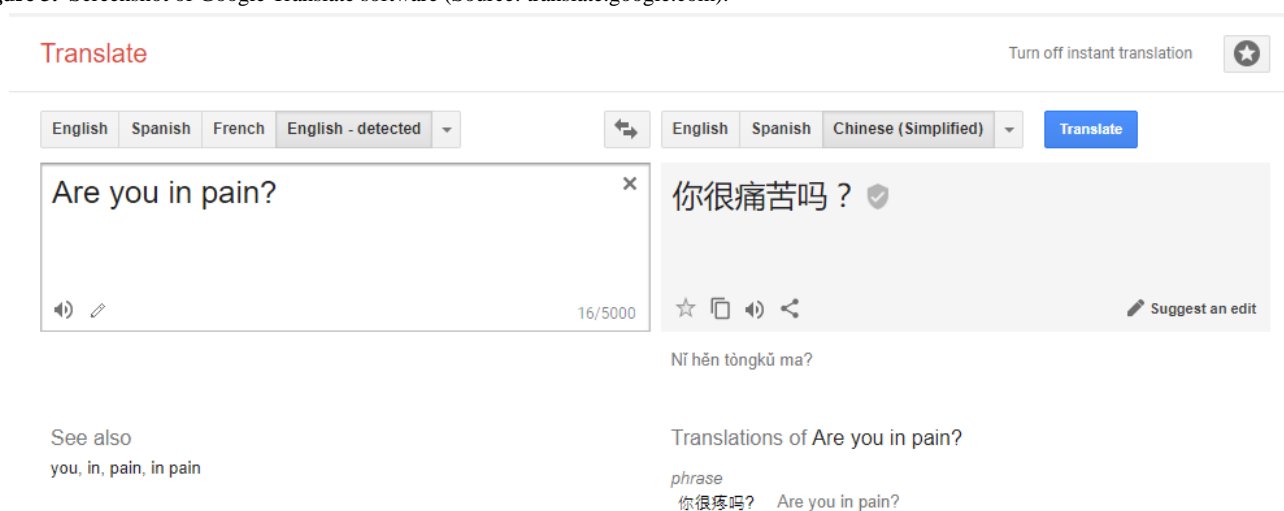


Figure 2. Screenshot of QuickSpeak translation software (Source: www.esosolutions.com).

Google Translate

Google Translate is a free, Web-based and app-based translation software that allows users to write free text in one language and have it converted to written or spoken text in another language. Google Translate utilizes a statistics-based translation (statistical

machine translation) method that produces translations based on their probability of being correct [25]. Currently, Google Translate can translate over 100 languages. Google Translate has been used in many machine translation studies for comparison [26,27]. Figure 3 presents a screenshot of Google Translate.

Figure 3. Screenshot of Google Translate software (Source: translate.google.com).

Study Design

Because field evaluation of new translation tools poses logistical and ethical issues, we drew from scenario-based design to guide this research. Scenario-based design is a key approach to testing and comparing the usefulness of new technologies under “controlled” but realistic conditions [28,29]. In scenario-based design, potential technology users assess the value of technology through participation as actors in realistic, scripted situations. The scripted situations, or scenarios, are developed based on knowledge of actual events, revealed through interviews, focus groups, or observations with the potential technology users.

For our study, 2 research team members (AT, a pediatrician and MT, an emergency medicine physician) created three pairs of scenarios, based on their experience and prior review of transcripts from real-life emergency calls involving LEP.

Scenarios were developed to illustrate a common situation occurring during an EMS response and described the information EMS responders need from LEP individuals. Each scenario described the emergency situation, the “patient” and “support person” (family or friend), precipitating events, the “patient’s” medications and allergies, and basic “patient” demographics, such as age and occupation. We created a pair of similar but not identical scenarios to test and compare the two translation tools.

A battalion chief from a local fire department reviewed the scenarios to ensure that they reflected realistic situations. In response to the review, we made minor modifications. Bilingual research team members translated completed scenarios into traditional Chinese and Spanish. [Multimedia Appendix 1](#) presents the scenario pairs created for this study.

To test the feasibility and usability of QuickSpeak and Google Translate, we held group simulation sessions with Spanish-speaking and Chinese-speaking (Mandarin) LEP participants and EMS personnel at locations convenient for the participants. There were 2 language-appropriate bilingual facilitators who recruited participants and organized the sessions. [Table 1](#) provides a summary of the sessions.

At the beginning of each session, the bilingual facilitator obtained informed consent from LEP participants; collected demographic data including age, education level, number of years in the United States, and self-identified spoken and written English proficiency levels; and explained the overall goals of the evaluation and the language technologies.

Prior to the session, we appraised the LEP participants of the scenario and their role. Working in pairs, one LEP participant

played the role of a “patient,” responding to the EMS provider’s questions with the assistance of either QuickSpeak or Google Translate; the other LEP participant served in the role of a friend or relative “support person.” We provided the EMS participants with information similar to what they would receive from a 911 dispatcher, such as the “patient’s” address, age, chief complaint, and primary language. We did not give the EMS personnel information regarding the underlying health issue that the LEP “patient” was acting out.

Each LEP participant acted in the role of a “patient” or “support person” in scenarios involving each of the two technologies. The EMS personnel sought answers to key questions, such as the chief complaint, symptoms, and medications.

Measures

At the end of the session, EMS personnel and LEP participants filled out postsession questionnaires, providing feedback on the translation technologies (see [Multimedia Appendix 2](#)). The EMS personnel questionnaire included qualitative feedback questions to gather their impressions and experiences with the translation technologies in their own words. For example, it asked them to compare the two technologies (Google Translate and QuickSpeak), identify problems they experienced, and suggest changes. It also collected information on their prior experiences using translation technologies during an emergency.

LEP questionnaires were translated into Spanish and Chinese by native-speaking bilingual research members. The LEP questionnaire asked similar qualitative questions about the participants’ experiences using the translation technology, the problems they encountered, and whether they had ever needed translation during a medical emergency.

The EMS participants were also asked to evaluate the usability of the technologies using a System Usability Scale (SUS) instrument [30]. The SUS generates a quantitative measure of usability through 10 5-point, Likert-type questions, where participants provide their level of agreement or disagreement. The SUS is employed widely for assessing the perceived usability of technologies including mobile apps and monitoring devices for health care [30-33], and it has demonstrated validity, reliability, and sensitivity in numerous studies [34-36]. Since the SUS measures usability, it was only administered to EMS participants, as they were the primary user group operating the translation tools, and LEP participants were not handling the translation technologies and driving the interactions.

LEP participants received a US \$25 honorarium for participation in our study, but as paid professionals, EMS participants could not accept honorariums.

Table 1. Overview of simulation sessions.

Sessions	Location	Limited English proficient participants, n (%)	Emergency Medical Services personnel, n (%)
Spanish #1	Local fire department	6 (23)	6 (35)
Spanish #2	Local fire department	8 (31)	6 (35)
Chinese #1	Research office	4 (15)	1 (6)
Chinese #2	Chinese group home residence	4 (15)	1 (6)
Chinese #3	Local fire department	4 (15)	4 (25)

Data Analysis

We used thematic analysis to examine qualitative responses to open-ended questions on the postsession questionnaires. The practical research question of whether automated language translation tools can facilitate LEP communication in emergency settings drove our thematic analysis. Researchers (YKC, KD, SW, DS) coded the questionnaire responses independently and then met to discuss identified codes and themes. Through several rounds of discussion, we reconciled differences and grouped similar codes to formulate meaningful thematic categories [37].

We conducted descriptive statistical analysis of quantitative data using R software (R Foundation for Statistical Computing) [38]. We also used a Mann-Whitney U test to investigate the relationship between SUS scores and the two technology tools evaluated.

Results

Participants

We held 5 group simulation sessions (3 Chinese and 2 Spanish) with 12 Chinese-speaking LEP participants, 14 Spanish-speaking LEP participants, and 17 EMS personnel. Each session lasted about 1.5-2.5 hours, depending on the number of people in the group. Table 2 summarizes the characteristics of participants. The EMS personnel in the study had a mean age of 44.2 years and an average of 17.8 years of experience.

Table 2. Participants' demographics.

Characteristics	Emergency Medical Services (n=17)	Chinese-speaking (n=12)	Spanish-speaking (n=14)
Age (years), mean (SD)	44.2 (9.6)	46.0 (25.1)	44.7 (16.1)
Years of Emergency Medical Services experience, mean (SD)	17.8 (11.9)	N/A ^a	N/A
Years in the United States, mean (SD)	N/A	7.3 (7.5)	13.9 (8.8)
Education, n (%)			
Less than high school or equivalent	0 (0)	0 (0)	5 (36)
High school graduate or equivalent	2 (12)	5 (42)	4 (29)
Some college or college graduate	13 (76)	2 (17)	3 (21)
Graduate or professional degree	2 (12)	3 (25)	2 (14)
Chose not to answer	0 (0)	2 (17)	0 (0)
Self-reported English level (spoken), n (%)			
Beginner	N/A	4 (33)	3 (21)
Intermediate	N/A	8 (67)	5 (29)
Advanced	N/A	0 (0)	3 (21)
Chose not to answer	N/A	0 (0)	3 (21)
Self-reported English level (written), n (%)			
Beginner	N/A	4 (33)	4 (29)
Intermediate	N/A	8 (67)	5 (36)
Advanced	N/A	0 (0)	2 (14)
Chose not to answer	N/A	0 (0)	3 (21)

^aN/A: not applicable.

The Chinese-speaking LEP participants had a mean age of 46 years and had lived in the United States for an average of 7.3 years. The Spanish-speaking LEP participants had a mean age of 44.7 years and lived in the United States for an average of 13.9 years. Over half of the Chinese-speaking participants and two-thirds of the Spanish-speaking participants identified themselves as having intermediate level English, both spoken and written.

Comparison of QuickSpeak and Google Translate

Postsession Questionnaire

Of 17 EMS respondents, 53% (n=9) indicated that they preferred QuickSpeak over Google Translate. Some EMS personnel (3/17, 18%) stated that they would like a tool that combines features of both technologies. There was 1 EMS participant who said they would not use either system. In the specific follow-up questions (summarized in Table 3), 76% (13/17) of EMS participants stated that QuickSpeak helped them to get the information needed during the simulation session. Fewer participants (10/17, 59%) reported that Google Translate provided the needed information. All but 1 EMS participant noted that QuickSpeak helped them to communicate with LEP participants. In contrast, only approximately two-thirds of respondents mentioned that Google Translate helped them with communication.

Table 3. Emergency Medical Services' ability to obtain the needed information (n=17).

Follow-up question to Emergency Medical Services	QuickSpeak, n (%)	Google Translate, n (%)
Able to get the information needed		
Yes	13 (76)	10 (59)
No	2 (12)	5 (29)
Maybe	0 (0)	1 (6)
Chose not to answer	2 (12)	1 (6)
Helped with communication		
Yes	16 (94)	10 (59)
No	0 (0)	3 (18)
Maybe	1 (6)	3 (18)
Chose not to answer	0 (0)	1 (6)

Table 4. Postsession questionnaire results from limited English proficient participants.

Criteria evaluated	Chinese-speaking (n=12), n (%)		Spanish-speaking (n=14), n (%)	
	QuickSpeak	Google Translate	QuickSpeak	Google Translate
Tool useful overall				
Yes	12 (100)	5 (42)	11 (79)	2 (14)
No	0 (0)	6 (50)	1 (7)	8 (57)
Maybe	0 (0)	1 (8)	2 (14)	4 (29)
Help to understand Emergency Medical Services				
Yes	11 (92)	3 (25)	12 (86)	4 (29)
No	1 (8)	7 (58)	2 (14)	5 (36)
Maybe	0 (0)	2 (17)	0 (0)	5 (36)
Help to speak to Emergency Medical Services				
Yes	9 (75)	3 (25)	11 (85)	6 (43)
No	0 (0)	6 (50)	1 (8)	8 (57)
Maybe	3 (25)	3 (25)	1 (8)	0 (0)

Table 4 summarizes findings from the postsession questionnaire administered to LEP participants. Similar to EMS, both Chinese-speaking and Spanish-speaking LEP participants clearly favored QuickSpeak. When asked about overall usefulness of the tools, all 12 Chinese-speaking LEP participants and 11 of the 14 Spanish-speaking LEP participants noted that QuickSpeak was useful. There was 1 participant who commented on the necessity of such a tool.

Yes, it was useful at the time. It is necessary when there is no interpreter. [Chinese-speaking, P3]

Relatively few LEP participants deemed Google Translate useful (5/12, 42% Chinese-speaking and 2/14, 14% Spanish-speaking). Some participants explained that it took too long for EMS personnel to use Google Translate, and they did not feel confident in the quality of the translation.

[Google Translate is] not useful nor pleasant...I could not communicate what I have or what I need. It takes too long for them to ask questions and it does not feel

safe or that one is being understood.
[Spanish-speaking, P2]

Again, similar to responses from EMS participants, the majority of Chinese-speaking and Spanish-speaking LEP participants (11/12, 92% and 12/14, 86%, respectively) thought QuickSpeak helped them understand the EMS personnel's questions. On the other hand, Google Translate was considered helpful by only 25% (3/12) Chinese-speaking and 29% (4/14) Spanish-speaking LEP participants.

When describing their experience with Google Translate, many LEP participants mentioned difficulty in understanding what was being said due to the poor translation quality, ambiguous meanings, and inappropriate wording.

The more basic ones [questions] yes, but the rest, no. The language, the words or the grammar is not appropriate. The words were not translated correctly.
[Spanish-speaking, P4]

Textbox 1. Summary of limited English proficient participants' feedback on problems encountered during simulation sessions.

QuickSpeak

- restriction on response format (yes or no)
- sound unclear
- slow communication process
- poor quality translation
- difficult to communicate temporal or body position information

Google Translate

- restriction on response format (yes or no)
- unclear questions
- slow communication process
- awkward interaction
- unsafe
- cannot communicate back
- guessing necessary to understand
- poor quality
- difficult to understand
- difficult to ask questions
- culturally inappropriate words
- inconsistent

In addition, some participants mentioned that they had to do a lot of guesswork to make connections between poorly translated words.

Basically I can understand these questions, but with my guesses and understandings. [Chinese-speaking, P6]

The use of inappropriate words was also mentioned in relation to QuickSpeak. Spanish-speaking LEP participants, in particular, mentioned the ambiguity of word choices such as “drinking” (beverages or alcohol) or “drug.”

Yes, but I believe that when it asks, “Were you drinking?” it is too general. The word “drug” can be used in a different manner by different people. It could mean: drug (illegal), medication, medicine, or remedy. [Spanish-speaking, P2]

The LEP participants also commented on whether the tools helped them to respond or speak to EMS personnel. There were 9 of 12 Chinese-speaking participants and 11 of 14 Spanish-speaking participants who mentioned that QuickSpeak helped them speak to EMS personnel. However, only 3 of 12 Chinese-speaking and 6 of 14 Spanish-speaking participants thought that Google Translate helped them. Many participants experienced difficulty communicating detailed responses with both technologies.

When we answered with more than a yes or no, they looked at us with a face of “What?” showing a question mark face. [Chinese-speaking, P8]

During sessions, the research staff noted that when EMS were using the Google Translate tool, they started by typing in open-ended questions. However, because they could not understand the responses given, they evolved to asking more “yes” or “no” questions, similar to QuickSpeak. When typing questions into Google Translate, EMS participants also frequently forgot to add question marks, which affected the interpretation of the translation. Textbox 1 provides a summary of LEP participants' feedback on the problems encountered during their simulation experience.

As primary users of the tools, the EMS participants were asked to provide feedback on usability and recommendations for improving each of the tools. For Google Translate, some EMS personnel mentioned that having a list of predefined questions (as with QuickSpeak) would be helpful. Specifically, they suggested that 4-5 essential questions be placed on the home screen. Some EMS personnel recommended increasing the size of the speaker button, which when clicked plays translated audio, for better usability. There was 1 EMS participant who mentioned that a better translation accent would aid comprehension. Another recommended that Google Translate create a medical domain-specific translation service.

For the QuickSpeak tool, many EMS personnel suggested adding the ability to type and verbalize their own questions, similar to the free-text ability of Google Translate. Some recommended that the list of predetermined questions follow a more logical flow of normal questioning. Some specific suggestions were: use a decision tree to assist with selecting appropriate questions, show a full-body image on the screen allowing EMS personnel

to click body parts and view relevant questions, remove questions from the list once they have been asked, allow EMS personnel to add or modify existing questions loaded in QuickSpeak, and expand the number of languages translated.

Some recommendations applied to both technologies. EMS personnel recommended that both services support voice-operated, two-way translation or communication. They also suggested that actual interactions with LEP patients be audio-recorded for record keeping, education, and accountability. A summary of EMS feedback is provided in [Textbox 2](#).

System Usability Scale Score Evaluation

[Table 5](#) shows the results of the SUS for QuickSpeak and Google Translate. The SUS was only administered to EMS participants, as they were the primary user group operating the translation tools. The mean SUS score for QuickSpeak was higher than the score for Google Translate. The difference between the two translation tools was statistically significant (Mann-Whitney U test, $z=-2.1$; $P=.04$). The results were similar for the Chinese and Spanish simulation sessions. However, EMS personnel who participated in the Chinese sessions rated Google Translate higher than EMS personnel who participated in the Spanish sessions.

Textbox 2. Emergency Medical Services personnel feedback on problems encountered during simulation sessions.

QuickSpeak	
•	restriction on response format (yes or no)
•	cannot create own questions
•	need more questions
•	difficult to find a question
•	not able to type own questions
•	poor organization of question flow
•	questions not specific
•	low sound volume
•	difficult to solicit temporal information
•	too many questions
Google Translate	
•	difficulty understanding limited English proficient responses unless questions posed in yes or no format
•	poor quality translation
•	slow communication process
•	difficult to compose questions
•	too much attention directed to the screen

Table 5. System Usability Scale scores.

Language group	Tool	System Usability Scale Score ^a		
		Mean (SD)	Median	P value
Overall	QuickSpeak	65.3 (13.7)	65.0	.04
	Google Translate	48.4 (25.6)	50.0	
Chinese-speaking participants	QuickSpeak	63.6 (9.8)	65.0	.8
	Google Translate	56.1 (28.1)	62.5	
Spanish-speaking participants	QuickSpeak	66.5 (16.3)	70.0	.02
	Google Translate	43.0 (23.7)	42.5	

^aMaximum possible score is 100.

Discussion

Principal Findings

Our scenario-based evaluation of the two translation tools confirmed a need for better tools to assist in communication

between LEP individuals and EMS personnel during medical emergencies. This is consistent with prior study findings showing that EMS personnel often experience frustration with telephone language interpreters and resort to talking with bystanders or using body language and keywords they happen

to know [39]. Studies show that LEP individuals who have ad hoc interpreters (bystanders or family members) experience more dissatisfaction than those with medically trained interpreters and are more likely to experience errors that could impact clinical care [40]. Improving translation options is key to overcoming communication barriers and increasing the quality of emergency care for LEP individuals.

Although neither translation tool was considered ideal for use in the field, in our scenarios, LEP and EMS participants clearly preferred the fixed question translations of QuickSpeak over the free-text translation of Google Translate. Both LEP and EMS participants thought the ideal translation tool would have the accuracy and clarity of prefixed questions but the flexibility and potential bidirectional communication of the free-text tools. Google Translate currently allows for bidirectional communication, but inaccuracies in translation were often compounded with the existing system. Communication between EMS and LEP individuals was marginally improved with Google Translate and QuickSpeak; however, inaccuracies and potential for miscommunication are great, and neither tool was considered ready for use in the field, where the risks of any miscommunication or delay are high. Our prior research investigating the potential use of freely available Web-based translation systems, such as Google Translate and Microsoft Translator, indicates that in the area of health, these tools require careful postediting by professional translators to improve accuracy [18,41]. Obviously, this kind of post-editing would not be possible in the field. However, machine translation technology is constantly improving. Google has recently updated their translation system to utilize sophisticated artificial intelligence to produce more accurate language translations [42]. Further evaluation of the use of these tools in health settings is needed as automated language translation technology evolves.

In addition to improving translation accuracy, future translation technologies for emergency response should give particular attention to the needs and design recommendations of LEP individuals and EMS personnel. The EMS personnel voiced concern about operating a device and using the translation technology in the fast-paced, real-life emergency setting. Hands-free, accurate speech recognition technology that can facilitate bidirectional communication would be ideal. EMS participants also expressed a desire for translation answers to be recorded for use in documenting encounters. Archiving of answers associated with personal health information, however, would require careful consideration of Health Insurance Portability and Accountability Act of 1996 policies and rules.

Currently, bilingual staff, interpreters, and language lines that can provide real-time, accurate translations are the gold standard for translation and interpretation in clinical settings. Prior studies have shown that LEP patients experience more satisfaction with telephone interpreters than with family, ad hoc, or no interpreters [40]. However, in informal interviews with EMS personnel, there is a hesitancy to use language lines because of concerns that they take too long and that phone-based interpreters may not accurately assess the situation at hand. Despite these perceived barriers, in the absence of in-person interpreters, language lines continue to be the best method of ensuring fast and accurate translations. As automated language translation technologies continue to evolve, further evaluations will be needed to assess whether they can provide safe, effective communication between English and non-English speakers in the field of emergency response.

Limitations

The simulation sessions were limited in number, and all took place in King County, Washington, which may limit the generalizability of our results. We evaluated Spanish and Mandarin Chinese, the two most common non-English languages spoken in our region. In general, Google Translate performs better with Spanish than with Chinese or lesser used languages [18,41,43], so it is likely that results may have been different if we had tested different languages. In addition, our study took place at one point in time, and translation tools using statistical machine translation are constantly evolving. Our evaluation of the translation tools was based on simulation sessions between EMS and LEP participants. Although we used scenarios based on actual EMS responses, these sessions took place in a controlled environment and do not accurately reflect the performance of these tools in the field.

New Contributions to the Literature

In the context of a growing LEP population in the United States and disparities in medical care resulting from language barriers, improving translation technologies is critical. In a recent review, Tate (2015) concludes that there are “substantial gaps in understanding the interaction between language barriers and prehospital care” [44]. Our study sheds light on the challenges of the use of new translation technologies in the prehospital, emergency care setting. While there is a significant need for translation tools to assist in translations in these settings, we need to continue to evaluate automated translation technologies as they evolve, to determine how they compare to more traditional phone interpreter services in terms of acceptability, accuracy, and efficiency.

Acknowledgments

The authors wish to acknowledge the Medic One Foundation grant that supported this research. This work was also supported in part by the National Institutes of Health, National Library of Medicine, and Biomedical and Health Informatics Training Program at the University of Washington (Grant No. T15LM007442). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors would like to thank the EMS personnel and the Spanish and Chinese native-speaking community members that generously participated in the study. The authors also thank Xiyao Yang for her help in conducting the research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Scenario pairs used for the simulation sessions.

[PDF File (Adobe PDF File), 24KB - [publichealth_v5i1e11171_app1.pdf](#)]

Multimedia Appendix 2

Post-Session Questionnaire.

[PDF File (Adobe PDF File), 20KB - [publichealth_v5i1e11171_app2.pdf](#)]

References

1. United States Census Bureau. 2015 Nov 3. Census Bureau Reports at Least 350 Languages Spoken in U.S. Homes URL: <http://www.census.gov/newsroom/press-releases/2015/cb15-185.html> [accessed 2018-04-30] [WebCite Cache ID 6z4bjXOI2]
2. United States Census Bureau. Selected Social Characteristics in the United States URL: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_DP02&prodType=table [accessed 2018-05-25] [WebCite Cache ID 6z4cRbOUb]
3. United States Census Bureau. American Community Survey: American Fact Finder URL: <https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk> [accessed 2018-05-25] [WebCite Cache ID 6z4cJ43pj]
4. Limited English Proficiency. 2016. URL: <https://www.lep.gov/faqs/faqs.html> [accessed 2018-12-04] [WebCite Cache ID 74PmG1dSH]
5. Pandya C, McHugh M, Batalova J. Migration Policy Institute. 2011 Dec. Limited English Proficient Individuals in the United States: Number, Share, Growth, and Linguistic Diversity URL: <https://www.migrationpolicy.org/research/limited-english-proficient-individuals-united-states-number-share-growth-and-linguistic> [accessed 2018-04-30] [WebCite Cache ID 6z4cYNw7M]
6. Agency FHR. Agency for Healthcare Research and Quality. 2012. Improving Patient Safety Systems for Patients with Limited English Proficiency URL: <https://www.ahrq.gov/professionals/systems/hospital/lepguide/lepguidesumm.html#what> [accessed 2018-05-25] [WebCite Cache ID 6z4cgb33t]
7. Flores G. Language barriers to health care in the United States. *N Engl J Med* 2006 Jul 20;355(3):229-231. [doi: [10.1056/NEJMp058316](https://doi.org/10.1056/NEJMp058316)] [Medline: [16855260](https://pubmed.ncbi.nlm.nih.gov/16855260/)]
8. Meischke HW, Calhoun RE, Yip M, Tu S, Painter IS. The effect of language barriers on dispatching EMS response. *Prehosp Emerg Care* 2013;17(4):475-480. [doi: [10.3109/10903127.2013.811565](https://doi.org/10.3109/10903127.2013.811565)] [Medline: [23952940](https://pubmed.ncbi.nlm.nih.gov/23952940/)]
9. Grow RW, Sztajnkrycer MD, Moore BR. Language barriers as a reported cause of prehospital care delay in Minnesota. *Prehosp Emerg Care* 2008 Aug;12(1):76-79. [doi: [10.1080/10903120701709878](https://doi.org/10.1080/10903120701709878)] [Medline: [18189182](https://pubmed.ncbi.nlm.nih.gov/18189182/)]
10. Tate RC, Hodkinson PW, Meehan-Coussee K, Cooperstein N. Strategies Used by Prehospital Providers to Overcome Language Barriers. *Prehosp Emerg Care* 2016 Jan;20(3):404-414. [doi: [10.3109/10903127.2015.1102994](https://doi.org/10.3109/10903127.2015.1102994)] [Medline: [26807873](https://pubmed.ncbi.nlm.nih.gov/26807873/)]
11. Yamada KC, Inoue S, Sakamoto Y. An effective support system of emergency medical services with tablet computers. *JMIR Mhealth Uhealth* 2015 Feb 27;3(1):e23 [FREE Full text] [doi: [10.2196/mhealth.3293](https://doi.org/10.2196/mhealth.3293)] [Medline: [25803096](https://pubmed.ncbi.nlm.nih.gov/25803096/)]
12. Taylor R, Crichton N, Moulton B, Gibson F. A prospective observational study of machine translation software to overcome the challenge of including ethnic diversity in healthcare research. *Nurs Open* 2015 Dec;2(1):14-23 [FREE Full text] [doi: [10.1002/nop2.13](https://doi.org/10.1002/nop2.13)] [Medline: [27708797](https://pubmed.ncbi.nlm.nih.gov/27708797/)]
13. Zeng-Treitler Q, Kim H, Rosembat G, Keselman A. Can multilingual machine translation help make medical record content more comprehensible to patients? *Stud Health Technol Inform* 2010;160(Pt 1):73-77. [Medline: [20841653](https://pubmed.ncbi.nlm.nih.gov/20841653/)]
14. Ehsani F, Kimsey J, Zuber E, Master D. Speech to speech translation for nurse patient interaction. 2008 Presented at: Proceedings of the 22nd International Conference on Computational Linguistics; 2008; Manchester, UK p. 54-59.
15. Chen X, Acosta S, Barry AE. Evaluating the Accuracy of Google Translate for Diabetes Education Material. *JMIR Diabetes* 2016 Jun 28;1(1):e3. [doi: [10.2196/diabetes.5848](https://doi.org/10.2196/diabetes.5848)]
16. Abreu R, Adriatico T. Spanish for the Audiologist: Is There an App for That? *Perspect Comm Dis Sci CLD Pop* 2015 Dec 01;22(3):122-128. [doi: [10.1044/cds22.3.122](https://doi.org/10.1044/cds22.3.122)]
17. Albrecht U, Behrends M, Schmeer R, Matthies HK, von Jan U. Usage of multilingual mobile translation applications in clinical settings. *JMIR Mhealth Uhealth* 2013 Apr 23;1(1):e4 [FREE Full text] [doi: [10.2196/mhealth.2268](https://doi.org/10.2196/mhealth.2268)] [Medline: [25100677](https://pubmed.ncbi.nlm.nih.gov/25100677/)]
18. Turner AM, Dew KN, Desai L, Martin N, Kirchhoff K. Machine Translation of Public Health Materials From English to Chinese: A Feasibility Study. *JMIR Public Health Surveill* 2015;1(2):e17 [FREE Full text] [doi: [10.2196/publichealth.4779](https://doi.org/10.2196/publichealth.4779)] [Medline: [27227135](https://pubmed.ncbi.nlm.nih.gov/27227135/)]

19. Randhawa G, Ferreyra M, Ahmed R, Ezzat O, Pottie K. Using machine translation in clinical practice. *Can Fam Physician* 2013 Apr;59(4):382-383 [FREE Full text] [Medline: 23585608]
20. Meischke H, Chavez D, Bradley S, Rea T, Eisenberg M. Emergency communications with limited-English-proficiency populations. *Prehosp Emerg Care* 2010;14(2):265-271. [doi: 10.3109/10903120903524948] [Medline: 20095823]
21. Zonng J, Batalova J. Migration Policy Institute. 2015 Jul 8. The limited English proficient population in the United States URL: <https://www.migrationpolicy.org/article/limited-english-proficient-population-united-states#Language%20Diversity> [accessed 2018-05-25] [WebCite Cache ID 74QlkVQUl]
22. Karliner LS, Napoles-Springer AM, Schillinger D, Bibbins-Domingo K, Pérez-Stable EJ. Identification of limited English proficient patients in clinical care. *J Gen Intern Med* 2008 Oct;23(10):1555-1560 [FREE Full text] [doi: 10.1007/s11606-008-0693-y] [Medline: 18618200]
23. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant* 2018 Sep 14;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s1135-017-0574-8] [Medline: 29937585]
24. Public Health- Seattle & King County. Vulnerable Populations Strategic Initiative URL: <https://www.kingcounty.gov/depts/health/emergency-medical-services/vulnerable-populations.aspx> [accessed 2018-05-26] [WebCite Cache ID 6zgQxySt9]
25. Google Translate. URL: <https://translate.google.com/> [accessed 2018-12-05] [WebCite Cache ID 74Qm4uaG9]
26. Dew KN, Turner AM, Choi YK, Bosold A, Kirchhoff K. Development of machine translation technology for assisting health communication: A systematic review. *J Biomed Inform* 2018 Sep;85:56-67. [doi: 10.1016/j.jbi.2018.07.018] [Medline: 30031857]
27. Chen X, Acosta S, Barry AE. Machine or Human? Evaluating the Quality of a Language Translation Mobile App for Diabetes Education Material. *JMIR Diabetes* 2017 Jun 29;2(1):e13 [FREE Full text] [doi: 10.2196/diabetes.7446] [Medline: 30291084]
28. Rosson M, Carroll J. Scenario-based usability engineering. In: *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, techniques*. New York, USA: ACM; 2002 Presented at: 4th conference on Designing interactive systems: processes, practices, methods, techniques; June 25-28th; London, UK p. 413. [doi: 10.1145/778712.778776]
29. Reeder B, Turner AM. Scenario-based design: a method for connecting information system design with public health operations and emergency management. *J Biomed Inform* 2011 Dec;44(6):978-988 [FREE Full text] [doi: 10.1016/j.jbi.2011.07.004] [Medline: 21807120]
30. Brooke J. SUS: A "quick and dirty" usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL. *Usability evaluation in industry*. London: Taylor & Francis; 1996:189-194.
31. Harte R, Quinlan L, Glynn L, Rodríguez-Molinero A, Baker P, Scharf T, et al. Human-Centered Design Study: Enhancing the Usability of a Mobile Phone App in an Integrated Falls Risk Detection System for Use by Older Adult Users. *JMIR Mhealth Uhealth* 2017 May 30;5(5):e71 [FREE Full text] [doi: 10.2196/mhealth.7046] [Medline: 28559227]
32. Mohamad Marzuki MF, Yaacob N, Yaacob N. Translation, Cross-Cultural Adaptation, and Validation of the Malay Version of the System Usability Scale Questionnaire for the Assessment of Mobile Apps. *JMIR Hum Factors* 2018 May 14;5(2):e10308 [FREE Full text] [doi: 10.2196/10308] [Medline: 29759955]
33. Alnasser A, Kyle J, Alkhalifah A, Marais D. Relationship Between Evidence Requirements, User Expectations, and Actual Experiences: Usability Evaluation of the Twazon Arabic Weight Loss App. *JMIR Hum Factors* 2018 Apr 17;5(2):e16 [FREE Full text] [doi: 10.2196/humanfactors.9765] [Medline: 29666042]
34. Sauro J. Measuring Usability with the System Usability Scale (SUS). 2011 Feb 2. URL: <https://measuringu.com/sus/> [accessed 2018-04-30] [WebCite Cache ID 6z4cpkcIb]
35. Lewis J, Sauro J. The factor structure of the system usability scale. NY: Springer; 2009 Presented at: *Proceedings of the International Conference on Human Centered Design*; July 19-24; San Diego, CA p. 94-103.
36. Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction* 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
37. Berg B, Lune H. *Qualitative Research Methods for the Social Sciences*. London, UK: Pearson; 2011.
38. The R Foundation. The R Project for Statistical Computing URL: <https://www.r-project.org/> [accessed 2018-05-25] [WebCite Cache ID 6z4cwtetJ]
39. Tate RC, Hodgkinson PW, Meehan-Coussee K, Cooperstein N. Strategies Used by Prehospital Providers to Overcome Language Barriers. *Prehosp Emerg Care* 2016 Jan;20(3):404-414. [doi: 10.3109/10903127.2015.1102994] [Medline: 26807873]
40. Flores G. The impact of medical interpreter services on the quality of health care: a systematic review. *Med Care Res Rev* 2005 Jun;62(3):255-299. [doi: 10.1177/1077558705275416] [Medline: 15894705]
41. Kirchhoff K, Turner AM, Axelrod A, Saavedra F. Application of statistical machine translation to public health information: a feasibility study. *J Am Med Inform Assoc* 2011;18(4):473-478 [FREE Full text] [doi: 10.1136/amiajnl-2011-000176] [Medline: 21498805]
42. Wu Y, Schuster M, Chen Z, Quoc V, Norouzi M, Macherey W, et al. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. *arXiv* 2016.

43. Wu C, Xia F, Deleger L, Solti I. Statistical machine translation for biomedical text: are we there yet. 2011 Presented at: AMIA Annual Symposium; Oct 22-26; Washington DC p. 1290-1299.
44. Tate R. The Need for More Prehospital Research on Language Barriers: A Narrative Review. West J Emerg Med 2015 Dec 17;16(7):1094-1105 [FREE Full text] [doi: [10.5811/westjem.2015.8.27621](https://doi.org/10.5811/westjem.2015.8.27621)] [Medline: [26759662](https://pubmed.ncbi.nlm.nih.gov/26759662/)]

Abbreviations

EMS: Emergency Medical Services

LEP: limited English proficient

SMT: statistical machine translation

SUS: System Usability Scale

Edited by T Sanchez; submitted 31.05.18; peer-reviewed by C Latulipe, X Chen; comments to author 04.09.18; revised version received 10.10.18; accepted 09.11.18; published 28.01.19.

Please cite as:

Turner AM, Choi YK, Dew K, Tsai MT, Bosold AL, Wu S, Smith D, Meischke H

Evaluating the Usefulness of Translation Technologies for Emergency Response Communication: A Scenario-Based Study

JMIR Public Health Surveill 2019;5(1):e11171

URL: <http://publichealth.jmir.org/2019/1/e11171/>

doi: [10.2196/11171](https://doi.org/10.2196/11171)

PMID: [30688652](https://pubmed.ncbi.nlm.nih.gov/30688652/)

©Anne M Turner, Yong K Choi, Kristin Dew, Ming-Tse Tsai, Alyssa L Bosold, Shuyang Wu, Donahue Smith, Hendrika Meischke. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 28.01.2019. 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/>