
JMIR Public Health and Surveillance

Impact Factor (2023): 3.5
Volume 5 (2019), Issue 2 ISSN 2369-2960 Editor in Chief: Travis Sanchez, PhD, MPH

Contents

Viewpoint

- Programmatic Mapping: Providing Evidence for High Impact HIV Prevention Programs for Female Sex Workers ([e12636](#))
Faran Emmanuel, Navindra Persaud, Sharon Weir, Parinita Bhattacharjee, Shajy Isac. 4

Original Papers

- Pneumococcal Vaccination Utilization Among Hispanic Long-Term Colorectal Cancer Survivors: Cross-Sectional Assessment of Claims ([e12603](#))
Ryan Moran, Jill Waalen, James Murphy, Vinit Nalawade, Melody Schiaffino. 12
- New York City HIV Care Continuum Dashboards: Using Surveillance Data to Improve HIV Care Among People Living With HIV in New York City ([e13086](#))
Sarah Braunstein, Karen Coeytaux, Charulata Sabharwal, Qiang Xia, Rebekkah Robbins, Beverly Obeng, Demetre Daskalakis. 25
- Social Media Surveillance for Outbreak Projection via Transmission Models: Longitudinal Observational Study ([e11615](#))
Anahita Safarishahrbijari, Nathaniel Osgood. 34
- Improved Real-Time Influenza Surveillance: Using Internet Search Data in Eight Latin American Countries ([e12214](#))
Leonardo Clemente, Fred Lu, Mauricio Santillana. 45
- Identifying Key Topics Bearing Negative Sentiment on Twitter: Insights Concerning the 2015-2016 Zika Epidemic ([e11036](#))
Ravali Mamidi, Michele Miller, Tanvi Banerjee, William Romine, Amit Sheth. 67
- Preliminary Flu Outbreak Prediction Using Twitter Posts Classification and Linear Regression With Historical Centers for Disease Control and Prevention Reports: Prediction Framework Study ([e12383](#))
Ali Alessa, Miad Faezipour. 85
- “Where There’s Smoke, There’s Fire”: A Content Analysis of Print and Web-Based News Media Reporting of the Philip Morris–Funded Foundation for a Smoke-Free World ([e14067](#))
Christina Watts, Becky Freeman. 102

Google for Sexual Relationships: Mixed-Methods Study on Digital Flirting and Online Dating Among Adolescent Youth and Young Adults (e10695)	
James Lykens, Molly Pilloton, Cara Silva, Emma Schlamm, Kate Wilburn, Emma Pence.	112
#PrEP4Love: An Evaluation of a Sex-Positive HIV Prevention Campaign (e12822)	
Jessica Dehlin, Ryan Stillwagon, Jim Pickett, Lance Keene, John Schneider.	140
Factors Associated With Willingness to Use Pre-Exposure Prophylaxis in Brazil, Mexico, and Peru: Web-Based Survey Among Men Who Have Sex With Men (e13771)	
Thiago Torres, Kelika Konda, E Vega-Ramirez, Oliver Elorreaga, Dulce Diaz-Sosa, Brenda Hoagland, Steven Diaz, Cristina Pimenta, Marcos Benedetti, Hugo Lopez-Gatell, Rebeca Robles-Garcia, Beatriz Grinsztejn, Carlos Caceres, Valdilea Veloso, ImPrEP Study Group.	153
Continuous Glucose Monitoring in the Real World Using Photosurveillance of #Dexcom on Instagram: Exploratory Mixed Methods Study (e11024)	
Michelle Litchman, Sarah Wawrzynski, Whitney Woodruff, Joseph Arrington, Quynh Nguyen, Perry Gee.	167
Early Detection of Adverse Drug Reactions in Social Health Networks: A Natural Language Processing Pipeline for Signal Detection (e11264)	
Azadeh Nikfarjam, Julia Ransohoff, Alison Callahan, Erik Jones, Brian Loew, Bernice Kwong, Kavita Sarin, Nigam Shah.	177
Detection of Spatiotemporal Prescription Opioid Hot Spots With Network Scan Statistics: Multistate Analysis (e12110)	
Arinjoy Basak, Jose Cadena, Achla Marathe, Anil Vullikanti.	188
Text-Based Illness Monitoring for Detection of Novel Influenza A Virus Infections During an Influenza A (H3N2)v Virus Outbreak in Michigan, 2016: Surveillance and Survey (e10842)	
Rebekah Stewart, John Rossow, Seth Eckel, Sally Bidol, Grant Ballew, Kimberly Signs, Julie Conover, Erin Burns, Joseph Bresee, Alicia Fry, Sonja Olsen, Matthew Biggerstaff.	199
Informing HIV Prevention Programs for Adolescent Girls and Young Women: A Modified Approach to Programmatic Mapping and Key Population Size Estimation (e11196)	
Eve Cheuk, Shajy Isac, Helgar Musyoki, Michael Pickles, Parinita Bhattacharjee, Peter Gichangi, Robert Lorway, Sharmistha Mishra, James Blanchard, Marissa Becker.	210
Prevalence of Schistosoma Haematobium Measured by a Mobile Health System in an Unexplored Endemic Region in the Subprefecture of Torrock, Chad (e13359)	
Didier Lalaye, Mirjam de Bruijn, Tom de Jong.	221
Capture-Recapture Among Men Who Have Sex With Men and Among Female Sex Workers in 11 Towns in Uganda (e12316)	
Kevin Apodaca, Reena Doshi, Moses Ogwal, Herbert Kiyingi, George Aluzimbi, Geoffrey Musinguzi, Ibrahim Lutalo, Evelyn Akello, Wolfgang Hladik.	230
Estimation of the Population Size of Men Who Have Sex With Men in Vietnam: Social App Multiplier Method (e12451)	
Vo Son, Ali Safarnejad, Nguyen Nga, Vu Linh, Le Tu, Pham Manh, Nguyen Long, Abu Abdul-Quader.	240
Wet Markets and Food Safety: TripAdvisor for Improved Global Digital Surveillance (e11477)	
Nicole Kogan, Isabelle Bolon, Nicolas Ray, Gabriel Alcoba, Jose Fernandez-Marquez, Martin Müller, Sharada Mohanty, Rafael Ruiz de Castañeda.	250
An Automated Text-Messaging Platform for Enhanced Retention and Data Collection in a Longitudinal Birth Cohort: Cohort Management Platform Analysis (e11666)	
Caroline Barry, Aditi Sabhlok, Victoria Saba, Alesha Majors, Julia Schechter, Erica Levine, Martin Streicher, Gary Bennett, Scott Kollins, Bernard Fuemmeler.	256

A Software Tool Aimed at Automating the Generation, Distribution, and Assessment of Social Media Messages for Health Promotion and Education Research (e11263) Katja Reuter, Alicia MacLennan, NamQuyen Le, Jennifer Unger, Elsi Kaiser, Praveen Angyan.....	266
Exploring Australian Hajj Tour Operators' Knowledge and Practices Regarding Pilgrims' Health Risks: A Qualitative Study (e10960) Amani Alqahtani, Mohamed Tashani, Anita Heywood, Robert Booy, Harunor Rashid, Kerrie Wiley.....	280

Tutorial

Google Trends in Infodemiology and Infoveillance: Methodology Framework (e13439) Amaryllis Mavragani, Gabriela Ochoa.....	52
--	----

Review

Tools for the Diagnosis of Herpes Simplex Virus 1/2: Systematic Review of Studies Published Between 2012 and 2018 (e14216) Zeeshaan Arshad, Abrar Alturkistani, David Brindley, Ching Lam, Kimberley Foley, Edward Meinert.....	120
--	-----

Corrigenda and Addenda

Correction: Google for Sexual Relationships: Mixed-Methods Study on Digital Flirting and Online Dating Among Adolescent Youth and Young Adults (e14815) James Lykens, Molly Pilloton, Cara Silva, Emma Schlam, Kate Wilburn, Emma Pence.....	228
---	-----

Viewpoint

Programmatic Mapping: Providing Evidence for High Impact HIV Prevention Programs for Female Sex Workers

Faran Emmanuel¹, MBBS, MSc (Epidemiology); Navindra Persaud², MBBS, MPH, PhD; Sharon S Weir³, PhD; Parinita Bhattacharjee⁴, MA; Shajy Isac¹, PhD

¹Centre for Global Public Health, Winnipeg, MB, Canada

²Family Health International 360, Washington DC, WA, United States

³University of North Carolina, Chapel Hill, NC, United States

⁴Partners for Health and Development in Africa, Nairobi, Kenya

Corresponding Author:

Faran Emmanuel, MBBS, MSc (Epidemiology)

Centre for Global Public Health

R070 Med Rehab Bldg, 771 McDermot Avenue

Winnipeg, MB, R3E 0T6

Canada

Phone: 1 204 975 7784

Email: faran.emmanuel@umanitoba.ca

Abstract

Programmatic mapping (PM) is a rapid and efficient mechanism to develop size estimates of key populations including female sex workers (FSWs) and geolocate them at physical locations in a systematic and scientific manner. At the macro level, this information forms the basis for allocating program resources, setting performance targets, and assess coverage. At a micro level, PM data provide specific information on hot spots, estimates of FSWs at those spots, and hot spot typology and days and times of operation, all of which provides targeted service delivery strategies. This information can provide a reliable platform to plan HIV prevention and treatment services to considerable scale and intensity. Above all, the entire PM process requires deep involvement of FSWs, which increases community ownership of the data and can lead to an increased uptake of services. Despite a few limitations, the approach is versatile and can be used in varied country contexts to generate important information about sex work and its dynamics. In this paper, we describe experiences and lessons learned from using evidence generated from PM of FSWs in multiple countries to develop HIV prevention programs at scale.

(*JMIR Public Health Surveill* 2019;5(2):e12636) doi:[10.2196/12636](https://doi.org/10.2196/12636)

KEYWORDS

programmatic mapping; PLACE; size estimation; female sex workers; HIV prevention; microplanning; key populations

Introduction

Female sex workers (FSWs) are disproportionately affected by the HIV epidemic [1]. Globally, 15% of the HIV infections among women may be attributed to sex work [2]. Primary prevention remains the key to epidemic control [3] as available evidence shows that targeted interventions implemented at scale among FSW can halt HIV transmission and reverse the epidemic [4-7]. Programs most likely to reduce new infections are community-owned, grounded in evidence and use a mix of biomedical, behavioral, and structural interventions [8,9].

Acknowledging local epidemic diversity, the Joint United Nations Programme on HIV/AIDS (UNAIDS) Practical Guidelines for Intensifying HIV Prevention recommend the use

of local data to inform programs [10,11]. This includes *knowing the epidemic* to develop a clear understanding of the epidemic and tailoring “the response” based on a combination prevention strategy. Moreover, HIV prevention programs tailored to FSWs should include active outreach, engagement, mobilization, and empowerment of FSWs to reduce their vulnerability [12,13]. Thus, for programs to be effective and accountable, they need to prioritize FSWs that can be directly reached through peer outreach and geographically located services [14]. The recently launched HIV prevention 2020 roadmap recommends a “location-population approach” that addresses the heterogeneity of the HIV epidemic and ensures effective and efficient planning and delivery of HIV prevention services [15].

More recently, The Global Fund, the World Bank, World Health Organization (WHO) and UNAIDS, as well as various national

programs have successfully used programmatic mapping (PM) approaches to estimate size, geo-distribution, operational typologies, and structural determinants of various key populations (KPs) [16]. Although various size estimation methods provide an absolute number of the key vulnerable population [17,18], PM produces estimates at the level of each hot spot and also assesses the amount of overlap between various spots. Thus, population size estimates derived from PM are adjusted for potential duplication and also provide a consequential distribution of KP members at different spots. In addition to identifying key locations and quantifying risk populations, mapping results can be used to identify existing programs and quickly institute services where they were unavailable. Thus, it provides site-level data in a timely manner to inform program design, monitor the effect of programs, and guide programs to meet the changing needs of FSWs.

The knowledge derived from PM has been used abundantly to provide a reliable platform to plan preventive and treatment services for FSWs. In this paper, we describe experiences and lessons learned from using evidence generated from PM of female sex workers in multiple countries to develop HIV prevention programs for effective coverage at scale.

What is Programmatic Mapping?

PM [16] is a systematic and scientific approach that identifies locations (often called sites, venues, or hot spots) where KPs engage in risk and provides the local and national population size estimation data necessary to plan prevention and treatment services at appropriate scale and intensity. Programmatic experience in diverse settings show that although a few FSWs operate through hidden networks, a larger proportion congregate and/or meet clients in definable geographic settings and targeting these spots for program delivery is an efficient prevention strategy [19]. To identify where to reach FSWs, PM engages the FSW community as coimplementers to estimate the number of FSWs, describe their geo-distribution and operational typologies, and inform the design and placement of HIV prevention programs to maximize coverage [20,21]. PM includes profiling locations where FSWs can be reached and assesses service availability in those locations or nearby. The term “programmatic” underscores PM’s focus on generating information useful to program design and improvement.

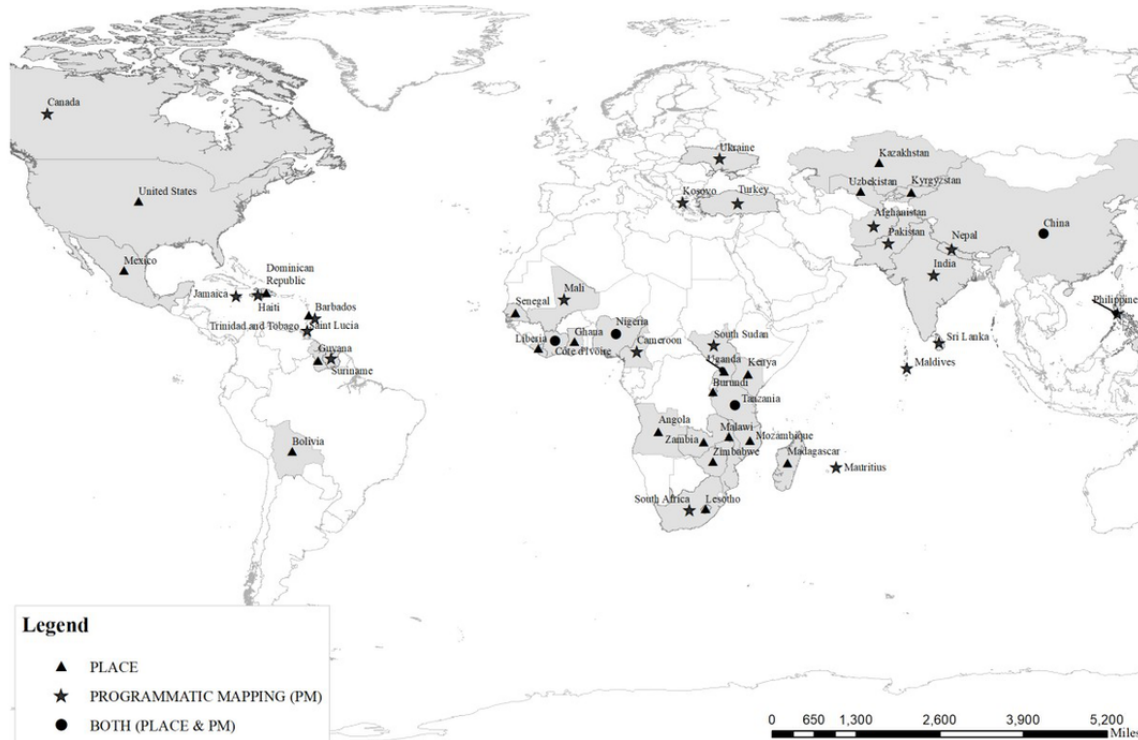
How Programmatic Mapping Works

During PM fieldwork, the study area is segmented into zones, and an exhaustive list of locations where FSWs congregate is systematically compiled within each zone. Subsequently, these spots are visited with the assistance of local community members (FSWs) to confirm onsite risk activities and the number of FSWs at each location. Population size is estimated at the site level, which is then adjusted to account for double counting and/or mobility across sites, to finally reach an estimated number of FSWs in the study area. Planning for and

calculating coverage requires an understanding of where sex work happens and an estimate of various subtypes in each of these locations and their characteristics. One of the key strengths of this approach lies not only in its development of size estimates but also in providing a distribution of FSWs at different hot spots. For planning services and subsequently monitoring coverage, this approach has been proven valuable in ensuring high level coverage [22].

PM includes 3 main approaches: basic geographic mapping (GM), the Priorities for Local AIDS Control Efforts (PLACE) method, and progressive mapping. Selecting the best approach depends on the context and the availability of time and resources. However, all adhere to the principles of generating location and population size information to design and implement programs, and community ownership and leadership. Although basic GM has been used primarily for mapping and estimating size of KPs, PLACE has been used pragmatically to identify locations where persons meet new partners, and thus, networks where HIV transmission is most likely to occur. Both GM and PLACE have been used where mapping and size estimates are not available and/or more rigorous baseline or follow-up estimates of program coverage are required. Progressive mapping approach is implemented as a component of a more mature program, where it is used quickly by program teams to generate information to scale up programs. A practical difference between the methods is that GM and PLACE include systematic interviews with key informants across all study zones to develop an exhaustive list of sites, whereas progressive mapping uses a crude list of sites from previous mapping as the starting point, available program data, and discussions with peer educators and outreach workers. Each method includes visits to identified spots to verify they are operational, describe their characteristics, and collect information on estimates, typology, and other operational characteristics of the hot spot. This is done by interviewing a key informant, preferably an FSW who operates from the identified spot, or someone knowledgeable about the spot (eg, a bar manager). Both GM and PLACE have extended the value of basic PM by using the venue list as a sampling frame for venue-based biobehavioral surveillance. Examples have been provided in subsequent sections on how countries with large FSW populations such as India, Kenya, and Haiti have used these approaches to gather sufficient information in a timely manner with enough geographic specificity and at low cost to plan KP program scale-up [23-25]. Other critical information such as condom availability or service availability for FSWs in the specific spots can be also collected in all the methods if such information is needed for effective program planning.

Development partners (eg, The Global Fund, the World Bank, the Bill and Melinda Gates Foundation, WHO/UNAIDS, and PEPFAR) have supported various PM approaches in a number of countries. [Figure 1](#) shows countries where various forms of PM have been implemented.

Figure 1. Countries where various forms of programmatic mapping have been implemented.

Using Programmatic Mapping for Development of Prevention Programs

There are a number of ways PM data can be used to develop strategic plans for targeted HIV prevention among KPs including FSWs and rapidly establish appropriate programs and basic services. In addition, PM has been used extensively to estimate the size of FSW populations, identify prevention coverage gaps, and quickly institute services where they are nonexistent. The following section presents some examples of how PM data have been used in various countries to strategize, develop, and implement FSW programs.

Use of Programmatic Mapping Data at a Macro Level to Develop National and Subnational Program Plans

Conceptually, high impact programs aim to reach the highest number of FSWs with quality services. At the macro level, national and subnational programming [23] focuses on establishing HIV prevention and care services, upscaling programs where needed, and launching new programs where they are nonexistent. Effective program coverage can only be measured when a realistic estimate of the number of FSWs is available.

PM data in Kenya estimated 103,298 FSWs in 7 provinces [26], which was used by National AIDS and STIs Control Program as a denominator to set national coverage targets for the minimum package of services defined in the national guidelines [27]. PM results guided decisions about where to prioritize services to improve program coverage. Thus, Nairobi (with 27% of the FSW population) was the top priority for prevention efforts, followed by Coast, Rift valley, Nyanza, and Western province. Within Nairobi, the subcounty of Starehe, with 25% of the Nairobi FSW population, was prioritized over counties

with lesser concentrations of FSWs [26]. PM data also inform progressive coverage based on resources and targets and provide insights into the heterogeneity of site typology and overall sex work so that intervention packages can be tailored to the local context [25,28]. For example, PM in Nigeria identified diverse sex work typologies based on the operational dynamics of sites and sex workers [29]. Thus, as 80% of the FSWs were venue-based in Abuja, the program focused on robust outreach and service delivery efforts at bars and night clubs and implemented less intensive programming with FSWs at other type of spots. On the basis of the number of FSWs (Lagos: 46,691 vs Anambra: 4846), 10-fold more resources were allocated for the FSW prevention program in Lagos. The geographical distribution of FSWs in Mauritius highlighted that expanding HIV prevention services in only 04 of the 09 districts would provide coverage to more than 80% of the FSW population in Mauritius, targeting specifically the high-intensity sites, thus, improving outreach and coverage within available resources [30].

Use of Programmatic Mapping Data at a Micro Level to Enhance Outreach and Services

Micro planning decentralizes outreach management and planning to a grass-root level of outreach workers (ORW) and peer educators (PEs) through a process that collects and uses data at an individual level to empower a community to make healthy decisions for themselves [31]. Although macro plans are made at the national or subnational levels, micro level planning takes place at the site level and becomes the cornerstone of peer-led programs [32].

Through PM, a list of all locations or hot spots where FSWs operate is developed, assigning a range of estimates for FSWs at each spot by subtypology (eg, street, venue, and public place). Furthermore, PM data characterize the spots by providing

information on timing of sex work activity at the spot (busiest time of the day or busiest day of the week), which is used as the basis for micro planning. Steps in micro planning include “hot spot validation” identified by PM, followed by “spot load mapping” where FSWs’ peers validate and record peak times and days of operation along with other characteristics of the spot. This is done to build ownership of the PM data by the peers, update spot lists, and estimate the number of PEs required. “Contact mapping” is then done to understand FSW networks and select PEs and determine their allocation across the target area, taking spot clustering in account. “Peer plans” are developed by PEs for each allocated FSW based on her risk and vulnerability, prioritizing those with higher risk and vulnerability. Finally, “Peer calendars” are used to document services provided to each FSW based on individual peer plans and track them at each visit outreach and micro planning [31,32]. The entire process of micro planning was conducted in Haiti using the PM data generated using the PLACE approach. These tools assisted the program to recognize each sex worker as an individual, assess her HIV risk (client volume, age, duration in sex work) and vulnerability (violence), and help plan a prevention pathway for her [33].

Likewise, in Mombasa, Kenya, PM data supported successful scale-up of HIV programs for FSWs. PM identified 9208 FSWs with the highest concentration (5809) in Kisauni and Mvita [26]. Hot spot validation and spot load mapping informed that 60% of FSWs operated through venues such as bars and clubs and another 30% through street spots. The 233 validated spots were plotted on a map, highlighting the number of FSW congregating at each spot. Location, spot size, and typology were used to allocate spots to PEs within their area of operation. On the basis of a ratio of 1 PE for 60 to 75 FSWs, 68 PEs were recruited, and 12 ORWs were hired to supervise them. Within areas of high FSW concentrations, 2 drops in centers and clinics were established, and linkage with government and NGO clinics were established for FSWs in dispersed or far away spots. Through peer plans and peer calendars, cumulative enrollment increased to 80% of the estimated FSW population within the first 10 months. An average of 35 to 40 condoms per month were distributed to each FSW during outreach as per the client volume calculated at registration [24]. As shown in the example, sex worker programs benefit tremendously from PM data in its finest details to estimate the specific number of condoms, lubes, outreach testing supplies, and other materials needed. Not only does this improve the efficiency and effectiveness of programs, it has proven its value in strengthening community networks through a community-based approach, which adds to the ownership of such programs [34].

Use of Programmatic Mapping Data to Estimate Population Size

PM has been used extensively in a number of countries to produce size estimates of sex worker populations [25,26,29,30]. Mapping produces estimates of FSWs at each spot, assesses the amount of overlap between various spots, and then adjusts the

total number for potential duplication to calculate a range of estimated number of FSWs in a focused geography. Although PM provides valid population size estimates in a timely manner to inform program design and supports monitoring and evaluation of ongoing interventions, a few limitations of the approach need to be discussed. As the inherent approach focusses on geolocation-based sex work, there is a probability to miss out FSWs who do not visit physical venues and solicit online or via social media. Thus, care needs to be taken while relying on size estimates derived from PM approaches alone. Data from PM are best used as a starting point for programs, and it is anticipated that as programs implement, they will validate and supplement the information and thereby periodically update the estimates [23,24,33]. Despite this limitation, PM has been used extensively across countries for estimating size of the KPs successfully as part of a National size estimation exercise [25,26,30].

Use of Programmatic Mapping Data for Monitoring and Research

Finally, PM data provides service delivery programs with denominators, which are crucial to be able to set goals and establish benchmarks for key outcomes indicators. Key indicators related to program coverage and utilization of programs and services by target populations serve as markers for program success. Coverage gaps at spot level can be evaluated through mapping of these spots on a continuous basis. PM data provide denominators for monitoring 90-90-90 targets and service uptake indicators across the HIV prevention, care, and treatment cascade among FSWs. In Kenya, the national program routinely uses PM data as a denominator to measure the KP program performance every quarter [35]. Under the LINKAGES project, the list of spots generated through PM contributed to the development of a sampling frame for conducting integrated bio-behavioral surveys (IBBS) of KPs in Haiti, Malawi, Angola, and Mozambique, which helped development of 90-90-90 indicators, including proportion of FSW on treatment with suppressed viral loads [25,36-38].

Likewise, Pakistan has extensively used PM data to sample KPs for IBBS. In each city, the IBBS sample was distributed based on weights derived from the respective number of KPs in each zone. Spots were randomly selected from spot lists generated by PM within each zone, and later, respondents were selected based on sampling weights assigned to each spot based on spot size, making the sample representative of the overall population [39,40]. A similar approach was employed across India, where PM data were used to select a random sample of hot spots and select respondents from the selected hot spots. As PM listed all hot spots in a selected geography, the list of spots served as a sampling frame, and it was possible to select a representative sample of KPs using a Time Location Cluster Sampling technique [41]. Thus, PM provides a sampling frame from which a representative sample of FSWs can be selected for survey. Table 1 provides the information available through mapping and its use for service delivery planning.

Table 1. Information available through programmatic mapping and its use for service delivery planning.

Information	How this information is used
National size estimates	<p>Develop national strategic plans for HIV prevention, care, and support for key populations</p> <p>Decide on national budgetary and resource allocations and costing exercises</p> <p>Set denominators for coverage targets for HIV prevention and treatment programs</p> <p>Evaluate prevention response at a national level and identify coverage gaps</p>
Subnational size estimates	<p>Help the country to decide where to prioritize programs to improve meaningful coverage</p> <p>Decide on resource allocations at a regional or subnational level</p> <p>Compare prevention response among regions and identify prevention gaps</p> <p>Draw representative samples for key population research by assigning sampling weights based on national distribution of study population</p>
Geo-distribution of spots	<p>Target locations/spots for localizing interventions</p> <p>Develop local maps and set a plan for coverage</p> <p>Prepare spot clusters and allocate spots to peer educators based on geo-proximity of spots</p> <p>Establish locations for health clinics, drop-in centers, HIV testing centers etc, as well as condom and lubricant distribution channels</p>
Number and size of spots	<p>Target locations/spots for localizing interventions</p> <p>Decide which spots to prioritize and focus to match coverage targets</p>
Operational dynamics of a spot; peak days and peak times of operations	<p>Determine human resource needs, that is, how many peer educators and outreach workers are needed to adequately cover the population. For example, 1 peer educator could work with 50-60 FSWs^a, whereas 1 outreach worker can manage working with 5-6 peer educators.</p> <p>Size estimates provided at the spot level can be used to estimate the specific number of condoms, lube, outreach testing supplies, and other materials needed</p> <p>The human resource and commodity plan needs to be based on peak estimates so that no one remains uncovered</p> <p>Peak times and peak days might be utilized to determine time of outreach</p>
Operational typology of spots/FSWs	<p>Inform intervention design based on the subtypology of spots and FSWs (eg, brothel, street, bar, night club, massage parlor based, and home-based FSWs).</p>
Use in research	<p>Mapping data is used as a sampling frame for national level surveys, including integrated bio-behavioral surveys.</p>

^aFSWs: female sex workers.

Conclusions

PM is a low-cost and reliable approach designed to generate the most critical information required for planning and implementing effective HIV prevention, care, and treatment programs. PM data have been used at national, provincial, district, subdistrict, and spot levels. At a national level, PM informs funding requirements and resource allocation, program scale-up, target setting, and coverage assessment. At a subnational level, PM prioritizes geographies where interventions should be scaled up to ensure the highest coverage possible within the available resources. Saturating coverage of KPs in high concentration geographies is preferable to spreading services thinly across a wider area. At the district or town level, PM assists programs to assess resources required for outreach and clinical staffs and for commodity requirements such as

number of condoms/lube and HIV testing kits. At a spot level, PM enhances outreach to ensure the right number of peer ORWs are available to saturate coverage as well as allocate spots to the most ideal PE. PM further improves program design by carefully evaluating typologies and functional timings of the spots and establishes denominator to monitor access to services and coverage at each spot.

Despite a few limitations, PM provides essential information about the size, distribution, and characteristics of FSWs in a systematic and scientific manner. Above all, the process relies substantially on the strength and involvement of civil society and community organizations that represent and are engaged with KPs. This practically increases the appropriateness and ownership of program design and implementation and results in increased uptake and more efficient and effective programs based on the needs and priorities of the FSW community.

Acknowledgments

The authors would like to acknowledge research teams in all countries where PM was conducted. They would like to thank the Ministries of Health in several countries where PM has been conducted in partnership. They thank international donors, that is, The World Bank, Bill and Melinda Gates Foundation, and GFATM for their funding support to conduct PM in various countries. This publication was made possible by the generous support of the American people through the United States Agency for International Development and the United States President's Emergency Plan for AIDS Relief (PEPFAR) through the LINKAGES project, cooperative agreement number AID-OAA-A-14-00045. The contents of this publication are the sole responsibility of LINKAGES and do not necessarily reflect the views of USAID, PEPFAR, or the US Government.

Authors' Contributions

FE conceptualized the whole paper, developed the first draft, and finalized the manuscript. SSW, NP, and PB wrote programmatic examples and provided additions in the discussion section. SI wrote up a large part of the conclusion section and added examples in the text.

Conflicts of Interest

None declared.

References

1. 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](https://doi.org/10.1016/S1473-3099(12)70066-X)] [Medline: [22424777](https://pubmed.ncbi.nlm.nih.gov/22424777/)]
2. Prüss-Ustün A, Wolf J, Driscoll T, Degenhardt L, Neira M, Calleja JM. HIV due to female sex work: regional and global estimates. *PLoS One* 2013;8(5):e63476 [FREE Full text] [doi: [10.1371/journal.pone.0063476](https://doi.org/10.1371/journal.pone.0063476)] [Medline: [23717432](https://pubmed.ncbi.nlm.nih.gov/23717432/)]
3. Ainsworth M, Teokul W. Breaking the silence: setting realistic priorities for AIDS control in less-developed countries. *Lancet* 2000 Jul 1;356(9223):55-60. [doi: [10.1016/S0140-6736\(00\)02440-5](https://doi.org/10.1016/S0140-6736(00)02440-5)] [Medline: [10892777](https://pubmed.ncbi.nlm.nih.gov/10892777/)]
4. Steen R, Wheeler T, Gorgens M, Mziray E, Dallabetta G. Feasible, efficient and necessary, without exception - Working with sex workers interrupts HIV/STI transmission and brings treatment to many in need. *PLoS One* 2015;10(10):e0121145 [FREE Full text] [doi: [10.1371/journal.pone.0121145](https://doi.org/10.1371/journal.pone.0121145)] [Medline: [26488796](https://pubmed.ncbi.nlm.nih.gov/26488796/)]
5. Steen R, Zhao P, Wi TE, PUNCHIHEWA N, Abeyewickreme I, Lo Y. Halting and reversing HIV epidemics in Asia by interrupting transmission in sex work: experience and outcomes from ten countries. *Expert Rev Anti Infect Ther* 2013 Oct;11(10):999-1015. [doi: [10.1586/14787210.2013.824717](https://doi.org/10.1586/14787210.2013.824717)] [Medline: [24124797](https://pubmed.ncbi.nlm.nih.gov/24124797/)]
6. UNAIDS. Female sex worker HIV prevention projects?: Lessons learnt from Papua New Guinea, India and Bangladesh URL: http://www.unaids.org/sites/default/files/media_asset/jc438-femsexwork_en_0.pdf [accessed 2019-04-23] [WebCite Cache ID 77rGasXlu]
7. Boily M, Pickles M, Lowndes CM, Ramesh BM, Washington R, Moses S, et al. Positive impact of a large-scale HIV prevention programme among female sex workers and clients in South India. *AIDS* 2013 Jun 1;27(9):1449-1460 [FREE Full text] [doi: [10.1097/QAD.0b013e32835fba81](https://doi.org/10.1097/QAD.0b013e32835fba81)] [Medline: [23462221](https://pubmed.ncbi.nlm.nih.gov/23462221/)]
8. UNAIDS. Combination HIV Prevention: tailoring and coordinating biomedical, behavioural and structural strategies to reduce new HIV infections URL: http://www.unaids.org/sites/default/files/media_asset/JC2007_Combination_Prevention_paper_en_0.pdf [accessed 2019-04-23] [WebCite Cache ID 77rGdz2Z]
9. Mishra S, Steen R, Gerbase A, Lo Y, Boily M. Impact of high-risk sex and focused interventions in heterosexual HIV epidemics: a systematic review of mathematical models. *PLoS One* 2012;7(11):e50691 [FREE Full text] [doi: [10.1371/journal.pone.0050691](https://doi.org/10.1371/journal.pone.0050691)] [Medline: [23226357](https://pubmed.ncbi.nlm.nih.gov/23226357/)]
10. UNAIDS. Practical Guidelines for Intensifying HIV Prevention URL: http://data.unaids.org/pub/manual/2007/20070306_prevention_guidelines_towards_universal_access_en.pdf [accessed 2019-04-23] [WebCite Cache ID 77rGgA1O5]
11. Jones A, Cremin I, Abdullah F, Idoko J, Cherutich P, Kilonzo N, et al. Transformation of HIV from pandemic to low-endemic levels: a public health approach to combination prevention. *Lancet* 2014 Jul 19;384(9939):272-279. [doi: [10.1016/S0140-6736\(13\)62230-8](https://doi.org/10.1016/S0140-6736(13)62230-8)] [Medline: [24740087](https://pubmed.ncbi.nlm.nih.gov/24740087/)]
12. Wilson D, Halperin DT. Know your epidemic, know your response: a useful approach, if we get it right. *Lancet* 2008 Aug 9;372(9637):423-426. [doi: [10.1016/S0140-6736\(08\)60883-1](https://doi.org/10.1016/S0140-6736(08)60883-1)] [Medline: [18687462](https://pubmed.ncbi.nlm.nih.gov/18687462/)]
13. Vassall A, Chandrashekar S, Pickles M, Beattie TS, Shetty G, Bhattacharjee P, CHARME India Group, et al. Community mobilisation and empowerment interventions as part of HIV prevention for female sex workers in Southern India: a cost-effectiveness analysis. *PLoS One* 2014;9(10):e110562 [FREE Full text] [doi: [10.1371/journal.pone.0110562](https://doi.org/10.1371/journal.pone.0110562)] [Medline: [25333501](https://pubmed.ncbi.nlm.nih.gov/25333501/)]
14. World Health Organization. Implementing Comprehensive HIV / STI Programmes with Sex Workers URL: <https://www.who.int/sorry/> [accessed 2019-04-23] [WebCite Cache ID 77rGs6fwZ]
15. UNAIDS. HIV Prevention 2020 Road Map URL: http://www.unaids.org/sites/default/files/media_asset/hiv-prevention-2020-road-map_en.pdf [accessed 2019-04-23] [WebCite Cache ID 77rGxVklX]

16. GFATM. Using Programmatic Mapping to Improve Program Access and Coverage for Key Populations: Guidelines for Countries URL: https://www.theglobalfund.org/media/1291/publication_genderequalitykeypopulationsreview_report_en.pdf?u=636727910290000000 [accessed 2019-04-23] [WebCite Cache ID 77rH28ujr]
17. Sabin K, Zhao J, Garcia CJ, Sheng Y, Arias GS, 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/)]
18. 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](https://pubmed.ncbi.nlm.nih.gov/24393694/)]
19. 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](https://pubmed.ncbi.nlm.nih.gov/23627850/)]
20. GFATM. The Global Fund position on size estimation and collection of spatial data, including programmatic mapping, of key populations URL: https://www.theglobalfund.org/media/1291/publication_genderequalitykeypopulationsreview_report_en.pdf?u=636727910290000000 [accessed 2019-04-23] [WebCite Cache ID 77rH8P4tx]
21. UNAIDS. 2015. Focus on Location and Populations URL: http://www.unaids.org/sites/default/files/media_asset/WAD2015_report_en_part01.pdf [accessed 2019-04-29] [WebCite Cache ID 77z7V3tXA]
22. Laga M, Vuylsteke B. Evaluating AVAHAN's design, implementation and impact: lessons learned for the HIV Prevention Community. *BMC Public Health* 2011 Dec 29;11 Suppl 6:S16 [FREE Full text] [doi: [10.1186/1471-2458-11-S6-S16](https://doi.org/10.1186/1471-2458-11-S6-S16)] [Medline: [22376320](https://pubmed.ncbi.nlm.nih.gov/22376320/)]
23. Blanchard JF, Bhattacharjee P, Kumaran S, Ramesh BM, Kumar NS, Washington RG, et al. Concepts and strategies for scaling up focused prevention for sex workers in India. *Sex Transm Infect* 2008 Oct;84(Suppl 2):ii19-ii23. [doi: [10.1136/sti.2008.033134](https://doi.org/10.1136/sti.2008.033134)] [Medline: [18799487](https://pubmed.ncbi.nlm.nih.gov/18799487/)]
24. McClarty LM, Bhattacharjee P, Isac S, Emmanuel F, Kioko J, Njiraini M, et al. Key Programme Science lessons from an HIV prevention 'Learning Site' for sex workers in Mombasa, Kenya. *Sex Transm Infect* 2018 Aug;94(5):346-352 [FREE Full text] [doi: [10.1136/sextrans-2017-053228](https://doi.org/10.1136/sextrans-2017-053228)] [Medline: [29242195](https://pubmed.ncbi.nlm.nih.gov/29242195/)]
25. FHI 360. 2018. Programmatic Mapping and Size Estimation of Key Populations in Haiti FINAL URL: <https://www.fhi360.org/sites/default/files/media/documents/resource-linkages-haiti-size-key-populations-april-2017.pdf> [accessed 2019-04-23] [WebCite Cache ID 77rHNCJB9]
26. 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/)]
27. NASCOP. National AIDS/STD Control Programme of Kenya. Standards for Peer-Education and Outreach Programs for Sex Workers URL: http://guidelines.health.go.ke:8000/media/Sex_Worker_Guidelines.pdf [accessed 2019-04-23] [WebCite Cache ID 77rHSr4qS]
28. Buzdugan R, Copas A, Moses S, Blanchard J, Isac S, Ramesh BM, et al. Devising a female sex work typology using data from Karnataka, India. *Int J Epidemiol* 2010 Apr;39(2):439-448. [doi: [10.1093/ije/dyp351](https://doi.org/10.1093/ije/dyp351)] [Medline: [19995861](https://pubmed.ncbi.nlm.nih.gov/19995861/)]
29. Ikpeazu A, Momah-Haruna A, Mari BM, Thompson LH, Ogungbemi K, Daniel U, et al. An appraisal of female sex work in Nigeria--implications for designing and scaling up HIV prevention programmes. *PLoS One* 2014;9(8):e103619 [FREE Full text] [doi: [10.1371/journal.pone.0103619](https://doi.org/10.1371/journal.pone.0103619)] [Medline: [25118691](https://pubmed.ncbi.nlm.nih.gov/25118691/)]
30. Pathack A, Sauntally A, Soobhany S, Comins CA, Kinoo SA, Emmanuel F. Programmatic mapping to determine the size and dynamics of sex work and injecting drug use in Mauritius. *Afr J AIDS Res* 2018 Jul;17(2):129-136. [doi: [10.2989/16085906.2018.1462216](https://doi.org/10.2989/16085906.2018.1462216)] [Medline: [29745291](https://pubmed.ncbi.nlm.nih.gov/29745291/)]
31. AIDSFree. Micro-Planning in Peer Led Outreach Programs: A Handbook Based on the Experiences of the Avahan India AIDS Initiative URL: <https://aidsfree.usaid.gov/resources/prevention-update/editions/september-2013/micro-planning-peer-led-outreach-programs> [accessed 2019-04-29] [WebCite Cache ID 77z8w0kiA]
32. Partners for Health and Development in Africa. Outreach and Micro-planning URL: <http://www.phdaf.org/docs/Learning%20Site%20Case%20Study,%20Outreach%20and%20Microplanning.pdf> [accessed 2019-04-23] [WebCite Cache ID 77rHdNlmm]
33. President's Emergency Plan For AIDS Relief. PEPFAR Haiti Country Operational Plan 2017 (COP 2017) Strategic Direction Summary URL: <https://www.pepfar.gov/documents/organization/272014.pdf> [accessed 2019-04-23] [WebCite Cache ID 77rHi1mgc]
34. KHPT. World Bank. A Systematic Approach to the Design and Scale-up of Targeted Interventions for HIV Prevention among Urban Female Sex Workers URL: <http://documents.worldbank.org/curated/en/336971468151769158/pdf/NonAsciiFileName0.pdf> [accessed 2019-04-23] [WebCite Cache ID 77rHmdmGw]
35. NASCOP. Programme Quarterly Data, December 2017 URL: <https://www.nascop.or.ke/?p=1884> [accessed 2019-04-23] [WebCite Cache ID 77rHqrSNU]
36. LINKAGES. FHI 360. 2018. PLACE report Malawi URL: <https://www.fhi360.org/sites/default/files/media/documents/resource-linkages-malawi-place-report.pdf> [accessed 2019-04-23] [WebCite Cache ID 77rIAwJcn]

37. Herce M, Miller W, Bula A, Edwards J, Sapalalo P, Lancaster K, et al. Achieving the first 90 for key populations in sub-Saharan Africa through venue-based outreach: challenges and opportunities for HIV prevention based on PLACE study findings from Malawi and Angola. *J Int AIDS Soc* 2018 Jul;21(Suppl 5):e25132 [FREE Full text] [doi: [10.1002/jia2.25132](https://doi.org/10.1002/jia2.25132)] [Medline: [30033589](https://pubmed.ncbi.nlm.nih.gov/30033589/)]
38. Miller W, Manuel J, Mbofana F, Herce M, Edwards J, Rocuts F, et al. HIV testing and treatment cascade among key populations in Mozambique 22nd International AIDS Conference. 2018 Presented at: Oral abstracts of the 22nd International AIDS Conference; July 23-27, 2018; Amsterdam, the Netherlands.
39. Emmanuel F, Salim M, Akhtar N, Arshad S, Reza TE. Second-generation surveillance for HIV/AIDS in Pakistan: results from the 4th round of Integrated Behavior and Biological Survey 2011-2012. *Sex Transm Infect* 2013 Nov;89(Suppl 3):iii23-iii28 [FREE Full text] [doi: [10.1136/sextans-2013-051161](https://doi.org/10.1136/sextans-2013-051161)] [Medline: [23912818](https://pubmed.ncbi.nlm.nih.gov/23912818/)]
40. Emmanuel F, Adrien A, Athar U, Imran M, Reza T, Blanchard J. Using surveillance data for action: lessons learnt from the second generation HIV/AIDS surveillance project in Pakistan. *East Mediterr Health J* 2011 Aug;17(8):712-718. [Medline: [21977576](https://pubmed.ncbi.nlm.nih.gov/21977576/)]
41. FHI 360. Integrated Behavioural and Biological Assessment: State Summary Report for Karnataka, Bangalore URL: <https://www.fhi360.org/sites/default/files/media/documents/Integrated> [accessed 2019-04-23] [WebCite Cache ID 77rIQ2BQa]

Abbreviations

FSWs: female sex workers
GM: geographic mapping
IBBS: integrated bio-behavioral survey
KPs: key populations
ORW: outreach worker
PE: peer educator
PLACE: Priorities for Local AIDS Control Efforts
PM: programmatic mapping
UNAIDS: Joint United Nations Programme on HIV/AIDS
WHO: World Health Organization

Edited by T Sanchez; submitted 29.10.18; peer-reviewed by A Altaf, T Saidel, M Parakh; comments to author 18.01.19; revised version received 08.03.19; accepted 22.03.19; published 06.06.19.

Please cite as:

Emmanuel F, Persaud N, Weir SS, Bhattacharjee P, Isac S
Programmatic Mapping: Providing Evidence for High Impact HIV Prevention Programs for Female Sex Workers
JMIR Public Health Surveill 2019;5(2):e12636
URL: <http://publichealth.jmir.org/2019/2/e12636/>
doi: [10.2196/12636](https://doi.org/10.2196/12636)
PMID: [31172964](https://pubmed.ncbi.nlm.nih.gov/31172964/)

©Faran Emmanuel, Navindra Persaud, Sharon S Weir, Parinita Bhattacharjee, Shajy Isac. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 06.06.2019. 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

Pneumococcal Vaccination Utilization Among Hispanic Long-Term Colorectal Cancer Survivors: Cross-Sectional Assessment of Claims

Ryan J Moran^{1,2}, MD, MPH; Jill Waalen^{1,3}, MD, MPH; James Murphy⁴, MD, MS; Vinit Nalawade^{2,5}, MIS; Melody Schiaffino^{2,5}, PhD, MPH

¹Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA, United States

²School of Public Health, San Diego State University, San Diego, CA, United States

³Scripps Translational Research Institute, San Diego, CA, United States

⁴Department of Radiation Medicine and Applied Sciences, University of California San Diego, La Jolla, CA, United States

⁵Institute for Behavioral and Community Health, San Diego State University, San Diego, CA, United States

Corresponding Author:

Ryan J Moran, MD, MPH

Department of Family Medicine and Public Health

University of California San Diego

9500 Gilman Drive, Mail Code #0811

La Jolla, CA, 92093

United States

Phone: 1 8582258535

Email: rjmoran@ucsd.edu

Abstract

Background: Colorectal cancer (CRC) is the second leading cancer-related cause of death in the United States. However, survivorship has been increasing. Both cancer survivors and underserved populations experience unique health-related challenges and disparities that may exist among long-term CRC survivors as it relates to routine preventive care, specifically pneumococcal (PNM) vaccination.

Objective: The aim of this study was to explore the relationship between long-term CRC survival and the receipt of PNM vaccine among Hispanic Medicare recipients compared with non-Hispanic populations.

Methods: This study is a cross-sectional analysis of the Surveillance, Epidemiology, and End Results (SEER)-Medicare claims data examining ethnic differences in the receipt of PNM vaccination among long-term CRC survivors. Multivariable logistic regression models considered Hispanic ethnicity while controlling for sociodemographic characteristics, comorbidity score, age, tumor stage, and SEER registry.

Results: Our sample revealed 32,501 long-term CRC survivors, and 1509 identified as Hispanic (4.64%) based on an established SEER algorithm. In total, 16,252 CRC survivors, or 50.00% of our sample, received a PNM vaccination. We found that Hispanic CRC survivors had 10% decreased odds of having received a PNM vaccine compared with non-Hispanic survivors ($P=.03$).

Conclusions: Disparities likely exist in the utilization of PNM vaccination among long-term CRC survivors. Among Medicare beneficiaries, the use of claims data regarding PNM vaccination highlights the relatively poor utilization of guideline-directed preventive care.

(*JMIR Public Health Surveill* 2019;5(2):e12603) doi:[10.2196/12603](https://doi.org/10.2196/12603)

KEYWORDS

Hispanic Americans; cancer survivors; medicare; preventive medicine; pneumococcal vaccines

Introduction

Background

Colorectal cancer (CRC) persists as the third most common cancer in the United States among both men and women [1-3]. Although it remains the second leading cause of death from cancers that affect both men and women [2], there has been a significant decrease in CRC-related mortality in the last 40 years [1,2]. Lifestyle modification, screening for early detection, and advances in treatment for individuals with CRC have led to an increase in 5-year survival rates from 50% to 66%. In the last decade—where data are available—the annual predicted rate of change for CRC mortality has decreased by 2.9% [1,3]. As a result of this, the number of long-term survivors has been steadily increasing. At present, an estimated 1.4 million CRC survivors are living in the United States. Of these, over 85% are aged older than 60 years [4,5].

Among cancer survivors, studies have found conflicting findings regarding routine preventive care utilization and health care utilization compared with noncancer survivors. For example, survivors were more likely to report pneumococcal (PNM) and influenza vaccination, blood pressure monitoring, cholesterol measurement, bone mineral scans, and lower endoscopy via the *Medicare Current Beneficiary Survey* [6]. However, a study using Surveillance, Epidemiology, and End Results (*SEER*)-*Medicare* data suggests decreased influenza vaccination, cholesterol screening, cervical cancer screening, and bone density screening [7]. Another study using a combined metric for overall appropriate care utilization from the *SEER-Medicare* database found overall decreased care in CRC survivors compared with noncancer cases [8,6].

Hispanics/Latinos represent the biggest minority racial or ethnic group in the United States, currently estimated at 17.6% of the population. By 2060, this group is predicted to represent over 28% of the population [9]. Although Hispanic/Latino individuals are less likely to be diagnosed with CRC and have a lower mortality rate than non-Hispanic whites [1,2], healthy Hispanic/Latinos are also less likely to receive routine preventive care [10,11] such as CRC screening. Furthermore, evidence suggests that under-represented minority populations who survive cancer may be less likely to receive other essential routine preventive care such as influenza vaccination [6], including long-term survivors of CRC [12]. Additional evidence from the *National Health Interview Survey* suggests that Hispanic/Latino cancer survivors are less likely to utilize health care services compared with non-Hispanic whites for those aged 65 years and younger, though this did not appear to be the case for survivors aged 65 years and older. Some explanations include less time in the health care system, lower socioeconomic capacity to access services, and barriers such as out-of-pocket expenses [13].

Objectives

Medicare represents the most significant public insurer in the United States, with over 55 million beneficiaries and more than

37 million of these on a traditional fee-for-service model [13]. Over 2 million beneficiaries are Hispanic/Latinos enrolled in a traditional Medicare plan [14], and Medicare Part B has been providing coverage for the PNM vaccine since 1981 with no out-of-pocket expenses [15]. Since 1984, vaccination has been advised for all adults aged 65 years and older [16]. In the general population, evidence suggests that Hispanic adults aged 65 years and older are less likely to have received a PNM vaccine and are less likely to know that it is advisable [10,17]. However, evidence evaluating how Hispanic or Latino long-term survivors of CRC fare with regard to PNM vaccination recommendations appears limited [6,7].

We sought to explore the relationship between long-term CRC survival and the receipt of PNM vaccine among Hispanic/Latino Medicare recipients compared with non-Hispanic/Latino populations, given that the majority of new CRC cases occur in those aged 65 years and older [2].

Methods

Study Design

Our study is a retrospective, cross-sectional analysis of patients diagnosed with CRC identified within an existing dataset of the SEER program linked with the *Center for Medicare and Medicaid Services* (CMS) administrative claims database for the years 2000 to 2011 (N=318,675).

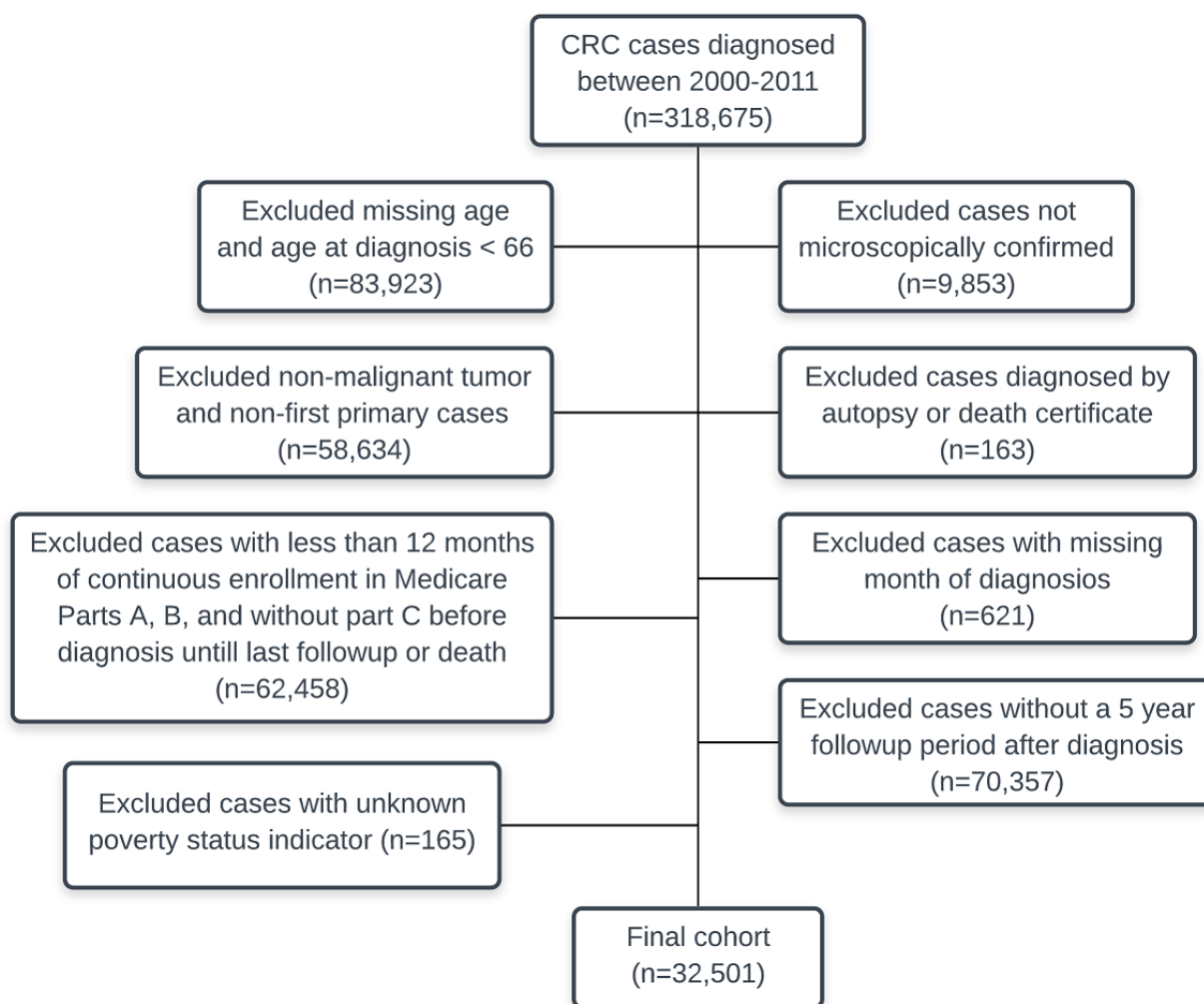
Population

SEER-Medicare data references information for patients with Medicare diagnosed with cancer, with a tumor registered at one of the National Cancer Institute's (NCI) 20 registries across the United States. These registries were then matched to CMS's administrative claim record enrollment file and had a 94% match. The dataset covers SEER diagnostic information for up to 10 diagnosed cancer occurrences for cases recorded in the SEER cancer registries [18]. Socioeconomic information derived from the US Census has been included for the years 1990 to 2000.

The Medicare data contain outpatient claims for diagnoses or specific procedures as defined by the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes and the Healthcare Common Procedure Coding System (HCPCS) codes along with their dates of service.

Eligibility

The cases were those patients aged 66 years and older who were diagnosed with a first-primary tumor located in the colon or rectum, CRC, at any stage. We defined CRC as those cancers using the site-recode ICD-0-3/World Health Organization 2008 codes: 21041, 21042, 21043, 21044, 21045, 21046, 21047, 21048, 21049, 21051, and 21052. Tumors had to be histologically confirmed and first-primaries. Survivors were included if they lived at least 5 years beyond the time of diagnosis.

Figure 1. Flow diagram defining cases of long-term colorectal cancer (CRC) survivors.

We excluded patients with missing ages or those who had been diagnosed by autopsy or death certificate. Patients with a missing month at diagnosis and patients with other primary cancers predating CRC diagnosis were also excluded. Cases with missing data such as the federal poverty category *unknown* were also excluded.

We selected patients who had continuous coverage of both Medicare Parts A and B but without Part C coverage from at least 1 year before diagnosis until the last follow-up or death to allow for 1 complete year of claims data before diagnosis to calculate a comorbidity score. To assure subjects were long-term survivors, we narrowed our study sample to those with at least 5 years of follow-up after cancer diagnosis. Time was calculated as the months from SEER-database entry to the point of death or end of follow-up (ending December 31, 2013), as has been described in a previously published cross-sectional approach that accounted for censorship [19]. The final SEER-Medicare dataset for analysis included 32,501 observations (Figure 1).

Exposure

SEER data were also used to extract some variables at diagnosis including patient age, sex, marital status, and race/ethnicity per the North American Association of Cancer Registries (NAACR)

Hispanic Identification Algorithm (NHIA) Derived Hispanic Origin. NHIA has an estimated sensitivity of 84% and a positive predictive value of over 94% for identifying those of Hispanic descent [20]. As we sought to explore vaccination status by those of ethnic identity, regardless of racial background, SEER data on race/ethnicity were used to create a dichotomous Hispanic/non-Hispanic variable.

Outcome

To assess for utilization of preventive services, specifically, PNM vaccination, claims data from Medicare identified the following ambulatory-based encounters wherever vaccination ICD-9-CM codes were involved: V03.82 and V06.6; using the Current Procedural Terminology (CPT) codes: 90670 and 90732; and the HCPCS code: G0009. As CRC survivors had a varied age of diagnosis, all data regarding immunization claims were included in our analysis regardless if this was before or after the time of diagnosis.

Covariates

Age at diagnosis was categorized into 5-year increments from 65 to 85 years. Although age at diagnosis and the date of PNM immunization claim varied chronologically, patients were included in the analysis as long as they reported a claim for an

immunization within the study period. For instances of multiple vaccinations, we used the earliest CPT/HCPCS claim and used the ICD-9 claim wherever a CPT/HCPCS claim date was absent. We also extracted the stage of disease (histology confirmed) given this likely greatly influenced visits to a health care setting, extracted the SEER registry site given the regional variation noted in preventive service utilization [21,22], and extracted a poverty indicator: the percent of the population who reside in the census-designated region at or below the federal poverty level (FPL). In addition, we included characteristics regarding the place of residence as either a metropolitan area or any other nonmetro area. Medicare date of death for the linked data from Medicare claims was used to confirm SEER month and year of death and the first day of the month when the Medicare date was not available (0.003% of the sample). Patient health status was assessed using the Charlson comorbidity scoring algorithm, an index to evaluate patient comorbidities with a SEER-Medicare-provided macro that included 10 comorbidities noted between the ages of 65 and 66 years: moderate/severe liver disease, cerebrovascular disease, peripheral vascular disease, renal disease, paralysis (hemiplegia or paraplegia), dementia, mild liver disease, congestive heart failure, chronic obstructive pulmonary disease, diabetes with complications, and diabetes [23-26]. The ICD-9 values for comorbidities were referenced from the comorbidity index found at the NCI [27,28]. Registries are geographically defined by SEER and correspond to different regions and cities to be representative of the American population. Of the original 20 SEER registry categories, 17 were available and recategorized if they had less than 50 Hispanic individuals, which was found to equate to less than 2% of the corresponding registry. All the remaining registries in our database were combined into 1 category.

To account for the variation of additional time since the 5-year survival cut-point of time, we included months after 5 years from diagnosis in our model (variable not reported in tables, available upon request).

Data Management and Statistical Analyses

Data exploration was conducted on all variables of interest to assess distribution, cut-points, and group differences. Descriptive statistics for all categorical variables and means for continuous variables are reported. Variables of interest were compared

based on our primary outcome variable, namely any receipt of PNM vaccine by the ethnic group using a chi-square analysis to consider the goodness-of-fit for modeling. Multivariable logistic regression models were fitted to independent variables of interest to our outcome. With a subanalysis using only those patients that did utilize a vaccine (n=16,171), we evaluated whether or not the place of service influenced receipt of vaccines among Hispanic/Latino versus non-Hispanic CRC survivors. We compared place of service by vaccine utilization across Hispanic/Latino versus non-Hispanic groups using a chi-square analysis to explore group differences. All data management and analyses were performed using SAS version 9.4 (SAS Institute Inc). The University of California-San Diego Human Research Protection Institutional Review Board Program approved this study, and data were utilized with agreement and approval from SEER-Medicare.

Results

Sample Characteristics

Our sample included 32,501 long-term CRC survivors, with a mean age of diagnosis of 76.6 years. The majority were diagnosed at a local stage (57.5%), were women (56.95%), were married (53.76%), and resided in a nonrural setting (87.79%). Over 50% of CRC survivors were between the ages of 66 and 75 years, and nearly two-thirds had an identified modified Charlson comorbidity score of 0 (67%). Over half (55%) resided in census-designated tracts where 10% or less of the population earned at or below 100% of the FPL. Overall, 50% (n=16,252) of CRC survivors received a PNM vaccination. Of these, approximately 1509 long-term CRC survivors were identified as Hispanic/Latino, compared with 30,992 who were identified as non-Hispanic (Table 1).

Statistically significant differences for both receipt of PNM vaccine and between Hispanic and non-Hispanic groups for gender, modified Charlson Comorbidity Score, categorized age of diagnosis, percent at or below the FPL, urban versus nonurban setting, marital status, and SEER registry location were noted. There was a statistically significant difference between the histology stage at diagnosis for PNM vaccine recipients but not by racial/ethnic group (Table 2).

Table 1. Characteristics of long-term colorectal cancer survivors in a Surveillance, Epidemiology, and End Results-Medicare database (2000 to 2011; N=32,501).

SEER ^a -Medicare characteristic and value	n (%)
Hispanic/Latino (SEER-North American Association of Cancer Registries Hispanic Identification Algorithm identity)	
Hispanic/Latino	1509 (4.64)
Non-Hispanic	30,992 (95.36)
Received a pneumococcal vaccine	
Yes	16,252 (50.00)
No	16,249 (50.00)
Age at diagnosis (years)	
66-69	5367 (16.51)
70-74	8207 (25.25)
75-79	8283 (25.49)
80-84	6407 (19.71)
85+	4237 (13.04)%
Gender	
Male	13993 43.05%
Female	18508 (56.95)
Urban status	
Urban/Metro	28,533 (87.79)
Rural	3968 (12.21)
Census tract federal poverty level (%)	
0 to <5	9464 (29.12)
5 to <10	9161 (28.19)
10 to <20	8795 (27.06)
>20	5081 (15.63)
Number of comorbidities (Charlson Index)	
0	21807 (67.10)
1	7339 (22.58)
2	2372 (7.30)
3 or more	983 (3.0)
Marital status	
Divorced	1800 (5.54)
Married	17473 (53.76)
Other	10,987 (33.81)
Single	2241 (6.90)
Tumor stage at diagnosis	
Localized	18,689 (57.50)
Regional	12,297 (37.84)
Distant	870 (2.7)
Unstaged	645 (1.98)
SEER registry region	
San Francisco	1119 (3.44)
Connecticut	2261 (6.96)

SEER ^a -Medicare characteristic and value	n (%)
San Jose	721 (2.2)
Los Angeles	2120 (6.52)
Greater California	5124 (15.77)
New Jersey	5620 (17.29)
New Mexico	744 (2.3)
Other ^b	14792 (45.51)

^aSEER: Surveillance, Epidemiology, and End Results.

^bCombined if less than 50 in the registry identified as Hispanic.

Table 2. Characteristics of Hispanic/Latino and non-Hispanic colorectal cancer survivors who received a pneumococcal vaccination.

Characteristics	Received PNM ^a vaccine, n (%)			Did not receive PNM vaccine, n (%)		
	Hispanic (n=727, 4.5%)	Non-Hispanic (n=15525, 95.53%)	P value	Hispanic (n=782, 4.8%)	Non-Hispanic (n=15,467; 95.19%)	P value
Age at diagnosis (years)						
66-69	158 (21.7)	2465 (15.88)	<.001	157 (20.1)	2587 (16.73)	<.001
70-74	210 (28.9)	4094 (26.37)	<.001	226 (28.9)	3677 (23.77)	<.001
75-79	195 (26.8)	4027 (25.94)	<.001	202 (25.8)	3859 (24.95)	<.001
80-84	106 (14.6)	3069 (19.77)	<.001	126 (16.1)	3106 (20.08)	<.001
85+	58 (8)	1870 (12.05)	<.001	71 (9)	2238 (14.47)	<.001
Gender						
Male	360 (49.5)	6541 (42.13)	<.001	394 (50.4)	6698 (43.31)	<.001
Female	367 (50.5)	8984 (57.87)	<.001	388 (49.6)	8769 (56.69)	<.001
Urban status						
Urban	678 (93.3)	13879 (89.40)	<.001	735 (94)	13241 (85.61)	<.0001
Rural	49 (7)	1646 (10.60)	<.001	47 (6)	2226 (14.39)	<.001
Federal Poverty Level (%)						
0 to <5	107 (14.7)	4956 (31.92)	<.001	125 (16.0)	4276 (27.65)	<.001
5 to <10	146 (20.1)	4471 (28.80)	<.001	159 (20.3)	4385 (28.35)	<.001
10 to <20	251 (34.5)	4002 (25.78)	<.001	269 (34.4)	4273 (27.63)	<.001
20 to 100	223 (30.7)	2096 (13.50)	<.001	229 (29.3)	2533 (16.38)	<.001
Charlson Comorbidity Index						
0	400 (55.0)	9947 (64.07)	<.001	510 (65.2)	10950 (70.8)	.006
1	200 (27.5)	3811 (24.55)	<.001	181 (23.2)	3147 (20.35)	.006
2	81 (11)	1261 (8.12)	<.001	65 (8)	965 (6.2)	.006
3 or more	46 (6)	506 (3.3)	<.001	26 (3)	405 (2.6)	.006
Marital status						
Divorced	46 (6)	812 (5.2)	.03	54 (7)	888 (5.7)	.04
Married	403 (55.4)	8606 (55.43)	.03	403 (51.5)	8061 (52.12)	.04
Other	215 (29.6)	5095 (32.82)	.03	253 (32.4)	5424 (35.07)	.04
Single	63 (9)	1012 (6.52)	.03	72 (9)	1094 (7.07)	.04
Tumor stage at diagnosis						
Localized	416 (57.2)	9088 (58.54)	.92	406 (51.9)	8779 (56.76)	.03
Regional	279 (38.4)	5781 (37.24)	.92	335 (42.8)	5902 (38.16)	.03
Distant	19 (3)	388 (2.5)	.92	19 (2)	444 (2.9)	.03
Unstaged	13 (2)	268 (1.7)	.92	22 (3)	342 (2.2)	.03
Surveillance, Epidemiology, and End Results registry region						
San Francisco	41 (6)	516 (3.3)	<.001	35 (4)	527 (3.4)	<.001
Connecticut	29 (4)	1151 (7.41)	<.001	24 (3)	1057 (6.83)	<.001
San Jose	39 (5)	364 (2.3)	<.001	31 (4)	287 (1.9)	<.001
Los Angeles	104 (14.3)	975 (6.3)	<.001	125 (15.98)	916 (5.9)	<.001
Greater California	230 (31.6)	2360 (15.20)	<.001	278 (35.6)	2256 (14.59)	<.001
New Jersey	113 (15.5)	2800 (18.04)	<.001	126 (16.1)	2581 (16.69)	<.001
New Mexico	104 (14.3)	288 (1.9)	<.001	89 (11)	263 (1.7)	<.001

Characteristics	Received PNM ^a vaccine, n (%)			Did not receive PNM vaccine, n (%)		
	Hispanic (n=727, 4.5%)	Non-Hispanic (n=15525, 95.53%)	P value	Hispanic (n=782, 4.8%)	Non-Hispanic (n=15,467; 95.19%)	P value
Other	67 (9)	7071 (45.55)	<.001	74 (9)	7580 (49.01)	<.001

^aPNM: pneumococcal.

Pneumococcal Vaccination Among Hispanic and Non-Hispanic Colorectal Cancer Survivors

The results from our adjusted multivariable model indicated 10% decreased odds of vaccination for Hispanic CRC survivors versus non-Hispanics, which is marginally statistically significant (adjusted odds ratio [AOR] 0.889, 95% CI 0.798 to 0.99; $P=.033$; [Table 3](#)).

Characteristics associated with increased odds of PNM vaccine receipt included being between ages 70 and 74 years and 75

and 79 years compared with those between aged 65 and 69 years, an increasing modified Charlson comorbidity index, being married, and being female ($P<.05$ for all). As the percent of a population in a census tract at the FPL increased, receipt of PNM vaccine odds decreased (less than 5% at FPL with AOR 1.31, 95% CI 1.213 to 1.408; $P=<.01$; between 10% to 19.9% at FPL AOR 1.108, 95% CI 1.033 to 1.189; $P=.004$). A subgroup comparison ([Table 4](#)) of long-term CRC survivors by Hispanic identify revealed differences in place of receipt of PNM vaccination ($P=.030$).

Table 3. Adjusted odds of receiving a pneumococcal vaccine among long-term colorectal cancer survivors (Surveillance, Epidemiology, and End Results-Medicare 2000 to 2011; N=32,501).

Characteristic and category	Receipt of pneumococcal vaccination		
	Adjusted odds ratio	95% CI	P value
Hispanic/Latino (SEER^a-North American Association of Cancer Registries Hispanic Identification Algorithm identity)			
No	Ref ^b	— ^c	—
Yes	0.888	0.798-0.990	.032
Age at diagnosis (years)			
66-69	Ref	—	—
70-74	1.14	1.063-1.222	<.001
75-79	1.055	0.984-1.132	.13
80-84	0.991	0.920-1.068	.82
>85	0.848	0.780-0.923	<.001
Charlson Comorbidity Score			
0	Ref	—	—
1	1.344	1.274-1.418	<.001
2	1.47	1.348-1.602	<.001
3 or more	1.437	1.262-1.636	<.001
Marital status			
Single	Ref	—	—
Married	1.148	1.049-1.255	.003
Divorced	0.967	0.853-1.096	.60
Other	1.004	0.914-1.101	.94
Urban status			
Urban or Metro	1.298	1.206-1.397	<.001
Rural	Ref	—	—
Census tract federal poverty level (%)			
0 to <5	1.307	1.213-1.408	<.001
5 to <10	1.181	1.100-1.267	<.001
10 to <20	1.108	1.033-1.189	.004
20 to 100	Ref	—	—
Gender			
Male	Ref	—	—
Female	1.139	1.086-1.194	<.001
Tumor stage at diagnosis			
Local	Ref	—	—
Regional	0.941	0.898-0.985	.009
Distant	0.856	0.746-0.982	.03
Unstaged	0.779	0.664-0.914	.002
SEER registry region			
San Francisco	0.964	0.851-1.086	.56
Connecticut	1.020	0.930-1.120	.67
San Jose	1.220	1.046-1.421	.01
Los Angeles	1.036	0.943-1.138	.46

Characteristic and category	Receipt of pneumococcal vaccination		
	Adjusted odds ratio	95% CI	P value
Greater California	1.039	0.972-1.110	.26
New Jersey	1.000	0.936-1.069	>.99
New Mexico	1.312	1.127-1.526	<.001
All Others	Ref	—	—

^aSEER: Surveillance, Epidemiology, and End Results.

^bServes as reference group, adjusted odds ratio not applicable.

^cNot applicable.

Table 4. Places of service utilized by Hispanic and non-Hispanic colorectal cancer survivors who received their first pneumococcal vaccination in a Surveillance, Epidemiology, and End Results-Medicare cohort (N=16,171; P=.030).

Vaccine site of administration ^a	Total, N (%)	Non-Hispanic, n (%)	Hispanic, n (%)
Physician's office	12,012 (74.28)	11,482 (74.33)	530 (73.3)
Hospital	1499 (9.27)	1415 (9.2)	84 (11.6)
Home/Home Health Agency, Nursing Facility or Custodial Care	749 (4.6)	720 (9.2)	29 (4)
Specialty centers	207 (1.3)	192 (1.2)	15 (2)
Community Care Centers	1704 (10.54)	1639 (10.61)	65 (9)

^aPlace of service was excluded if unlisted.

Discussion

Principal Findings

To our knowledge, this is the first study evaluating the utilization of preventive services by long-term survivors of CRC, utilizing claims data. A recent study assessed preventive medical care in long-term survivors of CRC utilizing similar methods but looked only at Asian Pacific Islanders [19] and another study looked at preventive services in long-term cancer survivors using survey data, finding lower rates in Hispanic and African Americans [29]. We noted a marginally significant difference in PNM vaccination utilization among those who identified as Hispanic versus non-Hispanic using a robust, large dataset and controlling for confounding variables. Clinically, although small, this difference is important given the anticipated growth in Hispanic Americans.

Importantly, we also found that claims data show that only approximately 50% of Medicare beneficiaries, overall, received at least 1 PNM vaccine. This estimate lags not only behind the *Healthy People 2020* goal of 90% utilization but also behind current estimates of PNM vaccination in populations aged over 65 years [30,25]. Furthermore, this number is markedly lower than results from the most recent Medicare Beneficiary Survey (2013) where 69.1% of respondents answered *yes* to the question "Have you ever had a shot for pneumonia?" [30]; however, it is more consistent with rates found for all beneficiaries with Medicare claims through part A and B [31].

Strengths and Limitations

One strength of our study was limiting our subject population to long-term CRC survivors—a group that presumptively has had precedented contact with the medical community; furthermore, Medicare claims data and the study limitation to

those aged 65 years and older should help mitigate medical need and payment ability, given our primary outcome should have no direct expense to individuals. Our study further controlled for socioeconomic status and morbidity burden. Identification of those who received a vaccine, either by CPT or ICD-9 classification provided a more robust ability to signify access to health systems, as previously cited concerns in the literature include lack of insurance, access, or knowledge of vaccine recommendation.

Our study had limitations. First, as we utilized any receipt of PNM vaccine at any time via Medicare claims data, many individuals may have opted to defer Medicare B enrollment until after the age of 65 years but may have received the vaccine with an alternative payer. Second, although these data were captured so long as Medicare was the payer, the data fail to establish the temporal relation of survivorship before preventive service updates, and thus our study fails to inform on the importance of this preventive service in this population. Third, claims data do not permit for understanding those who may have been offered vaccination and declined to receive it. Fourth, our data rely on claims from Medicare; although this is likely comprehensive, it may fail to include some individuals with alternative payment mechanisms, including cash. Finally, our comorbidity score was calculated for all patients based on their claims from 1 year before diagnosis until 1 day before diagnosis. It is possible that individuals had different health protectors that may have influenced their vaccine status.

Conclusions

Future studies should consider better metrics to distinguish vulnerable populations, specifically Hispanic-Americans. Claims data permit for well-powered studies to evaluate the impact and meaning of certain preventive services, for example, rates of

hospitalization for disseminated PNM disease, versus those without this service. Cost-effectiveness data are also lacking in these groups, which could help health care professionals, administrators, and policy-makers reach better decisions to allocate resources. Finally, adherence to guideline-directed care and overutilization of certain services should be explored, given some individuals in our study received more than 3 vaccines.

Previous studies citing health care provider recommendations as important mediators to vaccination utilization are limited in that they rely on survey data. Survey data can introduce recall bias, which may allow adverse selection for those who have received services, inappropriately denoting increased utilization. Those who have foregone this preventive service may have done so for a number of reasons, including lack of a provider visit, lack of recall of getting this vaccine or misunderstanding of specifics (eg, influenza vaccination being mistaken for PNM vaccine), lack of knowledge of vaccine significance or

importance, or actual lack of occurrence. In addition, we found numerous beneficiaries receiving more than 3 vaccines (not reported) highlight concern for overutilization of resources: although guidelines and best practices support certain high-risk individuals who receive additional vaccinations pending clinical circumstances, it is permissible that other vaccines were administered because of a lack of communication and a segregated health care system.

Providers, health care managers, and informational system specialists should all carry some concern and responsibility; utilization of preventive services (PNM vaccine) is relatively low in our study population, and system-wide recognition is needed to rectify this. Although shared decision making and patient-centered care remain important, directed initiatives and resource allocation may help improve these findings and increase the utilization of evidence-based routine preventive care.

Conflicts of Interest

None declared.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;66(1):7-30 [FREE Full text] [doi: [10.3322/caac.21332](https://doi.org/10.3322/caac.21332)] [Medline: [26742998](https://pubmed.ncbi.nlm.nih.gov/26742998/)]
2. Singh SD, Henley SJ, Ryerson AB. Surveillance for cancer incidence and mortality - United States, 2013. *MMWR Surveill Summ* 2017 Dec 27;66(4):1-36 [FREE Full text] [doi: [10.15585/mmwr.ss6604a1](https://doi.org/10.15585/mmwr.ss6604a1)] [Medline: [28125567](https://pubmed.ncbi.nlm.nih.gov/28125567/)]
3. Henley SJ, Singh SD, King J, Wilson RJ, O'Neil ME, Ryerson AB. Invasive cancer incidence and survival - United States, 2013. *MMWR Morb Mortal Wkly Rep* 2017 Jan 27;66(3):69-75 [FREE Full text] [doi: [10.15585/mmwr.mm6603a1](https://doi.org/10.15585/mmwr.mm6603a1)] [Medline: [28125576](https://pubmed.ncbi.nlm.nih.gov/28125576/)]
4. Miller KD, Siegel RL, Lin CC, Mariotto AB, Kramer JL, Rowland JH, et al. Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin* 2016 Jul;66(4):271-289 [FREE Full text] [doi: [10.3322/caac.21349](https://doi.org/10.3322/caac.21349)] [Medline: [27253694](https://pubmed.ncbi.nlm.nih.gov/27253694/)]
5. American Cancer Society. Atlanta: American Cancer Society; 2014. Key Statistics for Colorectal Cancer URL: <https://www.cancer.org/cancer/colon-rectal-cancer/about/key-statistics.html> [accessed 2018-10-19] [WebCite Cache ID 73HkwotD9]
6. Lowenstein LM, Ouellet JA, Dale W, Fan L, Mohile SG. Preventive care in older cancer survivors. *J Geriatr Oncol* 2015 Mar;6(2):85-92 [FREE Full text] [doi: [10.1016/j.jgo.2014.12.003](https://doi.org/10.1016/j.jgo.2014.12.003)] [Medline: [25547206](https://pubmed.ncbi.nlm.nih.gov/25547206/)]
7. Earle CC, Neville BA. Under use of necessary care among cancer survivors. *Cancer* 2004 Oct 15;101(8):1712-1719 [FREE Full text] [doi: [10.1002/cncr.20560](https://doi.org/10.1002/cncr.20560)] [Medline: [15386307](https://pubmed.ncbi.nlm.nih.gov/15386307/)]
8. Snyder CF, Frick KD, Herbert RJ, Blackford AL, Neville BA, Wolff AC, et al. Quality of care for comorbid conditions during the transition to survivorship: differences between cancer survivors and noncancer controls. *J Clin Oncol* 2013 Mar 20;31(9):1140-1148 [FREE Full text] [doi: [10.1200/JCO.2012.43.0272](https://doi.org/10.1200/JCO.2012.43.0272)] [Medline: [23401438](https://pubmed.ncbi.nlm.nih.gov/23401438/)]
9. United States Census Bureau. 2016. Hispanic Heritage Month 2016 Fact Sheet URL: <https://www.census.gov/newsroom/facts-for-features/2016/cb16-ff16.html> [accessed 2018-10-18] [WebCite Cache ID 73HIUe4ZT]
10. Lees KA, Wortley PM, Coughlin SS. Comparison of racial/ethnic disparities in adult immunization and cancer screening. *Am J Prev Med* 2005 Dec;29(5):404-411. [doi: [10.1016/j.amepre.2005.08.009](https://doi.org/10.1016/j.amepre.2005.08.009)] [Medline: [16376703](https://pubmed.ncbi.nlm.nih.gov/16376703/)]
11. Siegel RL, Fedewa SA, Miller KD, Goding-Sauer A, Pinheiro PS, Martinez-Tyson D, et al. Cancer statistics for Hispanics/Latinos, 2015. *CA Cancer J Clin* 2015;65(6):457-480 [FREE Full text] [doi: [10.3322/caac.21314](https://doi.org/10.3322/caac.21314)] [Medline: [26375877](https://pubmed.ncbi.nlm.nih.gov/26375877/)]
12. Snyder CF, Earle CC, Herbert RJ, Neville BA, Blackford AL, Frick KD. Preventive care for colorectal cancer survivors: a 5-year longitudinal study. *J Clin Oncol* 2008 Mar 1;26(7):1073-1079. [doi: [10.1200/JCO.2007.11.9859](https://doi.org/10.1200/JCO.2007.11.9859)] [Medline: [18309941](https://pubmed.ncbi.nlm.nih.gov/18309941/)]
13. Weaver KE, Rowland JH, Bellizzi KM, Aziz NM. Forgoing medical care because of cost: assessing disparities in healthcare access among cancer survivors living in the United States. *Cancer* 2010 Jul 15;116(14):3493-3504 [FREE Full text] [doi: [10.1002/cncr.25209](https://doi.org/10.1002/cncr.25209)] [Medline: [20549763](https://pubmed.ncbi.nlm.nih.gov/20549763/)]
14. Juliette C, Anothony D, Tricia N, Gretchen J. Henry J Kaiser Family Foundation. 2018 Nov 28. Sources of Supplemental Coverage Among Medicare Beneficiaries in 2016 URL: <https://www.kff.org/medicare/issue-brief/sources-of-supplemental-coverage-among-medicare-beneficiaries-in-2016/> [accessed 2018-12-14] [WebCite Cache ID 73HleSZSe]

15. McBean AM, Babish JD, Prihoda R. The utilization of pneumococcal polysaccharide vaccine among elderly Medicare beneficiaries, 1985 through 1988. *Arch Intern Med* 1991 Oct;151(10):2009-2016. [doi: [10.1001/archinte.1991.00400100085015](https://doi.org/10.1001/archinte.1991.00400100085015)] [Medline: [1929689](https://pubmed.ncbi.nlm.nih.gov/1929689/)]
16. Centers for Disease Control (CDC). Update: pneumococcal polysaccharide vaccine usage--United States. *MMWR Morb Mortal Wkly Rep* 1984 May 25;33(20):273-6, 281. [Medline: [6425629](https://pubmed.ncbi.nlm.nih.gov/6425629/)]
17. Singleton JA, Santibanez TA, Wortley PM. Influenza and pneumococcal vaccination of adults aged > or = 65: racial/ethnic differences. *Am J Prev Med* 2005 Dec;29(5):412-420. [doi: [10.1016/j.amepre.2005.08.012](https://doi.org/10.1016/j.amepre.2005.08.012)] [Medline: [16376704](https://pubmed.ncbi.nlm.nih.gov/16376704/)]
18. National Cancer Institute.: National Institutes of Health; 2018. Table Describing Changes in Codes for NCI Comorbidity Index URL: <https://healthcaredelivery.cancer.gov/seermedicare/considerations/comorbidity-table.html> [accessed 2018-10-18] [WebCite Cache ID 73HnqmicR]
19. Steele CB, Townsend JS, Tai E, Thomas CC. Physician visits and preventive care among Asian American and Pacific Islander long-term survivors of colorectal cancer, USA, 1996-2006. *J Cancer Surviv* 2014 Mar;8(1):70-79 [FREE Full text] [doi: [10.1007/s11764-013-0319-1](https://doi.org/10.1007/s11764-013-0319-1)] [Medline: [24190589](https://pubmed.ncbi.nlm.nih.gov/24190589/)]
20. Catherine GA, Peg B, Fracis B, Michael G, Mei-chin H, Andrew L, et al. North American Association of Central Cancer Registries. 2011 Sep. NAACCR Guideline for Enhancing Hispanic-Latino Identification: Revised NAACCR Hispanic/Latino Identification Algorithm URL: https://www.naacrr.org/wp-content/uploads/2016/11/NHIA_v2_2_1_09122011.pdf [accessed 2018-10-19] [WebCite Cache ID 73HnvF8PG]
21. Bolen J, Adams M, Shenson D. Routine preventive services for older women: a composite measure highlights gaps in delivery. *J Womens Health (Larchmt)* 2007 Jun;16(5):583-593. [doi: [10.1089/jwh.2007.CDC5](https://doi.org/10.1089/jwh.2007.CDC5)] [Medline: [17627396](https://pubmed.ncbi.nlm.nih.gov/17627396/)]
22. Hurley LP, Bridges CB, Harpaz R, Allison MA, O'Leary ST, Crane LA, et al. U.S. physicians' perspective of adult vaccine delivery. *Ann Intern Med* 2014 Feb 4;160(3):161 [FREE Full text] [doi: [10.7326/M13-2332](https://doi.org/10.7326/M13-2332)] [Medline: [24658693](https://pubmed.ncbi.nlm.nih.gov/24658693/)]
23. Klabunde CN, Potosky AL, Legler JM, Warren JL. Development of a comorbidity index using physician claims data. *J Clin Epidemiol* 2000 Dec;53(12):1258-1267. [doi: [10.1016/S0895-4356\(00\)00256-0](https://doi.org/10.1016/S0895-4356(00)00256-0)] [Medline: [11146273](https://pubmed.ncbi.nlm.nih.gov/11146273/)]
24. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. *J Clin Epidemiol* 1993 Oct;46(10):1075-9; discussion 1081. [Medline: [8410092](https://pubmed.ncbi.nlm.nih.gov/8410092/)]
25. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992 Jun;45(6):613-619. [Medline: [1607900](https://pubmed.ncbi.nlm.nih.gov/1607900/)]
26. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40(5):373-383. [doi: [10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8)] [Medline: [3558716](https://pubmed.ncbi.nlm.nih.gov/3558716/)]
27. National Cancer Institute.: National Institutes of Health; 2018. Table Describing Changes in Codes for NCI Comorbidity Index URL: <https://healthcaredelivery.cancer.gov/seermedicare/considerations/comorbidity-table.html> [accessed 2019-04-15] [WebCite Cache ID 77ekVNIZQ]
28. National Cancer Institute.: National Institutes of Health; 2018. NCI Comorbidity Index Overview URL: <https://healthcaredelivery.cancer.gov/seermedicare/considerations/comorbidity.html> [accessed 2018-10-19] [WebCite Cache ID 73HnimV15]
29. Palmer NR, Geiger AM, Felder TM, Lu L, Case LD, Weaver KE. Racial/Ethnic disparities in health care receipt among male cancer survivors. *Am J Public Health* 2013 Jul;103(7):1306-1313 [FREE Full text] [doi: [10.2105/AJPH.2012.301096](https://doi.org/10.2105/AJPH.2012.301096)] [Medline: [23678936](https://pubmed.ncbi.nlm.nih.gov/23678936/)]
30. Centers for Medicare and Medicaid Services. 2013. Medicare Current Beneficiary Survey (MCBS) URL: <http://download.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/MCBS-Public-Use-File/downloads/2013atccodebook.pdf> [accessed 2018-10-19] [WebCite Cache ID 73HosSrEn]
31. Black CL, Williams WW, Warnock R, Pilishvili T, Kim D, Kelman JA. Pneumococcal vaccination among Medicare beneficiaries occurring after the advisory committee on immunization practices recommendation for routine use of 13-valent pneumococcal conjugate vaccine and 23-valent pneumococcal polysaccharide vaccine for adults aged ≥65 years. *MMWR Morb Mortal Wkly Rep* 2017 Jul 14;66(27):728-733 [FREE Full text] [doi: [10.15585/mmwr.mm6627a4](https://doi.org/10.15585/mmwr.mm6627a4)] [Medline: [28704347](https://pubmed.ncbi.nlm.nih.gov/28704347/)]

Abbreviations

AOR: adjusted odds ratio

CMS: Center for Medicare and Medicaid Services

CRC: colorectal cancer

CPT: Current Procedural Terminology

FPL: federal poverty level

HCPCS: Healthcare Common Procedure Coding System

ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification

NAACR: North American Association of Cancer Registries

NCI: National Cancer Institute

NHIA: NAACR Hispanic Identification Algorithm
PNM: pneumococcal
SEER: Surveillance, Epidemiology, and End Results

Edited by G Eysenbach; submitted 24.10.18; peer-reviewed by R Vita, W Camelo Castillo; comments to author 16.11.18; revised version received 29.01.19; accepted 30.03.19; published 13.05.19.

Please cite as:

Moran RJ, Waalen J, Murphy J, Nalawade V, Schiaffino M

Pneumococcal Vaccination Utilization Among Hispanic Long-Term Colorectal Cancer Survivors: Cross-Sectional Assessment of Claims

JMIR Public Health Surveill 2019;5(2):e12603

URL: <http://publichealth.jmir.org/2019/2/e12603/>

doi: [10.2196/12603](https://doi.org/10.2196/12603)

PMID: [31094341](https://pubmed.ncbi.nlm.nih.gov/31094341/)

©Ryan J Moran, Jill Waalen, James Murphy, Vinit Nalawade, Melody Schiaffino. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 13.05.2019. 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

New York City HIV Care Continuum Dashboards: Using Surveillance Data to Improve HIV Care Among People Living With HIV in New York City

Sarah L Braunstein¹, MPH, PhD; Karen Coeytaux², MPH, PhD; Charulata J Sabharwal³, MPH, MD; Qiang Xia¹, MPH, MD; Rebekkah S Robbins¹, MPH; Beverly Obeng¹, MPH; Demetre C Daskalakis⁴, MPH, MD

¹HIV Epidemiology and Field Services Program, Bureau of HIV Prevention and Control, New York City Department of Health and Mental Hygiene, Long Island City, NY, United States

²Rare Disease Registries Epidemiology and Biostatistics Group, Global Medical Affairs, Sanofi Genzyme, Bridgewater, NJ, United States

³Global Staph aureus Program, Pfizer Vaccine Clinical Research & Development, Pearl River, NY, United States

⁴Division of Disease Control, New York City Department of Health and Mental Hygiene, Long Island City, NY, United States

Corresponding Author:

Sarah L Braunstein, MPH, PhD
HIV Epidemiology and Field Services Program
Bureau of HIV Prevention and Control
New York City Department of Health and Mental Hygiene
42-09 28th Street
Long Island City, NY, 11101
United States
Phone: 1 347 396 7760
Email: sbraunstein@health.nyc.gov

Abstract

Background: HIV surveillance data can be used to improve patient outcomes.

Objective: This study aimed to describe and present findings from the HIV care continuum dashboards (CCDs) initiative, which uses surveillance data to quantify and track outcomes for HIV patients at major clinical institutions in New York City.

Methods: HIV surveillance data collected since 2011 were used to provide high-volume New York City clinical facilities with their performance on two key outcomes: linkage to care (LTC), among patients newly diagnosed with HIV and viral load suppression (VLS), among patients in HIV care.

Results: The initiative included 21 facilities covering 33.78% (1135/3360) of new HIV diagnoses and 46.34% (28,405/61,298) of patients in HIV care in New York City in 2011 and was extended to a total of 47 sites covering 44.23% (1008/2279) of new diagnoses and 69.59% (43,897/63,083) of New York City patients in care in 2016. Since feedback of outcomes to providers began, aggregate LTC has improved by 1 percentage point and VLS by 16 percentage points.

Conclusions: Disseminating information on key facility-level HIV outcomes promotes collaboration between public health and the clinical community to end the HIV epidemic. Similar initiatives can be adopted by other jurisdictions with mature surveillance systems and supportive laws and policies.

(*JMIR Public Health Surveill* 2019;5(2):e13086) doi:[10.2196/13086](https://doi.org/10.2196/13086)

KEYWORDS

HIV; surveillance; quality of care; best practices

Introduction

Background

Over the last two decades, tremendous advances in HIV treatment have transformed HIV into a chronic, manageable condition [1]. Indeed, available HIV treatment and prevention

programs allow persons living with HIV (PLWH) to live healthy lives and significantly reduce the spread of the disease [2-5]. Although HIV-related care outcomes have improved in recent years, HIV remains a significant public health challenge, with nearly 40,000 new infections in the United States in 2015 and over 6000 HIV-related deaths [6]. Ensuring that all PLWH

receive high-quality medical care has become 1 of the top national priorities in bending the HIV prevalence curve and ending the epidemic [7].

The HIV care continuum has long been used as a framework for monitoring care-related outcomes for PLWH and informing the quality of HIV care delivered. It offers a comprehensive overview of the efficiency of HIV clinical management by highlighting stages between time of diagnosis and viral load suppression (VLS) that might need improvement. In particular, it provides an understanding of the performance on timely linkage to care (LTC), retention in care, and achievement of VLS [8,9]. In 2010, the National HIV/AIDS Strategy (NHAS) called for a more coordinated response to the HIV epidemic and set national goals to be reached by the end of 2015. NHAS was updated in 2015, and new targets to be reached by the end of 2020 were provided [10,11].

Many national and local strategies call for the use of HIV surveillance and other public health data to measure population-level progress toward goals and identify gaps along the HIV care continuum [12]. HIV surveillance data are widely available, population-based, and collected in a standardized way across jurisdictions; as such, they are uniquely suited for use in measuring care continuum outcomes at the local and national level. The HIV surveillance registry of the New York City Department of Health and Mental Hygiene (NYC DOHMH) represents a comprehensive and high-quality source of HIV data and includes detailed information on providers and clinical facilities associated with HIV diagnoses and HIV-related laboratory tests indicating HIV care. NYC DOHMH has used surveillance data extensively to guide its programmatic and field activities [13,14]. New York State public health law emphasizes (and allows) sharing of HIV surveillance data externally by the health department to enhance patient LTC and retention in care [15]. The HIV care continuum can be adapted to measure the effectiveness of individual clinical institutions, inform the quality of HIV care delivered by those institutions, and point to possible areas for improvement and intervention planning.

Objectives

In 2012, NYC DOHMH launched the HIV care continuum dashboard (CCD) initiative, which consisted of providing selected New York City clinical providers with facility-level aggregate data reflecting the HIV care performance of their own facilities, their peers, and New York City as a whole. The goals of the CCD initiative were to monitor local progress toward the NHAS goals, to identify potential low-performing facilities in need of intervention, and to encourage adoption of best practices from high-performing facilities. We report on the implementation of this initiative by NYC DOHMH and the novel use of outcomes data to improve clinical management of HIV care by New York City providers.

Methods

Data Sources

New York State Public Health law requires named reporting to the NYC DOHMH of all HIV/AIDS diagnoses, all HIV-related illness, and all cluster of differentiation 4 (CD4), viral load (VL), and genotype tests conducted for New York City PLWH [15]. The NYC DOHMH manages the New York City HIV surveillance registry, which is continuously updated with demographic, clinical, and other information on persons receiving HIV care in New York City and meeting the HIV surveillance case definitions of the Centers for Disease Control and Prevention [16]. New York City Vital Statistics Registry and national death data (ie, National Death Index and Social Security Death Master File data) are routinely used to update death information on vital status in the registry. The New York City registry contains a cumulative total of more than 10 million HIV-related laboratory test results for over 240,000 individuals.

Outcomes

The CCDs are facility-level performance reports derived from HIV surveillance data, displaying indicators on timely LTC, VLS, and VL below transmission threshold (BTT) for a specific 12-month period. In general, New York City HIV surveillance data are lagged to account for reporting delays as well as a standard dissemination timeline, such that data for the previous calendar year are released in December each year. The CCDs are released twice annually, in June and December. The June CCDs contain data for July to June of the previous year, and the December CCDs contain data for January to December of the previous year (eg, CCDs released in December 2012 contained data for January-December 2011, CCDs released in June 2013 contained data for July 2011-June 2012, and so on). CCDs are sent by email to each facility's leadership, and VLS data from the December CCDs only are published on a dedicated page on the DOHMH website. [17]

The CCDs contain eligible facilities' performance on LTC and VLS from the most recent analyzable 12-month period (Figure 1). The reports also include data from the site's previous CCD so that facilities can evaluate their progress over time. Facilities also receive data on the proportion of their newly diagnosed patients that were linked at their facility versus at other New York City facilities. National and local targets are highlighted with goal lines for each indicator, enabling facilities to evaluate whether they meet the goals and make them aware of the current recommendations, as these targets can evolve with changes in national and local policy. In addition, facilities are provided with rank plots reflecting the LTC and VLS performance for all CCD sites during the same period, which allows comparison against peers' performance (Figure 2). A frequently asked questions document is included in each release to assist facilities with interpreting the CCD. Finally, DOHMH has a dedicated email account to answer queries and provide assistance to sites.

Figure 1. Overview of New York City HIV care continuum dashboard, site-specific performance. CCD: care continuum dashboard; CD4: cluster of differentiation 4; NYC: New York City.

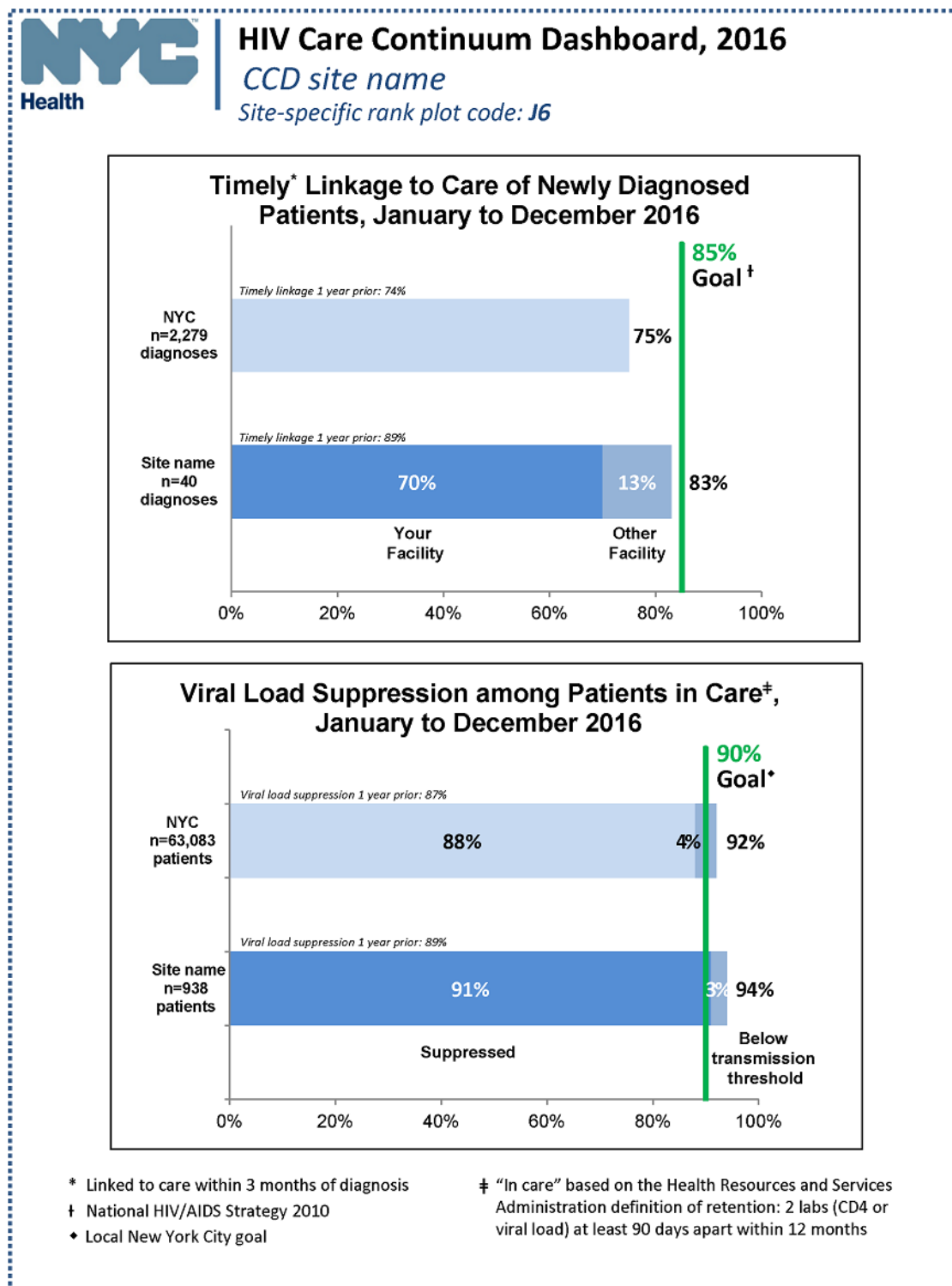
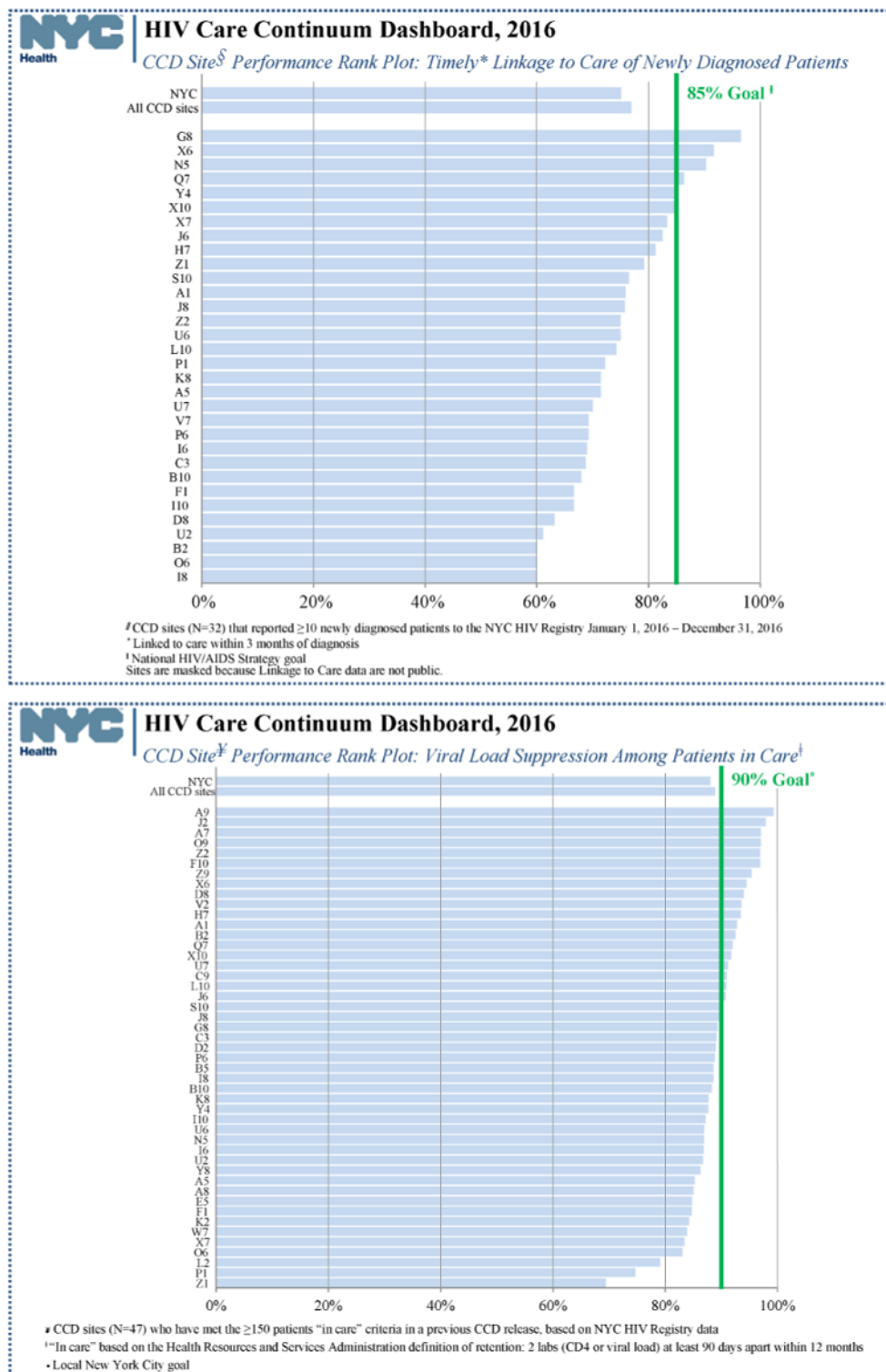


Figure 2. Overview of New York City HIV care continuum dashboard, linkage to care, and viral suppression rank plots. CCD: care continuum dashboard; CD4: cluster of differentiation 4.



Timely Linkage to Care

The timely LTC metric represents the proportion of newly diagnosed persons over 12 months at a given facility who linked to care within 3 months of their diagnosis. Linkage could have occurred either at the facility reporting the diagnosis or at any other facility in New York City. Timely LTC is considered to have occurred if any VL, CD4, or genotype test is ordered by

the provider within 8 to 91 days of HIV diagnosis and reported to the DOHMH. HIV-related laboratory tests drawn within 7 days of a patient’s HIV diagnosis are likely to be part of the diagnostic work-up and therefore, do not represent LTC [18]. The timely LTC indicator is generated for facilities with 10 or more new diagnoses in the time period considered. The facility-based goal of 85% aligns with the LTC goal set forth in the 2010 and 2015 NHAS [10,11].

Viral Load Suppression

The VLS metric represents the proportion of persons whose most recent quantitative HIV VL was ≤ 200 copies/mL among all persons in care at the facility within the 12-month period of interest [19]. A patient is considered to be in continuous care at a facility if they have at least 2 HIV laboratory reports (CD4 or VL) 90 days apart or more ordered by this facility during the given 12-month time period. A patient who meets this definition at more than 1 New York City facility is only included in the CCD for the ordering facility of the patient's last HIV laboratory test during the time period of interest. The VLS goal of 85% is based on examination of local New York City HIV surveillance data.

Viral Load Below Transmission Threshold

The "VL BTT" indicator represents the proportion of patients at a given facility whose most recent VL was $< 1,500$ copies/mL among all persons in care (using the same definition of continuous care as for viral suppression) at that facility during the 12-month time period of interest. This indicator has both public health significance (as a measure of the proportion of patients with reduced potential to transmit HIV to partners) and clinical significance (as a measure of the proportion of patients who have low if not yet suppressed VL) [20,21]. In addition, this indicator serves as an adjustment for facilities that may care for populations with more barriers to achieving progress through the continuum of care. In exploratory analyses to develop the "VL BTT indicator" as an adjusted measure of viral suppression, no single demographic characteristic (eg, gender, race/ethnicity, age, borough of residence, and HIV transmission risk) of patients in care was useful for distinguishing those facilities that might have more complex patient populations and therefore more challenges to viral suppression. We therefore decided not to make specific adjustments to VL based on specific clinic populations but to create a generic indicator to capture the proportion of patients at each clinic who have low but not suppressed VL. VL < 1500 copies/mL, which is lower than most VL set points in untreated individuals, implies an effort by a facility to treat patients with antiretroviral therapy (ART) and as such provides a crude estimate of the preventive impact of such treatment efforts in PLWH.

In more recent CCDs (those released in 2016 and 2017), the VLS target was moved to 90% from 85% in response to the substantial improvements over time in VLS among the CCD sites (see Figures 1 and 2). However, for easier comparison across years in this report, the VLS target was kept at 85%. A weighted average was used when assessing trends for a group of clinical sites to account for the relative size of the sites. In addition, low-performing facilities are defined as sites whose VLS or LTC is at least 10 percentage points below the 85% target.

Results

Program Implementation

The first CCD release in December 2012 included data for 21 high-volume facilities (those with ≥ 1000 patients, and all New York City public hospitals). All site-affiliated clinics are included in a single facility CCD. In the December 2015 release, NYC DOHMH included 26 additional facilities (all with ≥ 150 patients in care), added the "BTT indicator," and publicly released VLS data on DOHMH's website for the 21 original CCD facilities. VLS data on all 47 CCD facilities have been publicly released since 2015 (Figure 3).

The 21 original sites whose data were released in December 2012 collectively covered 33.78% (1135/3360) of all newly HIV-diagnosed persons and 46.34% (28,405/61,298) of PLWH in care in New York City in 2011. In the December 2012 release, all 21 sites were eligible to receive LTC data for their newly diagnosed patients and VLS data for their patients established in care. Out of the 47 clinical sites that received CCDs in December 2017, 32 sites representing 41% of all new diagnoses were eligible to receive LTC data and all 47 were eligible to receive VLS data. These 47 sites collectively covered 44.23% of all newly diagnosed persons (1008/2279) and 69.59% (43,897/63,083) of PLWH in care in New York City in 2016.

Impact on Outcomes

Table 1 shows the trends in VLS and LTC for the 21 original sites, which received all CCDs released during 2012 to 2017, reporting 2011 to 2016 data. The number of patients in care at the 21 sites remained relatively stable over the 4-year period. The weighted average VLS performance substantially increased from 73% in 2011 to 89% in 2016, with a sharp increase between 2012 and 2013 coinciding with the early roll-out of the recommendation of ART treatment for all. Improvements in VLS performance observed among the 5 lowest performing facilities were comparable with those among the 5 highest performing facilities, with a 13-point increase for both. The number of facilities meeting or exceeding the 85% target for VLS increased from 0 in 2011 to 16 in 2016. Moreover, there was a near-complete reduction in the number of low-performing sites performing at least 10 percentage points below target, from 13 in 2011 to 1 in 2016.

The number of newly diagnosed patients at the 21 original sites declined by about one-third between 2011 and 2016, from 1135 in 2011 to 720 in 2016. There was a modest increase in the weighted average performance on LTC (from 76%-77%) as well as a small increase in the number of facilities achieving the 85% national goal for linkage (from 2-4). The number of facilities performing at least 10 percentage points below target remained stable (from 11-12). In general, the declines in new HIV diagnoses, substantial increases in viral suppression, and more modest increases in LTC among the CCD sites are consistent with overall trends in HIV care outcomes in New York City during this period [22].

Figure 3. New York City (NYC) human immunodeficiency virus care continuum dashboards (CCDs) release timeline, 2012-2017. CCD releases are generated using data from a previous 12-month period to ensure completeness of the data (eg, the December 2012 release included data for January-December 2011; the June 2013 release included data for July 2011-June 2012, and so on). LTC: linkage to care; VLS: viral load suppression; BTT: below transmission threshold.

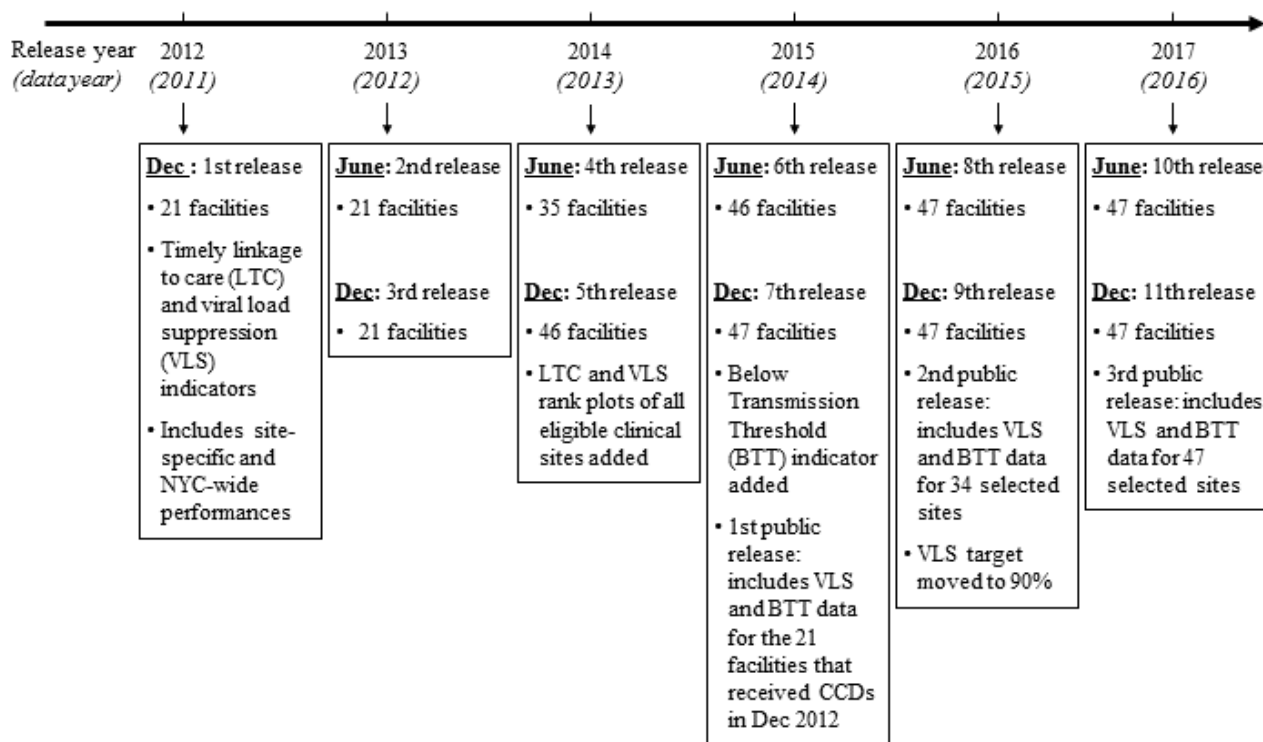


Table 1. 2011 to 2016 performance on viral load suppression and linkage to care for the 21 original care continuum dashboard sites.

Indicator	2011 ^a	2012	2013	2014	2015	2016
Viral load suppression (VLS)						
Patients in care, n	28,253	30,462	30,156	30,630	28,734	30,874
VLS, weighted average (%)	73	73	84	86	88	89
Sites above 85% VLS target, n	0	0	9	10	14	16
Sites at least 10% below target, n	13	13	1	1	0	1
Linkage to care (LTC)						
Diagnoses, n	1135	1024	991	914	769	720
LTC, weighted average (%)	76	74	77	79	78	77
Eligible sites (≥10 diagnoses), n	21	20	19	20	20	19
Sites above 85% LTC target, n	2	1	2	4	4	4
Sites at least 10% below target, n	11	10	7	8	6	12

^aRecommendation for antiretroviral therapy for all people living with diagnosed HIV infection made in New York City in December 2011.

Discussion

Principal Findings

NYC DOHMH used the wealth of data in the HIV surveillance registry to create facility-level reports—the CCDs—that provide clinicians with valuable data for monitoring and improving the clinical management of PLWH in care in New York City. The CCD initiative fosters collaboration among HIV-care stakeholders, including providers, the Health Department, and patients. Ideally, this effort empowers providers to continuously

raise their standards to ensure timely care of newly diagnosed patients, achieve undetectable VL for all PLWH, and ultimately prevent forward transmission of HIV in New York City.

Since the launch of the initiative, ongoing dialogue with some recipients has encouraged them to take ownership of the CCDs and use the reports to monitor their progress and address some clinical management issues rapidly. For example, 1 facility contacted the health department following the reporting of an unusually low VLS performance for their facility. This prompted an internal investigation by the facility, which revealed a lack

of adherence to treatment among some of their patients and underscored the need for closer follow-up. The VLS performance of this facility has since substantially improved, although it remains below target. Another facility noticed important discrepancies between the LTC performance reported in their 2016 CCD and their internal estimates. On the basis of their feedback, NYC DOHMH investigated these discrepancies and found that some patients who had had their first care visit within 7 days after diagnosis failed to come back for an additional visit within 3 months. As a result of this observation, the facility became more aware of the importance of following up within 3 months with recently diagnosed patients for ongoing care engagement. We will monitor potential improvements in LTC at this site because of this investigation going forward. To date, such valuable feedback from sites has been received on a case-by-case basis. In the future, we are considering conducting a more systematic evaluation of the utility of the CCDs to recipient sites. In general, it is difficult to determine the full impact of the CCDs on patient outcomes at the sites because of the ecologic nature of the data and reports.

Ongoing discussions with providers have also helped to refine the methodology used in creating the CCDs to capture HIV care as accurately as possible. For example, VL and CD4 count were initially the only tests considered when assessing LTC. After consulting with providers, it was decided in 2016 to also include genotype tests as a marker of LTC. Going forward, advancements in biomedical therapies (eg, tenofovir alafenamide fumarate instead of tenofovir) might lead providers to draw laboratory tests less frequently, which could influence the validity of our current definition of retention in care. For example, it may be more effective to consider patients with 1 laboratory test or more in the year retained in care versus requiring 2 laboratory tests at least 3 months apart. Similarly, providers are being encouraged to link patients to care immediately after HIV diagnosis, and so changes to our 3-month linkage indicator in the near future are being considered.

The CCD initiative relies on the programmatic use of CD4, VL, diagnostic, and genotype tests that are reportable to DOHMH under New York State law. The use of these data for measuring and monitoring HIV care-related outcomes at the population level has been validated in previous studies [18-21,23,24]. New York City DOHMH has developed sophisticated procedures to ensure the quality, completeness, and timeliness of its HIV registry data. In particular, the NYC DOHMH registry integrates high-quality provider/facility-level data. Surveillance data can therefore be used to assess care received by a patient across all New York City facilities as opposed to the single facility-level view that is typically available to providers. High-quality data, an expanded purview for using surveillance data for public health action under New York State law, and local and national strategies that emphasize data-to-care approaches, have been cornerstones in the development of the CCDs.

Limitations

The data used to generate the CCDs have inherent limitations. The New York City HIV surveillance registry contains

laboratory tests ordered by New York City providers only. Thus, it fails to systematically capture patients who move or receive care outside of New York City. Therefore, patients who moved out of jurisdiction shortly after being diagnosed in New York City might appear as not timely linked to care (or not linked to care ever) despite successfully linking to care outside of New York City. Follow-up discussion with facilities focusing on patients in need of engagement in care often results in identification of such cases. Furthermore, clinic visits not associated with laboratory tests are not captured. This is an inherent limitation of using lab-based surveillance data to assess care engagement. However, previous validation work showed this discrepancy was relatively small [18]. In addition, to ensure completeness of data, CCDs are generated using data for a 12-month period plus a reporting lag. Moving forward, this lag time would ideally be shortened and enable closer to real-time reporting. Furthermore, facility-level VLS performance in the CCDs applies only to patients considered to be established in care at a provider based on the specific and relatively conservative definition of retention in care used for the CCDs; the definition fails to capture patients who are transitioning between providers as well as individuals who may be engaged in care but receiving less frequent VL monitoring. Despite these limitations, this definition allows us to capture and report on the care status of approximately two-thirds of New York City PLWH in care in the CCDs. Finally, by design, the CCDs do not include information on patients who are not in care or who have been lost to care. However, DOHMH maintains the HIV Care Status Reports system, which is a secure, Web-based system New York City providers can use to query information about lost-to-care patients against the surveillance registry for a check on patients' current HIV care status in New York City.

Conclusions and Implications

The CCDs are a novel approach to sharing aggregate-level surveillance data with providers who are responsible for the medical care of HIV patients. They enable the DOHMH and individual recipients to monitor progress toward national and local HIV care goals and monitor the quality of care delivery, detect areas for improvement, and inform the development of interventions (eg, training and technical assistance). Through the public release of VLS data, providers can identify peers with strong outcomes and engage in discussions regarding best practices. The CCDs also enable the identification of lower performing sites. To assist these sites, NYC DOHMH is developing a program in collaboration with New York State Department of Health to overcome possible barriers to the delivery of high-quality HIV care. Finally, public availability of CCD data promotes patient choice in where to seek HIV care. HIV surveillance data are a rich data source for monitoring outcomes such as engagement in care and viral suppression among PLWH in a jurisdiction and can be used effectively to monitor outcomes at the facility level. Similar initiatives can be adopted by other jurisdictions with mature surveillance systems and capacity and laws that support sharing surveillance data with providers.

Acknowledgments

The authors would like to thank the staff of the HIV Epidemiology and Field Services Program for their work in collecting and managing the data that underpin this work. This work was supported in part by the Centers for Disease Control and Prevention (6 NU62PS003639-05-05 and 5 NU62PS003993-05).

Authors' Contributions

SLB directed the initiative and led the conceptualization of the analysis and writing of the manuscript. KC completed the analysis and drafted the manuscript and served as lead analyst on the initiative while at DOHMH. CJS helped develop and launch the initiative while at DOHMH and contributed to the manuscript. QX, RSR, and BO worked on the analysis and contributed to the manuscript and generally contribute to the surveillance data system. DCD provided oversight of the initiative and contributed to the manuscript.

Conflicts of Interest

None declared.

References

1. Patterson S, Cescon A, Samji H, Chan K, Zhang W, Raboud J, CANOC collaboration. Life expectancy of HIV-positive individuals on combination antiretroviral therapy in Canada. *BMC Infect Dis* 2015 Jul 17;15:274 [FREE Full text] [doi: [10.1186/s12879-015-0969-x](https://doi.org/10.1186/s12879-015-0969-x)] [Medline: [26183704](https://pubmed.ncbi.nlm.nih.gov/26183704/)]
2. TEMPRANO ANRS 12136 Study Group, Danel C, Moh R, Gabillard D, Badje A, Le Carrou J, et al. A trial of early antiretrovirals and isoniazid preventive therapy in Africa. *N Engl J Med* 2015 Aug 27;373(9):808-822. [doi: [10.1056/NEJMoa1507198](https://doi.org/10.1056/NEJMoa1507198)] [Medline: [26193126](https://pubmed.ncbi.nlm.nih.gov/26193126/)]
3. Lundgren J, Babiker AG, Neaton JD. Antiretroviral therapy in early HIV infection. *N Engl J Med* 2016 Dec 28;374(4):394. [doi: [10.1056/NEJMc1513311](https://doi.org/10.1056/NEJMc1513311)] [Medline: [26816019](https://pubmed.ncbi.nlm.nih.gov/26816019/)]
4. Palella FJ, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N Engl J Med* 1998 Mar 26;338(13):853-860. [doi: [10.1056/NEJM199803263381301](https://doi.org/10.1056/NEJM199803263381301)] [Medline: [9516219](https://pubmed.ncbi.nlm.nih.gov/9516219/)]
5. Samji H, Cescon A, Hogg RS, Modur SP, Althoff KN, Buchacz K, et al. Closing the gap: increases in life expectancy among treated HIV-positive individuals in the United States and Canada. *PLoS One* 2013;8(12):e81355 [FREE Full text] [doi: [10.1371/journal.pone.0081355](https://doi.org/10.1371/journal.pone.0081355)] [Medline: [24367482](https://pubmed.ncbi.nlm.nih.gov/24367482/)]
6. Centers for Disease Control and Prevention. 2016. HIV in the United States URL: <https://www.cdc.gov/hiv/statistics/overview/ataglance.html> [accessed 2018-08-31] [WebCite Cache ID 725fHrjVn]
7. New York State Department of Health. 2015. 2015 Blueprint for Ending the Epidemic URL: https://www.health.ny.gov/diseases/aids/ending_the_epidemic/docs/blueprint.pdf [accessed 2018-08-31] [WebCite Cache ID 725fZWVS4]
8. Centers for Disease Control and Prevention. 2014. Integrated Guidance for Developing Epidemiologic Profiles: HIV Prevention and Ryan White HIV/AIDS Programs Planning URL: https://www.cdc.gov/hiv/pdf/guidelines_developing_epidemiologic_profiles.pdf [accessed 2018-08-31] [WebCite Cache ID 725fqIKyU]
9. Gardner EM, McLees MP, Steiner JF, Del RC, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clin Infect Dis* 2011 Mar 15;52(6):793-800 [FREE Full text] [doi: [10.1093/cid/ciq243](https://doi.org/10.1093/cid/ciq243)] [Medline: [21367734](https://pubmed.ncbi.nlm.nih.gov/21367734/)]
10. White House Office of National AIDS Policy. 2010. National HIV/AIDS Strategy: Federal Implementation Plan URL: <https://files.hiv.gov/s3fs-public/nhas.pdf> [accessed 2018-08-31] [WebCite Cache ID 725gBfQwU]
11. White House Office of National AIDS Policy. 2015. National HIV/AIDS Strategy for the United States: Updated to 2020 URL: <https://www.hiv.gov/sites/default/files/nhas-2020-action-plan.pdf> [accessed 2018-08-31] [WebCite Cache ID 725g8Ji7Z]
12. Centers for Disease Control and Prevention. 2014. Monitoring selected national HIV prevention and care objectives by using HIV surveillance data: United States and 6 dependent areas, 2012 URL: <http://www.cdc.gov/hiv/library/reports/surveillance/> [accessed 2018-08-31] [WebCite Cache ID 725gLPWaH]
13. Udeagu CN, Webster TR, Bocour A, Michel P, Shepard CW. Lost or just not following up: public health effort to re-engage HIV-infected persons lost to follow-up into HIV medical care. *AIDS* 2013 Sep 10;27(14):2271-2279. [doi: [10.1097/QAD.0b013e328362fdde](https://doi.org/10.1097/QAD.0b013e328362fdde)] [Medline: [23669157](https://pubmed.ncbi.nlm.nih.gov/23669157/)]
14. Wiewel EW, Braunstein SL, Xia Q, Shepard CW, Torian LV. Monitoring outcomes for newly diagnosed and prevalent HIV cases using a care continuum created with New York city surveillance data. *J Acquir Immune Defic Syndr* 2015 Feb 01;68(2):217-226. [doi: [10.1097/QAI.0000000000000424](https://doi.org/10.1097/QAI.0000000000000424)] [Medline: [25394192](https://pubmed.ncbi.nlm.nih.gov/25394192/)]
15. State of New York. HIV Testing and Counseling. In: State of New York Laws. Albany, New York: State of New York; 2010.

16. Centers for Disease Control Prevention. Revised surveillance case definition for HIV infection--United States. 2014. Morbidity and Mortality Weekly Report URL: <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr6303a1.htm> [accessed 2019-05-25] [WebCite Cache ID 78drynMEf]
17. New York City Department of Health and Mental Hygiene. 2017. HIV Care Continuum Dashboard URL: <https://www1.nyc.gov/site/doh/health/health-topics/care-continuum-dashboard.page> [accessed 2018-08-31] [WebCite Cache ID 725jMnKF3]
18. Sabharwal CJ, Braunstein SL, Robbins RS, Shepard CW. Optimizing the use of surveillance data for monitoring the care status of persons recently diagnosed with HIV in NYC. *J Acquir Immune Defic Syndr* 2014 Apr 15;65(5):571-578. [doi: [10.1097/QAI.0000000000000077](https://doi.org/10.1097/QAI.0000000000000077)] [Medline: [24326601](https://pubmed.ncbi.nlm.nih.gov/24326601/)]
19. Health Resources and Services Administration. 2009. The HIV/AIDS Program: HAB Performance Measures Group URL: <https://hab.hrsa.gov/sites/default/files/hab/About/clinical-quality-management/coremeasures.pdf> [accessed 2018-08-31] [WebCite Cache ID 725ghBsfs]
20. Marks G, Gardner LI, Rose CE, Zinski A, Moore RD, Holman S, et al. Time above 1500 copies: a viral load measure for assessing transmission risk of HIV-positive patients in care. *AIDS* 2015 May 15;29(8):947-954 [FREE Full text] [doi: [10.1097/QAD.0000000000000640](https://doi.org/10.1097/QAD.0000000000000640)] [Medline: [25768835](https://pubmed.ncbi.nlm.nih.gov/25768835/)]
21. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, Wabwire-Mangen F, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. *N Engl J Med* 2000 Mar 30;342(13):921-929. [doi: [10.1056/NEJM200003303421303](https://doi.org/10.1056/NEJM200003303421303)] [Medline: [10738050](https://pubmed.ncbi.nlm.nih.gov/10738050/)]
22. HIV Epidemiology and Field Services Program. HIV Surveillance Annual Report, 2016. New York, NY: New York City Department of Health and Mental Hygiene; Dec 2017.
23. Centers for Disease Control and Prevention. National HIV Prevention Progress Report, 2013 URL: <https://www.cdc.gov/hiv/pdf/policies-nationalprogressreport-optim.pdf> [accessed 2018-08-31] [WebCite Cache ID 725hQYggK]
24. Pati R, Robbins RS, Braunstein SL. Validation of retention in HIV care status using the New York City HIV surveillance registry and clinical care data from a large HIV care center. *J Public Health Manag Pract* 2017;23(6):564-570 [FREE Full text] [doi: [10.1097/PHH.0000000000000515](https://doi.org/10.1097/PHH.0000000000000515)] [Medline: [28079643](https://pubmed.ncbi.nlm.nih.gov/28079643/)]

Abbreviations

ART: antiretroviral therapy

BTT: below transmission threshold

CCD: care continuum dashboard

CD4: cluster of differentiation 4

LTC: linkage to care

NHAS: National HIV/AIDS Strategy

NYC DOHMH: New York City Department of Health and Mental Hygiene

PLWH: persons living with HIV

VL: viral load

VLS: viral load suppression

Edited by T Sanchez; submitted 14.12.18; peer-reviewed by P Sweeney, K Brady, D Gruber, P Wortley; comments to author 01.02.19; revised version received 25.02.19; accepted 08.03.19; published 19.06.19.

Please cite as:

Braunstein SL, Coeytaux K, Sabharwal CJ, Xia Q, Robbins RS, Obeng B, Daskalakis DC

New York City HIV Care Continuum Dashboards: Using Surveillance Data to Improve HIV Care Among People Living With HIV in New York City

JMIR Public Health Surveill 2019;5(2):e13086

URL: <http://publichealth.jmir.org/2019/2/e13086/>

doi: [10.2196/13086](https://doi.org/10.2196/13086)

PMID: [31219053](https://pubmed.ncbi.nlm.nih.gov/31219053/)

©Sarah L Braunstein, Karen Coeytaux, Charulata J Sabharwal, Qiang Xia, Rebekkah S Robbins, Beverly Obeng, Demetre C Daskalakis. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 19.06.2019. 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

Social Media Surveillance for Outbreak Projection via Transmission Models: Longitudinal Observational Study

Anahita Safarishahrbijari¹, MSc; Nathaniel D Osgood¹, PhD

University of Saskatchewan, Saskatoon, SK, Canada

Corresponding Author:

Anahita Safarishahrbijari, MSc

University of Saskatchewan

506, 101 Cumberland Ave S

Saskatoon, SK,

Canada

Phone: 1 3067139797

Email: anahita.safari@usask.ca

Abstract

Background: Although dynamic models are increasingly used by decision makers as a source of insight to guide interventions in order to control communicable disease outbreaks, such models have long suffered from a risk of rapid obsolescence due to failure to keep updated with emerging epidemiological evidence. The application of statistical filtering algorithms to high-velocity data streams has recently demonstrated effectiveness in allowing such models to be automatically regrounded by each new set of incoming observations. The attractiveness of such techniques has been enhanced by the emergence of a new generation of geospatially specific, high-velocity data sources, including daily counts of relevant searches and social media posts. The information available in such electronic data sources complements that of traditional epidemiological data sources.

Objective: This study aims to evaluate the degree to which the predictive accuracy of pandemic projection models regrounded via machine learning in daily clinical data can be enhanced by extending such methods to leverage daily search counts.

Methods: We combined a previously published influenza A (H1N1) pandemic projection model with the sequential Monte Carlo technique of particle filtering, to reground the model by using confirmed incident case counts and search volumes. The effectiveness of particle filtering was evaluated using a norm discrepancy metric via predictive and dataset-specific cross-validation.

Results: Our results suggested that despite the data quality limitations of daily search volume data, the predictive accuracy of dynamic models can be strongly elevated by inclusion of such data in filtering methods.

Conclusions: The predictive accuracy of dynamic models can be notably enhanced by tapping a readily accessible, publicly available, high-velocity data source. This work highlights a low-cost, low-burden avenue for strengthening model-based outbreak intervention response planning using low-cost public electronic datasets.

(*JMIR Public Health Surveill* 2019;5(2):e11615) doi:[10.2196/11615](https://doi.org/10.2196/11615)

KEYWORDS

machine learning; infectious disease transmission; disease models; system dynamics analysis; social media; outbreaks; infodemiology; infoveillance

Introduction

The capacity to accurately project communicable disease outbreak evolution is of great value in public health planning for prevention and control strategies. Use of such information can inform resource allocation, including surge-capacity planning and planning of the timing of outbreak response immunization campaigns, and, when applied across distinct scenarios, provide a basis for evaluating tradeoffs between intervention strategies. Although dynamic models are

increasingly widely used to conduct such scenario projection, the construction of such models for new and rapidly evolving pathogens commonly faces significant barriers due to uncertainties regarding important factors governing the natural history of the disease, such as duration of latent, incubation, and infectious phases; the probability of asymptomatic carriage; rates of waning immunity; contact rates; and per-discordant-contact transmission probabilities. Moreover, even the most intricate models face strict limitations in their ability to project evolution of factors treated as stochastic, such as weather-related variables and the timing of arrival of

exogenous infections due to global travel. Using computational statistical estimation methods such as sequential Monte Carlo techniques, in recent years, researchers have contributed approaches to elevate the predictive accuracy of dynamic transmission models by updating their state estimates at the time of appearance of each new observation. The predictive accuracy of methods has thus far been evaluated purely in the context of models that make use of traditional surveillance data sets, such as laboratory and clinically confirmed case reports [1-6].

Although such traditional surveillance data sets offer high-quality, rich information about individuals who present for medical care, they suffer from notable shortcomings, including delayed reporting and a failure to include counts of infective individuals who choose not to present for care. In a separate stream of work from the dynamic modeling work noted above, in recent years, researchers have sought to compensate for the limitations of traditional epidemiological data sources more generally by exploiting information related to online communicational behavior, particularly, the growing tendency of many users to search, post, and tweet about their illnesses. Specifically, such researchers have assessed the health insights that can be gained from public health surveillance applications employing a variety of online sources of information.

A prominent line of this work has focused on time sequences of search query volumes, such as those previously captured in Google Flu Trends (GFT) [7] and (on a more generic and continuing basis) Google Trends [8]. Within this sphere, a wide variety of investigations have used statistical and machine learning methods to perform classification and analysis on such Google search volume data and volumes of social media postings, including those for communicable illnesses [9-12]. Many researchers have investigated biomedical and health-related knowledge obtained from the Twitter platform, suggesting opportunities and limitations associated with different machine learning classifiers and training models for tweet mining [13-15]. Other case studies have reported a significant correlation between Tweets and clinical reports and concluded that social media text mining can improve public health communication efforts by providing insight into major themes of public concerns in the health sphere [16,17].

An important subset of research in this area has leveraged data obtained from Google to develop statistical forecast models and evaluated the degree to which GFT data in combination with statistical models can support accurate predictions [18-20] and correlations with real-time empirical data [21]. Some investigators jointly used multiple data sources, including GFT and Twitter, and compared the performance of statistical prediction models using each data source and in scenarios where different data sources complement one another [22,23].

The prediction of epidemic outbreaks by dynamic models often involves significant error and generally needs to consider both underlying dynamics and noise related to both measurement and process evolution. Although older techniques based on Kalman Filtering and variants [24] have long provided a computationally frugal means of filtering stochastic dynamic models, such maximum likelihood estimation-based approaches

are impaired by strong distributional assumptions concerning measurement and process noise and limited accommodation for nonlinearity in characterization of the system. This challenge in handling nonlinearity is experienced most in terms of an inability to capture the effects of probability distributions across multiple basins of attraction and a requirement for model linearization that is problematic for important modeling formalisms, such as agent-based models. For these and other reasons, recent research has increasingly turned to stronger filtering methods. Several authors have applied the sequential Monte Carlo technique of particle filtering as an effective tool in support of both model estimation and predictions from real-world data. Ong et al established a real-time surveillance system in Singapore to feed data into a stochastic model of influenza-like disease dynamics, which was refitted daily using particle filtering [1]. Osgood and Liu used a synthetic ground truth model to evaluate the effectiveness of particle filtering for an H1N1-like infection in the presence of noisy data and systematic model simplifications [2]. Safarishahrbijari et al evaluated the effectiveness of particle filtering subject to specifics of the configuration, such as frequency of data sampling and representation of behavior change in the form of an evolving contact rate for H1N1 [3,5]. Oraji et al developed a system dynamics model for studying the tuberculosis transmission and applied particle filtering to estimate the latent state of the system, including many epidemiological quantities that are not directly measured. Their results suggested an improvement in model accuracy using particle filtering and high additional value extending from consideration of additional epidemiological quantities in the probabilistic model [4]. Li et al applied particle filtering to a measles compartmental model using reported measles incidence for Saskatchewan. They also performed particle filtering on an age-structured adaptation of their model by dividing the population into age groups for children and adults. According to their results, particle filtering can offer high predictive capacity for measles outbreak dynamics in a low-vaccination context [6].

Epstein et al explored the effect of adaptive behaviors such as social distancing based on fear and contact behavior in models of epidemic dynamics. They used nonlinear dynamic systems and agent-based computation and integrated disease and fear of the disease contagion processes. Based on their models, individuals anxious (“scared”) about or infected by a pathogen can transfer fear through contact with other individuals who are not scared, and scared individuals may isolate themselves, thereby influencing the contact rate dynamic, which is a key parameter in governing outbreak evolution. The authors studied flight as a behavioral response and concluded that even small levels of fear-inspired flight can have a dramatic impact on spatiotemporal epidemic dynamics [25].

Despite the fact that both high-velocity search volume and social media data and transmission models share a temporal perspective, data drawn from such internet series has not, to our knowledge, been previously used as a source of information for filtering (via recurrent regrouping) compartmental transmission models with the arrival of new data.

In this work, we sought to address that gap by combining the transmission model from the study by Epstein et al [25] with

the sequential Monte Carlo method of particle filtering, considering the interaction between disease and fear of disease contagion processes for the 2009-2010 H1N1 influenza pandemic. The particle filtered model used time series of both clinically observed data and daily Google search query volumes to automatically and recurrently re-ground the model as successive data points became available. Based on lessons learned from previous studies [3,5] about the importance of incorporating higher-velocity data rather than time-averaged data, we made use of daily data. In contrast to past particle filtering work on grounding transmission models, which have used empirical data purely as a comparison with model results reflecting the natural history of infection, the model presented here engaged in such comparisons for the clinical data and further compared the search query volume data with ideation-related model state (individuals with fear).

Methods

Particle Filtered Model

In the first stage of characterization of the particle filtered model, we present the formulation of the existing Epstein compartmental model from a previous study [25], which characterizes the population into states according to both their natural history of infection and presence of anxiety regarding influenza. The state variables of the model are as follows: Susceptible to pathogen and fear (S), Infected with fear (I_F), Infected with pathogen (I_P), Infected with pathogen and fear (I_{FP}), Removed due to fear (R_F), Removed due to fear and pathogen (R_{FP}), and Recovered (R). We used an adaptation of the model that included an Exposed (E) state variable (Figure 1). In this model, λ_F is the (hazard) rate of removal due to self-isolation of those in fear only, λ_P refers to the rate of recovery from infection with pathogen, λ_{FP} represents the rate of removal due to self-isolation of the infected who are also afraid, and H is the rate of recovery from fear (alone) and return to circulation [25]. The parameters α and β denote transmissibility of fear and pathogen, respectively. Specifically, α represents the probability that a contact between an individual A who is currently without fear but who is susceptible or infected purely with the pathogen and an individual B with either fear or the pathogen will cause individual A to become afraid. In contrast, β denotes the probability that a contact between an

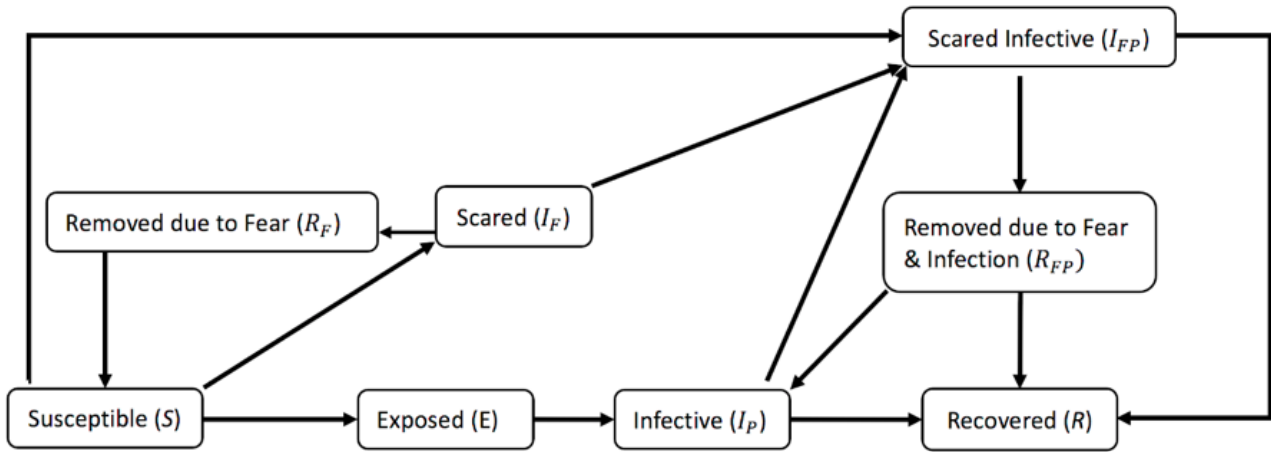
individual A who has never been infected with the pathogen and an individual B who is specifically infected with the pathogen will infect individual A with the pathogen. Given that α and β are probabilities (and are thus of unit dimension), it bears emphasis that simple dimensional analysis demonstrates that the original authors assume an effective per-person-per-unit time mixing rate with a value of unity. Although not considered within the scope of the original article, this mixing rate can itself be characterized in accordance with long-time mathematical epidemiology practice as the product of a per-unit-time contact rate c and disease transmissibility divided by the (constant) total population N . Because we consider changes to the value of c within this work, this quantity is shown explicitly in the equations below. To explain this term, which is required for dimensional consistency, we note that each transmission term, such as:



can be considered to characterize the rate of transmission (in terms of persons per unit time) from possible transmitters in category Y (here, I_{FP}) to persons in at-risk category X (here, S). Each such at-risk person X is assumed to engage in an average of c contacts per unit time. Those overall contacts are then assumed to be spread proportionally among the compartments in the population, with the fraction taking place with those in a category Y of possible transmitters, which is the count of people in Y divided by the total population N . The probability in the prefix of the term (here, $\beta\alpha$) indicates the probability that each such potentially transmitting contacts leads to the type of transmission being considered in that term (either fear, pathogen, or, as in this example, both).

When adapting the model, we took advantage of the previously demonstrated [3,5] capacity of particle filtering to support stochastic evolution of designated parameters (captured as state variables). One of the stochastic parameters included in this model represents the fraction of reported incidents (f_P), which is the fraction of people who are reported to public health authorities when emerging from the latent state and is both uncertain and evolving over time. Likewise, the fraction of people becoming afraid who search Google upon infection, named the fraction of Google search incidents (f_F), is further treated as a dynamic uncertain parameter.

Figure 1. System dynamics model.



Other parameters also treated as stochastic are the contact rate (c), removal rate from those with fear to self-isolation (λ_F), and removal rate from those with fear who are also infected (λ_{FP}). To support this, such dynamic parameters are associated with state variables evolving over time according to stochastic differential equations. Because variable c is a nonnegative quantity, we performed a log-transform on this variable according to the Brownian Motion, so that it varied over the real numbers. The stochastic differential equation of contact rate c is described as:

$$dc = \dots$$

(1), where dW_t is a standard Wiener process following a normal distribution with mean of 0 and variance of 1. Thus,

$$c = \dots$$

follows a normal distribution with mean of 0 and variance of s_c^2 . We also performed a log-transform on λ_F ; the stochastic differential equation of λ_F is formulated as:

$$d\lambda_F = \dots$$

The initial values of c and λ_F are drawn uniformly from the interval between 0 and 100 per day and between 0.4 and 1 per day, respectively. The SDs of s_c and s_{λ_F} were both selected to be 1.

In contrast, reflecting the fact that f_p and f_F represent fractions, such parameters were logit-transformed, with the initial value for each varying between 0 and 0.2. We described the stochastic differential equations of fractions f_p and f_F according to Brownian Motion as:

$$df_p = \dots$$

Within the model, the parameter f_p is multiplied by inflows to state variables Infected (I) and Scared Infective (I_{FP}) to account for fractional actual reporting. Similarly, the parameter f_F is multiplied by inflows to state variables Scared (I_F), Scared Infective (I_{FP}), Removed due to Fear and Infection (R_{FP}), and Removed due to Fear (R_F) and accounts for the fractional of the actual scared population.

We treated λ_{FP} as:

$$d\lambda_{FP} = \dots$$

and then considered λ'_{FP} as a fraction and performed a logit-transform on it. This parameter varies over the range from 0 to 1 and the dynamic process for λ'_{FP} is similar to f_p and f_F , specifically,







$$d\lambda'_{FP} = \dots$$

The SDs s_{f_p} , s_{f_F} , and $s_{\lambda'_{FP}}$ are selected to be 5, 5 and 1, respectively. The initial values of f_p , f_F , and λ'_{FP} are set on the intervals [0, 0.2), [0, 0.2) and [0, 0.5), respectively.

By applying random walks to these parameters, a more accurate estimate was achieved during model simulation. As such, in our model, each particle at each point in time is associated with all state variables and state variables associated with stochastic parameters ($S, E, I_F, I_P, I_{FP}, R_F, R_{FP}, R, c, f_p, f_F, \lambda_F$, and λ'_{FP}) (Table 1).

$$df_F = \dots$$

Table 1. Parameters used in the model.

Parameter name	Notation	Value for Quebec	Value for Manitoba	Unit
Probability of infection transmission given exposure	β	0.04	0.04	Unit
Probability of fear transmission given exposure	α	0.02	0.02	Unit
Mean latent time	τ	3	3	Day
Mean time to recovery	μ	7	7	Day
Total population of province	N	7843475	1214403	Person
Rate of recovery from fear	H	0.2	0.2	One per day
Rate of removal to self-isolation from fear	λ_F	Dynamic	Dynamic	One per day
Fraction of mean time to recovery of going from “Scared Infected” to “Recovered” via “Removed Due to Fear & Infection”	λ'_{FP}	Dynamic	Dynamic	Unit
Rate of removal to self-isolation from fear and pathogen	λ_{FP}			One per day
Rate of recovery from infection with pathogen	λ_P			One per day
Rate of recovery from removal due to fear and infection	λ'_P			One per day

Description of Data Sources

We evaluated the prediction of the above-described dynamic model assisted by particle filtering against two publicly available empirical datasets. The first was from Manitoba Health - Healthy Living and Seniors and included daily laboratory-confirmed case counts of pandemic H1N1 influenza for the period of October 6, 2009, through January 18, 2010, for the province of Manitoba [26]. The second dataset was from the Institut National de Santé Publique du Québec, a public health expertise and reference center in Quebec, and included daily confirmed case counts of pandemic H1N1 influenza between October 6, 2009, and December 19, 2010 [27].

In addition to the daily clinical case count data noted above, we obtained normalized daily Google search counts from Google trends and weekly normalized data from GFT for Manitoba and Quebec during the second pandemic wave. Reflecting the linguistic differences between the two provinces, the search terms used for each were distinct. In Manitoba, we used search terms “flu” and “H1N1,” while for Quebec, we used “flu,” “Influenza A virus sub-type H1N1,” “h1n1 vaccination,” “ah1n1,” “ah1n1 vaccin,” “grippe,” and “grippe ah1n1,” which are the most frequent search queries related to this topic suggested by Google during that period.

Particle Values and Parameter Values

When defining the likelihood function for observing empirical data, given the state of a given particle, the exact variant of the likelihood used varied across three different scenarios examined. The first scenario evaluated the impact of assuming a likelihood formulation that considered purely clinical data, termed $L_{infection}$ with pathogen. The likelihood being used in the second scenario considered only the likelihood of observing the empirical data regarding Google search counts for the appropriate province in

light of the count of individuals posited to be currently in fear within the model, a likelihood denoted as $L_{infection}$ with fear.

Following several past contributions [2-4,28], we assume that each epidemiological quantity follows a Pascal distribution function. Thus, y_t and i_t represent observed individuals per day and particle-posed daily rate (count per day) of new cases, respectively:

$$y_t \sim \text{Pascal}(r, p_P)$$

In the formulation for the likelihood function, r is a dispersion parameter

$$i_t \sim \text{Pascal}(r, p_F)$$

and

$$y_{Pt} \sim \text{Pascal}(r_P, p_{Pt})$$

where y_{Pt} and y_{Ft} represent number of laboratory-confirmed incident cases reported for day t and number of Google search incidents for that day, respectively. The probabilities p_P and p_F follow

$$p_P = \frac{1}{1 + r_P}$$

and

$$p_F = \frac{1}{1 + r_F}$$

respectively, where i_{Pt} is a fraction of the flow of new cases of infection and i_{Ft} is a fraction of the flow of new cases of scared. The dispersion parameter $L_{infection}$ with pathogen (r_P) was considered as 40, while $L_{infection}$ with fear (r_F) was considered as 25. This reflects the larger noise that we believed to be associated with Google search data, in light of the fact that a larger dispersion parameter leads to a more narrowly dispersed distribution.

The third scenario considered a total likelihood function L_T consisting of a combination of $L_{infection\ with\ pathogen}$ and $L_{infection\ with\ fear}$. For defining the total likelihood function, the simplifying assumption was made that deviations with respect to one measure were independent of the other, and thus, the total multivariate likelihood function could be treated as a multiplication of two univariate likelihood functions, given as $L_T=L_{infection\ with\ pathogen} \times L_{infection\ with\ fear}$

The purpose of running this third scenario was to compare the effectiveness of a univariate likelihood function with that of the multivariate likelihood function, when evaluated in terms of a calculated discrepancy of model predictions against the epidemiologically confirmed case count.

The three scenarios noted above were conducted using particle filtering, employing 1000 particles. For each such scenario, reflective of the need to make decisions in light of uncertainty about the evolution of an unfolding outbreak, in which only information about time points up to the present is available, we sought to examine the impact of right censoring the empirical data at certain time point T^* , representing the current time (ie, the time from which the model is forecasting outbreak evolution). Thus, as the model ran, particle weights were updated based on observations from day one until and including day T^* ; after day T^* , particle filtering ceased, particle weights were no longer updated using historic data, and no further particles were resampled. Each scenario included a sequence of subscenarios that employed the following distinct values of T^* : {25, 30, 35, 40, 45, 50}.

To judge the accuracy of particle filter – informed projections for future times against the standard of the reported case counts for those times, we defined a discrepancy metric as the expected value of the L^2 norm of the difference between sampled particles (reporting rate coefficient \times [infected state+scared infected state]) and reported case count observations calculated after

time T^* . We sampled n particles ($n=700$) according to their weights and obtained the discrepancy value using the following equation:

$$d_i = \sum_{j=1}^n (x_{ij} - y_i)^2$$

where

$$x_{ij}$$

is the value associated with sampled particle j at observation i and

$$y_i$$

is the respective reported clinical cases at observation i . T_f is the final observation time, and T^* indicates the time from which the projection is being made (ie, the time up to which the particles' weights were updated based on observation, where $0 \leq t \leq T^*$). Using this formulation, we evaluated how well projections forward predicted the empirical data after T^* , the time at which particle filtering was completed.

Results

In this work, for each scenario (each associated with a particular likelihood function), we plotted the graphs associated with $T^*=30$ for Manitoba and Quebec. We characterize the results below according to the scenario.

Particle Filtering Using Two Likelihood Functions

Figures 2 and 3 depict the empirical data (red and magenta points) superimposed on samples (blue and green) from the model-generated distribution of particles for the model output of the number of reported cases (left panel) and number of searches (right panel) for Manitoba (Figure 2) and Quebec (Figure 3). For $T^*=30$, the high posterior density for the projection period is quite localized for the cases of pathogen and the number of searches.

Figure 2. Empirical data (red and magenta points) superimposed on samples (blue and green) from the model-generated distribution of particles for the model output of the count of reported cases (left panel) and number of searches (right panel) using two likelihood functions, $T^*=30$ for Manitoba.

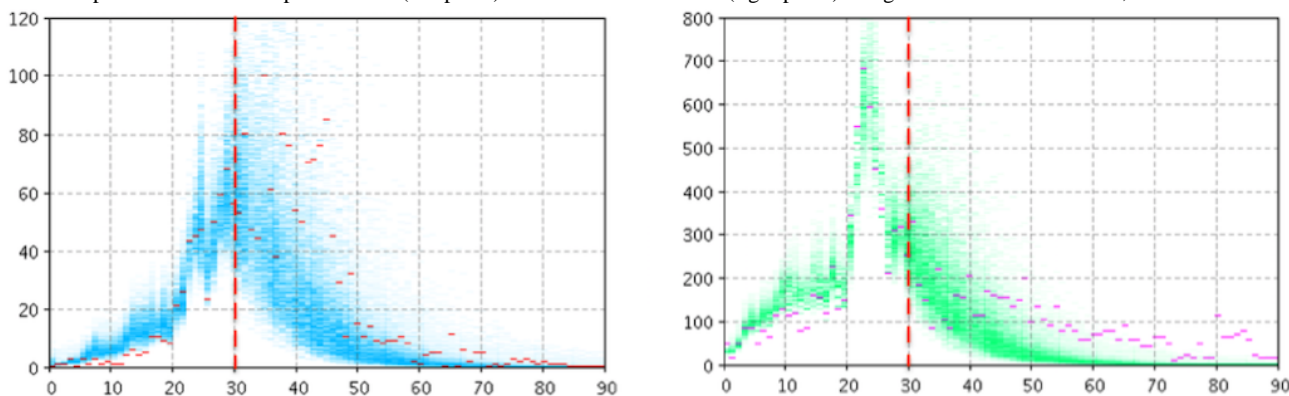
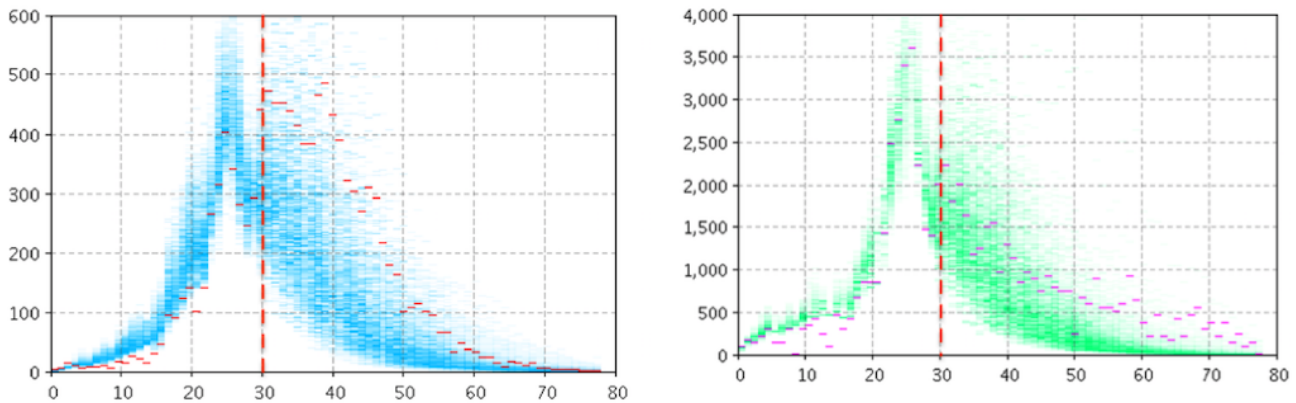


Figure 3. Empirical data (red and magenta points) superimposed on samples (blue and green) from the model-generated distribution of particles for the model output of the count of reported cases (left panel) and number of searches (right panel) using two likelihood functions, $T^*=30$ for Quebec.



Particle Filtering Using the Likelihood Function Associated With Clinical Data Alone

In this configuration, particle filtering was performed using $L_{infection\ with\ pathogen}$ as the sole likelihood function. Figures 4 and 5 depict empirical data (red and magenta points) superimposed on samples (blue and green) from the model-generated distribution of particles for the model output of the number of reported cases (left panel) and number of searches (right panel)

for Manitoba (Figure 4) and Quebec (Figure 5). Despite the fact that the particle filtering employs reasonably high-resolution clinical data, the system exhibits great difficulty both in accurately projecting the number of clinical case reports forward from the point where particle filtering ceases (T^*) and in doing so in a fashion where the high posterior density region is localized. Unsurprisingly, the model informed by the reported clinical case counts alone is unable to accurately characterize the search volume within the population.

Figure 4. Empirical data (red and magenta points) superimposed on samples (blue and green) from the model-generated distribution of particles for the model output of the count of reported cases (left panel) and count of searches (right panel) using the likelihood function associated with clinical data alone, $T^*=30$ for Manitoba.

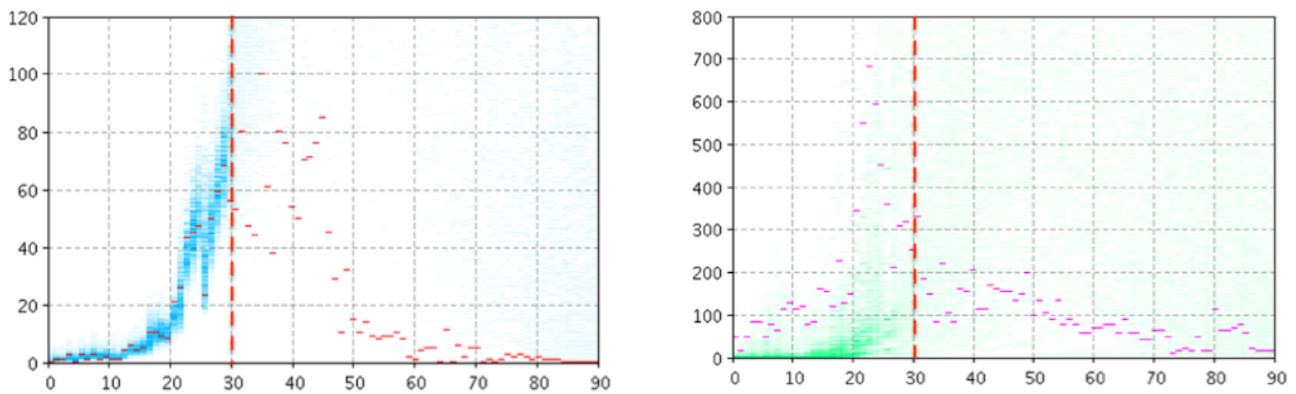
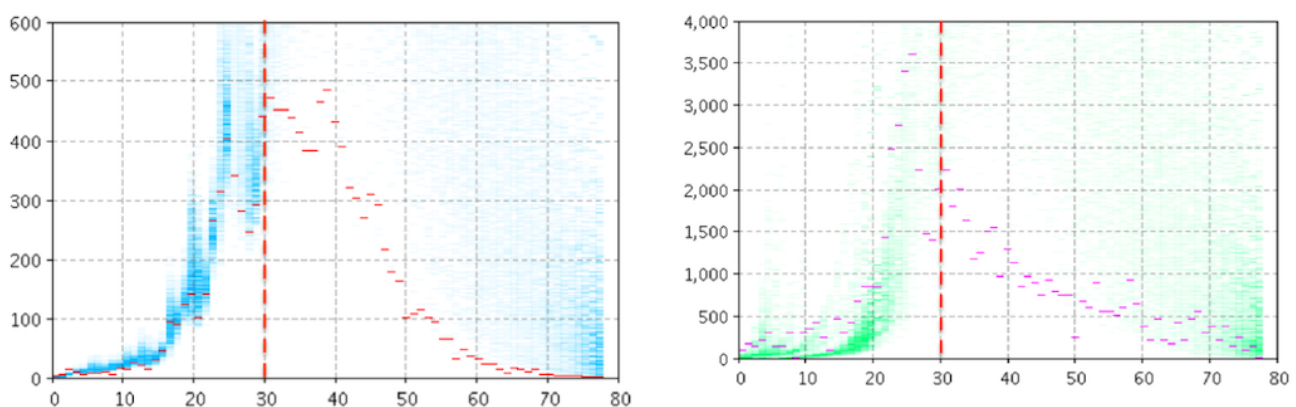


Figure 5. Empirical data (red and magenta points) superimposed on samples (blue and green) from the model-generated distribution of particles for the model output of the count of reported cases (left panel) and count of searches (right panel) using the likelihood function associated with clinical data alone, $T^*=30$ for Quebec.



Particle Filtering Using the Likelihood Function Associated With Search Volume Data Alone

In this configuration, particle filtering was performed using $L_{infection\ with\ fear}$ as the sole likelihood function. Figures 6 and 7 depict empirical data (red and magenta points) superimposed on samples (blue and green) from the model-generated distribution of particles for the model output of the number of

reported cases (left panel) and number of searches (right panel) for Manitoba (Figure 6) and Quebec (Figure 7). Although the results for both jurisdictions show some localization in the projections of the prevalent case count of those living in fear, the failure to consider the clinical case count in particle filtering (and to accordingly update the model estimates for the current number of infectives, susceptibles, and the contact rate) leads to poor projection accuracy for the reported clinical case count.

Figure 6. Empirical data (red and magenta points) superimposed on samples (blue and green) from the model-generated distribution of particles for the model output of the count of reported cases (left panel) and count of searches (right panel) when using the likelihood function associated with search volume data alone, $T^*=30$ for Manitoba.

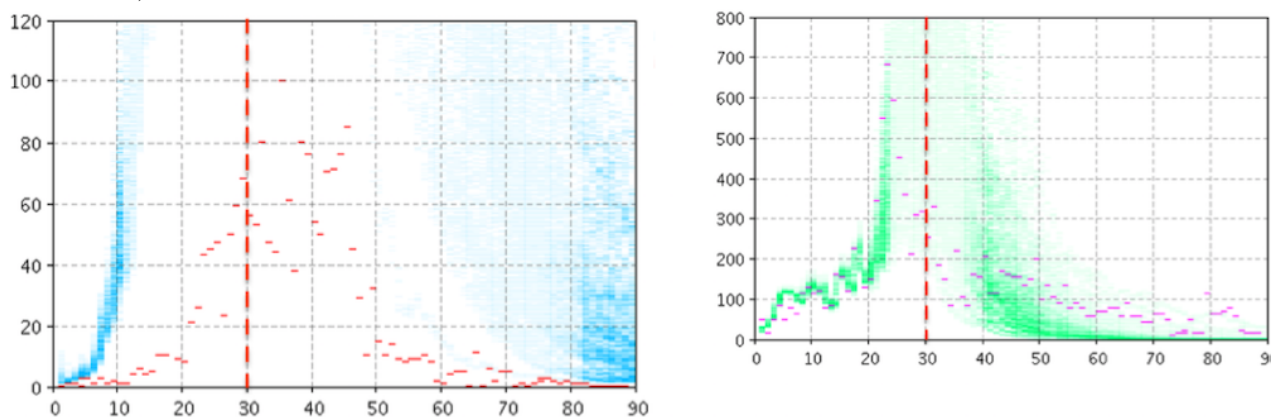


Figure 7. Empirical data (red and magenta points) superimposed on samples (blue and green) from the model-generated distribution of particles for the model output of the count of reported cases (left panel) and count of searches (right panel) when using the likelihood function associated with search volume data alone, $T^*=30$ for Quebec.

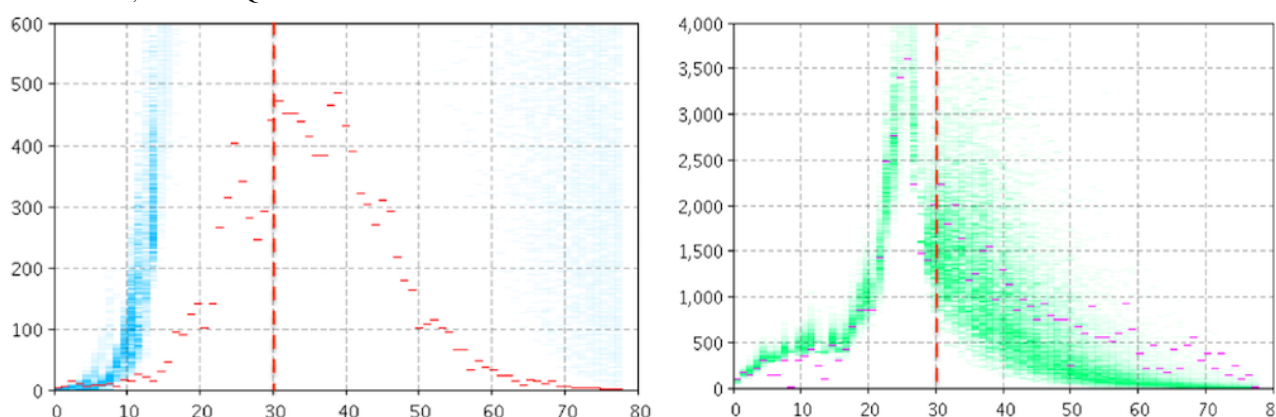


Table 2. Discrepancies associated with different scenarios and T^* values for Manitoba and Quebec.

Scenario	T^*					
	25	30	35	40	45	50
Manitoba						
Google Likelihood	7,846,178	2,896,092	1,941,998	695,330	192,819	13,569
Clinical Likelihood	956,021	604,749	564,651	469,159	106,307	3275
Two Likelihoods	545	361	174	158	60	12
Quebec						
Google Likelihood	577,919,468	437,577,329	290,486,216	108,993,972	29,645,905	9,179,791
Clinical Likelihood	31,571,941	3,544,611	461,804	55,938	4862	751
Two Likelihoods	535,927	17,386	8338	3322	1071	336

Comparison of Results Associated With Different Scenarios

Table 2 depicts the discrepancies between model clinical case predictions and empirical data for different check times (T^*) for Manitoba and Quebec. Unsurprisingly, given the results above, the discrepancy associated with particle filtering informed by both clinical and search volume data sets (in Particle Filtering Using Two Likelihood Functions) is smaller than the discrepancy associated with either data set in isolation. In addition, the discrepancy when using particle filtering informed by the (higher-quality) clinical case count data alone is lower than that informed purely by search volume. However, there is a marked difference between Manitoba and Quebec in the levels of discrepancy seen when using clinical case data alone as compared to using search volume data. For Manitoba, there is consistently less than an order of magnitude of difference in discrepancies between these two results. In contrast, for Quebec, using the clinical data alone within particle filtering yields a level of discrepancy several orders of magnitude below that resulting from search volume data. Intriguingly, for Manitoba, combining both data yields a reduction of discrepancy many orders of magnitude below either, despite the fact that discrepancy is calculated with respect to clinical case reports. This advantage of adding information from the search volume data to that from clinical case counts presumably reflects the fact that the added search volume information supports particle filtering in more accurately localizing the model state estimates than was the case using purely the reported clinical case counts—a factor manifested in the projections for both clinical case counts. In contrast, for Quebec, using both sources of information reduces the discrepancy significantly, typically by at least one order of magnitude, with the exception of time points $T^*=45$ and $T^*=50$.

Discussion

Principal Findings

In this contribution, we investigated the predictive accuracy gains from applying particle filtering using both traditional and search volume data to estimate latent states of a compartmental transmission model (including time evolution of stochastic parameters involved in that model). The capacity to perform this estimation then provides support for projection and scenario evolution using the model.

To be able to use search data effectively when particle filtering a transmission model, we found it helpful to move beyond the traditional scope of compartmental transmission models and to adopt a more articulated model of the outbreak, reflecting the fact that causal drivers promoting Web searches are not restricted to stages in the natural history of infection, but are additionally driven by factors with distinct but coupled dynamics, such as fluctuations in perceived risk on the part of the population. Responsive to this consideration, we have adapted a previously published model with an explicit consideration of the coupled dynamics of fear and pathogen. Although there are challenges associated with assessing perceived risk and anxiety on the part of the population during an outbreak, we found here that projection of outbreak dynamics

can be materially enhanced through inclusion of a surprisingly accessible source of data: Daily relative search query volumes for defined geographic regions on the widely used Google search engine. The reliable and timely public availability of such data across many areas of the world raises the prospects for significantly enhancing effective outbreak projection using combinations of dynamic modeling and machine learning techniques such as the particle filter.

Limitations

The work presented here has significant limitations. Although search trend data provide some indication of topic-specific interest over time in a defined spatial region, from the standpoint of “big data,” it is often available only with modest (daily) temporal resolution and frequently coarse geographic resolution. It is also affected by many unobserved confounders. Such search trend data are further limited by providing little sense of count of distinct users and no sense of longitudinal progression of a single user. In this regard, the Google search query volume time series compare unfavorably to the richness of information present in other publicly available types of online data, such as region-specific Twitter feeds.

In addition to the shortcomings in the data sources employed, there are notable methodological limitations of our study. The likelihood function employing two distinct data sources was simplistic in its design, merely serving to multiply each of the dataset-specific likelihood functions. The use of a random walk during particle filtering for no fewer than five distinct parameters likely contributes to a rapid divergence in the model’s estimates, compared to the behavior observed in previous particle filtered models of influenza [1,3]. Further experimentation is required with the parameters governing such random walks. A more significant yet large gain in accuracy, given the limited volatility likely for some of such parameters, may result from treating such parameters as unknown constants to be sampled for a given simulation from a posterior distribution within Particle Markov Chain Monte Carlo (PMCMC) techniques [29].

Such limitations point to natural avenues for future work. We expect that the prospects for the sorts of projections explored here will be significantly elevated by combining such data with other public data sources containing distinct sources of information, such as daily or finer resolution time series from Twitter and Tumblr. We further expect the accuracy of the projections to be improved by more powerful machine learning techniques, such as through the use of PMCMC techniques, ensemble techniques supporting inclusion of multiple models, and potential PMCMC techniques employing multiple models using reverse-jump MCMC strategies.

Conclusions

Pandemic forecasting is important for public health policy making due to its support for judicious planning involving resource allocation. Official statistics typically capture only subsets of the epidemiological burden (eg, the subset of individuals who engage in care seeking). Prospects for rapid use of such data to understand outbreak evolution are often further handicapped by reporting delays and a lack of capacity to project epidemiological case count time series forward.

Traditional outbreak data have been complemented in recent years by high-resolution data sets from public social media such as Twitter, Tumblr, and time series provided by the Google search application programming interface via Google trends and Google flu, which can be retrieved programmatically and analyzed over time. The results presented in this work suggest

that, when combined with traditional epidemiological data sources, social media-driven data sets, machine learning, and dynamic modeling can offer powerful tools for anticipating future evolution of and assessing intervention tradeoffs with respect to infectious disease outbreaks, particularly for emerging pathogens.

Conflicts of Interest

None declared.

References

1. Ong JBS, Chen MIC, Cook AR, Lee HC, Lee VJ, Lin RTP, et al. Real-time epidemic monitoring and forecasting of H1N1-2009 using influenza-like illness from general practice and family doctor clinics in Singapore. *PLoS One* 2010 Apr 14;5(4):e10036 [FREE Full text] [doi: [10.1371/journal.pone.0010036](https://doi.org/10.1371/journal.pone.0010036)] [Medline: [20418945](https://pubmed.ncbi.nlm.nih.gov/20418945/)]
2. Osgood N, Liu J. Towards closed loop modelingvaluating the prospects for creating recurrently regounded aggregate simulation models using particle filtering. 2014 Presented at: Proc Winter Simul Conf Internet; 2014; Huntington Beach p. E.
3. Safarishahrbijari A, Teyhouee A, Waldner C, Liu J, Osgood N. Predictive accuracy of particle filtering in dynamic models supporting outbreak projections. *BMC Infect Dis* 2017 Dec 26;17(1):648 [FREE Full text] [doi: [10.1186/s12879-017-2726-9](https://doi.org/10.1186/s12879-017-2726-9)] [Medline: [28950831](https://pubmed.ncbi.nlm.nih.gov/28950831/)]
4. Oraji R, Hoepfner V, Safarishahrbijari A, Osgood N. Combining Particle Filtering and Transmission Modeling for TB Control. 2016 Oct 4 Presented at: IEEE Int Conf Healthc Informatics Internet; 2016; Chicago.
5. Safarishahrbijari A, Lawrence T, Lomotey R, Liu J, Waldner C, Osgood N. Particle filtering in a SEIRV simulation model of H1N1 influenza. 2015 Presented at: Winter Simul Conf Internet; 2015; Huntington Beach.
6. Li X, Doroshenko A, Osgood N. Applying particle filtering in both aggregated and age-structured population compartmental models of pre-vaccination measles. *PLoS One* 2018;13(11):e0206529 [FREE Full text] [doi: [10.1371/journal.pone.0206529](https://doi.org/10.1371/journal.pone.0206529)] [Medline: [30388138](https://pubmed.ncbi.nlm.nih.gov/30388138/)]
7. Google Flu Trends - Canada. URL: <https://www.google.org/flutrends/about/data/flu/ca/data.txt> [accessed 2019-05-19] [WebCite Cache ID 78UBci6tp]
8. Google Trends. URL: <https://trends.google.com/trends/?geo=CA> [accessed 2019-05-19] [WebCite Cache ID 78UAFaKlz]
9. Chen B, Shao J, Liu K, Cai G, Jiang Z, Huang Y, et al. Does Eating Chicken Feet With Pickled Peppers Cause Avian Influenza? Observational Case Study on Chinese Social Media During the Avian Influenza A (H7N9) Outbreak. *JMIR Public Health Surveill* 2018 Mar 29;4(1):e32 [FREE Full text] [doi: [10.2196/publichealth.8198](https://doi.org/10.2196/publichealth.8198)] [Medline: [29599109](https://pubmed.ncbi.nlm.nih.gov/29599109/)]
10. Berlinberg EJ, Deiner MS, Porco TC, Acharya NR. Monitoring Interest in Herpes Zoster Vaccination: Analysis of Google Search Data. *JMIR Public Health Surveill* 2018 May 02;4(2):e10180 [FREE Full text] [doi: [10.2196/10180](https://doi.org/10.2196/10180)] [Medline: [29720364](https://pubmed.ncbi.nlm.nih.gov/29720364/)]
11. Xu C, Wang Y, Yang H, Hou J, Sun L, Zhang X, et al. Association Between Cancer Incidence and Mortality in Web-Based Data in China: Infodemiology Study. *J Med Internet Res* 2019 Jan 29;21(1):e10677 [FREE Full text] [doi: [10.2196/10677](https://doi.org/10.2196/10677)] [Medline: [30694203](https://pubmed.ncbi.nlm.nih.gov/30694203/)]
12. Mavragani A, Sampri A, Sypsa K, Tsagarakis KP. Integrating Smart Health in the US Health Care System: Infodemiology Study of Asthma Monitoring in the Google Era. *JMIR Public Health Surveill* 2018 Mar 12;4(1):e24 [FREE Full text] [doi: [10.2196/publichealth.8726](https://doi.org/10.2196/publichealth.8726)] [Medline: [29530839](https://pubmed.ncbi.nlm.nih.gov/29530839/)]
13. Chen S, Xu Q, Buchenberger J, Bagavathi A, Fair G, Shaikh S, et al. Dynamics of Health Agency Response and Public Engagement in Public Health Emergency: A Case Study of CDC Tweeting Patterns During the 2016 Zika Epidemic. *JMIR Public Health Surveill* 2018 Nov 22;4(4):e10827 [FREE Full text] [doi: [10.2196/10827](https://doi.org/10.2196/10827)] [Medline: [30467106](https://pubmed.ncbi.nlm.nih.gov/30467106/)]
14. Farhadloo M, Winneg K, Chan MS, Hall Jamieson K, Albarracin D. Associations of Topics of Discussion on Twitter With Survey Measures of Attitudes, Knowledge, and Behaviors Related to Zika: Probabilistic Study in the United States. *JMIR Public Health Surveill* 2018 Feb 09;4(1):e16 [FREE Full text] [doi: [10.2196/publichealth.8186](https://doi.org/10.2196/publichealth.8186)] [Medline: [29426815](https://pubmed.ncbi.nlm.nih.gov/29426815/)]
15. Chen T, Dredze M. Vaccine Images on Twitter: Analysis of What Images are Shared. *J Med Internet Res* 2018 Apr 03;20(4):e130 [FREE Full text] [doi: [10.2196/jmir.8221](https://doi.org/10.2196/jmir.8221)] [Medline: [29615386](https://pubmed.ncbi.nlm.nih.gov/29615386/)]
16. Lazard A, Scheinfeld E, Bernhardt J, Wilcox G, Suran M. Detecting themes of public concern: a text mining analysis of the Centers for Disease Control and Prevention's Ebola live Twitter chat. *Am J Infect Control* 2015 Oct 01;43(10):1109-1111. [doi: [10.1016/j.ajic.2015.05.025](https://doi.org/10.1016/j.ajic.2015.05.025)] [Medline: [26138998](https://pubmed.ncbi.nlm.nih.gov/26138998/)]
17. Allen C, Tsou M, Aslam A, Nagel A, Gawron J. Applying GIS and Machine Learning Methods to Twitter Data for Multiscale Surveillance of Influenza. *PLoS One* 2016 Jul 1;11(7):e0157734 [FREE Full text] [doi: [10.1371/journal.pone.0157734](https://doi.org/10.1371/journal.pone.0157734)] [Medline: [27455108](https://pubmed.ncbi.nlm.nih.gov/27455108/)]

18. Dugas A, Jalalpour M, Gel Y, Levin S, Torcaso F, Igusa T, et al. Influenza Forecasting with Google Flu Trends. *Online J Public Health Inform* 2013 Mar 01;23. [doi: [10.5210/ojphi.v5i1.4470](https://doi.org/10.5210/ojphi.v5i1.4470)]
19. Pollett S, Boscardin W, Azziz-Baumgartner E, Tinoco Y, Soto G, Romero C, et al. Evaluating Google Flu Trends in Latin America: Important Lessons for the Next Phase of Digital Disease Detection. *Clin Infect Dis* 2017 Jan 01;64(1):34-41 [FREE Full text] [doi: [10.1093/cid/ciw657](https://doi.org/10.1093/cid/ciw657)] [Medline: [27678084](https://pubmed.ncbi.nlm.nih.gov/27678084/)]
20. Araz O, Bentley D, Muelleman R. Using Google Flu Trends data in forecasting influenza-like-illness related ED visits in Omaha, Nebraska. *Am J Emerg Med* 2014 Sep;32(9):1016-1023. [doi: [10.1016/j.ajem.2014.05.052](https://doi.org/10.1016/j.ajem.2014.05.052)] [Medline: [25037278](https://pubmed.ncbi.nlm.nih.gov/25037278/)]
21. Thompson LH, Malik MT, Gumel A, Strome T, Mahmud SM. Emergency department and 'Google flu trends' data as syndromic surveillance indicators for seasonal influenza. *Epidemiol Infect* 2014 Nov;142(11):2397-2405. [doi: [10.1017/S0950268813003464](https://doi.org/10.1017/S0950268813003464)] [Medline: [24480399](https://pubmed.ncbi.nlm.nih.gov/24480399/)]
22. Santillana M, Nguyen A, Dredze M, Paul M, Nsoesie E, Brownstein J. Combining Search, Social Media, and Traditional Data Sources to Improve Influenza Surveillance. *PLoS Comput Biol* 2015 Oct;11(10):e1004513 [FREE Full text] [doi: [10.1371/journal.pcbi.1004513](https://doi.org/10.1371/journal.pcbi.1004513)] [Medline: [26513245](https://pubmed.ncbi.nlm.nih.gov/26513245/)]
23. Sharpe J, Hopkins R, Cook R, Striley C. Evaluating Google, Twitter, and Wikipedia as Tools for Influenza Surveillance Using Bayesian Change Point Analysis: A Comparative Analysis. *JMIR Public Health Surveill* 2016 Oct 20;2(2):e161 [FREE Full text] [doi: [10.2196/publichealth.5901](https://doi.org/10.2196/publichealth.5901)] [Medline: [27765731](https://pubmed.ncbi.nlm.nih.gov/27765731/)]
24. Gelb A. *Applied Optimal Estimation*. Cambridge: MIT Press; 1974.
25. Epstein JM, Parker J, Cummings D, Hammond RA. Coupled contagion dynamics of fear and disease: mathematical and computational explorations. *PLoS One* 2008;3(12):e3955 [FREE Full text] [doi: [10.1371/journal.pone.0003955](https://doi.org/10.1371/journal.pone.0003955)] [Medline: [19079607](https://pubmed.ncbi.nlm.nih.gov/19079607/)]
26. H1N1 Flu in Manitoba. 2009 URL: <https://www.gov.mb.ca/health/documents/h1n1.pdf> [accessed 2019-05-19] [WebCite Cache ID 78UUht7Bv]
27. Brousseau N. Bilan épidémiologique de la pandémie d'influenza A(H1N1). URL: https://www.inspq.qc.ca/sites/default/files/publications/1212_bilana_h1n1_2009.pdf [accessed 2019-05-19] [WebCite Cache ID 78UVE7rGH]
28. Dorigatti I, Cauchemez S, Pugliese A, Ferguson N. A new approach to characterising infectious disease transmission dynamics from sentinel surveillance: application to the Italian 2009-2010 A/H1N1 influenza pandemic. *Epidemics* 2012 Mar;4(1):9-21 [FREE Full text] [doi: [10.1016/j.epidem.2011.11.001](https://doi.org/10.1016/j.epidem.2011.11.001)] [Medline: [22325010](https://pubmed.ncbi.nlm.nih.gov/22325010/)]
29. Andrieu C, Doucet A, Holenstein R. Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2010;1. [doi: [10.1111/j.1467-9868.2009.00736.x](https://doi.org/10.1111/j.1467-9868.2009.00736.x)]

Abbreviations

GFT: Google Flu Trends

PMCMC: Particle Markov Chain Monte Carlo

Edited by G Eysenbach; submitted 18.07.18; peer-reviewed by E André, S Ghosh; comments to author 23.12.18; revised version received 17.02.19; accepted 18.02.19; published 26.05.19.

Please cite as:

Safarishahrbijari A, Osgood ND

Social Media Surveillance for Outbreak Projection via Transmission Models: Longitudinal Observational Study

JMIR Public Health Surveill 2019;5(2):e11615

URL: <http://publichealth.jmir.org/2019/2/e11615/>

doi: [10.2196/11615](https://doi.org/10.2196/11615)

PMID: [31199339](https://pubmed.ncbi.nlm.nih.gov/31199339/)

©Anahita Safarishahrbijari, Nathaniel D Osgood. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 26.05.2019. 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

Improved Real-Time Influenza Surveillance: Using Internet Search Data in Eight Latin American Countries

Leonardo Clemente^{1,2*}, BSc; Fred Lu^{2*}, BA; Mauricio Santillana^{2,3}, MSc, PhD

¹School of Engineering and Sciences, Tecnológico de Monterrey, Monterrey, Mexico

²Computational Health Informatics Program, Boston Children's Hospital, Boston, MA, United States

³Department of Pediatrics, Harvard Medical School, Boston, MA, United States

*these authors contributed equally

Corresponding Author:

Mauricio Santillana, MSc, PhD

Computational Health Informatics Program

Boston Children's Hospital

1 Autumn St

Boston, MA, 02215

United States

Phone: 1 617 919 1795

Email: msantill@g.harvard.edu

Abstract

Background: Novel influenza surveillance systems that leverage Internet-based real-time data sources including Internet search frequencies, social-network information, and crowd-sourced flu surveillance tools have shown improved accuracy over the past few years in data-rich countries like the United States. These systems not only track flu activity accurately, but they also report flu estimates a week or more ahead of the publication of reports produced by healthcare-based systems, such as those implemented and managed by the Centers for Disease Control and Prevention. Previous work has shown that the predictive capabilities of novel flu surveillance systems, like Google Flu Trends (GFT), in developing countries in Latin America have not yet delivered acceptable flu estimates.

Objective: The aim of this study was to show that recent methodological improvements on the use of Internet search engine information to track diseases can lead to improved retrospective flu estimates in multiple countries in Latin America.

Methods: A machine learning-based methodology that uses flu-related Internet search activity and historical information to monitor flu activity, named ARGO (AutoRegression with Google search), was extended to generate flu predictions for 8 Latin American countries (Argentina, Bolivia, Brazil, Chile, Mexico, Paraguay, Peru, and Uruguay) for the time period: January 2012 to December of 2016. These retrospective (out-of-sample) Influenza activity predictions were compared with historically observed flu suspected cases in each country, as reported by Flunet, an influenza surveillance database maintained by the World Health Organization. For a baseline comparison, retrospective (out-of-sample) flu estimates were produced for the same time period using autoregressive models that only leverage historical flu activity information.

Results: Our results show that ARGO-like models' predictive power outperform autoregressive models in 6 out of 8 countries in the 2012-2016 time period. Moreover, ARGO significantly improves on historical flu estimates produced by the now discontinued GFT for the time period of 2012-2015, where GFT information is publicly available.

Conclusions: We demonstrate here that a self-correcting machine learning method, leveraging Internet-based disease-related search activity and historical flu trends, has the potential to produce reliable and timely flu estimates in multiple Latin American countries. This methodology may prove helpful to local public health officials who design and implement interventions aimed at mitigating the effects of influenza outbreaks. Our methodology generally outperforms both the now-discontinued tool GFT, and autoregressive methodologies that exploit only historical flu activity to produce future disease estimates.

(*JMIR Public Health Surveill* 2019;5(2):e12214) doi:[10.2196/12214](https://doi.org/10.2196/12214)

KEYWORDS

google flu trends; influenza monitoring; real-time disease surveillance; digital epidemiology; influenza, human; developing countries; machine learning

Introduction

Background

With the highest mortality of any respiratory infectious disease in the young and elderly in Latin America, influenza poses significant health and economic challenges to low- and middle-income countries in the region [1]. The World Health Organization (WHO) maintains a health care-based disease surveillance system that collects information on flu activity from local ministries of health around the world. Unfortunately, these reports have a common delay of at least a week in Latin America, limiting the ability for a timely response to unexpected epidemic outbreaks. Reliable surveillance systems that monitor flu activity in real time in this region would help public health institutions deploy timely vaccination campaigns and optimally allocate resources during epidemic outbreaks. Multiple research teams have proposed complementary methods to estimate and forecast flu activity in real time in data-rich countries such as the United States, using techniques ranging from statistical [2,3] to mechanistic [4,5] and incorporating a variety of data sources, such as internet search information, flu-related Twitter microblogs [6,7], crowdsourced flu surveillance [8,9], clinician search activity [10], electronic health records [11], and Wikipedia access [12,13], as summarized in a study by Santillana [14]. However, a reliable system that leverages internet search activity to monitor flu activity in multiple developing nations is not yet available.

An early large-scale implementation of real-time disease surveillance started in 2008 with Google Flu Trends (GFT), a Web-based tool that used Google search activity to produce flu activity estimates in multiple locations around the world [15]. Although GFT was initially perceived as a technological innovation, its large prediction errors during the 2009 H1N1 flu pandemic and the 2013 flu season in the United States raised methodological concerns from multiple researchers [16-18]. A recent study by Pollet et al showed that GFT's flu estimates in Latin America had yielded poor results [19].

Objectives

The discontinuation of GFT in 2015 led many to believe that internet search trends were too noisy to track disease activity, a problem exacerbated in developing countries with limited internet access. However, recent research has shown that robust and dynamically self-correcting machine learning methodologies can extract meaningful signals from real-time search activity to track Zika and Dengue activity in low- to middle-income countries around the world [11,20,21]. We apply lessons learned from these studies and successfully extend a state-of-the-art modeling approach for flu surveillance to 8 Latin American countries.

Methods

Data Acquisition

We built a predictive methodology aimed at estimating suspected flu activity, as reported by FluNet, an online surveillance tool maintained by the WHO. FluNet collects and aggregates multiple indicators of flu activity at the country level.

For this study, we selected the number of processed specimens (NPSs) as the ground truth. As these specimens were taken from patients with flu-like symptoms and then sent to a laboratory for testing, we interpreted them as an indicator of suspected flu activity in the population. Weekly aggregated NPS reports were collected from January 5, 2009, to December 25, 2016, for Argentina, Bolivia, Brazil, Chile, Mexico, Paraguay, Peru, and Uruguay.

Given their near real-time availability via the online tool, Google Trends, we selected influenza-related internet search activity to be used in our models as proxies or predictors for flu activity. Where available and based on country-specific historical flu indicators (during the training time period of our models), we used the online tool, Google Correlate, to identify flu-related search term trends, leading to a total of 285 Spanish terms and 96 Portuguese terms. See [Multimedia Appendix 1](#) for further description of this process.

Models and Benchmarks

We extended AutoRegression with Google search information (ARGO), a methodology originally conceived and tested to track flu activity in the United States in multiple spatial scales as a way to produce retrospective and strictly out-of-sample flu estimates individually for each country [20,22]. This methodology is based on a multivariable regularized linear model that is dynamically recalibrated every week as new flu activity information becomes available. Besides online search information, ARGO incorporates short-term and seasonal historical flu information to improve the accuracy of predictions and mitigate the undesired effect of spikes in search activity (induced perhaps by overreaction in the population during potential health threats reported by the news). More details on this approach can be found in a study by Yang et al [22].

Given a weekly as-yet-unseen NPS report to estimate, we used historical NPS and Google Trends information from the previous most recent 2 years (104 weeks) of data to calibrate ARGO and predict the given week's NPS report. To assess ARGO's predictive power, we built autoregressive models separately for each country (named AR52 throughout this paper) that only use historical flu activity from the 52 weeks before predictions and generated retrospective out-of-sample estimates over the same time period. All models were built using the glmnet package on MATLAB (MathWorks) version 2014a [3,23].

Metrics

To compare the predictive ability of ARGO and AR52, we calculated Pearson correlations and the root mean square error between model predictions and the subsequently observed suspected flu cases. The added value of using Google search activity as a predictor was tested via an efficiency metric [22] that quantifies the improvement of ARGO over a simple autoregressive model. This efficiency metric is calculated as the ratio between the mean square errors of AR52 and ARGO. For the efficiency metric, 90% CIs were generated using the stationary block bootstrap method [24]. We report 2 additional metrics aimed at evaluating our method's ability to correctly identify the timing of peaks and timing of the onset of epidemic outbreaks. These metrics are referred to as ΔP and ΔO ,

respectively, and measure the distance, in weeks, between the observed peak (or onsets) and the predicted one. See [Multimedia Appendix 1](#) for further explanation.

Results

AutoRegression With Google Search Models Outperform AR52 and Google Flu Trends

Retrospective out-of-sample estimates of flu activity were produced, for each of the 8 countries, from January 1, 2012, to December 25, 2016, and compared with the FluNet reported suspected cases (NPSs). Brazil's NPS data were only available until October 9, 2016. Note that because of FluNet's reporting delays, our models, which rely on past available values of FluNet and current internet search activity, estimate current flu activity at least 1 week ahead of official reports.

In [Figure 1](#), we show our real-time flu estimates and the subsequently observed suspected flu cases for each country. Contextually, historical GFT values (scaled to be displayed alongside with NPS values) and autoregressive estimates are also shown. Our models (ARGO and AR52) accurately predict NPS values in each country. GFT shows consistently large discrepancies when compared with the observed values, consistent with the findings reported by Pollett et al [19].

ARGO displays improvement in 6 countries in terms of the efficiency metric ([Figure 2](#)), reaching significant error reductions compared with AR52 in Brazil (155 to 104 or 33%), Mexico (243 to 184 or 24%), Peru (48 to 40 or 16%), and Chile (131 to 119 or 9%). ARGO consistently outperforms GFT on Pearson correlations during the time period when GFT was active in every country and improves upon AR52 in all countries except Bolivia and Uruguay, over the whole study period, reaching significant correlation increases in Brazil (from 0.891 to 0.957), Mexico (from 0.86 to 0.92), and Peru (from 0.84 to 0.89).

Figure 1. Graphical representation of the number of processed specimens (NPSs) as reported by WHO's FluNet (black), along with the NPS estimates generated by ARGO (red), AR (light blue), and Google Flu Trends (GFT; blue), over the whole study period of January 1, 2012 to December 25, 2016.

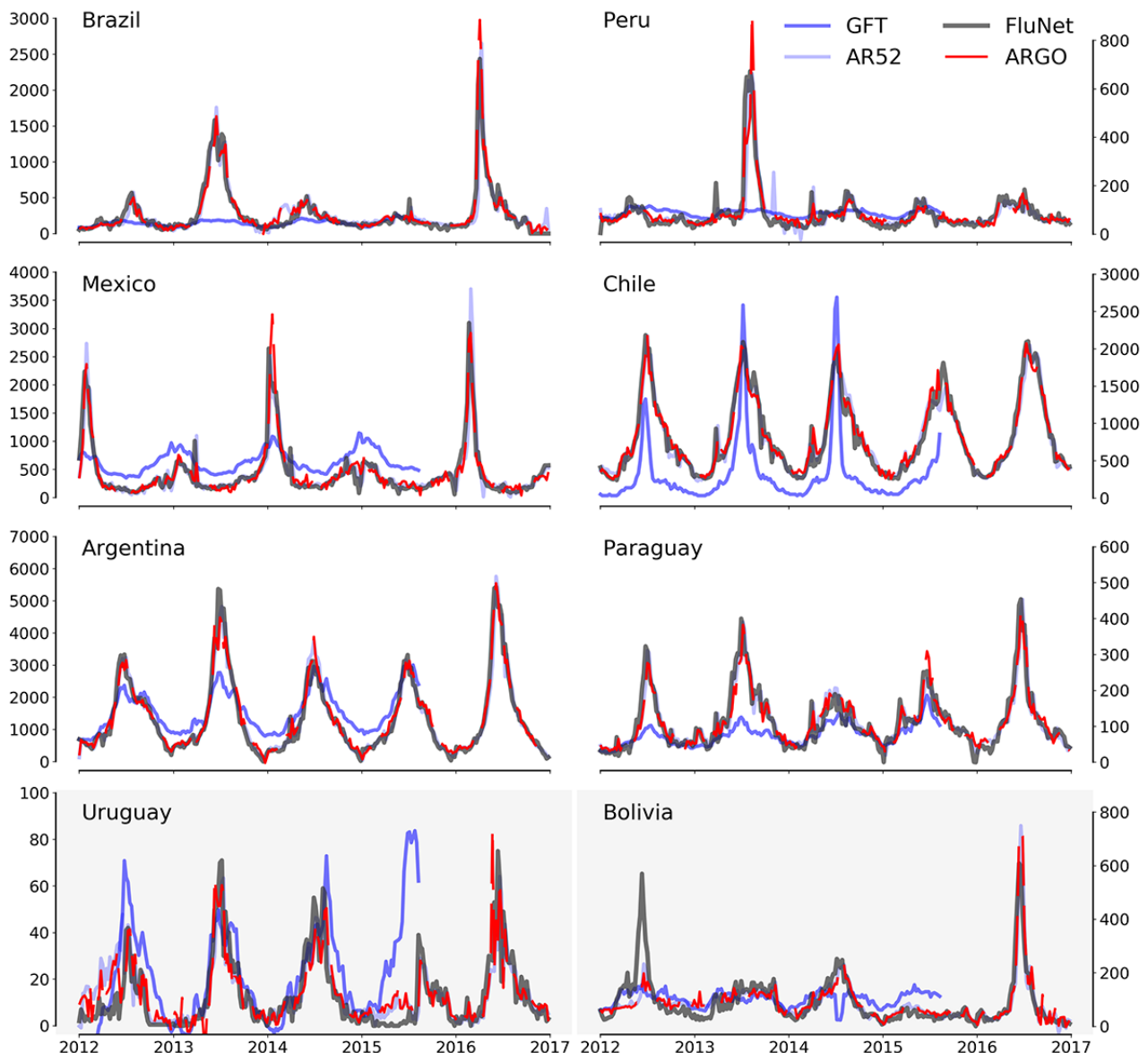
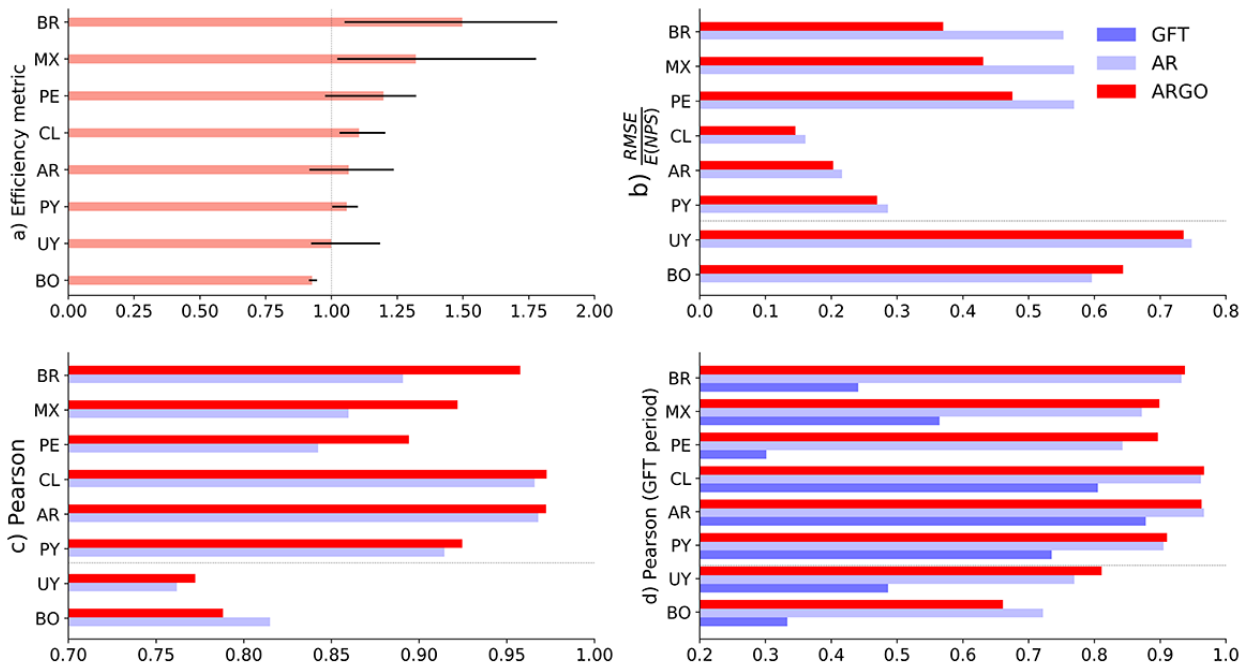


Figure 2. Set of bar graphs that show our performance metrics to assess the predictive power of AutoRegression with Google search (ARGO). (a) Efficiency metric values (salmon color) for each individual country with their respective 90% CI (solid black line). (b) Root mean square error (RMSE) values for ARGO (red) and AR52 (light blue) during the whole study period. Each country's RMSE value is normalized by their respective average number of processed specimen (NPS) over the whole study period to avoid scale differences in visualization. (c) Pearson correlation scores of ARGO and AR during the full period. (d) Pearson correlation values for ARGO (red), AR (gray), and Google Flu Trends (GFT; blue) during the period in our study where GFT was active.



Peaks and Onsets

Multimedia Appendix 1 shows our method's ability to identify the timing of peaks and timing of the onset as captured by the metrics ΔP and ΔO , respectively. Out of a total of 39 peaks, ARGO predicts 10 on the observed week, 5 ahead of time, and 24 with a lag. In contrast, AR52 predicted 7 peaks exactly, 2 ahead of time, and 30 with a lag. Out of 39 measured onsets, ARGO predicted 11 at the exact date, 12 ahead of time, and 13 with a lag, with the rest of the onsets incorrectly estimated with more than a month's distance. AR52 predicted 8 onsets at the exact time, 8 ahead of time, 20 with a lag, and the rest were estimated with more than a month's distance. As expected, the baseline autoregressive model, AR52, predicted more outbreak peaks late, with a lag of at least a week. ARGO, in contrast, is slightly more responsive and predicts outbreak peaks on time or weeks before the observed peak. Furthermore, even when ARGO's peak timing is not accurate, the magnitude of the peak is captured better than its AR52 counterpart. For more detailed information, see Multimedia Appendix 1.

Discussion

Combining Historical Flu Activity and Google Search Data

ARGO's prediction performance shows that internet search volumes and historical flu activity, when combined with dynamic machine learning techniques, can effectively detect real-time suspected flu cases in several Latin American countries. Our results considerably outperform the historical predictive performance of GFT highlighting (1) the importance

of moving away from one-size-fits-all approaches such as those used by GFT and (b) the value of combining local flu epidemiological information with influenza-related internet search trends. The overall improvement of ARGO over the baseline autoregressive model indicates that internet search engine data, even in middle-income countries, provide increased responsiveness to changing disease trends. This improvement is clear in Brazil, Chile, Mexico, Peru, Paraguay, and Argentina, whereas in Uruguay and Bolivia, the inclusion of Google search data does not seem to improve the baseline model.

The availability of an online tool to select relevant flu-related terms (Google Correlate) that track historical flu activity was found to be a critical element for ARGO to improve performance over the autoregressive benchmark (Argentina, Chile, Mexico, Peru, and Brazil), suggesting that the most meaningful flu-related search queries are country-specific. In countries, such as Uruguay, where many weekly data points were missing on FluNet, ARGO's predictive ability was reduced. Our best performance was seen in Brazil, Mexico, and Peru, where flu data were collected consistently every week during this study's time period (see Multimedia Appendix 1).

Number of Processed Specimens as Our Gold Standard

On the basis of our previous research findings monitoring Dengue and Zika activity in Latin America [20,21], we chose the number of suspected influenza cases (as captured by the FluNet's NPSs) as our gold standard for our prediction tasks. Our choice was based on the intuitive fact that flu-related Google search activity is higher when more people "suspect" they may be affected by flu-like symptoms, regardless of the outcome of any lab test. As such, our models may prove useful

to improve the timely allocation of resources in health care facilities in situations when increased numbers of people, with flu-like symptoms and respiratory needs, may need to be seen. It is relevant to point out that using NPS case counts as a gold standard implies that our models are not directly estimating confirmed influenza case counts but suspected Influenza-like Illness activity trends. Our choice of gold standard is meaningful as it may help health care providers prepare for traffic fluctuations of patients presenting with symptoms of influenza. However, from an epidemiological perspective, more standard test-positive influenza proportions reported on previous Latin America studies [19,25] should also be considered in future studies.

Limitations and Future Work

Combining internet search volumes and historical flu activity via ARGO shows strong potential for the development of timely

flu surveillance in low- or middle-income countries. However, flu estimates at the national level may not be reflective of outbreak conditions at the local level, especially in countries with significant geographical heterogeneity. At present, FluNet only provides national-level flu data. In the future, as more fine-grained epidemiological information becomes available in developing countries, studies should evaluate the feasibility of deploying disease surveillance platforms at finer spatial scales. Successful extensions of our methodologies at the city and state levels in data-rich environments such as the United States [26,27] indicate that ARGO-like methodologies can accurately monitor influenza at these spatial resolutions. Moreover, there is strong evidence that internet access is rapidly increasing in many Latin American countries [28], leading us to hypothesize that the performance of methodologies using internet-based data will increase over time following an increase of the quality of Google search data.

Acknowledgments

MS and FL were partially funded by the Centers for Disease Control and Prevention's Cooperative 375 Agreement PPHF 11797-998G-15. LC was partially funded by a grant 734557 by the Mexico National Council of Science and Technology. MS was partially supported by the National Institute of General Medical Sciences, of the National Institutes of Health, under Award Number R01GM130668. 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.

Multimedia Appendix 1

Supplementary materials. This document includes several Tables and Figures, as well as more information regarding our Methods and Results section.

[[PDF File \(Adobe PDF File\), 5MB - publichealth_v5i2e12214_app1.pdf](#)]

References

1. Savy V, Ciapponi A, Bardach A, Glujovsky D, Aruj P, Mazzoni A, et al. Burden of influenza in Latin America and the Caribbean: a systematic review and meta-analysis. *Influenza Other Respir Viruses* 2013 Nov;7(6):1017-1032 [[FREE Full text](#)] [doi: [10.1111/irv.12036](#)] [Medline: [23210504](#)]
2. Brooks LC, Farrow DC, Hyun S, Tibshirani RJ, Rosenfeld R. Flexible modeling of epidemics with an empirical Bayes Framework. *PLoS Comput Biol* 2015 Aug;11(8):e1004382 [[FREE Full text](#)] [doi: [10.1371/journal.pcbi.1004382](#)] [Medline: [26317693](#)]
3. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. *J R Stat Soc Series B Stat Methodol* 2011;73(3):273-282. [doi: [10.1111/j.1467-9868.2011.00771.x](#)]
4. Shaman J, Karspeck A. Forecasting seasonal outbreaks of influenza. *Proc Natl Acad Sci U S A* 2012 Dec 11;109(50):20425-20430 [[FREE Full text](#)] [doi: [10.1073/pnas.1208772109](#)] [Medline: [23184969](#)]
5. Zhang Q, Perra N, Perrota D, Tizzoni M, Paolotti D, Vespignani A. Forecasting seasonal influenza fusing digital indicators and a mechanistic disease model. In: Proceedings of the 26th International Conference on World Wide Web. 2017 Presented at: WWW'17; April 3-7, 2017; Perth, Australia p. 311. [doi: [10.1145/3038912.3052678](#)]
6. 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](#)]
7. Paul MJ, Dredze M. You are what you Tweet: Analyzing Twitter for public health. In: International AAAI Conference on Weblogs and Social Media. 2011 Presented at: Fifth International ICWSM; July 17-21, 2011; Barcelona, Spain URL: http://www.cs.jhu.edu/~mpaul/files/2011.icwsm.twitter_health.pdf
8. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wójcik O, et al. Flu near you: crowdsourced symptom reporting spanning 2 influenza seasons. *Am J Public Health* 2015 Oct;105(10):2124-2130. [doi: [10.2105/AJPH.2015.302696](#)] [Medline: [26270299](#)]

9. Paolotti D, Carnahan A, Colizza V, Eames K, Edmunds J, Gomes G, et al. Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience. *Clin Microbiol Infect* 2014 Jan;20(1):17-21 [FREE Full text] [doi: [10.1111/1469-0691.12477](https://doi.org/10.1111/1469-0691.12477)] [Medline: [24350723](https://pubmed.ncbi.nlm.nih.gov/24350723/)]
10. Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. Using clinicians' search query data to monitor influenza epidemics. *Clin Infect Dis* 2014 Nov 15;59(10):1446-1450. [doi: [10.1093/cid/ciu647](https://doi.org/10.1093/cid/ciu647)] [Medline: [25115873](https://pubmed.ncbi.nlm.nih.gov/25115873/)]
11. Yang S, Santillana M, Brownstein JS, Gray J, Richardson S, Kou SC. Using electronic health records and Internet search information for accurate influenza forecasting. *BMC Infect Dis* 2017 May 8;17(1):332 [FREE Full text] [doi: [10.1186/s12879-017-2424-7](https://doi.org/10.1186/s12879-017-2424-7)] [Medline: [28482810](https://pubmed.ncbi.nlm.nih.gov/28482810/)]
12. Generous N, Fairchild G, Deshpande A, Del Valle SY, Priedhorsky R. Global disease monitoring and forecasting with Wikipedia. *PLoS Comput Biol* 2014 Nov;10(11):e1003892 [FREE Full text] [doi: [10.1371/journal.pcbi.1003892](https://doi.org/10.1371/journal.pcbi.1003892)] [Medline: [25392913](https://pubmed.ncbi.nlm.nih.gov/25392913/)]
13. Hickmann KS, Fairchild G, Priedhorsky R, Generous N, Hyman JM, Deshpande A, et al. Forecasting the 2013-2014 influenza season using Wikipedia. *PLoS Comput Biol* 2015 May;11(5):e1004239 [FREE Full text] [doi: [10.1371/journal.pcbi.1004239](https://doi.org/10.1371/journal.pcbi.1004239)] [Medline: [25974758](https://pubmed.ncbi.nlm.nih.gov/25974758/)]
14. Santillana M. Editorial commentary: perspectives on the future of internet search engines and biosurveillance systems. *Clin Infect Dis* 2017 Dec 1;64(1):42-43. [doi: [10.1093/cid/ciw660](https://doi.org/10.1093/cid/ciw660)] [Medline: [27678086](https://pubmed.ncbi.nlm.nih.gov/27678086/)]
15. 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](https://doi.org/10.1038/nature07634)] [Medline: [19020500](https://pubmed.ncbi.nlm.nih.gov/19020500/)]
16. 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](https://doi.org/10.1016/j.amepre.2014.05.020)] [Medline: [24997572](https://pubmed.ncbi.nlm.nih.gov/24997572/)]
17. 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](https://doi.org/10.1126/science.1248506)] [Medline: [24626916](https://pubmed.ncbi.nlm.nih.gov/24626916/)]
18. 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](https://doi.org/10.1371/journal.pcbi.1003256)] [Medline: [24146603](https://pubmed.ncbi.nlm.nih.gov/24146603/)]
19. Pollett S, Boscardin WJ, Azziz-Baumgartner E, Tinoco YO, Soto G, Romero C, et al. Evaluating Google Flu Trends in Latin America: important lessons for the next phase of digital disease detection. *Clin Infect Dis* 2017 Jan 1;64(1):34-41. [doi: [10.1093/cid/ciw657](https://doi.org/10.1093/cid/ciw657)] [Medline: [27678084](https://pubmed.ncbi.nlm.nih.gov/27678084/)]
20. Yang S, Kou SC, Lu F, Brownstein JS, Brooke N, Santillana M. Advances in using Internet searches to track dengue. *PLoS Comput Biol* 2017 Jul;13(7):e1005607 [FREE Full text] [doi: [10.1371/journal.pcbi.1005607](https://doi.org/10.1371/journal.pcbi.1005607)] [Medline: [28727821](https://pubmed.ncbi.nlm.nih.gov/28727821/)]
21. McGough SF, Brownstein JS, Hawkins JB, Santillana M. Forecasting Zika incidence in the 2016 Latin America outbreak combining traditional disease surveillance with search, social media, and news report data. *PLoS Negl Trop Dis* 2017 Dec;11(1):e0005295 [FREE Full text] [doi: [10.1371/journal.pntd.0005295](https://doi.org/10.1371/journal.pntd.0005295)] [Medline: [28085877](https://pubmed.ncbi.nlm.nih.gov/28085877/)]
22. Yang S, Santillana M, Kou SC. Accurate estimation of influenza epidemics using Google search data via ARGO. *Proc Natl Acad Sci U S A* 2015 Nov 24;112(47):14473-14478 [FREE Full text] [doi: [10.1073/pnas.1515373112](https://doi.org/10.1073/pnas.1515373112)] [Medline: [26553980](https://pubmed.ncbi.nlm.nih.gov/26553980/)]
23. Qian J, Hastie T, Tibshirani R, Simon N. WebCite. 2013. Glnnet in Matlab URL: <http://www.webcitation.org/76YjrZOGa> [accessed 2019-02-28] [WebCite Cache ID 76Yk0dh0J]
24. Politis DN, Romano JP. The stationary bootstrap. *J Am Stat Assoc* 2012 Feb 27;99(428):1303-1313. [doi: [10.1080/01621459.1994.10476870](https://doi.org/10.1080/01621459.1994.10476870)]
25. Durand LO, Cheng P, Palekar R, Clara W, Jara J, Cerpa M, et al. Timing of influenza epidemics and vaccines in the American tropics, 2002-2008, 2011-2014. *Influenza Other Respir Viruses* 2016 May;10(3):170-175 [FREE Full text] [doi: [10.1111/irv.12371](https://doi.org/10.1111/irv.12371)] [Medline: [26701079](https://pubmed.ncbi.nlm.nih.gov/26701079/)]
26. Lu FS, Hou S, Baltrusaitis K, Shah M, Leskovec J, Sosic 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 9;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/)]
27. Lu FS, Hattab MW, Clemente CL, Biggerstaff M, Santillana M. Improved state-level influenza nowcasting in the United States leveraging internet-based data and network approaches. *Nat Commun* 2019 Dec 11;10(1):147 [FREE Full text] [doi: [10.1038/s41467-018-08082-0](https://doi.org/10.1038/s41467-018-08082-0)] [Medline: [30635558](https://pubmed.ncbi.nlm.nih.gov/30635558/)]
28. International Telecommunication Union. Statistics URL: <https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx> [accessed 2019-02-08] [WebCite Cache ID 761ThSiLh]

Abbreviations

- ARGO:** AutoRegression with Google search
- GFT:** Google Flu Trends
- NPS:** number of processed specimens
- WHO:** World Health Organization

Edited by T Sanchez; submitted 14.09.18; peer-reviewed by B McLeod, S Pollett; comments to author 18.01.19; revised version received 11.02.19; accepted 15.02.19; published 04.04.19.

Please cite as:

Clemente L, Lu F, Santillana M

Improved Real-Time Influenza Surveillance: Using Internet Search Data in Eight Latin American Countries

JMIR Public Health Surveill 2019;5(2):e12214

URL: <https://publichealth.jmir.org/2019/2/e12214/>

doi: [10.2196/12214](https://doi.org/10.2196/12214)

PMID: [30946017](https://pubmed.ncbi.nlm.nih.gov/30946017/)

©Leonardo Clemente, Fred Lu, Mauricio Santillana. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 04.04.2019. 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.

Tutorial

Google Trends in Infodemiology and Infoveillance: Methodology Framework

Amaryllis Mavragani¹, BSc, MSc; Gabriela Ochoa¹, BSc, MSc, PhD

Department of Computing Science and Mathematics, Faculty of Natural Sciences, University of Stirling, Stirling, United Kingdom

Corresponding Author:

Amaryllis Mavragani, BSc, MSc

Department of Computing Science and Mathematics

Faculty of Natural Sciences

University of Stirling

University Campus

Stirling, FK94LA

United Kingdom

Phone: 44 7523782711

Email: amaryllis.mavragani1@stir.ac.uk

Abstract

Internet data are being increasingly integrated into health informatics research and are becoming a useful tool for exploring human behavior. The most popular tool for examining online behavior is Google Trends, an open tool that provides information on trends and the variations of online interest in selected keywords and topics over time. Online search traffic data from Google have been shown to be useful in analyzing human behavior toward health topics and in predicting disease occurrence and outbreaks. Despite the large number of Google Trends studies during the last decade, the literature on the subject lacks a specific methodology framework. This article aims at providing an overview of the tool and data and at presenting the first methodology framework in using Google Trends in infodemiology and infoveillance, including the main factors that need to be taken into account for a strong methodology base. We provide a step-by-step guide for the methodology that needs to be followed when using Google Trends and the essential aspects required for valid results in this line of research. At first, an overview of the tool and the data are presented, followed by an analysis of the key methodological points for ensuring the validity of the results, which include selecting the appropriate keyword(s), region(s), period, and category. Overall, this article presents and analyzes the key points that need to be considered to achieve a strong methodological basis for using Google Trends data, which is crucial for ensuring the value and validity of the results, as the analysis of online queries is extensively integrated in health research in the big data era.

(*JMIR Public Health Surveill* 2019;5(2):e13439) doi:[10.2196/13439](https://doi.org/10.2196/13439)

KEYWORDS

big data; health; infodemiology; infoveillance; internet behavior; Google Trends

Introduction

The use of internet data has become an integral part of health informatics over the past decade, with online sources becoming increasingly available and providing data that can be useful in analyzing and predicting human behavior. This use of the internet has formed two new concepts: “Infodemiology,” first defined by Eysenbach as “the science of distribution and determinants of information in an electronic medium, specifically the Internet, or in a population, with the ultimate aim to inform public health and public policy” [1], and “Infoveillance,” defined as “the longitudinal tracking of infodemiology metrics for surveillance and trend analysis” [2].

The main limitation of validating this line of research is the general lack of openness and availability of official health data. Data collection and analysis of official health data on disease occurrence and prevalence involve several health officials and can even take years until the relevant data are available. This means that data cannot be accessed in real time, which is crucial in health assessment. In several countries, official health data are not publicly available, and even in countries where data are available, they usually consist of large time-interval data (eg, annual data), which makes the analysis and forecasting of diseases and outbreaks more difficult.

Nevertheless, data from several online sources are being widely used to monitor disease outbreaks and occurrence, mainly from Google [3-7] and social media [8-12]. Twitter has become increasingly popular over the past few years [13-19], while

several other studies have combined data from different online sources such as Facebook and Twitter [20] or Google, Twitter, and electronic health records [21].

Currently, the most popular tool in addressing health issues and topics with the use of internet data is Google Trends [22], an open online tool that provides both real-time and archived information on Google queries from 2004 on. The main advantage of Google Trends is that it uses the revealed and not stated users' preferences [23]; therefore, we can obtain information that would be otherwise difficult or impossible to collect. In addition, as data are available in real time, it solves issues that arise with traditional, time-consuming survey methods. Another advantage is that, as Web searches are performed anonymously, it enables the analysis and forecasting of sensitive diseases and topics, such as AIDS [24], mental illnesses and suicide [25-27], and illegal drugs [28,29].

Despite the limitations of data from traditional sources and owing to the fact that online data have shown to be valuable in predictions, the combination of traditional data and Web-based data should be explored, as the results could provide valid and interesting results. Over the past few years, the diversity of online sources used in addressing infodemiology topics is increasing. Indicative recent publications of online sources and combinations of sources are presented in Table 1.

As discussed above, many studies have used Google Trends data to analyze online behavior toward health topics and to forecast prevalence of diseases. However, the literature lacks a methodology framework that provides a concise overview and detailed guidance for future researchers. We believe such a framework is imperative, as the analysis of online data is based on empirical relationships, and thus, a solid methodological basis of any Google Trends study is crucial for ensuring the value and validity of the results.

Table 1. Recent indicative infodemiology studies.

Author(s)	Keywords	Google Trends	Twitter	Facebook	Other social media (eg, YouTube)	Blogs, forums, news outlets, Wikipedia	Databases, electronic health records	Other search engines (Baidu)
Abdellaoui et al [30]	Drug treatment					✓		
Allen et al [31]	Tobacco waterpipe		✓					
Berlinger et al [32]	Herpes, Vaccination	✓						
Bragazzi and Mahroum [33]	Plague, Madagascar	✓						
Chen et al [18]	Zika epidemic		✓					
Forounghi et al [34]	Cancer	✓						
Gianfredi et al [35]	Pertussis	✓						
Hswen et al [36]	Psychological analysis, Autism		✓					
Jones et al [37]	Cancer					✓		
Kandula et al [38]	Influenza	✓						
Keller et al [39]	Bowel disease, Pregnancy, Medication			✓	✓	✓		
Mavragani et al [7]	Asthma	✓						
Mejova et al [40]	Health monitoring			✓				
Odlum et al [41]	HIV/AIDS		✓					
Phillips et al [42]	Cancer	✓						
Poirier et al [43]	Influenza, Hospitals	✓					✓	
Radin et al [44]	Systematic Lupus Erythematosus	✓						
Rocchetti et al [20]	Crohn's disease		✓	✓				
Tana et al [25]	Depression, Finland	✓						
Vsconcellos-Silva et al [45]	Cancer	✓						
Wakamiya et al [46]	Influenza		✓					
Wang et al [47]	Obesity	✓						
Watah et al [48]	West Nile Virus	✓			✓	✓		
Xu et al [49]	Cancer, China							✓

We proceed in a step-by-step manner to develop the methodology framework that should be followed when using Google Trends in infodemiology. First, we provide an overview of how the data are retrieved and adjusted along with the available features, followed by the methodology framework for choosing the appropriate keyword(s), region(s), period, and category. Finally, the results are discussed, along with the limitations of the tool and suggestions for future research.

Methodology Framework

Data Overview

Google Trends is an open online tool that provides information on what was and is trending, based on actual users' Google queries. It offers a variety of choices, such as Trending Searches, Year in Search, and Explore. [Table 2](#) describes the features offered by Google Trends and their respective descriptions.

When using Google Trends for research, data are retrieved from the "Explore" feature, which allows download of real-time data from the last week and archived data for specific keywords and topics from January 2004 up to 36 hours before the search is conducted. The data are retrieved directly from the Google Trends Explore page in .csv format after the examined keyword(s) is entered and the region, period, and category are selected. By default, the period is set to "Worldwide," the time frame is set to "past 12 months," and the category is set to "All categories."

The data are normalized over the selected time frame, and the adjustment is reported by Google as follows:

Search results are proportionate to the time and location of a query by the following process: Each data point is divided by the total searches of the geography and time range it represents to compare relative popularity. Otherwise, places with the most search volume would always be ranked highest. The resulting numbers are then scaled on a range of 0 to 100 based on a topic's proportion to all searches on all topics. Different regions that show the same search interest for a term don't always have the same total search volumes [50]

The normalization of data indicates that the values vary from 0 to 100. The value 0 does not necessarily indicate no searches, but rather indicates very low search volumes that are not included in the results. The adjustment process also excludes

queries that are made over a short time frame from the same internet protocol address and queries that contain special characters. Google does not have a filter for controversial topics, but it excludes related search terms that are sexual. However, it allows retrieval of queries' normalized hits for any keyword entered, independent of filters.

Google Trends allows one to explore the online interest in one term or the comparison of the online interest for up to five terms. It allows a variety of combinations to compare different terms and regions as follows:

- For one term in one region over a specific period, such as for "Asthma" in the United States from January 2004 to December 2014 ([Figure 1a](#))
- For the same term in different regions over the same period, such as for "Tuberculosis" in the United States and United Kingdom from March 24, 2007, to April 7, 2011 ([Figure 1b](#))
- For different terms (up to five) in the same region for the same period, such as for the terms "Chlamydia," "Tuberculosis," and "Syphilis" in Australia from October 5, 2012, to December 18, 2012 ([Figure 1c](#))
- For different terms (up to five) for different regions over the same period, such as comparing the term "Asthma" in the United States, "AIDS" in the United Kingdom, and "Measles" in Canada from June 1, 2017, to July 15, 2018 ([Figure 1d](#))

When the term(s), region(s), period(s), and category are defined, the outputs are a graph of the variations of all examined terms in the online interest over the selected time frame ([Figure 1](#)) and their respective heat maps, which are presented separately for all examined regions ([Figure 2](#)); all datasets can be downloaded in .csv format.

Apart from the graph, the .csv with the relative search volumes, and the interest heat maps, Google Trends also shows and allows one to download .csv files of (1) the "Top related queries", defined as "Top searches are terms that are most frequently searched with the term you entered in the same search session, within the chosen category, country, or region" ([Figure 3a](#)); (2) the "Rising related queries", defined as "terms that were searched for with the keyword you entered...which had the most significant growth in volume in the requested time period" ([Figure 3b](#)); (3) the "Top Related Topics" ([Figure 3c](#)); and (4) the "Rising Related Topics" ([Figure 3d](#)).

Table 2. Google Trends Features and Descriptions.

Feature	Description
Homepage	Provides an overview of what is searched for in a selected region (default: United States)
Explore	Allows exploration of the online interest for specific keywords over selected periods and regions (default: worldwide, 12 months)
Trending Searches	Shows the trending queries for (1) daily search trends and (2) real-time search trends in a selected region (default: United States)
Year in Searches	Show what was trending in a specific region in a specific year (default: United States, previous year)
Subscriptions	Allows subscription for (1) a specific topic in a specific region and sends updates for noteworthy events (via email either once a week or once a month) and (2) trending searches and sends updates about trending searches (via email either as it happens, or once a day, or once a week and includes either “Top Daily Searches,” “Majority of Daily Search Trends,” or “All Daily Search Trends”)

Figure 1. Graphs of the variations in the online interest for the examined terms over the selected time frame in Google Trends.

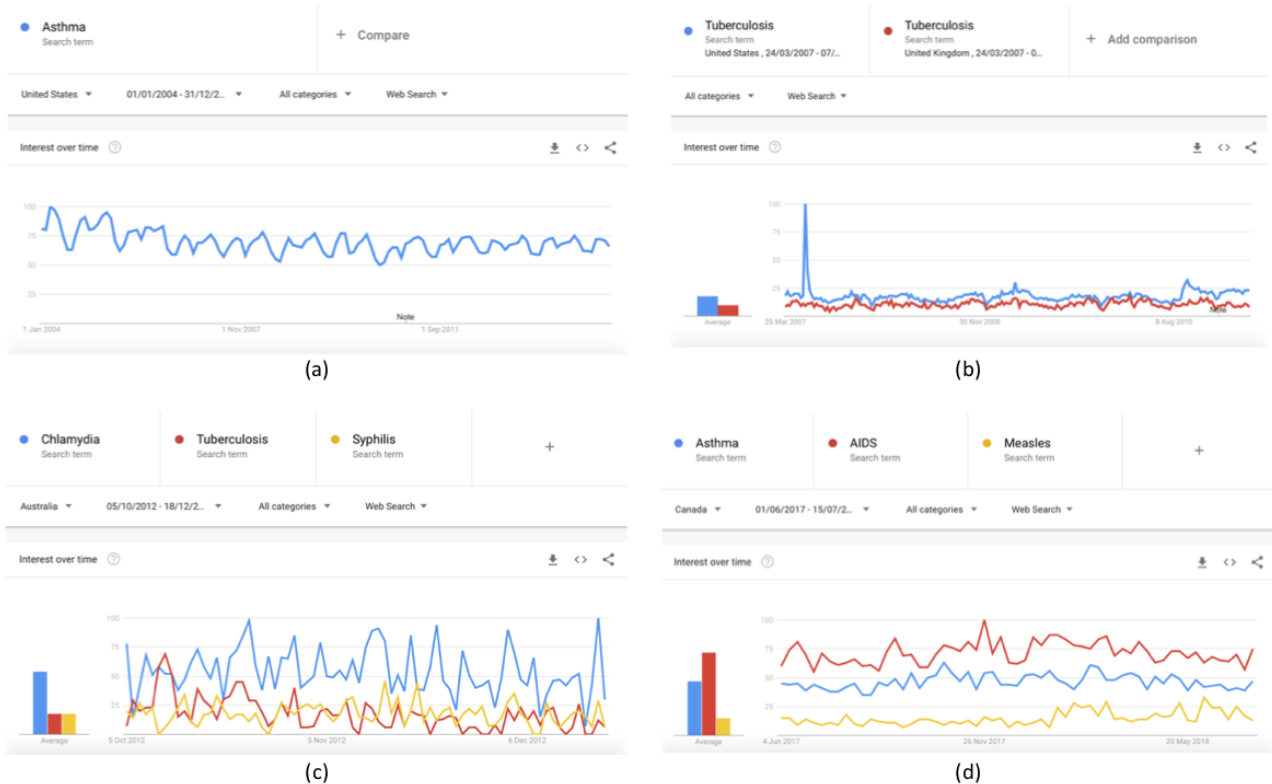


Figure 2. Heat map for (a) “Asthma” in the United States from Jan 2004 to Dec 2014; (b) “Tuberculosis” in the United States and United Kingdom from March 24, 2007, to April 7, 2011; (c) “Chlamydia,” “Tuberculosis,” and “Syphilis” in Australia from Oct 5, 2012, to Dec 18, 2012; (d) “Asthma” in the United States, “AIDS” in the United Kingdom, and “Measles” in Canada from June 1, 2017, to July 15, 2018.

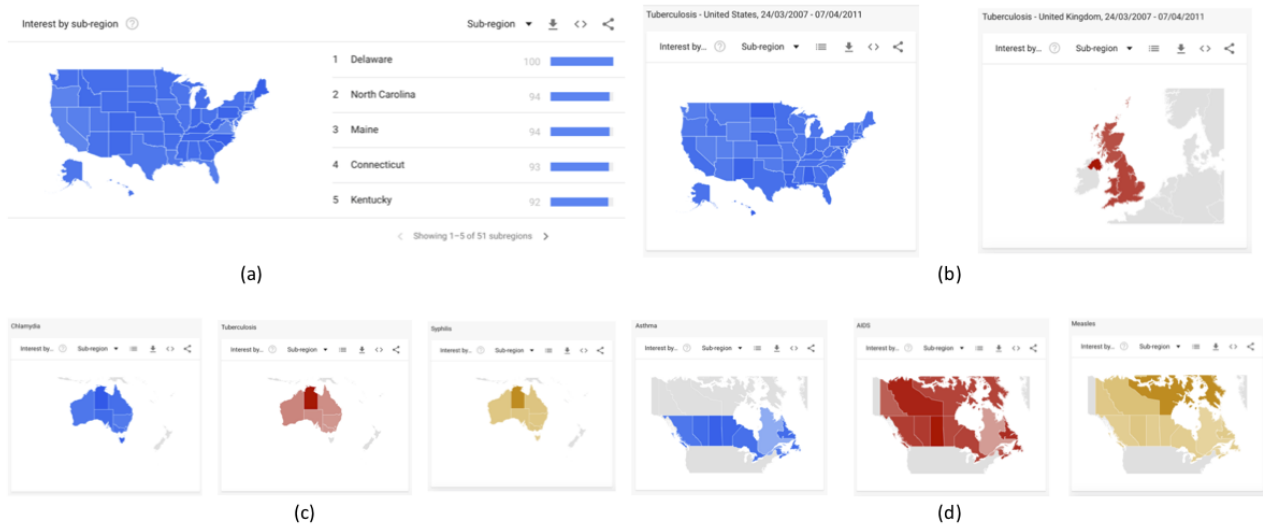
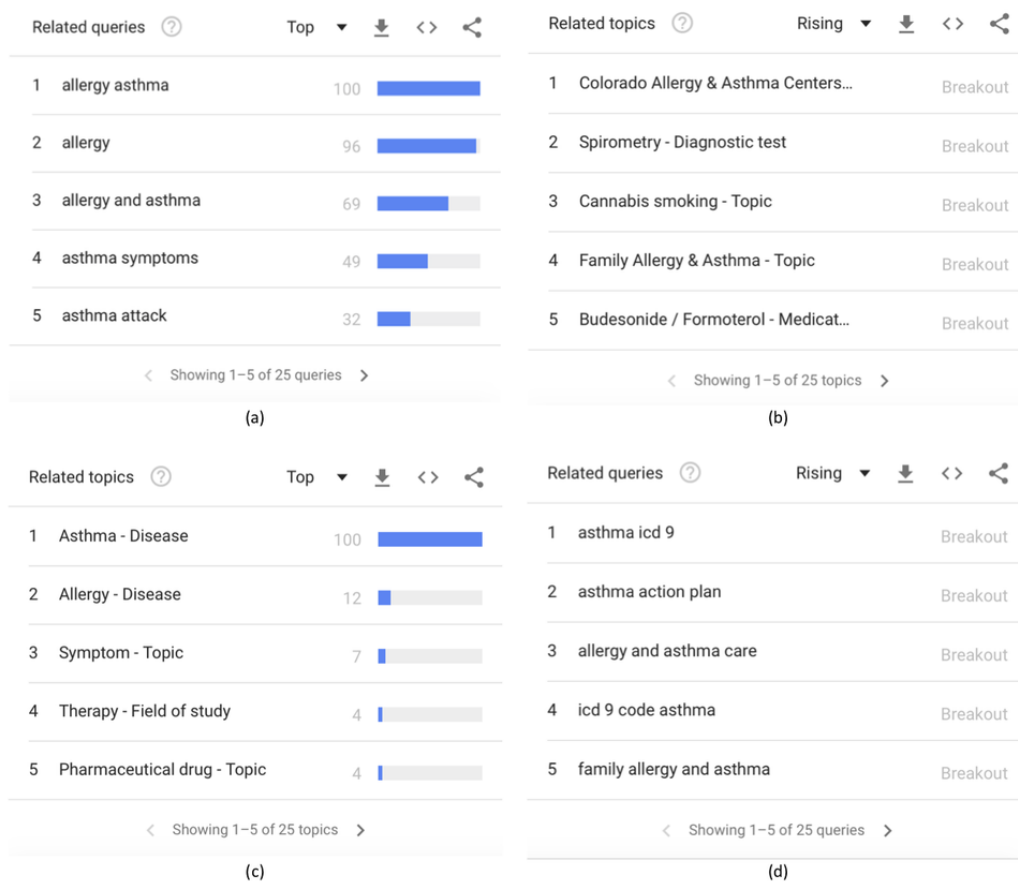


Figure 3. Google Trends’ (a) top related queries, (b) rising related topics, (c) top related topics, and (d) rising related queries for “Asthma” in the United States from Jan 1, 2004, to Dec 31, 2014.



Keyword Selection

The selection of the correct keyword(s) when examining online queries is key for valid results [51]. Thus, many factors should be taken into consideration when using Google Trends data in order to ensure a valid analysis.

Google Trends is not case sensitive, but it takes into account accents, plural or singular forms, and spelling mistakes. Therefore, whatever the choice of keywords or combination of keywords, parts of the respective queries will not be considered for further analysis.

To partly overcome this limitation, the “+” feature can be used to include the most commonly encountered misspellings, which

are selected and entered manually; however, we should keep in mind that some results will always be missing, as all possible spelling variations cannot be included. In addition, incorrect spellings of some words could be used even more often than the correct one, in which case, the analysis will not be trivial. However, in most of the cases, the correct spelling is the most commonly used, and therefore, the analysis can proceed as usual. For example, gonorrhoea is often misspelled, mainly as "Gonorrhea," which is also the Spanish term for the disease. As depicted in Figure 4a, both terms have significantly high volumes. Therefore, to include more results, both terms could be entered as the search term by using the "+" feature (Figure 4b). In this way, all results including the correct and the incorrect spellings are aggregated in the results. Note that this is not limited to only two terms; the "+" feature can be used for multiple keywords or for results in multiple languages in a region.

In the case of accents, before choosing the keywords to be examined, the variations in interest between the terms with and those without accents and special characters should be explored. For example, measles translates into "Sarampión," "ošpice," "mässling," and "Ίλαρά" in Spanish, Slovenian, Swedish, and Greek, respectively. As depicted in Figure 5, in Spanish and Greek, the term without the accent is searched for in higher volumes; in Slovenian, the term with the accent is mostly used; and in Swedish, the term without the accent is almost nonexistent. Thus, in Greek searches, the term without accent should be selected, in Slovenian and Swedish searches, terms with accents should be used, while for Spanish, as both terms yield significant results, either both terms using the "+" feature or the term without the accent should be selected.

Another important aspect is the use of quotation marks when selecting the keyword. This obviously applies only to keywords with two or more words. For example, breast cancer can be searched online by using or not using quotes. To elaborate, the term "breast cancer" without quotes will yield results that include the words "breast" and "cancer" in any possible combination and order; for example, keywords "breast cancer screening" and "breast and colon cancer" are both included in the results. However, when using quotes, the term "breast cancer" is included as is; for example, "breast cancer screening," "living with breast cancer," and "breast cancer patient." As shown in Figure 6a, the results are almost identical in this case. However, this is not always the case. As depicted in Figure 6b, this is clearly different for "HIV test." When searching for HIV test with and without quotes, the results differ in volumes of searches, despite the trend being very similar but not exactly the same.

Finally, when researching with Google Trends, the options of "search term" and "disease" (or "topic") are available when entering a keyword. Although the "search term" gives results for all keywords that include the selected term, "disease" includes various keywords that fall within the category, or, as Google describes it, "topics are a group of terms that share the same concept in any language."

Therefore, it is imperative that keyword selection is conducted with caution and that the available options and features are carefully explored and analyzed. This will ensure validity of the results.

Figure 4. Use of the "+" feature for including misspelled terms for (a) "Gonorrhoea" compared to "Gonorrhea"; (b) both terms by using the "+" feature.

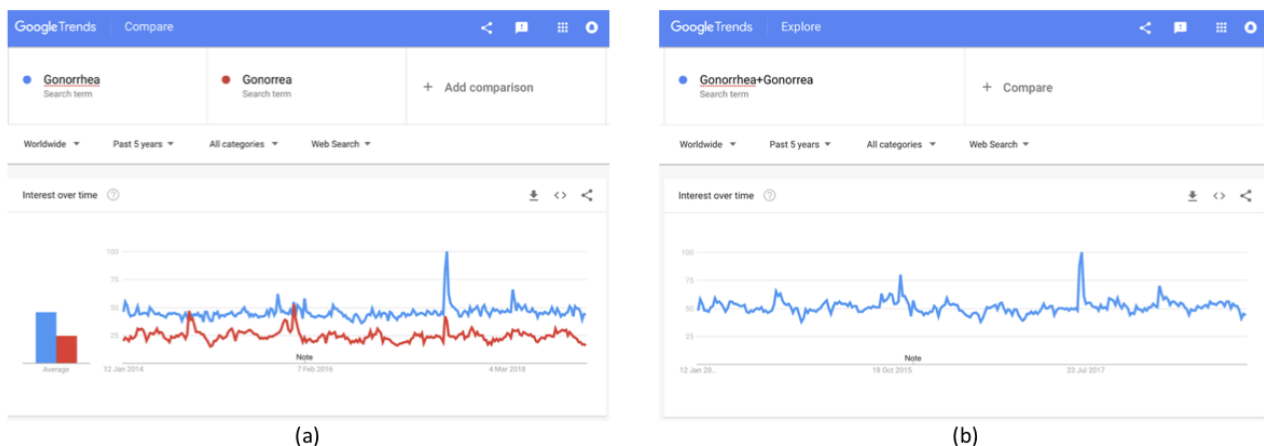


Figure 5. Selection of the correct keyword for measles based on the use of accents in the respective translated terms in (a) Spanish, (b) Slovenian, (c) Swedish, and (d) Greek.

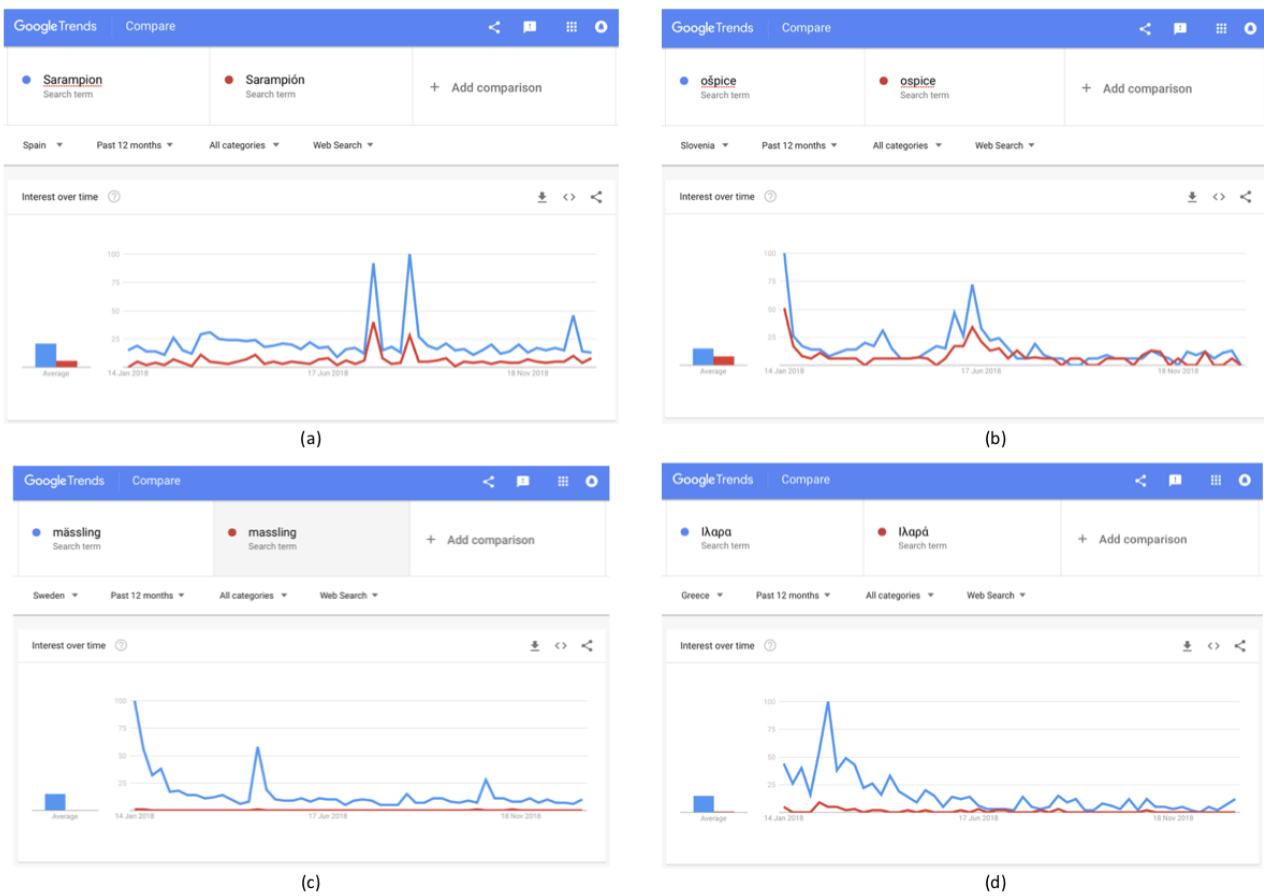
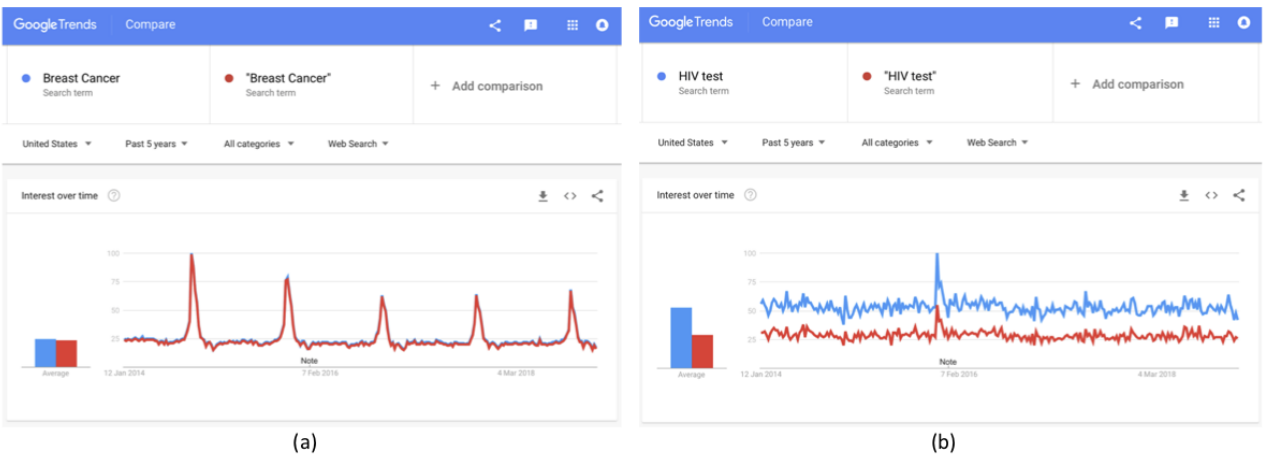


Figure 6. Differences in results with and without quotation marks for (a) “Breast Cancer” and (b) “HIV test.”.



Region Selection

The next step is to select the geographical region for which query data are retrieved. The first level of categorization allows data download for the online interest of one or more terms worldwide or by country. The list available includes all countries, in most of which interest in smaller regions can be explored.

For example, in the United States, it is possible to compare results even at metropolitan and city levels. Figure 7a shows the regional online interest in the term “Flu” worldwide, where

the United States is the country with the highest online interest in the examined term, followed by the rest of the 33 countries in which the examined term is most popular. Figure 7b shows the heat map of the interest by state in the United States in the term “Flu” over the past 5 years; either as a new independent search or by clicking on the country “USA” in the worldwide map. As shown in the right bottom corner of Figure 7, Google Trends provides the relative interest for all 50 US states plus Washington DC.

In the case of the United States, it is possible to examine the online interest by metropolitan area, as depicted in Figure 8

with the examples of California, Texas, New York, and Florida. The option for examining the online interest at the metropolitan level is not available for all countries, where from the state (or county) level, the interest changes directly to the city level. This includes fewer cities than regions with available metropolitan area data, as, for example, in countries with very large populations like India (Figure 9e) or with smaller populations like Greece (Figure 9f).

Figure 9 depicts the online interest by city in the selected metropolitan areas of Los Angeles in California, Dallas in Texas, New York in New York, and Miami in Florida.

At metropolitan level, by selecting the “include low search volume regions,” the total of the included cities is 123 in Los Angeles, 67 in Texas, 110 New York, and 50 in Miami, while in India and Greece, the number of cities remains 7 and 2, respectively.

Figure 7. Online interest in the term “Flu” over the past 5 years (a) worldwide and (b) in the United States.

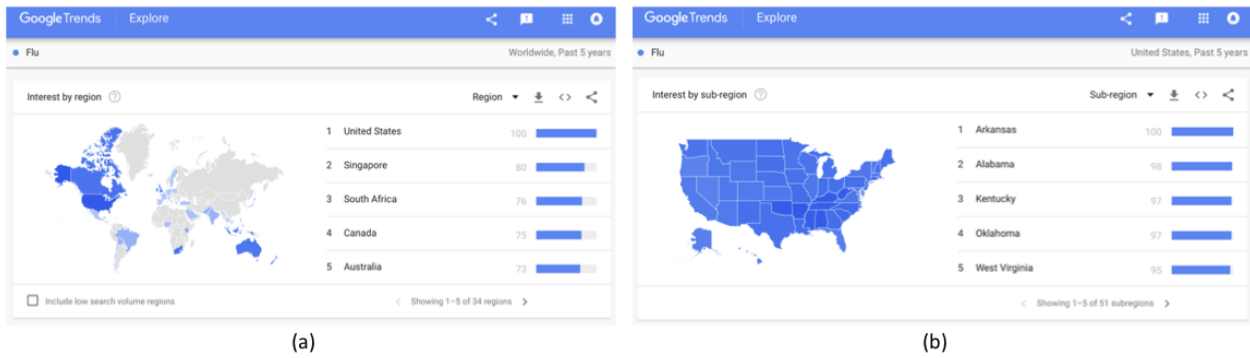


Figure 8. Regional online interest in the term “Flu” at metropolitan level over the past 5 years in (a) California, (b) Texas, (c) New York, and (d) Florida.

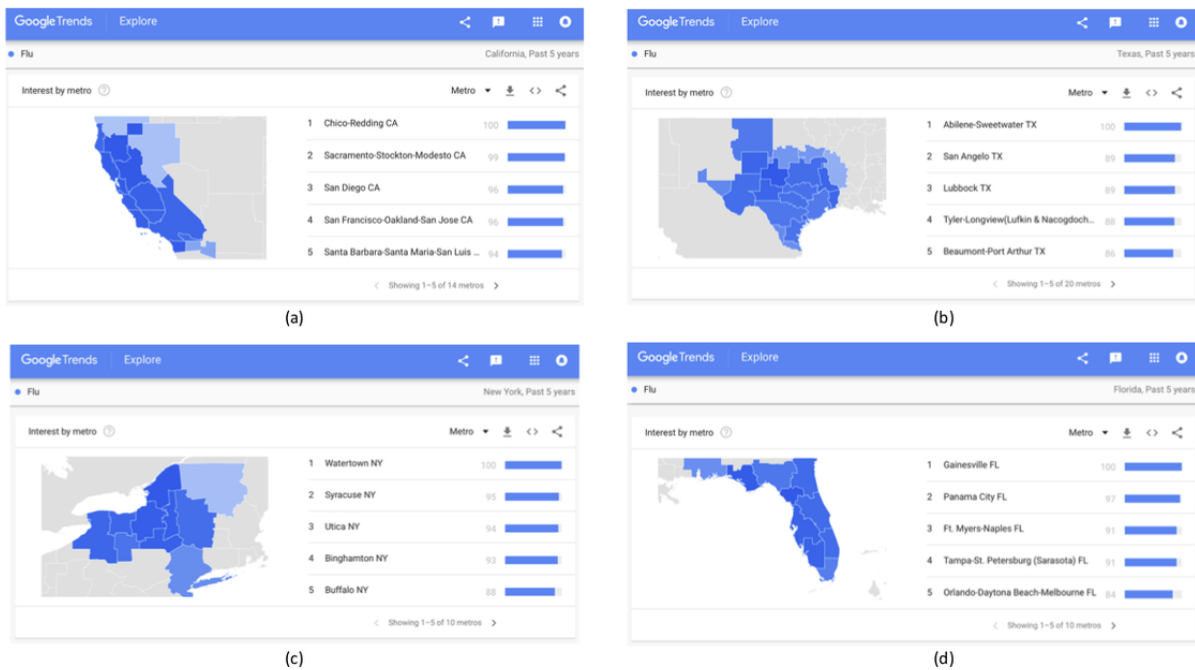
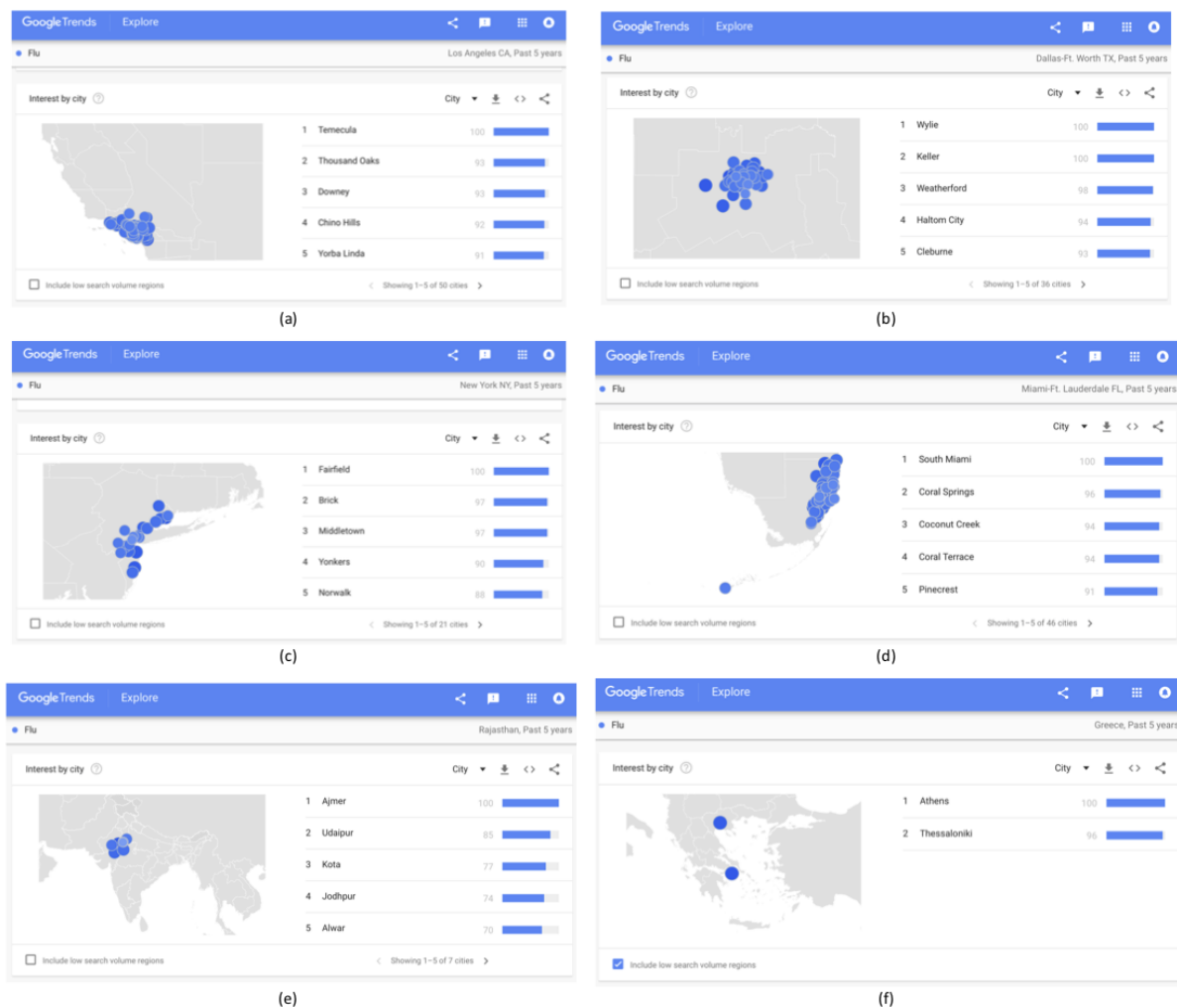


Figure 9. Regional online interest in the term “Flu” at city level over the past 5 years in (a) Los Angeles, (b) Dallas, (c) New York, (d) Miami, (e) India, and (f) Greece.



Period Selection

As the data are normalized over the selected period, the time frame for which Google Trends data are retrieved is crucial for the validity of the results. The selection of the examined time frame is one of the most common mistakes in Google Trends research. The main guideline is that the period selected for Google data should be exactly the same as the one for which official data are available and will be examined. For example, if monthly (or yearly) official data from January 2004 to December 2014 are available, then the selected period for retrieving Google Trends data should be January 2004 to December 2014. Neither 15 datasets for each individual year nor a random number of datasets arbitrarily chosen should be used; a single dataset should be compiled including the months from January 2004 to December 2014. Note that data may slightly vary depending on the time of retrieval; thus, the date and time of downloading must be reported.

Depending on the time frame, the interval for which data are available varies significantly (Table 3), which includes the data

intervals for the preselected time frames in Google Trends. Note that the default selection is 12 months.

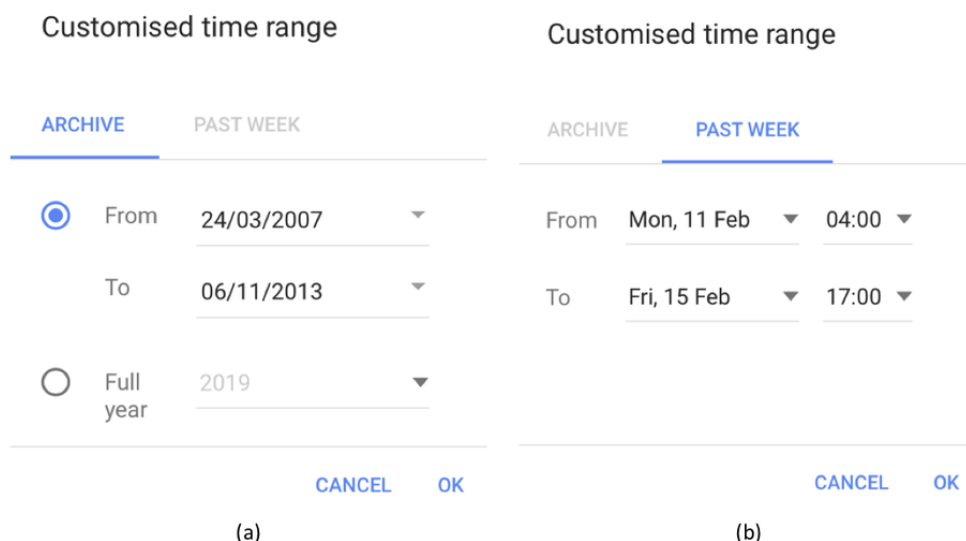
The time frame can be customized at will; for example, March 24, 2007, to November 6, 2013 (Figure 10 a). Furthermore, there is an option to select the exact hours for which data are retrieved, but only over the past week; for example, from February 11, 4 am, to February 15, 5 pm (Figure 10 b).

Finally, an important detail in the selection of the time frame is when the data retrieval changes from monthly to weekly and weekly to daily. For example, from April 28, 2013, to June 30, 2018, the data are retrieved in weekly intervals, while from April 27, 2013, to June 30, 2018, the data are retrieved in monthly intervals. Hence, the data from monthly to weekly changes in (roughly) 5 years and 2 months. For daily data, we observe that, for example, from October 4, 2017, to June 30, 2018, the data are retrieved in daily intervals, while from October 3, 2017, to June 30, 2018, the data are retrieved in weekly intervals; as such, the data interval changes from daily to weekly in (roughly) 10 months.

Table 3. Data intervals and number of observations for the default options in period selection.

Selected period	Data intervals	Number of observations
2004 to present	Monthly	>187
Past 5 years	Weekly	260
Full year (eg, 2004 or 2008)	Weekly	52
Past 12 months	Weekly	52
Past 90 days	Daily	90
Past 30 days	Daily	30
Past 7 days	Hourly	168
Past day	8 min	180
Past 4 hours	1 min	240
Past hour	1 min	60

Figure 10. Customized time range (a) from archive and (b) over the past week.



Search Categories

When exploring the online interest, the selected term can be analyzed based on a selected category. This feature is important to eliminate noisy data, especially in cases where the same word is used or can be attributed to different meanings or events. For example, the terms “yes” and “no” are very commonly searched for, so, when aiming at predicting the results of a referendum race, the search must be limited to the category “Politics” or “Campaign and elections” in order to retrieve the data that are attributed to the event. However, selecting a category is not required when the keyword searched is specific and not related to other words, meanings, and events.

The available categories are listed in Table A1 of [Multimedia Appendix 1](#). Note that most of these categories have subcategories, which, in turn, have other subcategories, allowing the available categories to be as broad or as narrow as required.

In this paper, we focus on the category of “Health” (first level of categorization). The main available subcategories (second level of categorization) of “Health” along with all available subcategories (third and fourth levels) are presented in Table A2 of [Multimedia Appendix 1](#).

Finally, another feature is the type of search conducted when entering a keyword, which consists of the options of “Web Search,” “Image Search,” “News Search,” “Google Shopping,” and “YouTube Search.” Apart from very specific cases, the “Web Search,” which is also the default option, should be selected.

Discussion

Over the past decade, Web-based data are used extensively in digital epidemiology, with online sources playing a central role in health informatics [1,2,52]. Digital disease detection [53] consists of detecting, analyzing, and predicting disease occurrence and spread, and several types of online sources are used, including mainly digital platforms [54,55]. When addressing infodemiology topics, a concept first introduced by Eysenbach [1], Google Trends is an important tool, and research on the subject is constantly expanding [56]. Most studies on Google Trends research are in health and medicine, focusing mainly on the surveillance and analysis of health topics and the forecasting of diseases, outbreaks, and epidemics. As Google Trends is open and user friendly, it is accessed and used by several researchers, even those who are not strictly related to

the field of big data, but use it as a means of exploring behavioral variations toward selected topics. The latter has resulted in differences in methodologies followed, which, at times, involve mistakes.

Despite the large number of studies in this line of research, there was a lack of a methodology framework that should be followed. This has produced differences in presentation, and, more importantly, in crucial mistakes that compromise the validity of the results. In this article, we provided a concise overview of the how the tool works and proposed a step-by-step methodology (ie, the four steps of selecting the correct/appropriate keyword, region, period, and category) to ensure the validity of the results in Google Trends research. We also included research examples to provide guidance not only to the experienced eye, but also to new researchers.

As is evident by the findings of this study, there are several limitations to the use of Google Trends data. First, despite the evident potential that Google data have to offer in epidemiology and disease surveillance, there have been some issues in the past, where online search traffic data at some point failed to accurately predict disease spreading, as in the case of Google Flu Trends [57], a Google tool for the surveillance of influenza-like illness (the flu) that is no longer available. Regardless, Google Flu Trends has been accurate in the past in predicting the spread of flu, as suggested by several studies and reports [58-60].

The latter could be partly attributed to the fact that, when researching with Google Trends, the sample is unknown and it cannot be shown to be representative. Despite this and considering the increasing internet penetration, previous studies have suggested that Web-based data have been empirically shown to provide valuable and valid results in exploring and predicting behavior and are correlated with actual data [61-66]. However, recent research has suggested that online queries do not provide valid results in regions with low internet penetration or low scorings in freedom of speech [67].

Furthermore, the data that are retrieved are normalized over the selected period; thus, the exact volumes of queries are not

known, limiting the way that the data can be processed and analysis can be performed. Therefore, the data should be analyzed in the appropriate way, and the results should be carefully interpreted.

In addition, the selection of keyword(s) plays a very important role in ensuring the validity of the results. In some cases, the noisy data (ie, queries not attributed to the examined term) must be excluded, which are not always trivial. This can be partly overcome by selecting a specific category, which always bares the risk of excluding results that are needed for analysis.

The analysis of Google Trends data has several other limitations, as examining Web data can bear threats to validity. Careful analysis should be performed to ensure that news reporting and sudden events do not compromise the validity of the results. In addition, as the sample is unknown, several other demographic factors such as age and sex cannot be included in the analysis.

Finally, as this field of research is relatively new, there is no standard way of reporting, resulting in the same meaning of different terms, different meanings of the same term, and different abbreviations. For example, Google Trends data are referred to as relative search volumes, search volumes, online queries, online search traffic data, normalized hits, and other terms. Thus, future research should focus on developing specific coding for Google Trends research, so that a unified way of reporting is followed by all researchers in the field.

In the era of big data, the analysis of Google queries has become a valuable tool for researchers to explore and predict human behavior, as it has been suggested that online data are correlated with actual health data. The methodology framework proposed in this article for researching with Google Trends is much needed to provide guidance for using Google Trends data in health assessment, and, more importantly, to help researchers and health officials and organizations avoid common mistakes that compromise the validity of the results. As research on the subject is expanding, future work should include the coding in Google Trends research and extend this framework along with changes in the tool and the analysis methods.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Google Trends categories.

[PDF File (Adobe PDF File), 51KB - [publichealth_v5i2e13439_app1.pdf](#)]

References

1. 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;11(1):e11 [FREE Full text] [doi: [10.2196/jmir.1157](#)] [Medline: [19329408](#)]
2. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med* 2011 May;40(5 Suppl 2):S154-S158. [doi: [10.1016/j.amepre.2011.02.006](#)] [Medline: [21521589](#)]

3. 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](https://doi.org/10.1371/journal.pone.0109583)] [Medline: [25337815](https://pubmed.ncbi.nlm.nih.gov/25337815/)]
4. 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](https://doi.org/10.2196/jmir.9366)] [Medline: [30401664](https://pubmed.ncbi.nlm.nih.gov/30401664/)]
5. Gamma A, Schleifer R, Weinmann W, Buadze A, Liebrez M. Could Google Trends Be Used to Predict Methamphetamine-Related Crime? An Analysis of Search Volume Data in Switzerland, Germany, and Austria. *PLoS One* 2016;11(11):e0166566 [FREE Full text] [doi: [10.1371/journal.pone.0166566](https://doi.org/10.1371/journal.pone.0166566)] [Medline: [27902717](https://pubmed.ncbi.nlm.nih.gov/27902717/)]
6. Mavragani A, Ochoa G. Infoveillance of infectious diseases in USA: STDs, tuberculosis, and hepatitis. *J Big Data* 2018 Sep 6;5(1). [doi: [10.1186/s40537-018-0140-9](https://doi.org/10.1186/s40537-018-0140-9)]
7. Mavragani A, Sampri A, Sypsa K, Tsagarakis KP. Integrating Smart Health in the US Health Care System: Infodemiology Study of Asthma Monitoring in the Google Era. *JMIR Public Health Surveill* 2018 Mar 12;4(1):e24 [FREE Full text] [doi: [10.2196/publichealth.8726](https://doi.org/10.2196/publichealth.8726)] [Medline: [29530839](https://pubmed.ncbi.nlm.nih.gov/29530839/)]
8. Simpson SS, Adams N, Brugman CM, Conners TJ. Detecting Novel and Emerging Drug Terms Using Natural Language Processing: A Social Media Corpus Study. *JMIR Public Health Surveill* 2018 Jan 08;4(1):e2 [FREE Full text] [doi: [10.2196/publichealth.7726](https://doi.org/10.2196/publichealth.7726)] [Medline: [29311050](https://pubmed.ncbi.nlm.nih.gov/29311050/)]
9. Wongkoblap A, Vadillo MA, Curcin V. Researching Mental Health Disorders in the Era of Social Media: Systematic Review. *J Med Internet Res* 2017 Dec 29;19(6):e228 [FREE Full text] [doi: [10.2196/jmir.7215](https://doi.org/10.2196/jmir.7215)] [Medline: [28663166](https://pubmed.ncbi.nlm.nih.gov/28663166/)]
10. Park SH, Hong SH. Identification of Primary Medication Concerns Regarding Thyroid Hormone Replacement Therapy From Online Patient Medication Reviews: Text Mining of Social Network Data. *J Med Internet Res* 2018 Oct 24;20(10):e11085 [FREE Full text] [doi: [10.2196/11085](https://doi.org/10.2196/11085)] [Medline: [30355555](https://pubmed.ncbi.nlm.nih.gov/30355555/)]
11. Ricard BJ, Marsch LA, Crosier B, Hassanpour S. Exploring the Utility of Community-Generated Social Media Content for Detecting Depression: An Analytical Study on Instagram. *J Med Internet Res* 2018 Dec 06;20(12):e11817 [FREE Full text] [doi: [10.2196/11817](https://doi.org/10.2196/11817)] [Medline: [30522991](https://pubmed.ncbi.nlm.nih.gov/30522991/)]
12. Huesch M, Chetlen A, Segel J, Schetter S. Frequencies of Private Mentions and Sharing of Mammography and Breast Cancer Terms on Facebook: A Pilot Study. *J Med Internet Res* 2017 Jun 09;19(6):e201 [FREE Full text] [doi: [10.2196/jmir.7508](https://doi.org/10.2196/jmir.7508)] [Medline: [28600279](https://pubmed.ncbi.nlm.nih.gov/28600279/)]
13. Chen T, Dredze M. Vaccine Images on Twitter: Analysis of What Images are Shared. *J Med Internet Res* 2018 Apr 03;20(4):e130 [FREE Full text] [doi: [10.2196/jmir.8221](https://doi.org/10.2196/jmir.8221)] [Medline: [29615386](https://pubmed.ncbi.nlm.nih.gov/29615386/)]
14. Farhadloo M, Winneg K, Chan MS, Hall Jamieson K, Albarracin D. Associations of Topics of Discussion on Twitter With Survey Measures of Attitudes, Knowledge, and Behaviors Related to Zika: Probabilistic Study in the United States. *JMIR Public Health Surveill* 2018 Feb 09;4(1):e16 [FREE Full text] [doi: [10.2196/publichealth.8186](https://doi.org/10.2196/publichealth.8186)] [Medline: [29426815](https://pubmed.ncbi.nlm.nih.gov/29426815/)]
15. van Lent LG, Sungur H, Kunneman FA, van de Velde B, Das E. Too Far to Care? Measuring Public Attention and Fear for Ebola Using Twitter. *J Med Internet Res* 2017 Dec 13;19(6):e193 [FREE Full text] [doi: [10.2196/jmir.7219](https://doi.org/10.2196/jmir.7219)] [Medline: [28611015](https://pubmed.ncbi.nlm.nih.gov/28611015/)]
16. Du J, Tang L, Xiang Y, Zhi D, Xu J, Song HY, et al. Public Perception Analysis of Tweets During the 2015 Measles Outbreak: Comparative Study Using Convolutional Neural Network Models. *J Med Internet Res* 2018 Jul 09;20(7):e236 [FREE Full text] [doi: [10.2196/jmir.9413](https://doi.org/10.2196/jmir.9413)] [Medline: [29986843](https://pubmed.ncbi.nlm.nih.gov/29986843/)]
17. Sewalk KC, Tuli G, Hswen Y, Brownstein JS, Hawkins JB. Using Twitter to Examine Web-Based Patient Experience Sentiments in the United States: Longitudinal Study. *J Med Internet Res* 2018 Oct 12;20(10):e10043 [FREE Full text] [doi: [10.2196/10043](https://doi.org/10.2196/10043)] [Medline: [30314959](https://pubmed.ncbi.nlm.nih.gov/30314959/)]
18. Chen S, Xu Q, Buchenberger J, Bagavathi A, Fair G, Shaikh S, et al. Dynamics of Health Agency Response and Public Engagement in Public Health Emergency: A Case Study of CDC Tweeting Patterns During the 2016 Zika Epidemic. *JMIR Public Health Surveill* 2018 Nov 22;4(4):e10827 [FREE Full text] [doi: [10.2196/10827](https://doi.org/10.2196/10827)] [Medline: [30467106](https://pubmed.ncbi.nlm.nih.gov/30467106/)]
19. Alvarez-Mon MA, Asunsolo Del Barco A, Lahera G, Quintero J, Ferre F, Pereira-Sanchez V, et al. Increasing Interest of Mass Communication Media and the General Public in the Distribution of Tweets About Mental Disorders: Observational Study. *J Med Internet Res* 2018 May 28;20(5):e205 [FREE Full text] [doi: [10.2196/jmir.9582](https://doi.org/10.2196/jmir.9582)] [Medline: [29807880](https://pubmed.ncbi.nlm.nih.gov/29807880/)]
20. Rocchetti M, Marfia G, Salomoni P, Prandi C, Zagari RM, Gningaye KFL, et al. Attitudes of Crohn's Disease Patients: Infodemiology Case Study and Sentiment Analysis of Facebook and Twitter Posts. *JMIR Public Health Surveill* 2017 Aug 09;3(3):e51 [FREE Full text] [doi: [10.2196/publichealth.7004](https://doi.org/10.2196/publichealth.7004)] [Medline: [28793981](https://pubmed.ncbi.nlm.nih.gov/28793981/)]
21. Lu FS, Hou S, Baltrusaitis K, Shah M, Leskovec J, Sosic 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/)]
22. Google Trends. URL: <https://trends.google.com/trends/?geo=US> [accessed 2019-01-10] [WebCite Cache ID 75LXQbBW2]
23. Mavragani A, Tsagarakis KP. YES or NO: Predicting the 2015 GReferendum results using Google Trends. *Technological Forecasting and Social Change* 2016 Aug;109:1-5. [doi: [10.1016/j.techfore.2016.04.028](https://doi.org/10.1016/j.techfore.2016.04.028)]
24. Mavragani A, Ochoa G. Forecasting AIDS prevalence in the United States using online search traffic data. *J Big Data* 2018 May 19;5(1). [doi: [10.1186/s40537-018-0126-7](https://doi.org/10.1186/s40537-018-0126-7)]

25. Tana JC, Kettunen J, Eirola E, Paakkonen H. Diurnal Variations of Depression-Related Health Information Seeking: Case Study in Finland Using Google Trends Data. *JMIR Ment Health* 2018 May 23;5(2):e43 [FREE Full text] [doi: [10.2196/mental.9152](https://doi.org/10.2196/mental.9152)] [Medline: [29792291](https://pubmed.ncbi.nlm.nih.gov/29792291/)]
26. Solano P, Ustulin M, Pizzorno E, Vichi M, Pompili M, Serafini G, et al. A Google-based approach for monitoring suicide risk. *Psychiatry Res* 2016 Dec 30;246:581-586. [doi: [10.1016/j.psychres.2016.10.030](https://doi.org/10.1016/j.psychres.2016.10.030)] [Medline: [27837725](https://pubmed.ncbi.nlm.nih.gov/27837725/)]
27. Arora VS, Stuckler D, McKee M. Tracking search engine queries for suicide in the United Kingdom, 2004-2013. *Public Health* 2016 Aug;137:147-153. [doi: [10.1016/j.puhe.2015.10.015](https://doi.org/10.1016/j.puhe.2015.10.015)] [Medline: [26976489](https://pubmed.ncbi.nlm.nih.gov/26976489/)]
28. Zheluk A, Quinn C, Meylaks P. Internet search and krokodil in the Russian Federation: an infoveillance study. *J Med Internet Res* 2014 Sep 18;16(9):e212 [FREE Full text] [doi: [10.2196/jmir.3203](https://doi.org/10.2196/jmir.3203)] [Medline: [25236385](https://pubmed.ncbi.nlm.nih.gov/25236385/)]
29. Zhang Z, Zheng X, Zeng DD, Leischow SJ. Tracking Dabbing Using Search Query Surveillance: A Case Study in the United States. *J Med Internet Res* 2016 Sep 16;18(9):e252 [FREE Full text] [doi: [10.2196/jmir.5802](https://doi.org/10.2196/jmir.5802)] [Medline: [27637361](https://pubmed.ncbi.nlm.nih.gov/27637361/)]
30. Abdellaoui R, Foulquié P, Texier N, Faviez C, Burgun A, Schück S. Detection of Cases of Noncompliance to Drug Treatment in Patient Forum Posts: Topic Model Approach. *J Med Internet Res* 2018 Mar 14;20(3):e85 [FREE Full text] [doi: [10.2196/jmir.9222](https://doi.org/10.2196/jmir.9222)] [Medline: [29540337](https://pubmed.ncbi.nlm.nih.gov/29540337/)]
31. Allem J, Dharmapuri L, Leventhal AM, Unger JB, Boley Cruz T. Hookah-Related Posts to Twitter From 2017 to 2018: Thematic Analysis. *J Med Internet Res* 2018 Nov 19;20(11):e11669 [FREE Full text] [doi: [10.2196/11669](https://doi.org/10.2196/11669)] [Medline: [30455162](https://pubmed.ncbi.nlm.nih.gov/30455162/)]
32. Berlinberg EJ, Deiner MS, Porco TC, Acharya NR. Monitoring Interest in Herpes Zoster Vaccination: Analysis of Google Search Data. *JMIR Public Health Surveill* 2018 May 02;4(2):e10180 [FREE Full text] [doi: [10.2196/10180](https://doi.org/10.2196/10180)] [Medline: [29720364](https://pubmed.ncbi.nlm.nih.gov/29720364/)]
33. 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 Mar 08;5(1):e13142 [FREE Full text] [doi: [10.2196/13142](https://doi.org/10.2196/13142)] [Medline: [30763255](https://pubmed.ncbi.nlm.nih.gov/30763255/)]
34. Foroughi F, Lam AK, Lim MSC, Saremi N, Ahmadvand A. "Googling" for Cancer: An Infodemiological Assessment of Online Search Interests in Australia, Canada, New Zealand, the United Kingdom, and the United States. *JMIR Cancer* 2016 May 04;2(1):e5 [FREE Full text] [doi: [10.2196/cancer.5212](https://doi.org/10.2196/cancer.5212)] [Medline: [28410185](https://pubmed.ncbi.nlm.nih.gov/28410185/)]
35. 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/)]
36. Hswen Y, Gopaluni A, Brownstein JS, Hawkins JB. Using Twitter to Detect Psychological Characteristics of Self-Identified Persons With Autism Spectrum Disorder: A Feasibility Study. *JMIR Mhealth Uhealth* 2019 Feb 12;7(2):e12264 [FREE Full text] [doi: [10.2196/12264](https://doi.org/10.2196/12264)] [Medline: [30747718](https://pubmed.ncbi.nlm.nih.gov/30747718/)]
37. Jones J, Pradhan M, Hosseini M, Kulanthaivel A, Hosseini M. Novel Approach to Cluster Patient-Generated Data Into Actionable Topics: Case Study of a Web-Based Breast Cancer Forum. *JMIR Med Inform* 2018 Nov 29;6(4):e45 [FREE Full text] [doi: [10.2196/medinform.9162](https://doi.org/10.2196/medinform.9162)] [Medline: [30497991](https://pubmed.ncbi.nlm.nih.gov/30497991/)]
38. 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/)]
39. Keller MS, Mosadeghi S, Cohen ER, Kwan J, Spiegel BMR. Reproductive Health and Medication Concerns for Patients With Inflammatory Bowel Disease: Thematic and Quantitative Analysis Using Social Listening. *J Med Internet Res* 2018 Jun 11;20(6):e206 [FREE Full text] [doi: [10.2196/jmir.9870](https://doi.org/10.2196/jmir.9870)] [Medline: [29891471](https://pubmed.ncbi.nlm.nih.gov/29891471/)]
40. Mejova Y, Weber I, Fernandez-Luque L. Online Health Monitoring using Facebook Advertisement Audience Estimates in the United States: Evaluation Study. *JMIR Public Health Surveill* 2018 Mar 28;4(1):e30 [FREE Full text] [doi: [10.2196/publichealth.7217](https://doi.org/10.2196/publichealth.7217)] [Medline: [29592849](https://pubmed.ncbi.nlm.nih.gov/29592849/)]
41. Odlum M, Yoon S, Broadwell P, Brewer R, Kuang D. How Twitter Can Support the HIV/AIDS Response to Achieve the 2030 Eradication Goal: In-Depth Thematic Analysis of World AIDS Day Tweets. *JMIR Public Health Surveill* 2018 Nov 22;4(4):e10262 [FREE Full text] [doi: [10.2196/10262](https://doi.org/10.2196/10262)] [Medline: [30467102](https://pubmed.ncbi.nlm.nih.gov/30467102/)]
42. Phillips CA, Barz LA, Li Y, Schapira MM, Bailey LC, Merchant RM. Relationship Between State-Level Google Online Search Volume and Cancer Incidence in the United States: Retrospective Study. *J Med Internet Res* 2018 Jan 08;20(1):e6 [FREE Full text] [doi: [10.2196/jmir.8870](https://doi.org/10.2196/jmir.8870)] [Medline: [29311051](https://pubmed.ncbi.nlm.nih.gov/29311051/)]
43. Poirier C, Lavenu A, Bertaud V, Campillo-Gimenez B, Chazard E, Cuggia M, et al. Real Time Influenza Monitoring Using Hospital Big Data in Combination with Machine Learning Methods: Comparison Study. *JMIR Public Health Surveill* 2018 Dec 21;4(4):e11361 [FREE Full text] [doi: [10.2196/11361](https://doi.org/10.2196/11361)] [Medline: [30578212](https://pubmed.ncbi.nlm.nih.gov/30578212/)]
44. Radin M, Sciascia S. Infodemiology of systemic lupus erythematosus using Google Trends. *Lupus* 2017 Jul;26(8):886-889. [doi: [10.1177/0961203317691372](https://doi.org/10.1177/0961203317691372)] [Medline: [28162030](https://pubmed.ncbi.nlm.nih.gov/28162030/)]
45. Vasconcellos-Silva PR, Carvalho DBF, Trajano V, de La Rocque LR, Sawada ACMB, Juvanhol LL. Using Google Trends Data to Study Public Interest in Breast Cancer Screening in Brazil: Why Not a Pink February? *JMIR Public Health Surveill* 2017 Apr 06;3(2):e17 [FREE Full text] [doi: [10.2196/publichealth.7015](https://doi.org/10.2196/publichealth.7015)] [Medline: [28385679](https://pubmed.ncbi.nlm.nih.gov/28385679/)]

46. Wakamiya S, Kawai Y, Aramaki E. Twitter-Based Influenza Detection After Flu Peak via Tweets With Indirect Information: Text Mining Study. *JMIR Public Health Surveill* 2018 Sep 25;4(3):e65 [FREE Full text] [doi: [10.2196/publichealth.8627](https://doi.org/10.2196/publichealth.8627)] [Medline: [30274968](https://pubmed.ncbi.nlm.nih.gov/30274968/)]
47. Wang H, Chen D. Economic Recession and Obesity-Related Internet Search Behavior in Taiwan: Analysis of Google Trends Data. *JMIR Public Health Surveill* 2018 Apr 06;4(2):e37 [FREE Full text] [doi: [10.2196/publichealth.7314](https://doi.org/10.2196/publichealth.7314)] [Medline: [29625958](https://pubmed.ncbi.nlm.nih.gov/29625958/)]
48. Watad A, Watad S, Mahroum N, Sharif K, Amital H, Bragazzi NL, et al. 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 Feb 28;5(1):e9176 [FREE Full text] [doi: [10.2196/publichealth.9176](https://doi.org/10.2196/publichealth.9176)] [Medline: [30601755](https://pubmed.ncbi.nlm.nih.gov/30601755/)]
49. Xu C, Wang Y, Yang H, Hou J, Sun L, Zhang X, et al. Association Between Cancer Incidence and Mortality in Web-Based Data in China: Infodemiology Study. *J Med Internet Res* 2019 Jan 29;21(1):e10677 [FREE Full text] [doi: [10.2196/10677](https://doi.org/10.2196/10677)] [Medline: [30694203](https://pubmed.ncbi.nlm.nih.gov/30694203/)]
50. Google Trends. How Trends data is adjusted URL: https://support.google.com/trends/answer/4365533?hl=en-GB&ref_topic=6248052 [accessed 2019-01-10] [WebCite Cache ID [75LXW9Qoh](https://www.webcitation.org/75LXW9Qoh)]
51. Scharkow M, Vogelgesang J. Measuring the Public Agenda using Search Engine Queries. *International Journal of Public Opinion Research* 2011 Mar 01;23(1):104-113. [doi: [10.1093/ijpor/edq048](https://doi.org/10.1093/ijpor/edq048)]
52. Salathé M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, Buckee C, et al. Digital epidemiology. *PLoS Comput Biol* 2012 Jul;8(7):e1002616 [FREE Full text] [doi: [10.1371/journal.pcbi.1002616](https://doi.org/10.1371/journal.pcbi.1002616)] [Medline: [22844241](https://pubmed.ncbi.nlm.nih.gov/22844241/)]
53. Brownstein J, Freifeld CC, Madoff LC. Digital disease detection--harnessing the Web for public health surveillance. *N Engl J Med* 2009 May 21;360(21):2153--5,2157 [FREE Full text] [doi: [10.1056/NEJMp0900702](https://doi.org/10.1056/NEJMp0900702)] [Medline: [19423867](https://pubmed.ncbi.nlm.nih.gov/19423867/)]
54. Chunara R, Goldstein E, Patterson-Lomba O, Brownstein JS. Estimating influenza attack rates in the United States using a participatory cohort. *Sci Rep* 2015 Apr 2;5(1). [doi: [10.1038/srep09540](https://doi.org/10.1038/srep09540)] [Medline: [25835538](https://pubmed.ncbi.nlm.nih.gov/25835538/)]
55. Leal-Neto OB, Dimech GS, Libel M, Oliveira W, Ferreira JP. Digital disease detection and participatory surveillance: overview and perspectives for Brazil. *Rev Saude Publica* 2016;50:17 [FREE Full text] [doi: [10.1590/S1518-8787.2016050006201](https://doi.org/10.1590/S1518-8787.2016050006201)] [Medline: [27191153](https://pubmed.ncbi.nlm.nih.gov/27191153/)]
56. Bernardo TM, Rajic A, Young I, Robiadek K, Pham MT, Funk JA. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. *J Med Internet Res* 2013;15(7):e147 [FREE Full text] [doi: [10.2196/jmir.2740](https://doi.org/10.2196/jmir.2740)] [Medline: [23896182](https://pubmed.ncbi.nlm.nih.gov/23896182/)]
57. Google Flu Trends Data. URL: <https://www.google.org/flutrends/about/> [accessed 2019-02-09] [WebCite Cache ID [763FITMeo](https://www.webcitation.org/763FITMeo)]
58. Carneiro H, 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](https://doi.org/10.1086/630200)] [Medline: [19845471](https://pubmed.ncbi.nlm.nih.gov/19845471/)]
59. Dugas AF, Jalalpour M, Gel Y, Levin S, Torcaso F, Igusa T, et al. Influenza forecasting with Google Flu Trends. *PLoS One* 2013 Feb;8(2):e56176 [FREE Full text] [doi: [10.1371/journal.pone.0056176](https://doi.org/10.1371/journal.pone.0056176)] [Medline: [23457520](https://pubmed.ncbi.nlm.nih.gov/23457520/)]
60. Eurosurveillance editorial team. Google Flu Trends includes 14 European countries. *Euro Surveill* 2009;14(40) [FREE Full text] [doi: [10.2807/ese.14.40.19352-en](https://doi.org/10.2807/ese.14.40.19352-en)]
61. Mavragani A, Ochoa G. The Internet and the Anti-Vaccine Movement: Tracking the 2017 EU Measles Outbreak. *BDCC* 2018 Jan 16;2(1):2. [doi: [10.3390/bdcc2010002](https://doi.org/10.3390/bdcc2010002)]
62. Preis T, Moat HS, Stanley HE, Bishop SR. Quantifying the advantage of looking forward. *Sci Rep* 2012;2:350 [FREE Full text] [doi: [10.1038/srep00350](https://doi.org/10.1038/srep00350)] [Medline: [22482034](https://pubmed.ncbi.nlm.nih.gov/22482034/)]
63. Preis T, Moat HS, Stanley HE. Quantifying trading behavior in financial markets using Google Trends. *Sci Rep* 2013;3:1684 [FREE Full text] [doi: [10.1038/srep01684](https://doi.org/10.1038/srep01684)] [Medline: [23619126](https://pubmed.ncbi.nlm.nih.gov/23619126/)]
64. Mavragani A, Sypsa K, Sampri A, Tsagarakis K. Quantifying the UK Online Interest in Substances of the EU Watchlist for Water Monitoring: Diclofenac, Estradiol, and the Macrolide Antibiotics. *Water* 2016 Nov 18;8(11):542. [doi: [10.3390/w8110542](https://doi.org/10.3390/w8110542)]
65. 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](https://pubmed.ncbi.nlm.nih.gov/17238340/)]
66. Jun S, Yoo HS, Choi S. Ten years of research change using Google Trends: From the perspective of big data utilizations and applications. *Technological Forecasting and Social Change* 2018 May;130:69-87. [doi: [10.1016/j.techfore.2017.11.009](https://doi.org/10.1016/j.techfore.2017.11.009)]
67. Mavragani A, Tsagarakis KP. Predicting referendum results in the Big Data Era. *J Big Data* 2019 Jan 14;6:3. [doi: [10.1186/s40537-018-0166-z](https://doi.org/10.1186/s40537-018-0166-z)]

Edited by G Eysenbach; submitted 18.01.19; peer-reviewed by SH Hong, A Zheluk, O Leal Neto, C Geoghegan, MS Aslam; comments to author 07.02.19; revised version received 17.02.19; accepted 23.03.19; published 29.05.19.

Please cite as:

Mavragani A, Ochoa G

Google Trends in Infodemiology and Infoveillance: Methodology Framework

JMIR Public Health Surveill 2019;5(2):e13439

URL: <http://publichealth.jmir.org/2019/2/e13439/>

doi: [10.2196/13439](https://doi.org/10.2196/13439)

PMID: [31144671](https://pubmed.ncbi.nlm.nih.gov/31144671/)

©Amaryllis Mavragani, Gabriela Ochoa. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 29.05.2019. 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

Identifying Key Topics Bearing Negative Sentiment on Twitter: Insights Concerning the 2015-2016 Zika Epidemic

Ravali Mamidi^{1*}, MS; Michele Miller^{2*}, MS; Tanvi Banerjee^{1,3}, PhD; William Romine², PhD; Amit Sheth^{1,3}, PhD

¹Computer Science and Engineering, Wright State University, Dayton, OH, United States

²Department of Biological Sciences, Wright State University, Dayton, OH, United States

³Kno.e.sis, Computer Science and Engineering, Wright State University, Dayton, OH, United States

*these authors contributed equally

Corresponding Author:

Michele Miller, MS

Department of Biological Sciences

Wright State University

3640 Colonel Glenn Hwy.

Dayton, OH, 45435

United States

Phone: 1 5742613969

Fax: 1 9377753320

Email: millerme91@gmail.com

Abstract

Background: To understand the public sentiment regarding the Zika virus, social media can be leveraged to understand how positive, negative, and neutral sentiments are expressed in society. Specifically, understanding the characteristics of negative sentiment could help inform federal disease control agencies' efforts to disseminate relevant information to the public about Zika-related issues.

Objective: The purpose of this study was to analyze the public sentiment concerning Zika using posts on Twitter and determine the qualitative characteristics of positive, negative, and neutral sentiments expressed.

Methods: Machine learning techniques and algorithms were used to analyze the sentiment of tweets concerning Zika. A supervised machine learning classifier was built to classify tweets into 3 sentiment categories: positive, neutral, and negative. Tweets in each category were then examined using a topic-modeling approach to determine the main topics for each category, with focus on the negative category.

Results: A total of 5303 tweets were manually annotated and used to train multiple classifiers. These performed moderately well (F1 score=0.48-0.68) with text-based feature extraction. All 48,734 tweets were then categorized into the sentiment categories. Overall, 10 topics for each sentiment category were identified using topic modeling, with a focus on the negative sentiment category.

Conclusions: Our study demonstrates how sentiment expressed within discussions of epidemics on Twitter can be discovered. This allows public health officials to understand public sentiment regarding an epidemic and enables them to address specific elements of negative sentiment in real time. Our negative sentiment classifier was able to identify tweets concerning Zika with 3 broad themes: *neural defects*, *Zika abnormalities*, and *reports and findings*. These broad themes were based on domain expertise and from topics discussed in journals such as *Morbidity and Mortality Weekly Report* and *Vaccine*. As the majority of topics in the negative sentiment category concerned symptoms, officials should focus on spreading information about prevention and treatment research.

(*JMIR Public Health Surveill* 2019;5(2):e11036) doi:[10.2196/11036](https://doi.org/10.2196/11036)

KEYWORDS

social media; machine learning; natural language processing; epidemiology; Zika; infodemiology; infoveillance; twitter; sentiment analysis

Introduction

Background

Zika was discovered in 1947 in Uganda [1]. From the 1960s to 1980s, only 14 cases were diagnosed across Asia and Africa, and it typically caused mild symptoms [2]. The first large outbreak occurred in 2007, with the virus spreading from Yap across the Pacific with cases reporting mild symptoms. However, cases were likely underreported from 1947 to 2008 because the symptoms were similar to chikungunya and dengue. It was not until this most recent outbreak that Zika was linked to Guillain-Barré syndrome and microcephaly [1]. Owing to the new-found association of Zika and neurological disorders, people started expressing concern with the Zika virus, especially after an article in the British Broadcasting Corporation (BBC) stated that the United States declared the Zika virus scarier than first thought [3].

In our previous exploratory study [4], we collected 1.2 million tweets over a period of 2 months and developed a 2-stage classifier to categorize relevant tweets as concerning 4 disease categories: symptoms, treatment, transmission, and prevention. Tweets in each disease category were then examined using topic modeling to ascertain the top 5 themes for each category. We demonstrated how discussions on Twitter can be discovered to aid public health officials' understanding of societal concerns. Our previous work focused on identifying relevant tweets with little emphasis on public sentiment. Much of the fear around Zika concerns the symptoms it causes [3]. Therefore, in this study, we turn our focus toward an in-depth analysis of the symptoms of Zika and undertake an analysis of specific positive, negative, and neutral sentiments expressed about the Zika virus.

Related Works

Identifying sentiment on a specific topic was pioneered by Chen et al [5,6]. Since then, several studies have looked at sentiment analysis on a variety of topics. Overall, 2 studies focused on personal communication tweets only [7,8]. The study by Daniulaityte et al [7] collected 15,623,869 tweets from May to November 2015 using keywords related to synthetic cannabinoids, marijuana concentrates, marijuana edibles, and cannabis. They found that using personal communication tweets only, compared with all tweets, improved binary sentiment classification (negative and positive) but not multiclass classification (positive, negative, and neutral). A study by Ji et al [8] collected tweets concerning listeria from September 26 to 28 and October 9 to 10 in 2011. They also focused on personal communication tweets only for sentiment classification (negative and not negative) and also found that the classifiers performed well after excluding nonpersonal communication (with a classification of F1 score=0.82-0.88). Instead of focusing on personal communication tweets alone, we included all relevant tweets after the BBC article about scientists declaring

Zika scarier than initially thought [3] in our previous study [4]. A study by Househ collected approximately 26 million tweets and Google News Trends concerning the Ebola virus from September 30 to October 29, 2014 [9]. This study also influenced the decision to use all tweets and not just personal communication when they found that news feeds were the largest Twitter influencers during the Ebola outbreak.

Ghenai and Mejova [10] collected 13,728,215 tweets concerning Zika from January to August 2016. Tweets were annotated as debunking a rumor, supporting a rumor, or neither. They concluded that mainstream news websites may help spread misinformation and fear. A study by Seltzer et al [11] collected 500 images from Instagram from May to August 2016 using the keyword *Zika*. Of those 500 images, only 342 were related to Zika. Of those 342 images, 193 were coded as *health* and 299 were coded as *public interest*. Of the *health* images, the majority related to transmission and prevention, which is similar to what we found in our previous study on Twitter [4]. This shows results can be corroborated across different social media platforms. Seltzer et al also found that many of the images portrayed negative sentiment and fear. Their study was limited to using images and was only concerned with negative sentiment. Our study will use tweets and will include positive, neutral, and negative sentiment.

In many of these studies, the main topical content within each sentiment category was not explored. We take this additional step in our study to determine the topics of public concern regarding the Zika virus. We also used all tweets including personal communication as well as news articles because news articles can go viral and include negative sentiment, as seen with the BBC article briefly described in the background section [3]. The phenomenon of news articles going viral and including negative sentiment is also discussed in our previous study [4].

Purpose of the Research

In this study, public sentiment concerning the Zika virus symptoms was explored to determine important topical subcategories for positive, neutral, and negative tweets. Using the framework shown in Figure 1, 2 main research questions (RQs) were addressed:

RQ1a: Data Annotation Analysis: What was the distribution of positive, neutral, and negative tweets in the gold standard dataset? What was the agreement between the 2 annotators' labels used as the gold standard for the sentiment classification?

RQ1b: Classification Performance: How well can we categorize tweets as positive, neutral, and negative in an automated fashion?

RQ2: Topical Analysis: What were the main topics discussed in the 3 sentiment categories with a focus on the negative sentiment category?

Figure 1. Block diagram of content retrieval using two-stage supervised classification followed by unsupervised analysis for characteristics of sentiment content.

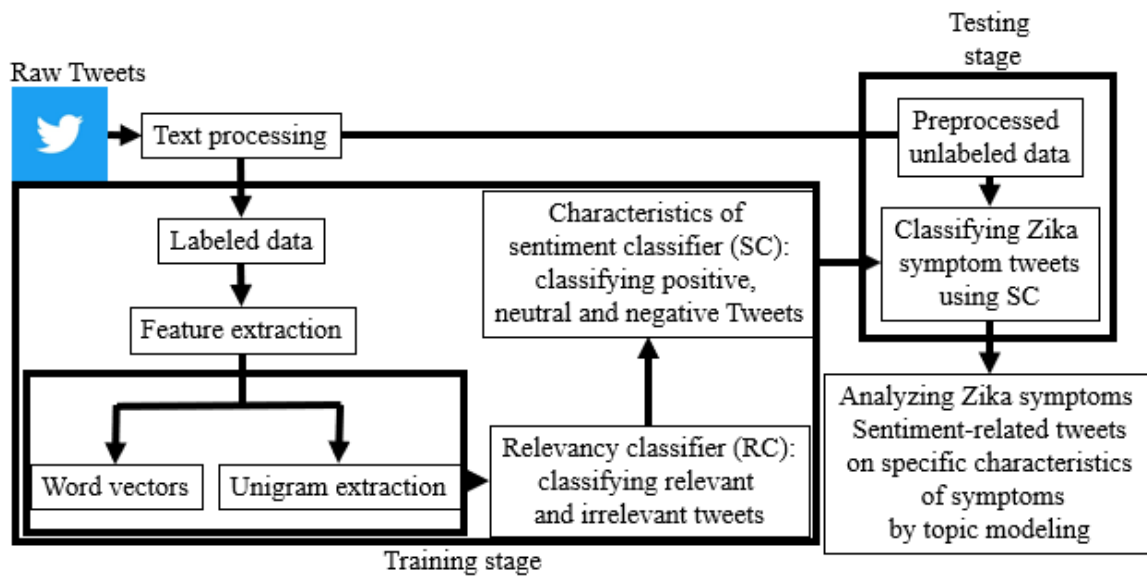
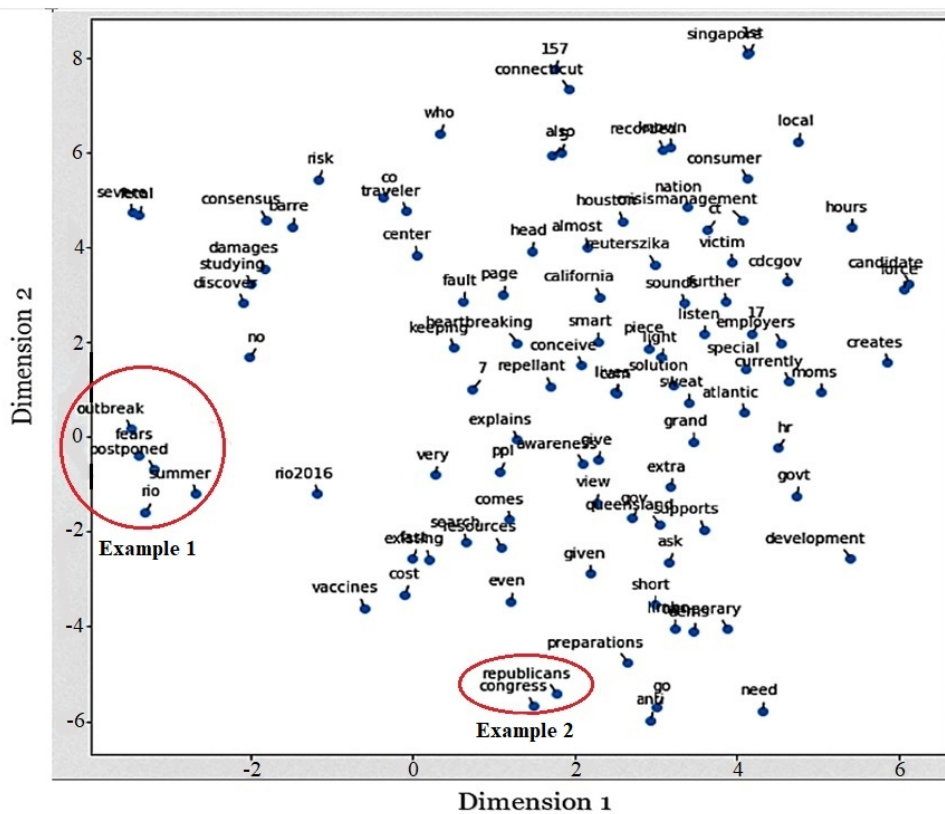


Figure 2. Visualization of Zika word embedding using t-SNE which shows clusters of related word groups within the context of Zika tweets.



Methods

Data Collection

This study utilizes data obtained in a previous study [4] using Twitris 2.0, a semantic Web application that aids comprehension of social perceptions by semantics-based processing of massive volumes of event-centric data on social media [12]. In the previous study, 1.2 million tweets were collected between February 24, 2016, and April 27, 2016, using the keywords *Zika*, *Zika virus*, and *Zika virus treatment* [4]. Before analysis,

tweets were preprocessed by removing non-American Standard Code for Information Interchange (ASCII) characters, capital letters, retweet indicators, numbers, screen handles (@username), punctuation, URLs, whitespaces, single characters such as *p* that do not convey any meaning about topics in the corpus, and stop words such as *and*, *so*, etc. A random sample of 1467 tweets was annotated as relevant versus not by 3 microbiology and immunology experts and used as the relevancy ground truth. All tweets were then classified as relevant or not using the relevancy ground truth and several supervised

classification techniques along with bootstrapping (bagging) techniques. The performance of the classifiers was assessed using tenfold cross-validation with average precision, recall, F1 score, and area under the curve being reported. The multinomial Naive Bayes classifier performed best with an area under the curve of 0.94. Another random sample of 1135 relevant tweets was annotated by the same 3 microbiology and immunology experts to use as the disease characteristics (DC; symptoms, treatment, transmission, and prevention) ground truth. The relevant tweets were then classified into 1 of the 4 DC categories using the DC ground truth and the same supervised classification techniques and performance measures used for relevancy classification. The multinomial Naive Bayes classifier performed best again with areas under the curve ranging from 0.83 to 0.94. This resulted in 48,734 tweets being classified as symptoms, 9937 tweets as treatment, 101,539 tweets as transmission, and 101,456 tweets as prevention [4]. As the Zika symptoms were of public concern, this study focuses on determining the sentiment of those 48,734 tweets collected and classified as discussing Zika symptoms in our previous study.

We have built upon that model described in [4] to explore the sentiments associated with the symptoms category. In this study, we used n-grams-based logistic regression to classify tweets as positive, negative, or neutral. The top themes in each sentiment category were then determined using latent Dirichlet allocation. This allowed us to better explore the themes in each sentiment category so public health officials can address the topics of public concern, such as neurological defects.

To address the RQs, we built the following methodological framework in Figure 1. The 48,734 tweets were preprocessed and labeled as positive, negative, and neutral. Features were then extracted using word embeddings and n-grams. A 2-staged classifier was built using the extracted features to identify the relevant tweets and then categorize them into the 3 sentiment categories. Preprocessed unlabeled tweets in each sentiment category were then analyzed using topic modeling techniques to find the top 10 topics for each of the 3 sentiment categories. This process is useful for discovering public sentiment regarding disease outbreaks and addressing apprehensions in real time.

Data Annotation Analysis (Addressing RQ1a)

A total of 5303 random tweets selected from a total of 48,734 tweets were annotated as positive, neutral, or negative by 2 annotators with domain knowledge related to Zika epidemics. A tweet was considered positive if it mentioned research discoveries related to Zika, as seen in this tweet: “#Zika structure discovered, raising hopes for new ways to combat virus” or reflected a positive attitude toward treatments, preventions, or funding for Zika as seen in this tweet “#Bayer scientists aiding in fight against #Zika virus.” A tweet was considered negative if it discussed the defects/disorders caused by Zika such as “CDC confirms Zika virus causes severe birth defects #business”, discusses the spread of Zika as seen in this tweet “#news Zika virus may spread to Europe in coming months, WHO warns #til_now #Reuters.” Tweets were considered neutral if they gave information with no emotionally charged wording such as *hope*, *combat*, and *severe* or the overall

sentiment of the tweet was neutral. Examples of neutral tweets are “Zika symptoms, diagnosis and treatment, from the CDC #ZikaVirus” and “WHO: #Zika situation report, March 31.” Agreement was found using the Cohen's kappa, which is a robust statistic useful for either interrater or intrarater reliability testing and accounts for the possibility of guessing [13]. These tweets became known as the gold standard dataset once significant agreement was reached (Kappa >.81) [13].

Preprocessing

Before data analysis could begin, tweets had to be preprocessed by removing screen handles (@username), URLs, non-ASCII characters, and retweet indicators. Tweets were then further processed by removing single letters such as *a*, *e*, and *i*; extra spaces; and stop words. Stop words are the most commonly used words in the English language such as *and*, *in*, and *for*. This preprocessed tweet corpus was used for extracting features using the word embeddings and n-grams. These features were extracted similarly to our earlier studies [4,14].

Word Embedding (Feature Extraction)

Machine learning algorithms are incapable of handling raw text or strings and require numeric data to extract knowledge from textual data and build applications. Word embedding is a technique that maps individual words to a predefined vector space in such a way that the semantic relation between words is preserved [15].

In addition, words or phrases from the tweets were embedded into the n-dimensional space where n is the number of words in the corpus. After word embedding, a sentence can be considered as a sequence of points that are grouped according to a semantic criterion so that 2 similar words are close to each other. It captures the context of words, while reducing the number of features in the data. To provide a better understanding of word embedding, we provide an example from a sample of our dataset. For visualizing the high-dimensional data, we used a technique called t-distributed stochastic neighbor embedding, which maps each data point to a lower dimensional space (of size 2) [16]. From Figure 2, we see the spatial distribution of a random sample of 100-word embeddings generated from the Word2vec model [17]. This figure is based on a subset of random tweets and is included purely to show how words used in the same context are close to each other in the vector space. We see that words that are similar eventually come spatially closer in the vector space. For example, words such as *outbreak*, *fears*, *postponed*, and *summer* (example 1) are spatially close because they are used in the same context in the case of the Rio Olympics and words such as *republicans* and *congress* (example 2) are spatially close together as they are used in the context of Zika funding. The word embedding algorithm was used to generate features to help classify tweets as positive, negative, or neutral.

Models

We used 2 different main models for classification. One was Word2vec [18] and the other was an n-gram model [14].

***n*-Gram Model**

In this model, features were extracted from tweets using the Stanford Natural Language Processing Part of Speech tagger [19] and *n*-grams [20], where an *n*-gram represents a sequence of words treated as a single entity or feature. Initially, features were identified from the tweets and the count for each feature was determined. Only the top 20 unigrams and bigrams were used for classification because the corpus was large, and we only wanted to capture the most frequently used text features. In total, there were 61 features. Examples include *AT_Mention*, *Zika*, *Discourse marker*, *microcephaly*, *fetal*, *Pronoun*, *health*, *birthdefects*, *Zika infection*, *Hashtag*, and *brain damage*.

Word2vec Model

Word2vec comprises 2 different methods: continuous bag of words (CBOW) and skip-gram [21]. In the CBOW method, the goal is to predict a word given the surrounding words, that is, the words before and after it [21]. Skip-gram is the opposite: we want to predict surrounding words given a single word [21]. The skip-gram method with negative sampling works best with the medium- or large-sized datasets [15]. As our dataset was considered medium sized [15], we used the skip-gram model with a negative sampling rate of 10.

For the word embeddings, we used the Gensim library version 2.2.0 of Python version 3.5.4 [22] for converting all the words to an *n*-dimensional space before training the classifiers. The tokenized words were then fed to the Word2vec tool and trained with the skip-gram model. We considered a window size of 4 because the average length of the tweets was less than 10 words, which means 4 tokens apart from the target words are considered as adjacent words.

With these collective parameters, we generated the word vectors of size 300 and tested the learned vectors using the similarity functionality of the Word2vec. To evaluate the vectors generated using the tool, we selected 2 words *dengue* and *Zika*, which are mosquito-borne diseases, to assess similarity. Similarity is used to find the distance between 2 vectors. The closer the similarity is to 1, the more closely related the words are [23]. The similarity was 0.92, which indicates the words are closely related or used in a similar context. When words like *microcephaly* and *pregnant* were used, it gave related words such as *woman*, *women*, and *infected*, among others.

Vector operations such as sum and mean were used to build the final feature vector. The following are the operations performed on the word vectors:

Sum of Word Embeddings: This is the sum of all word vectors in the tweet. $FV_{Sum} = \sum W$

Mean of word Embeddings: Average of all the word vectors in the tweet. $FV_{Mean} = 1/n \sum W$

W represents a single word in a tweet and FV_{Sum} and FV_{Mean} represent the feature vector of the tweets.

Classification Performance (Addressing RQ1b)

Supervised classification algorithms, including logistic regression, support vector machines with radial basis function kernel, and random forest, were used for classifying the tweets

into the 3 sentiment categories. These methods rely on labeled data, in this case, the 5303 randomly selected tweets that were annotated as positive, neutral, or negative by the 2 annotators from a total of 48,734 tweets. These classifiers were trained to categorize tweets into the specified categories based on the gold standard derived by the annotators.

The performance of each classifier was assessed using the stratified *k*-fold cross-validation as we had an unbalanced dataset. We report *k*=7 because there was no improvement in the result with increase of *k* and also it saves computation time. The stratified *k*-fold maintains equal number of samples for each annotator-labeled class [24]. In this method, 1 subsample (fold) of tweets was used for a testing set and the remaining 6 for training. This was repeated 7 times, with each subsample being used as the testing subsample once [24]. This study reports average recall (indication of category tweets not missed by the classifier), precision (correctly categorized tweets), and F1 scores (weighted average of precision and recall) as measures of classification performance for each classifier.

Topical Analysis (Addressing RQ2)

Previous studies, such as the one by Lau, Collier, and Baldwin [25], have shown the usefulness of LDA for grouping text into themes in short text documents such as tweets. In this study, we used LDA topic modeling to identify the underlying topics discussed within each of the sentiment categories. In LDA, documents (tweets in this case) are represented as random mixtures over hidden topics, where each topic is characterized by a distribution over words that occur most frequently within that topic [26]. More specifically, LDA is a 3-level hierarchical Bayesian model, in which each word in a corpus is modeled as a finite mixture over an underlying set of topics. Each topic is then modeled as an infinite mixture over an underlying set of topic probabilities. The top words belonging to each topic are given as an output, and it is up to the researcher to interpret the topic's meaning. This aids better qualitative exploration of the subtopics in each of the 3 categories.

To determine the number of topics required for topic modeling, we used perplexity, a measure used to evaluate topic models generated by LDA where the smaller the perplexity score, the better the generalization performance [22,26]. We used this measure to evaluate the topic modeling results by testing a range of 2 to 100 topic models for the 3 sentiment categories. For calculating the perplexity measure, preprocessed tweets were used. Words that occurred only once or twice in the corpus were removed as they increase the number of topics but will not give generalizable information [26].

Results

In this section, the distribution of tweets in the gold standard dataset is discussed. The performance of 3 different classifiers using the Word2vec and *n*-gram models is also explained. Finally, the topic modeling results for the positive, neutral, and negative categories is explored with a focus on the themes that emerged in the negative sentiment category.

Data Annotation Analysis (Addressing RQ1a)

To train the classifiers, the gold standard dataset had to be created as described in the methods section above. The kappa value for the level of agreement between the 2 annotators was 0.95, indicating near-perfect agreement [13]. The distribution of the tweets in the gold standard dataset is shown in Figure 3. The majority of tweets displayed negative sentiment (2423; 46% of the total tweets) and the fewest displayed positive sentiment (1010; 19%). As can be seen in Figure 3, there is high class imbalance in the 3 sentiment categories.

Classification Performance (Addressing RQ1b)

Table 1 provides the performance of the 2 text-processing models and the corresponding classifiers. The n-gram model

performed slightly better than the word-embedding model. For this dataset, classifiers performed reasonably well, with F1 scores ranging from 0.48 to 0.68. However, the logistic regression classifier used with the n-gram model performed the best with an F1 score of 0.68. This performance is comparable with that in similar studies [7,18].

Using the n-gram-based logistic regression sentiment classifier, we categorized all 48,734 tweets obtained from our previous study (Figure 4) [4]. The total number of negative tweets was almost 4 times larger than the positive and neutral categories combined. We can clearly see from Figure 4 that this is a highly unbalanced dataset, with the majority of tweets belonging to the negative sentiment category.

Figure 3. Distribution of tweets in three sentiment categories.

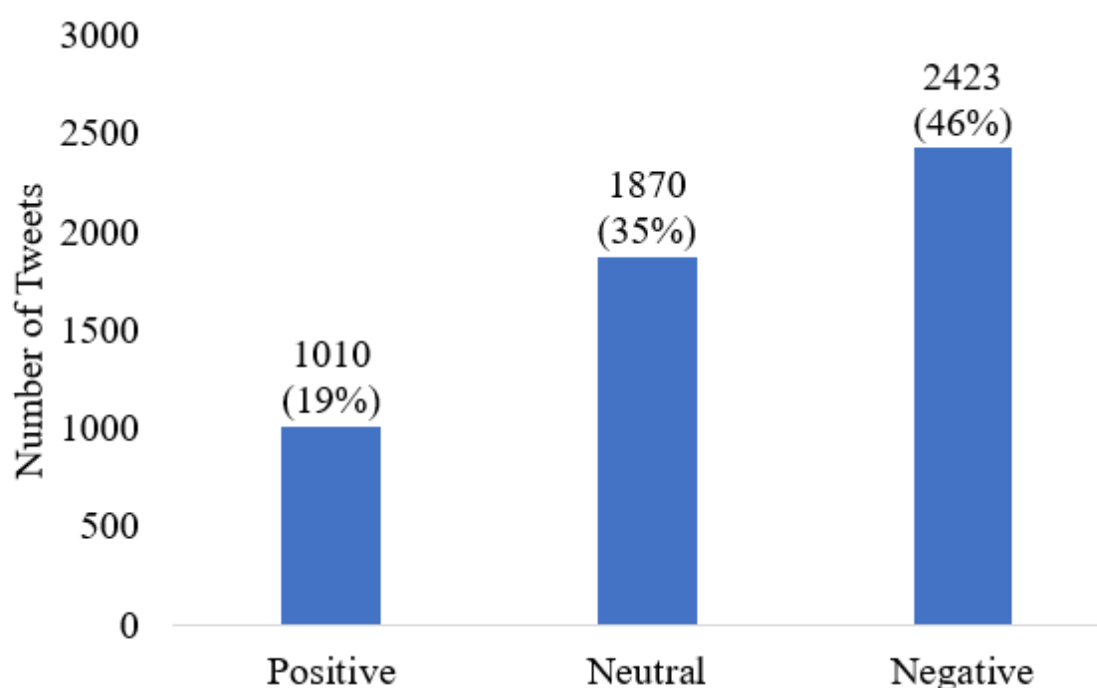
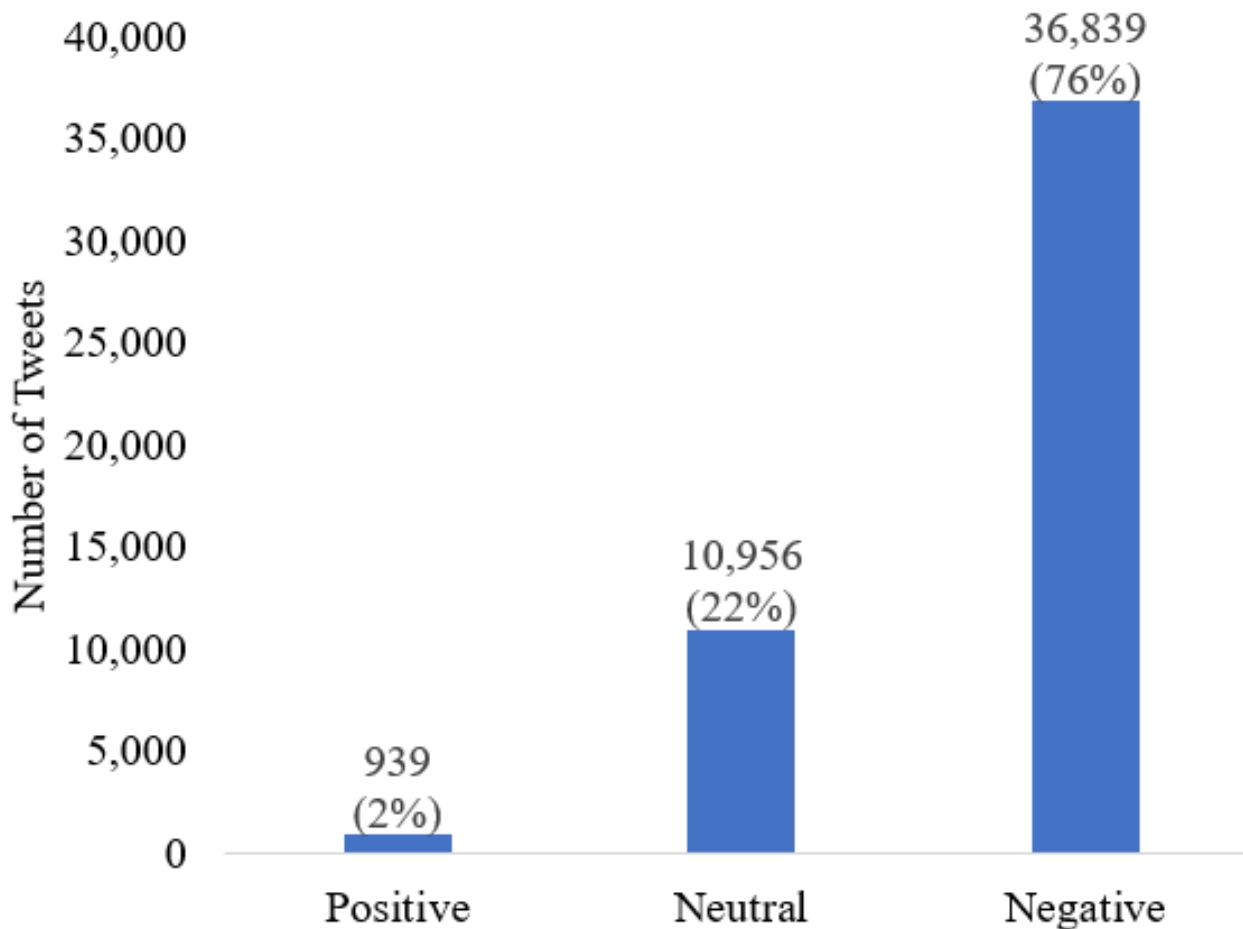


Table 1. Classifier performance for sentiment analysis using sevenfold cross-validation. The classifiers used are logistic, support vector machine, and random forest.

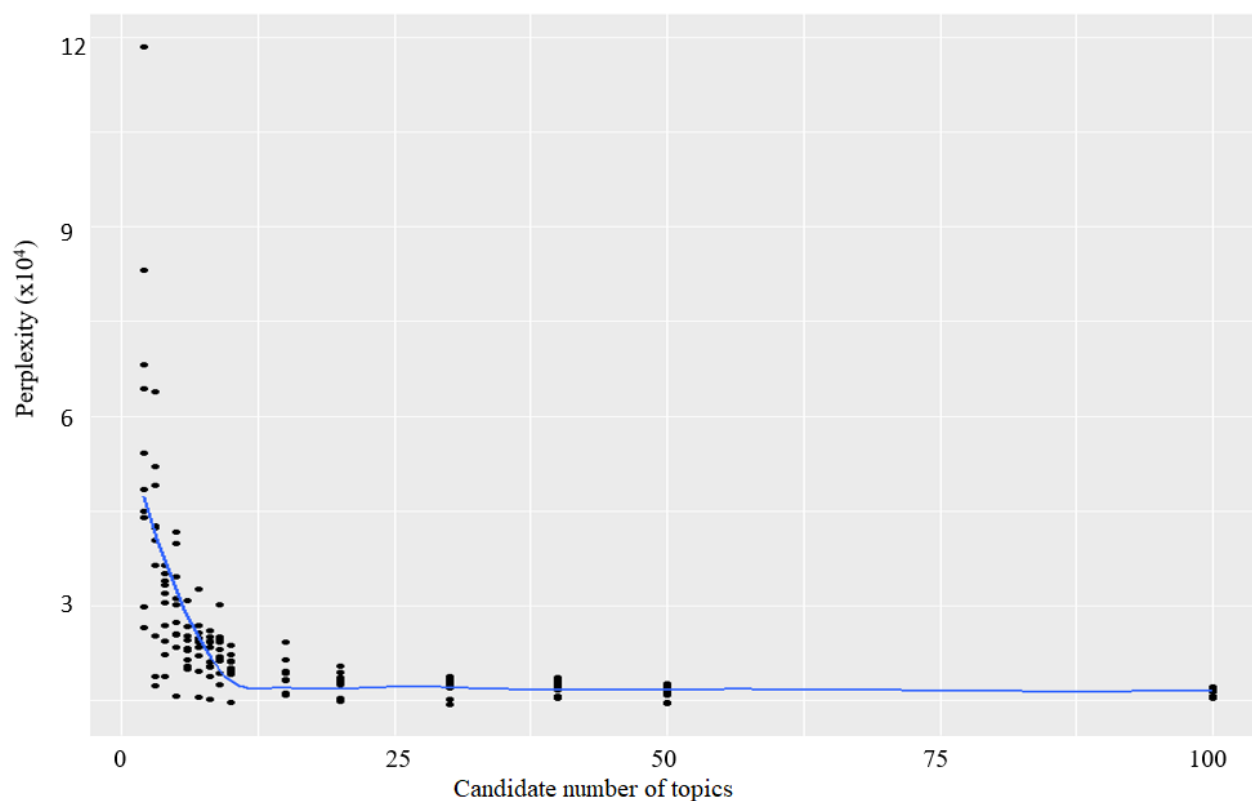
Classifier	Precision	Recall	F1 score
Word2vec FV_{Sum} model			
Logistic regression	.68	.66	0.66
Support vector machines	.67	.65	0.65
Random forest	.55	.53	0.48
Word2vec FV_{Mean} model			
Logistic regression	.63	.63	0.63
Support vector machines	.66	.65	0.65
Random forest	.50	.50	0.50
n-gram model			
Logistic regression	.69	.68	0.68
Support vector machines	.65	.65	0.65
Random forest	.68	.67	0.67

Figure 4. Number of tweets in three categories of the symptoms dataset (obtained from the n-gram based sentiment classifier).

Topical Analysis (Addressing RQ2)

Within the tweets with negative sentiment, the perplexity decreased rapidly until about 10 topics and then leveled off (Figure 5). The perplexity graph for the positive and neutral category are available online [27]. This indicates that increasing the number of topics after 10 will not significantly improve the generalizability of the LDA models [26]. Therefore, 10 topics per sentiment were extracted.

The results of the LDA are discussed below for the positive, neutral, and negative categories. Themes and topics for all 3 sentiment categories were determined by an epidemiology expert based on the words given for each theme and some sample tweets containing those words. First, the topics for the positive and neutral categories will be briefly discussed. The tables, including the theme names, topic words, and example tweets for the positive and neutral topic models, are available online [27]. Then, a more detailed explanation of the negative sentiment topics will be presented.

Figure 5. Perplexity plot measures for the 7-fold cross-validation of topic modeling for the negative sentiment category.

Topics From Positive and Neutral Sentiment

Within the positive sentiment themes, there were 4 broad qualitative topics within the 10 topics chosen using the perplexity measure with LDA: mosquito-killing methods, models to help understand the Zika virus, detection of the Zika virus in cells, and treatment and prevention discoveries (Table 2). These broader themes were labeled based on domain expertise and from journals such as Vaccine and MMWR, allowing further categorization of the 10 topics. For the broader theme of models that help understand the Zika virus, topic #1 contained tweets concerning a new model researchers were developing to study Zika pathogenesis and topic #2 described 3-dimensional (3D)-printed minibrains used for understanding the Zika virus. For the mosquito-killing methods theme, topic #4 contained tweets concerning sweat-emitting Brazilian billboards killing the Zika-carrying mosquitoes and topic #10 addressed other ways of killing Zika-carrying mosquitoes. In the treatment and prevention discoveries theme, topic #3

comprised tweets regarding the discovery of how Zika stunts the development of a fetus, topic #5 characterized the development of vaccines to treat Zika, and topic #8 reported about the IBM magic bullet to destroy all killer viruses. This *magic bullet* is actually a macromolecule that will attach to the surface of any virus and prevent it from attaching to a human cell [28]. If the virus cannot attach and enter a cell, infection is prevented. The macromolecule is also basic, neutralizing the acidity of an infected cell in case the virus is already infecting human cells by the time the *magic bullet* is used [28]. In the broader theme of detection of the Zika virus in cells, topic #6 regarded different types of tests for identifying Zika infection, topic #7 outlined the detection of Zika using fetal tissue, and topic #9 detailed the detection of Zika accumulations in the brain.

Table 2 Positive sentiment topic modeling results grouped together based on the broader themes. The numbers reflect the relative size of the theme. For example, the topic mouse model had more tweets than 3D-printed minibrains.

Table 2. Positive sentiment topic modeling results grouped together based on the broader themes. The numbers reflect the relative size of the theme. For example, the topic mouse model had more tweets than 3D-printed minibrains.

Topic	Words	Tweet
Model broader theme		
#1 Mouse model	Researcher, mouse, model, develop, health, and research	new #zika <i>mouse model researchers develop</i> another <i>mouse >model</i> of zika infection that mimics the disease in humans
#2 3-dimensional–printed Minibrains	Scientist, test, brain, mystery, and help	mini 3d printed <i>brains help scientists</i> understand zika virus
Mosquito broader theme		
#4 Brazilian billboards	Rapid, billboard, emit, Brazilian, and structure	sweat- <i>emitting Brazilian billboards</i> lure zika-carrying mosquitoes to their death mnn—mother nature network
#10 Killing mosquitoes	Mosquito, infect, kill, insight, and biomolecular	researchers develop #algae to <i>kill #mosquitoes</i> carrying viruses like #zika
Virus discovery broader theme		
#3 Fetal brain development	Fetus, human, discover, and help	how zika virus stunts <i>foetal</i> brain development researchers have <i>discovered</i> how hijacking a <i>human</i> immune mole...
#5 Vaccines	Model, infect, vaccine, provide, and develop	mouse <i>models</i> of zika virus <i>infection</i> in pregnancy <i>provide</i> basis to <i>develop</i> <i>vaccines</i> , treatments
#8 IBM magic bullet	Kill, develop, and understand	IBM research IBM announces magic bullet to zap all kinds of <i>killer</i> viruses, like #zika by seancaptain
Detection broader theme		
#6 Zika tests	Urine, discover, pattern, Jamaica, programmable, and molecular	#salingfollow interim cdc guidance finds <i>urine</i> specimen better than serum for rapid and specific zika testing—cdc
#7 Fetal tissue research	Fetal, tissue, infect, detect, equip, and test	last month, <i>fetal tissue</i> research helped doctors' understand how the zika virus <i>infects fetus</i> & how to <i>detect</i> its presence much
#9 Zika accumulation	Reveal, accumulate, Zika, virus, examine, pregnancy, and report	one of the first mouse models of #zika <i>reveals</i> the <i>virus accumulates</i> in the brain

Overall, the broader themes in Table 3 (model, mosquito, virus discovery, and detection) were present in the positive sentiment category because they all have to do with helping prevent transmission or research that could lead to treatments. Both of these topics reflect positive public perception because they help prevent the defects that have become associated with Zika. For example, tweets in the mosquito theme discussed ways to kill mosquitoes, which would help prevent the spread of Zika [29]. Tweets in the model and viral discovery themes addressed discoveries that could help lead to treatments, such as the IBM magic bullet [28]. Virus discovery tweets were positive because they pointed to faster ways to detect Zika. Knowing where Zika accumulates would help with developing treatments [30]. Tweets in the positive category also used words with positive connotations such as *understand*, *develop*, *hope*, *discover*, *benefit*, and *reveal*, among others. While themes in the positive sentiment category mainly addressed research to treat Zika and

prevention methods, themes in the neutral category mostly comprised posts from news agencies stating facts.

Within the neutral sentiment topics, there were 3 broader qualitative themes: public health messages, knowledge gaps, and Zika characteristics (Table 3). In the public health messages, topic #1 explained how scientists were trying to unravel the Zika mystery, topic #2 cautioned about the dangers of Zika infection to pregnant mothers, topic #3 declared that Zika is a mosquito-borne disease, topic #4 specified the laws regarding birth control and abortion, topic #5 discussed fighting the mosquitoes, and topic #6 regarded the officials warning the public to be careful not to be bitten at work. Knowledge gaps consisted of topic #7, which discussed knowledge gaps concerning the Zika virus. In the Zika characteristics theme, topic #8 affirmed Zika symptoms, topic #9 included comparisons between dengue and Zika, and topic #10 described fetal brain damage from Zika infection.

Table 3. Neutral sentiment topic modeling results grouped together based on the broader themes. The numbers reflect the relative size of the theme.

Topic	Words	Tweet
Public health messages broader theme		
#1 Zika mystery	Brazil, common, unravel, question, important, disease, and issue	#voanews <i>brazil</i> scientists seek to <i>unravel</i> mystery of zika twins scientists struggling to unravel t...
#2 Aedes mosquito	Mosquito, infect, pregnancy, outbreak, women, and child	the zika virus and the dengue <i>mosquito</i> have a common nature. very resistant ones, and very dangerous too. <i>infects</i> mothers with <i>pregnancy</i> !
#3 Mosquito-borne illness	Symptom, today, health, born, mosquito, and effect	zika is a <i>mosquito borne</i> illness that does not present <i>symptoms</i> in many people. that is a very dangerous thing.
#4 Abortion	Abortion, learn, worse, survive, guideline, and paper	zika virus, birth control and <i>abortion</i> our anti-woman laws will make this <i>worse</i> .
#5 Fight the bite	Infect, fight, bite, affect, and death	only 1 in 4 people <i>infected</i> w/ #zika will show symptoms. <i>fight the bite</i> , destroy mosquito breeding sites #nobitenozika
#6 Officials' warning	Officials, control, disease, center, and researcher	health <i>officials</i> warn against exposure to zika at work the <i>centers for disease control</i> and prevention #atlanta
Knowledge gap broader theme		
#7 Knowledge gap	People, expert, relate, Ebola, and cure	various ' <i>experts</i> ' need to get up to speed on the zika+ front now. time is of an essence. many <i>people</i> are 'behind the curve'.
Zika characteristics broader theme		
#8 Symptoms	Fever, scarier, infect, eye, and first	zika symptoms— <i>fever</i> , rash, joint pain, and/or red <i>eyes</i> . most people <i>infected</i> typically don't have symptoms though.
#9 Dengue	Dengue, flu, rash, compare, cause, and malaria	<i>dengue</i> & zika have a <i>rash</i> , fever etc. 4 dengue strains increasing in ja. docs need to be careful #testedorsuspected
#10 Fetal brain damage	Fetus, information, prevent, symptom, damage, and fetal	"why <i>fetal</i> tissue research is crucial to saving babies from zika new study uncovers ' <i>alarming</i> ' <i>information</i> ..."

In this case, the broader themes in Table 3 (public health messages, knowledge gaps, and Zika characteristics) highlight the neutral sentiments because the tweets in these themes were from public health experts and news agencies informing the public and thus are more likely to state facts than opinions. For example, the tweet "Officials: Zika-Infected Couples Should Postpone Pregnancy" is a statement from officials about postponing pregnancy during a Zika outbreak to help prevent babies born with birth defects. Some tweets were neutral even though they contained words with both positive and negative connotations because the sentiment of the tweet overall is neutral, such as this tweet "#voanews brazil scientists seek to unravel mystery of zika twins scientists struggling to unravel t..." Topics 1 through 6 all contained messages from public health agencies and were therefore labeled as public health messages. Topics 8 through 10 concerned characteristics of the Zika virus and thus were grouped together. Topic 7 did not belong in either category and was therefore made a separate theme. In summary, the neutral topics contained tweets from news agencies and public health officials. The negative sentiment topics also contained some tweets from news agencies and public health officials but additionally contained opinion tweets from the public.

Topics from the Negative Sentiment

Before data analysis, we had chosen to focus on the topics from the negative sentiment category specifically in the symptoms category from our previous paper [4] as that was found to be critical for public health officials [31-35]. We chose to focus on negative sentiment tweets as this is what health officials will

be most concerned with as there is greater need for intervention and information dissemination in these topics [31-35]. For example, a study by Glowacki et al [34] found that the Centers for Disease Control and Prevention (CDC) and the public expressed concerns about the spread of the Zika virus and that the CDC also focused on symptoms and education during a 1-hour live chat between the CDC and the public. Intense media focus on a topic, similar to the media focus during the Zika epidemic, causes concern among the general public [31]. Therefore, physicians and public health officials must address these concerns before they become entrenched in public discourse. The failure to act to the 2015 Ebola outbreak by the World Health Organization (WHO) and Centers for Disease Control (CDC) cost thousands of lives [32]. To prevent a similar failure, an intermediate-level response was needed to prevent overreaction while still taking adequate measures to respond to the Zika outbreak [32]. For example, during the Ebola outbreak, it was found that failure to engage communities had detrimental effects, whereas engaging communities helped curtail the outbreak [33]. The main ways to engage a community included involving family members in the care of loved ones in ways that did not put them at risk, tailoring global policies to local settings, using varied methods of communication, organizing regular meetings with the community, and identifying female and male community leaders to spread key messages. This is why public health officials in the CDC had the live chat with the public and posted information on social media as they gained new information concerning the Zika virus. The nature of the new symptoms associated with Zika could have encouraged fear and anxiety among the public [35]. Therefore, public health

officials need to continue to disseminate preventative methods and information on how to address symptoms to help mitigate the panic. In addition, this was the category with the majority of the tweets (Figure 4). By understanding what is of concern to the public, officials can focus on targeting their messages to addressing these concerns. A methodology that seems to be effective based on our previous study [4] and the current LDA results is creating catchy phrases such as “Fight the Bite” or using phrases that elicit emotion such as the BBC article stating “Zika is scarier than initially thought.” Public health officials can focus on creating similar phrases to address all the topics of negative concern. The topic model results for negative sentiment are shown in Table 4. In the negative sentiment topics, there were 3 broader topics: *neural defects caused by Zika infection, abnormalities because of Zika infection, and reports and findings concerning the Zika virus.* Topics #1 *brain defects*, #2 *neurological effects*, #5 *fetal effects*, and #8 *Guillain-Barré syndrome* all concerned the nervous system. Topics #6 *Zika abnormalities* and #9 *Zika effects* were both related to abnormalities resulting from Zika infection. Topics #3 *initial reports*, #4 *Zika impact*, #7 *ultrasounds*, and #10 *dengue association* all concerned reports and findings concerning the Zika virus. There was significant overlap between topics #3 and

#4 because they both addressed reports and findings concerning the Zika virus. However, topic #3 *initial reports* included tweets stating the locations where Zika is spreading, whereas topic #4 *Zika impact* included tweets concerning the BBC article that describes Zika as scarier than initially thought [3].

The broader themes in Table 4 (neural defects, Zika abnormalities, and reports and findings) were all negative because they addressed topics of concern for the general public. Before this outbreak, Zika was considered a mild illness with only 14 reported cases [2]. It was not until this most recent outbreak that Zika became associated with microcephaly, Guillain-Barré syndrome, and congenital Zika syndrome, all of which caused fear and concern across the globe [1,4,36,37].

Table 5 shows the percentage distribution of tweets belonging to each theme of the negative sentiment category. The tweets were evenly distributed across the topics, with the exception of topic #10 (dengue association). This is because people discussing this association are most likely epidemiologists and others in the public health field that understand antibodies, as seen in this tweet, “lab findings hint that #dengue antibodies intensify #zika infection=>leading to #microcephaly & gbs^a? Evidence.”

Table 4. Negative sentiment topic modeling results grouped together based on the broader themes. The numbers reflect the relative size of the theme.

Topic	Words	Tweet
Neural defects broader theme		
#1 Brain defects	Brain, microcephaly, baby, disorder, confirm, and cause	#zika virus confirmed zika causes brain damage in babies born with microcephaly brain abnormalities in babies
#2 Neurological effects	Severe, problem, immune, neural, death, and birth	human neural stem cells infected by #zika subsequently trigger an innate immune response that leads to cell death
#5 Fetal effects	Brazil, fetus, shrink, development, disrupt, outbreak, and pregnancy	in #brazil zika eats away at fetal brain, shrinks or destroys lobes controlling thought & prevents development.
#8 Guillain-Barré syndrome	Syndrome, rare, case, associate, cause, and microcephaly	cases of rare nervous disorder guillain-barre syndrome may increase if zika spreads via
Zika abnormalities broader theme		
#6 Zika abnormalities	Brain, eye, abnormality, scientific, consensus, confirm, and relate	a9 zika associated complications for pregnancy include miscarriage, stillbirth, brain abnormalities and eye abnormalities. #reuterszika
#9 Zika repercussion	Zikavirus, infect, child, adult, and fetal	researchers says that zika virus infection can stunt growth of children
Reports and Findings broader theme		
#3 Initial reports	Report, puerto rico, infect, link, and defect	puerto rico reports first zika-linked birth defect {3.1} puerto rico reports first zika-linked birth defect
#4 Zika impact	impact, spread, reuters, mosquito, and scarier	#reuters zika spread, impact 'scarier than we initially thought' u.s. health official
#7 Ultrasounds	ultrasound, doctor, baby, unborn, and infect	#chevy car ultra sounds missed zika infection until the one showing serious harm to her baby
#10 Dengue association	Expert, warn, sound, dengue, causal, fetus, spread, and microcephaly	lab findings hint that #dengue antibodies intensify #zika infection=>leading to #microcephaly & gbs ^a ? evidence

^aGBS: Guillain-Barré syndrome.

Table 5. Percent distribution of tweets belonging to the ten themes in the negative sentiment category.

Theme	Distribution of tweets, %
Brain defects	12
Neurological effects	12
Initial reports	11
Zika impact	11
Fetal effects	11
Zika abnormalities	10
Ultrasounds	10
Guillain-Barré syndrome	9
Zika repercussion	9
Dengue association	5

Discussion

In the discussion section, we will address one cause of tweets being misclassified with some examples. The 3 negative sentiment broader themes, *neural defects*, *Zika abnormalities*, and *reports and findings*, will then be explored and discussed in more detail.

Classification Analysis

As seen in [Table 2](#), classification is not 100% accurate, implying that some tweets were misclassified. We will focus on the negative tweets as those were the focus of our discussion. Some tweets were misclassified because of words such as *active*, *saliva*, *feds*, *busted*, *beast*, and *prenatal*, which were not seen by the model because the count of these words is less than the minimum count (set to 5) parameter given in the Word2vec model and hence were discarded. The minimum count was set to 5 (the default setting in Gensim) as words used fewer than 5 times do not add significant information to the analysis [38]. Adding more training data could improve these results; however, a study by Nakov et al annotated 6000 tweets and had similar F1 scores to our study [39]. As these words occurred fewer than 5 times, the algorithm was not able to identify these tweets as negative as it was not able to determine the words closer to these words. Examples of tweets that were incorrectly identified as negative are “#3tking Zika virus makes Rio Olympics a threat in #Brazil and abroad, #health expert says” and “#ap breaking cdc no longer any doubt that zika virus causes birth defects.” Examples of tweets incorrectly identified as positive are “#seattle major zika fail! feds busted for lazy response ...” and “@DrFriedenCDC Scary how you could substitute prenatal alcohol in place of Zika!Same symptoms,hidden—YetCDC quiet.”

Topic Model

In this section, we focus on the negative sentiment topics of neural defects, Zika abnormalities, ultrasounds, and dengue association. These themes and topics were chosen for discussion because they were topics of public concern, have been addressed by the CDC or WHO [36,40-44], and can be addressed by officials to help mitigate the concern. Zika impact was not addressed because it is the focus of our previous work [4]. Initial

reports were not addressed as it is specific to this outbreak and officials and the public cannot wholly prevent the spread of the Zika virus.

Neural Defects

Neural defects is a broader theme of concern for the public that needs to be addressed by public health officials to mitigate fear and concern because of the defects to the nervous system caused by Zika virus infection. For [Table 4](#), topics #1 (*brain defects*), #2 (*neurological effects*), #5 (*fetal effects*), and #8 (*Guillain-Barré syndrome*) all concern the neural system. For example, topic #1, *brain defects*, points to brain damage in babies because of microcephaly as seen in this tweet “scans show extent of brain damage in babies with microcephaly associated with zika...” Microcephaly has been a topic of concern for the CDC as babies born with microcephaly will require assistance throughout their lifetime [40,45]. The topic *neurological effects* (#2) includes tweets discussing the death of neural stem cells, which leads to neurological disorders in humans [46], as seen in this tweet, “zika virus targets human cortical neural progenitors causing cell death & attenuated neural cell growth.” The topic *fetal effects* (#5) also addresses brain shrinking or brain damage but additionally the tweets discuss the destruction of the brain lobes that control thought, vision, and other functions in fetuses as seen in this tweet, “scans & autopsies show that zika eats away at the fetal brain. it shrinks or destroys lobes that control thought, vision & other functions.” *Guillain-Barré syndrome* (topic #8) is a sickness caused by damage to nerve cells. The tweet “human neural stem cells infected by #zika subsequently trigger an innate immune response that leads to cell death” includes information on how Zika can lead to damage of neural stem cells and causes a disease such as Guillain-Barré syndrome [47]. The reader can see how topics #1, #2, #5, and #8 all include information on neural issues following Zika infection but all focus on different issues and are, therefore, 3 separate topics. By looking at these tweets, public health officials can see the public is concerned about the neurological defects caused by Zika. Therefore, the next steps officials need to take is to focus on how to prevent mosquito bites, especially when pregnant, to prevent these neurological defects. The “Fight the Bite” campaign is an example of such an effort [44].

Zika Abnormalities

Zika abnormalities is also an important broader theme to address because of the fear and concern of abnormalities and defects in infants because of Zika virus infection during pregnancy. In Table 4, the topics #6 (*Zika abnormalities*) and #9 (*Zika effects*) are both related to abnormalities because of Zika infection but include diverse problems. The topic *Zika abnormalities* (#6) describes various anomalies associated with the fetus and babies born with Zika infection as seen in this tweet, “birth defects linked to #zika now also incl hearing loss, vision problems, impaired growth, abnormalities in limbs.” These types of abnormalities are termed as congenital Zika syndrome by the CDC and includes a collapsed skull, eye scarring, severe muscle tension, and brain calcification [36,37]. The topic *Zika effects* (#9) focuses on the stunt in growth and development of children. Again, both of these topics concern abnormalities because of Zika infection but focus on 2 different abnormalities and are therefore kept as 2 distinct topics. By pushing prevention such as the “Fight the Bite” campaign, officials can help ease fears concerning these abnormalities.

Ultrasounds

Ultrasounds is another important topic to address because initial ultrasounds fail to reveal microcephaly and other birth defects, leading to a false sense of security for a couple [41,42,48,49]. As previously stated, Zika is linked to microcephaly; however, ultrasounds were found to have high false-negative predictions regarding the presence of microcephaly during the first and second trimesters of a woman’s gestational period [48]. Therefore, the topic of *ultrasounds* is important to discuss because pregnant women may have a false sense of security after getting an ultrasound and Zika not being detected in their fetus in the early stages of pregnancy. The CDC states on their website that microcephaly is more readily detected late in the second trimester to early in the third trimester [41]. Researchers are also recommending that parents have a magnetic resonance imaging (MRI) procedure on their newborn’s head performed because some abnormalities are not apparent at birth but may be detected in an MRI [42]. To address the concern of detecting microcephaly before a baby is born, officials need to keep providing up-to-date information on ways to detect microcephaly and to keep striving to improve detection methods to help the public make informed decisions regarding their fetus.

Dengue Association

Dengue association may explain why this Zika outbreak is associated with abnormalities and defects and previous infections were not, which is why it is an important topic to address [43,50-52]. Dengue is in the same family of viruses as Zika and is also spread by the same 2 mosquitoes as Zika [43]. If a person has been previously infected with 1 strain of dengue and then later gets infected with a different strain, they are at risk of developing severe dengue symptoms because of antibody-dependent enhancement (ADE) [50]. In the topic *dengue associations* (#10), scientists suspected and are starting to confirm that earlier illness of dengue enhances the chances of Zika infection also because of ADE [51,52]. The fact that this is in the negative sentiment category shows that the public is concerned with dengue interacting with Zika, which informs

public health officials that their messages concerning this topic are being heard and causing adequate concern. Now that there is evidence that previous dengue infection enhances the chances of more severe Zika infection, public health officials need to proliferate this message across social media sites and encourage those with past dengue infection to continue to take precautions against mosquito bites.

How to Address These Concerns

Now that public health officials know what the public is concerned about, they can focus on addressing these concerns. When an incident occurs, the normal tendency is to seek more information on the topic of interest [53]. This can be done by reading or listening to the news, performing internet searches, or communicating with others. Through this search for knowledge, concern can be diminished or enhanced, depending on the information gathered [54].

Complications related to processing can include the accuracy of the information shared, as at times the media is quick to report information without having all the facts or the reader may interpret the facts incorrectly [53]. Therefore, news agencies need to be more careful about what they publish and not use titles such as the BBC article did [3] that are meant to instill concern in the public. Deficiencies in communications among the media, the public, politicians, and scientists heightens concern [55]. For example, when nonexperts express views different from experts, public fear is heightened [56]. This is a difficult problem to address, as evidenced by the debate on vaccines and autism [57]. Experts need to keep putting factual information out there and also keep peer reviewing each other to make sure studies such as the one by Wakefield suggesting vaccines cause autism do not occur in the future [57]. Another common example is the level of information presented to the public. Scientists tend to use words the public does not understand, such as the word asymptomatic, causing a discrepancy between what is stated by public health officials and what the reader interprets. This can be addressed by scientists better explaining their work at an elementary school level.

The authors understand all of these suggestions are already being followed at some capacity by public health officials. However, there is always room for improvement.

Limitations

The tweets in our analysis were limited to the English language, which limits the generalizability of the study. This is critical as South American countries were the first and hardest hit countries. Future studies can address this limitation by analyzing tweets in Portuguese and Spanish. The keywords used in data collection were Zika, Zika virus, Zika virus treatment, and Zika treatment. Therefore, tweets that refer to this disease in another language would be overlooked. Tweets that refer to the disease without mentioning it by name would also be overlooked.

Without prosody, contextual, and spectral cues, sarcasm is difficult to detect [58], all 3 of which are impossible to determine in a tweet. Some research has been done using lexical and pragmatic factors [59]; however, even the human annotators had less than 50% agreement on whether a tweet was sarcastic

in this study. Clearly, if the ground truth is inconsistent, it cannot be modeled reliably with machine learning. The annotators in this study coded the tweets based on the sentiment they believed it expressed, with sarcasm being one of the causes of disagreement. However, very few tweets were considered to be possibly sarcastic in our dataset, thus limiting the effect.

Due to the short length of tweets and the large number of tweets collected, LDA has been previously shown to have some issues with overfitting, with the number of revealed topics exceeding the true number of topics [60]. We attempted to address both of these concerns in our study by combining positive tweets into a document, negative into another document, and neutral into a third document, thus making the datasets smaller and the topical domains more specific.

Conclusions

Overall, the negative sentiment topics focused on neural defects and abnormalities caused by the Zika virus. As these tweets were categorized as negative sentiments, officials could see that the public was concerned with the symptoms caused by the Zika virus. As the public was concerned, officials could focus on spreading information encouraging prevention. Officials could also see that the top themes all concerned actual symptoms and defects and did not focus on misconceptions or misinformation that they needed to address. Moving forward, officials can also start informing the public that studies are providing evidence for the Zika-dengue interaction hypothesis. They should focus these messages in areas where dengue is endemic as they are the ones most at risk of the interaction causing more severe Zika infection.

When another Zika outbreak occurs, we predict similar concerns (such as microcephaly) about the neurological defects will be expressed on social media. Although our current framework would still be applicable, the unsupervised topics within the tweets would change. Specifically, the relevancy and sentiment classifiers (the supervised part of the system) would still be effective in detecting tweets specific to Zika and specific to the particular topic such as symptoms. However, when a preventative vaccine for Zika virus infection is created and/or new symptoms arise that are associated with Zika virus infection, the topics of concern would change depending on the current issue of concern at a particular time. As of August 2018, no licensed vaccines were available; however, several candidates are in various stages of development, and clinical trials have begun [61]. Once a licensed vaccine is available, we predict negative sentiment concerning Zika virus symptoms will decrease but most likely will not disappear. At that time, the methods utilized in this study will still be relevant, but the major topics in the negative sentiment category will likely change because of the decrease in concern, which would also be indicated by the increase in tweets in the positive or neutral categories.

On the contrary, if new symptoms for Zika were to develop or further complications for those born with neurological defects were discovered, the topics of concern in the negative sentiment category would change to reflect concerns specific to the new symptoms. During the most recent outbreak, scientists suspected that people who previously had a dengue infection experienced worse symptoms from Zika than those who had not been previously infected with dengue [62]. Previous infection with a similar virus to Zika, such as West Nile, may cause new symptoms like we saw with dengue and Zika [51,52].

Our study is also useful for those that want to perform sentiment analysis with an epidemic, pandemic, or bioterrorism attack. Sentiment analysis is complex as most sentiment analysis tools just use the individual word polarities for measuring sentiment and generate an automated scoring mechanism based on these polarities to rate the sentiment levels of each tweet. This fails to incorporate the contextual information that needs to be incorporated for topic-specific sentiment analysis in this domain [5]. Scientific topics especially require manual labeling as science words with negative sentiment can actually have a positive context as seen in this tweet, "Obama diverts Ebola funds to fight Zika; Florida leads nation in case..." The word *fight* would typically have a negative connotation but has a positive one in this tweet. Some examples of other words that are typically considered negative but are actually positive when discussed under the context of epidemics are *combat*, *prevent*, and *impair*. If tweets containing these words were categorized using a sentiment word bank, they would have been incorrectly categorized as negative. This is an important issue because it does not correctly represent the public's feelings and may cause experts to believe the public is not as concerned about Zika symptoms if some of the negative tweets were misclassified as positive/neutral. Therefore, we used a manual labeling process where an entire tweet was assigned to a sentiment category by 2 domain experts. We believe that this need for a combination of data science and domain expertise is what makes our study challenging and interesting.

This is one of the first studies to address Zika sentiment classification using Twitter. Using such a system allows public health officials to ascertain public sentiment concerning disease outbreaks and address concerns in real time.

Future Work

Future studies could analyze the change in sentiment over time to see if the number of negative tweets decreases as the outbreak subsides and more advances in treatments are discovered. Studies could also look at sentiment by gender or geographic location. Both are prudent because of Zika's effect on fetuses and its comparative prevalence in equatorial regions, respectively. We would also suggest future studies to leverage other sources of information, such as other social media sites, newspapers, and blogs. Similar methodologies could also be applied generally to future pandemics and epidemics to ascertain public sentiment.

Conflicts of Interest

None declared.

References

1. World Health Organization (WHO). 2016. The history of Zika Virus URL: <http://www.who.int/emergencies/zika-virus/timeline/en/> [accessed 2018-05-13] [WebCite Cache ID 6zOOaR1TF]
2. Centers for Disease Control and Prevention. 2016. Zika Virus URL: <https://www.cdc.gov/zika/about/overview.html> [WebCite Cache ID 6zOP6OiXE]
3. British Broadcasting Corporation (BBC). Zika virus scarier than thought? says US URL: <http://www.bbc.com/news/> [WebCite Cache ID 6zOSoZHYC]
4. Miller M, Banerjee T, Muppalla R, Romine W, Sheth A. What Are People Tweeting About Zika? An Exploratory Study Concerning Its Symptoms, Treatment, Transmission, and Prevention. *JMIR Public Health Surveill* 2017 Jun 19;3(2):e38 [FREE Full text] [doi: [10.2196/publichealth.7157](https://doi.org/10.2196/publichealth.7157)] [Medline: [28630032](https://pubmed.ncbi.nlm.nih.gov/28630032/)]
5. Chen L, Wang W, Nagarajan M, Wang S, Sheth A. AAAI. 2012. Extracting diverse sentiment expressions with target-dependent polarity from twitter URL: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4609> [WebCite Cache ID 6zOPbKiLs]
6. Sheth A, Wang W, Chen L. Google Patents. Topic-specific sentiment extraction URL: <https://patents.google.com/patent/US20140358523> [accessed 2018-05-13] [WebCite Cache ID 6zOPlu9tx]
7. Daniulaityte R, Chen L, Lamy FR, Carlson RG, Thirunarayan K, Sheth A. When Bad is Good?: identifying personal communication and sentiment in drug-related tweets. *JMIR Public Health Surveill* 2016 Oct 24;2(2):e162 [FREE Full text] [doi: [10.2196/publichealth.6327](https://doi.org/10.2196/publichealth.6327)] [Medline: [27777215](https://pubmed.ncbi.nlm.nih.gov/27777215/)]
8. Ji X, Chun S, Geller J. Monitoring public health concerns using Twitter sentiment classifications. In: *Healthcare Informatics (ICHI)*. 2013 Presented at: IEEE International Conference; 2013; Karlsruhe p. 335-344. [doi: [10.1109/ICHI.2013.47](https://doi.org/10.1109/ICHI.2013.47)]
9. Househ M. Communicating Ebola through social media and electronic news media outlets: a cross-sectional study. *Health Informatics J* 2016 Dec;22(3):470-478. [doi: [10.1177/1460458214568037](https://doi.org/10.1177/1460458214568037)] [Medline: [25656678](https://pubmed.ncbi.nlm.nih.gov/25656678/)]
10. Ghenai A, & MY. Research Gate. 2017. Catching Zika fever: Application of crowdsourcing and machine learning for tracking health misinformation on Twitter URL: <https://arxiv.org/pdf/1707.03778.pdf> [accessed 2019-05-14] [WebCite Cache ID 78Mq0WJyq]
11. Seltzer EK, Horst-Martz E, Lu M, Merchant RM. Public sentiment and discourse about Zika virus on Instagram. *Public Health* 2017 Sep;150:170-175. [doi: [10.1016/j.puhe.2017.07.015](https://doi.org/10.1016/j.puhe.2017.07.015)] [Medline: [28806618](https://pubmed.ncbi.nlm.nih.gov/28806618/)]
12. Sheth A, Purohit H, Smith GA, Brunn J, Jadhav A, Kapanipathi P, et al. Twitris- A System for Collective Social Intelligence. In: Alhajj R, Rokne J, editors. *Encyclopedia of Social Network Analysis and Mining (ESNAM)*. New York: Springer; 2017:1-23.
13. McHugh M. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-282 [FREE Full text] [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]
14. Muppalla R, Miller M, Banerjee T, Romine W. Discovering explanatory models to identify relevant tweets on Zika. In: *Conf Proc IEEE Eng Med Biol Soc*. 2017 Dec Presented at: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Seogwipo, , pp; July 11-15, 2017; Jeju Island, Korea p. 1194-1197. [doi: [10.1109/EMBC.2017.8037044](https://doi.org/10.1109/EMBC.2017.8037044)]
15. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Neural Information Processing Systems. Distributed representations of words and phrases and their compositionality URL: <https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf> [accessed 2018-05-13] [WebCite Cache ID 6zOQfF3Oj]
16. Van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res* 2008 Nov 09;9:2579-2605 [FREE Full text]
17. Word2vec. Google Code. 2013. Word2vec URL: <https://code.google.com/archive/p/word2vec/> [WebCite Cache ID 6zOQuDXw0]
18. Wijeratne S, Balasuriya L, Doran D, Sheth A. Knoesis. 2016. Word embeddings to enhance Twitter gang member profile identification URL: <http://knoesis.org/node/2753> [WebCite Cache ID 6zOQyfi6J]
19. Gimpel K, Schneider N, O'Connor B, Das D, Mills D, Eisenstein J, et al. Part-of-speech tagging for Twitter: Annotation, features, and experiments. 2011 Presented at: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies short papers-Volume 2; June; 2011; Portland, Oregon.
20. Brown P, Desouza P, Mercer R, Pietra V, Lai JC. Class-based n-gram models of natural language. *Comput Linguist* 2016;18(4):467-479.
21. Rong X. Arxiv. word2vec Parameter Learning Explained URL: <https://arxiv.org/pdf/1411.2738.pdf> [accessed 2019-05-10] [WebCite Cache ID 78GvgchEO]
22. Python. Gensim Library URL: <https://pypi.org/project/gensim/2.2.0/> [accessed 2018-05-13] [WebCite Cache ID 6zORES7tr]
23. Bamler R, Mandt S. Dynamic word embeddings. 2017 Presented at: International Conference on Machine Learning; 2017; Sydney, Australia p. 380-389 URL: <https://pdfs.semanticscholar.org/8a33/8e6068f2d9dd4630c3f0967f7b566f0819ba.pdf>
24. Forman G, Scholz M. Apples-to-apples in cross-validation studies: pitfalls in classifier performance measurement. *ACM SIGKDD Explorations Newsletter* 2010;12(1):49-57 [FREE Full text]
25. Lau J, Collier N, Baldwin T. On-line trend analysis with topic models: # twitter trends detection topic model online. 2012 Presented at: Proceedings of COLING; 2012; Mumbai, India p. 1519-1534.

26. Blei D, Ng A, Jordan M. Latent dirichllocation allocation. *J Mach Learn Res* 2003 Jan 03;993-1022 [FREE Full text]
27. BitBucket. Zika Sentiment URL: <http://ravali-mamidi.info/sentiment-topics> [accessed 2018-05-13] [WebCite Cache ID 6zORLqaQH]
28. International Business Machines. IBM Journal. Say goodbye to viral infections with IBM URL: <https://www.ibmjournal.com/internet-of-things/ibm-cure-for-viruses> [accessed 2018-05-13] [WebCite Cache ID 6zORXBpw9]
29. Mahyoub J, Aziz AT, Panneerselvam C, Murugan K, Roni M, Trivedi S, et al. Springer. 2017. Seagrasses as Sources of Mosquito Nano-Larvicides? Toxicity and Uptake of Halodule uninervis-Biofabricated Silver Nanoparticles in Dengue and Zika Virus Vector *Aedes aegypti* URL: <https://link.springer.com/article/10.1007%2Fs10876-016-1127-3> [accessed 2019-05-10] [WebCite Cache ID 78GwUiYEI]
30. Retallack H, Di Lullo E, Arias C, Knopp KA, Laurie MT, Sandoval-Espinosa C, et al. Zika virus cell tropism in the developing human brain and inhibition by azithromycin. *Proc Natl Acad Sci U S A* 2016 Dec 13;113(50):14408-14413 [FREE Full text] [doi: [10.1073/pnas.1618029113](https://doi.org/10.1073/pnas.1618029113)] [Medline: [27911847](https://pubmed.ncbi.nlm.nih.gov/27911847/)]
31. Dredze M, Broniatowski D, Hilyard KM. Zika vaccine misconceptions: a social media analysis. *Vaccine* 2016 Dec 24;34(30):3441-3442 [FREE Full text] [doi: [10.1016/j.vaccine.2016.05.008](https://doi.org/10.1016/j.vaccine.2016.05.008)] [Medline: [27216759](https://pubmed.ncbi.nlm.nih.gov/27216759/)]
32. Lucey DR, Gostin LO. The emerging Zika pandemic. *J Am Med Assoc* 2016 Mar 01;315(9):865-866. [doi: [10.1001/jama.2016.0904](https://doi.org/10.1001/jama.2016.0904)] [Medline: [26818622](https://pubmed.ncbi.nlm.nih.gov/26818622/)]
33. Coltart CE, Lindsey B, Ghinai I, Johnson AM, Heymann DL. The Ebola outbreak, 2013-2016: old lessons for new epidemics. *Philos Trans R Soc Lond B Biol Sci* 2017 May 26;372(1721):- [FREE Full text] [doi: [10.1098/rstb.2016.0297](https://doi.org/10.1098/rstb.2016.0297)] [Medline: [28396469](https://pubmed.ncbi.nlm.nih.gov/28396469/)]
34. Glowacki E, Lazard A, Wilcox G, Mackert M, Bernhardt JM. Identifying the public's concerns and the Centers for Disease Control and Prevention's reactions during a health crisis: An analysis of a Zika live Twitter chat. *Am J Infect Control* 2016 Dec 01;44(12):1709-1711. [doi: [10.1016/j.ajic.2016.05.025](https://doi.org/10.1016/j.ajic.2016.05.025)] [Medline: [27544795](https://pubmed.ncbi.nlm.nih.gov/27544795/)]
35. Southwell B, Dolina S, Jimenez-Magdaleno K, Squiers L, Kelly BJ. Zika virus-related news coverage and online behavior, United States, Guatemala, and Brazil. *Emerg Infect Dis* 2016 Dec;22(7):1320-1321 [FREE Full text] [doi: [10.3201/eid2207.160415](https://doi.org/10.3201/eid2207.160415)] [Medline: [27100826](https://pubmed.ncbi.nlm.nih.gov/27100826/)]
36. Centers for Disease Control and Prevention (CDC). Congenital Zika syndrome & other birth defects URL: <https://www.cdc.gov/pregnancy/zika/testing-follow-up/zika-syndrome-birth-defects.html> [WebCite Cache ID 6zORj8yIC]
37. The Wall Street Journal. 2016 Nov 3. The Effects of Zika on Babies' Brains Go Beyond Microcephaly URL: <https://www.wsj.com/articles/the-effects-of-zika-on-babies-brains-go-beyond-microcephaly-1478191331> [accessed 2018-05-13] [WebCite Cache ID 6zOT2qDFX]
38. Rare Technologies. Gensim default parameters URL: <https://rare-technologies.com/word2vec-tutorial/> [accessed 2018-05-13] [WebCite Cache ID 6zORuGbVi]
39. Nakov P, Ritter A, Rosenthal S, Sebastiani F, Stoyanov V. SemEval-2016 Task 4: Sentiment Analysis in Twitter. 2016 Presented at: Proceedings of the 10th international workshop on semantic evaluation; 2016; San Diego p. 1-18.
40. Centers for Disease Control and Prevention. Microcephaly & other birth defects URL: https://www.cdc.gov/zika/healtheffects/birth_defects.html [WebCite Cache ID 6zOS4TMAC]
41. Centers for Disease Control and Prevention. Prenatal Care URL: <https://www.cdc.gov/pregnancy/zika/testing-follow-up/prenatal-care.html> [WebCite Cache ID 6zOSBuus6]
42. Centers for Disease Control and Prevention. Care for babies with congenital Zika Syndrome URL: <https://www.cdc.gov/pregnancy/zika/family/care-for-babies-with-congenital-zika.html> [WebCite Cache ID 6zOSL6Afu]
43. World Health Organization (WHO). 2017. Dengue and severe dengue URL: <http://www.who.int/mediacentre/factsheets/fs117/en/> [WebCite Cache ID 6zOSSZ64P]
44. Centers for Disease Control and Prevention. Fight the bite URL: https://www2c.cdc.gov/podcasts/media/pdf/FighttheBite2_transcript.pdf [WebCite Cache ID 6zOSWjUSx]
45. Wellcome Trust. Science Daily. New insights into how the Zika virus causes microcephaly URL: <https://www.sciencedaily.com/releases/2017/06/170601151903.htm> [WebCite Cache ID 6zOScEzx6]
46. National Institute of Health (NIH). 2017. Zika virus persists in the central nervous system and lymph nodes of Rhesus monkeys URL: <https://www.nih.gov/news-events/news-releases/zika-virus-persists-central-nervous-system-lymph-nodes-rhesus-monkeys> [accessed 2018-05-13] [WebCite Cache ID 6zOTyBbiF]
47. Cao-Lormeau V, Blake A, Mons S, Lastere S, Roche C, Vanhomwegen J, et al. Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet* 2016 Apr 09;387(10027):1531-1539 [FREE Full text] [doi: [10.1016/S0140-6736\(16\)00562-6](https://doi.org/10.1016/S0140-6736(16)00562-6)] [Medline: [26948433](https://pubmed.ncbi.nlm.nih.gov/26948433/)]
48. Chibueze EC, Parsons AJ, Lopes KS, Yo T, Swa T, Nagata C, et al. Diagnostic accuracy of ultrasound scanning for prenatal microcephaly in the context of Zika virus infection: a systematic review and meta-analysis. *Sci Rep* 2017 Dec 23;7(1):2310 [FREE Full text] [doi: [10.1038/s41598-017-01991-y](https://doi.org/10.1038/s41598-017-01991-y)] [Medline: [28536443](https://pubmed.ncbi.nlm.nih.gov/28536443/)]
49. Sun LH. The Washington Post. Ultrasounds missed her Zika infection - until one showed serious harm to her fetus URL: <https://www.washingtonpost.com/news/to-your-health/wp/2016/03/30/>

- [why-ultrasounds-may-give-mothers-with-zika-a-false-sense-of-security/?utm_term=.b991518f47aa](#) [WebCite Cache ID 6zOU4mrHJ]
50. Littaua R, Kurane I, Ennis FA. Human IgG Fc receptor II mediates antibody-dependent enhancement of dengue virus infection. *J Immunol* 1990 Apr 15;144(8):3183-3186. [Medline: [2139079](#)]
 51. Schnirring L. Center for Infectious Disease Research and Policy. Lab findings hint that dengue antibodies intensify Zika infection URL: <http://www.cidrap.umn.edu/news-perspective/2016/04/lab-findings-hint-dengue-antibodies-intensify-zika-infection> [accessed 2018-05-13] [WebCite Cache ID 6zOUAoYz5]
 52. Paul LM, Carlin ER, Jenkins MM, Tan AL, Barcellona CM, Nicholson CO, et al. Dengue virus antibodies enhance Zika virus infection. *Clin Transl Immunology* 2016 Dec;5(12):e117 [FREE Full text] [doi: [10.1038/cti.2016.72](https://doi.org/10.1038/cti.2016.72)] [Medline: [28090318](#)]
 53. Pollard WE. Public perceptions of information sources concerning bioterrorism before and after anthrax attacks: an analysis of national survey data. *J Health Commun* 2003;8(Suppl 1):93-103. [doi: [10.1080/713851974](https://doi.org/10.1080/713851974)] [Medline: [14692574](#)]
 54. Slovic P, Finucane M, Peters E, MacGregor DG. Science Direct. 2007. The affect heuristic URL: <https://linkinghub.elsevier.com/retrieve/pii/S0377221705003577> [accessed 2019-05-10] [WebCite Cache ID 78GySyyyh]
 55. Devaux CA. The hidden face of academic researches on classified highly pathogenic microorganisms. *Infect Genet Evol* 2015 Jan;29:26-34 [FREE Full text] [doi: [10.1016/j.meegid.2014.10.028](https://doi.org/10.1016/j.meegid.2014.10.028)] [Medline: [25445654](#)]
 56. Malet D, Korbitz M. Public Risk Communications in disaster recovery: results from a biological decontamination experiment. 2013 Presented at: Annual Conference of the Australian Political Studies Association; 2013; Perth, Australia.
 57. Wakefield A, Murch S, Anthony A, Linnell J, Casson D, Malik M, et al. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet* 1998 Dec 28;351(9103):637-641. [Medline: [9500320](#)]
 58. Tepperman J, Traum D, Narayanan S. University of Southern California. 2006. Yeah Right: Sarcasm Recognition for Spoken Dialogue Systems URL: https://sail.usc.edu/publications/files/tepperman_interspeech_2006b.pdf [accessed 2019-05-14] [WebCite Cache ID 78MvcSLRh]
 59. González-Ibáñez R, Muresan S, Wacholder N. Identifying sarcasm in Twitter: a closer look. 2011 Presented at: 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologieshort Papers; 2011; Portland, Oregon.
 60. Tang J, Meng Z, Nguyen X, Mei Q, Zhang M. Understanding the limiting factors of topic modeling via posterior contraction analysis. 2014 Presented at: International Conference on Machine Learning; June 21-26, 2014; Beijing, China.
 61. National Institute of Health. 2018. NIH begins clinical trial of live, attenuated Zika vaccine URL: <https://www.nih.gov/news-events/news-releases/nih-begins-clinical-trial-live-attenuated-zika-vaccine> [accessed 2019-03-12] [WebCite Cache ID 76p188iCm]
 62. Bardina SV, Bunduc P, Tripathi S, Duehr J, Frere JJ, Brown JA, et al. Enhancement of Zika virus pathogenesis by preexisting ant flavivirus immunity. *Science* 2017 Dec 14;356(6334):175-180 [FREE Full text] [doi: [10.1126/science.aal4365](https://doi.org/10.1126/science.aal4365)] [Medline: [28360135](#)]

Abbreviations

- ADE:** antibody-dependent enhancement
ASCII: American Standard Code for Information Interchange
CBOW: continuous bag of words
CDC: Centers for Disease Control and Prevention
LDA: latent Dirichlet allocation
MRI: magnetic resonance imaging
RQ: research question
WHO: World Health Organization
3D: 3-dimensional

Edited by T Sanchez; submitted 13.05.18; peer-reviewed by M Farhadloo, D Ghosh; comments to author 14.09.18; revised version received 08.11.18; accepted 16.04.19; published 04.06.19.

Please cite as:

Mamidi R, Miller M, Banerjee T, Romine W, Sheth A
 Identifying Key Topics Bearing Negative Sentiment on Twitter: Insights Concerning the 2015-2016 Zika Epidemic
JMIR Public Health Surveill 2019;5(2):e11036
 URL: <http://publichealth.jmir.org/2019/2/e11036/>
 doi:[10.2196/11036](https://doi.org/10.2196/11036)
 PMID:[31165711](https://pubmed.ncbi.nlm.nih.gov/31165711/)

©Ravali Mamidi, Michele Miller, Tanvi Banerjee, William Romine, Amit Sheth. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 04.06.2019. 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

Preliminary Flu Outbreak Prediction Using Twitter Posts Classification and Linear Regression With Historical Centers for Disease Control and Prevention Reports: Prediction Framework Study

Ali Alessa^{1,2}, PhD; Miad Faezipour^{1,3}, BSc, MSc, PhD

¹Department of Computer Science and Engineering, University of Bridgeport, Bridgeport, CT, United States

²Institute of Public Administration, Riyadh, Saudi Arabia

³Department of Biomedical Engineering, University of Bridgeport, Bridgeport, CT, United States

Corresponding Author:

Miad Faezipour, BSc, MSc, PhD

Department of Computer Science and Engineering

University of Bridgeport

221 University Avenue

Bridgeport, CT, 06604

United States

Phone: 1 203 576 4702

Email: mfaezipo@bridgeport.edu

Abstract

Background: Social networking sites (SNSs) such as Twitter are widely used by diverse demographic populations. The amount of data within SNSs has created an efficient resource for real-time analysis. Thus, data from SNSs can be used effectively to track disease outbreaks and provide necessary warnings. Current SNS-based flu detection and prediction frameworks apply conventional machine learning approaches that require lengthy training and testing, which is not the optimal solution for new outbreaks with new signs and symptoms.

Objective: The objective of this study was to propose an efficient and accurate framework that uses data from SNSs to track disease outbreaks and provide early warnings, even for newest outbreaks, accurately.

Methods: We presented a framework of outbreak prediction that included 3 main modules: text classification, mapping, and linear regression for weekly flu rate predictions. The text classification module used the features of sentiment analysis and predefined keyword occurrences. Various classifiers, including FastText (FT) and 6 conventional machine learning algorithms, were evaluated to identify the most efficient and accurate one for the proposed framework. The text classifiers were trained and tested using a pre-labeled dataset of flu-related and unrelated Twitter postings. The selected text classifier was then used to classify over 8,400,000 tweet documents. The flu-related documents were then mapped on a weekly basis using a mapping module. Finally, the mapped results were passed together with historical Centers for Disease Control and Prevention (CDC) data to a linear regression module for weekly flu rate predictions.

Results: The evaluation of flu tweet classification showed that FT, together with the extracted features, achieved accurate results with an *F*-measure value of 89.9% in addition to its efficiency. Therefore, FT was chosen to be the classification module to work together with the other modules in the proposed framework, including a regression-based estimator, for flu trend predictions. The estimator was evaluated using several regression models. Regression results show that the linear regression-based estimator achieved the highest accuracy results using the measure of Pearson correlation. Thus, the linear regression model was used for the module of weekly flu rate estimation. The prediction results were compared with the available recent data from CDC as the ground truth and showed a strong correlation of 96.29% .

Conclusions: The results demonstrated the efficiency and the accuracy of the proposed framework that can be used even for new outbreaks with new signs and symptoms. The classification results demonstrated that the FT-based framework improves the accuracy and the efficiency of flu disease surveillance systems that use unstructured data such as data from SNSs.

(*JMIR Public Health Surveill* 2019;5(2):e12383) doi:[10.2196/12383](https://doi.org/10.2196/12383)

KEYWORDS

FastText; influenza; machine learning; social networking site; text classification

Introduction

Background

According to the Centers for Disease Control and Prevention (CDC), flu is a serious contagious respiratory illness that can lead to hospitalization and sometimes death. About 250,000 to 500,000 deaths occur worldwide each year because of flu. Flu is common during some seasons, but there can be deadly outbreaks that spread suddenly in a community.

Social networking sites (SNSs) are tools that include big data about users and their shared thoughts and ideas, in addition to real-time data of users' conversations and statuses. The amount of data, aside from the growth of SNS users, represents the important role of SNSs in real-time analysis and predictions in many areas, including the area of public health [1]. SNSs provide an efficient resource to conduct disease surveillance and a communication tool to prevent disease outbreaks [2].

To produce outbreak reports, typical disease surveillance systems depend on official statistics based on patient visits [3]. In the United States, these reports are produced by the CDC to inform health care providers about certain disease outbreaks such as influenza outbreaks. CDC publishes flu-related reports using the US Influenza Like Illness Surveillance Network (ILINet) that gathers flu-related information of outpatients from hundreds of health care providers around the United States. ILINet shows accurate results in detecting flu outbreaks, but it is costly and takes a long time to issue the required reports. It is crucial for any disease surveillance system to collect related data and provide the reports as early as possible to prevent the spread of the disease. To this end, many solutions have been proposed to generate earlier outbreak warnings. Examples include volumes of telephone calls, over-the-counter drug sales [3], search engine logs [4-9], and data from SNSs that can be used for real-time analysis for better services [10-14]. Analysis of search engine logs, such as Google Flu Trend (GFT), estimates the percentage of ILI cases using flu correlated queries. In 2013, GFT overpredicted the percentage of the ILI cases by the double [15]. Compared with different resources used for surveillance, that is, search engine logs, data from SNSs are more descriptive and available to the public. Because SNSs provide certain information about users, the collected data can be used to simulate the spread of disease outbreaks in connected geographic areas with temporal analysis [15].

In this study, we relied on the Twitter microblog to conduct minute-by-minute analysis to track the high frequency of posted messages. We present a framework to track influenza trends through Twitter postings. The framework includes preprocessing, feature extraction, Twitter documents classification, documents weekly-mapping, and weekly flu rate predictions. The preprocessing phase includes stemming and removal of stop words and ineffective characters, which are nonalphanumeric tokens. Thereafter, the preprocessed data are used to extract features to be passed to a tweet classifier to distinguish between flu-related tweets and unrelated ones. The

flu-related documents are then mapped on a weekly basis. Finally, the mapped results are passed together with historical CDC data to an estimator for flu trend predictions.

The data generated from SNSs are valuable for real-time analysis and outbreak predictions, but its volume is huge. Therefore, one of the main challenges in analyzing this huge volume of data is to find the best approach for accurate analysis in a time-efficient manner. Current Twitter-based flu detection and prediction frameworks apply conventional machine learning approaches that require lengthy training and testing, which is not the optimal solution to be used for a new outbreak with new signs and symptoms. Regardless of the analysis time, many studies only report the accuracy of different machine learning approaches. Thus, more efficient solutions are required for accurate results with less processing time. In this study, we demonstrate that using FastText (FT) can enhance the efficiency of Twitter-based flu outbreak prediction models. Originally, FT became an efficient text classifier that was proposed by Facebook. FT performs more quickly than deep learning classifiers for training and testing procedures and produces comparably accurate results. The FT classifier can train more than a billion words in about 10 min and then predict multiple classes within half a million sentences in less than a minute [16].

The aim of this study was to develop an efficient Twitter-based model that provides accurate results with less processing time to predict seasonal and serious outbreaks such as H5N1. This study presents an accurate and efficient FT-based framework to generate influenza trend predictions from Twitter. In addition to the typical textual features, the proposed framework uses the features of text sentiment analysis and the occurrences of predefined topic keywords to distinguish between flu-related tweets and unrelated ones to be passed together with historical CDC data to an estimator module for weekly flu rate predictions. The main contributions of this study can be summarized as follows: (1) demonstrating that FT classifier can improve the efficiency of tweet classification; (2) including sentiment analysis of the analyzed posts as a feature to improve the accuracy of the classification results; (3) examining various conventional machine learning algorithms for flu-related tweets; (4) proposing a weekly flu rate estimator based on the linear regression model that uses a combination of the classification results and historical CDC data; and (5) examining, in addition to the linear regression model, several regression techniques for weekly flu rate estimation.

Problem Definition

SNS postings can be seen as triggers for different event predictions such as disease outbreaks. Discovering knowledge from the posts for flu surveillance models requires an efficient approach of text processing. It includes gathering the related text (posts) about the disease and then issuing necessary reports at an early stage, which is crucial for outbreak prevention. Because the gathered data is unstructured, the first step is to preprocess the unstructured content to analyze the data and

produce the results in an understandable way. The second step is feature extraction, which is key to performance enhancement. The third step is knowledge extraction using machine learning techniques for text classification, which includes model training and testing. A post on a microblogging site is then classified into either related or unrelated classes, as can be seen from the following example:

Related: *I'm sick, I got flu yesterday.*

Unrelated: *I'm sick of school.*

Our literature survey indicates that most of the existing frameworks use conventional machine learning classifiers [17]. These approaches require a longer time duration for the training process. A new outbreak may require retraining the used prediction model with its new signs and symptoms to consider the related posts. Thus, such approaches are not optimal solutions for new deadly flu outbreaks.

The proposed framework using FT classifier, together with the extracted features, which have not been previously used for Twitter-based flu surveillance models, aims to extract related posts faster with a comparable accuracy. Thus, it can be used for urgent cases to stop the spread of a new deadly outbreak. Improving the efficiency, along with the accuracy of text classification is important for text-based surveillance systems for generating early reports. To develop a flu outbreak prediction framework, the classified tweets are then passed together with historical CDC data to an estimator module for weekly flu rate predictions.

Previous Work

Previous works about Twitter-based flu surveillance systems include machine learning methods to filter unrelated flu posts. A selected classifier is trained with an annotated dataset using a set of features. The literature discusses various detection and prediction models that used different classification methods with different feature extraction techniques.

Broniatowski et al [18] and Lamb et al [19] proposed a multilevel classification model that included a binary classifier to distinguish between flu-related and unrelated flu tweets. The preclassifiers were used to filter unwanted posts such as health-irrelevant posts to increase the efficiency of the flu-related/unrelated classifier in further stages/levels. Here, the researchers demonstrated that multilevel classification can improve classification accuracy.

Aramaki et al [20] proposed a framework that consisted of 2 parts: a tweet crawler and a support vector machine (SVM)-based classifier that was used to extract only the actual influenza tweets and excluded the unrelated ones such as news and questions. The initial dataset for their study was collected between November 2008 and June 2010. It included 300 million general tweets. The dataset was then filtered using the "Influenza" keyword to get a set of only flu-related tweets that contained 400,000 tweets. The flu-related dataset was divided into 2 parts: a training dataset that contained 5000 tweets (November 2008) and a test dataset that contained all the remaining tweets between December 2008 and June 2010. The training dataset was assigned to a human annotator to label each

tweet as either positive or negative. A tweet was labeled positive if it met 2 conditions. First, the flu tweet should be about the person who posted the tweet or about another person in a nearby area (maximum an area of the city). If the distance is unknown, the tweet is considered negative. Second, the flu tweet should be an affirmative sentence and in present tense or past tense with maximum period of 24 hours, which can be checked using specific keywords such as "yesterday." The SVM classifier was implemented using the bag-of-words (BoW) feature representation. The authors compared the accuracy of the SVM-based classifier with 6 other different machine learning methods and found that SVM was the most accurate method.

Santos et al [21] also applied SVM-based classification to detect FLI in Portugal using twitter posts. For the purpose of training and testing, a dataset with 2704 posts was manually annotated with 650 textual features. A subset of the annotated dataset was used to train the classifier. The classified tweets, together with search queries, were applied to a regression model as predictors. The classifier was implemented using the BoW feature representation, and the feature selection process was based on a mutual information (MI) value that was used to pick the best set of features. In this approach, each feature is applied to a true class, and then the MI value is assigned to the feature. The value of MI is based on how the feature is related to the true class. A feature with high MI value represents being more related to the true class.

Yang et al [22] proposed the first SVM-based method to predict flu trends from Chinese microblogging sites in Beijing. The collected data for their study included 3,505,110 posts between September 2013 and December 2013. Among those, 5000 random posts were selected for manual annotation (sick and not sick labels) to be used for training and testing purposes. Of these, 285 of the sick posts and 285 of the not sick posts were picked for training. For higher accuracy, word-based features were used instead of character-based features. In addition, the term frequency-inverse document frequency (TF-IDF) method was considered for weighting. Different classifiers were compared to decide which classifier would be best for the problem. The authors found that SVM was the best classifier for big data problems.

Byrd et al [23] proposed a framework based on the Naïve Bayes classifier. The framework consisted of preprocessing and flu tweet classification based on sentiment analysis. Three machine learning algorithms were evaluated. The results indicated that the highest accuracy algorithm was the Naïve Bayes classifier. The classifier was implemented using the Stanford CoreNLP (a natural language processing [NLP] software) and trained using the OpenNLP training dataset, which included 100 annotated tweets. Sentiment analysis is considered accurate when there is a match between the predicted sentiment polarity and the manual assigned opinion of the sentiment. The researchers found that Naïve Bayes was the most accurate classifier with a rate of 70% match.

Methods

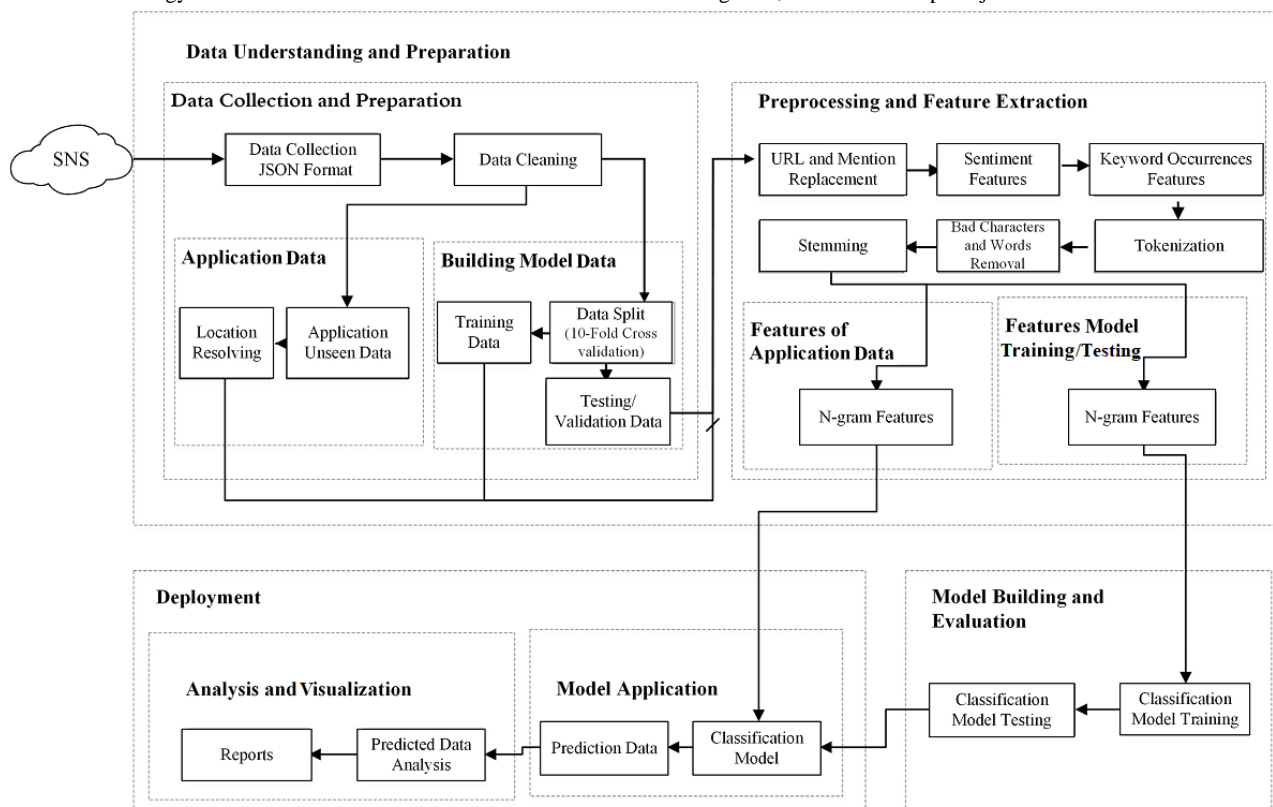
Proposed Framework

The proposed framework, which includes a classification model for flu posts, published on the Twitter microblogging site, is implemented using the Cross Industry Standard Process for Data Mining (CRISP-DM). It is a well-known standard for implementing data mining frameworks. This standard includes the following 6 steps [24]:

- Business understanding
- Data understanding
- Data preparation
- Modeling
- Evaluation
- Deployment

On the basis of the CRISP-DM standard, the methodology for this study is presented in [Figure 1](#).

Figure 1. Methodology for text classification of flu tweets. SNS: social networking sites; JSON: JavaScript Object Notation.



Data Collection and Preparation

Classification Model Data

For classification model training and testing, we prepared a labeled dataset that is a combination of multiple manually labeled datasets obtained from [19,25]. This makes the total instances of the merged dataset 10,592 tweets (5249 flu-related and 5343 flu-unrelated posts). Due to Twitter guidelines, the tweets in the obtained datasets were released with tweet IDs instead of the text of the tweets. Therefore, we developed a script that works together with the Twitter application programming interface (API) to retrieve the corresponding tweet texts using the given IDs. The collected tweets were cleaned to include only the texts for training and testing purposes. Then, we divided the merged dataset into 2 parts: training set and testing set.

Twitter Influenza Surveillance Dataset

The labeled dataset obtained from Lamb et al [19] was initially filtered to contain any posts that have flu-related keywords. Then, every post in the dataset was labeled manually. It was prepared to train and test 3 flu-related classifiers that were used

as a part of an algorithm for seasonal flu predictions. The dataset was divided into 3 sets, 1 for each classifier. The first set consisted of tweets that were labeled as either flu-related tweets or unrelated. The second one had tweets with labels of flu infections or flu awareness. The tweets in the last set were labeled as either the flu tweet being about the author or about someone else. For our training dataset, we considered the tweets in the second and third datasets as flu-related tweets and combined all of them with only 2 labels: flu-related or unrelated.

Sanders Dataset

The labeled dataset obtained from Sanders [25] was prepared manually to train and test sentiment analysis algorithms. Each record in the dataset was annotated with a sentiment label, indicating a feeling toward either Google, Twitter, Microsoft or Apple. The labels were as follows: positive, neutral, negative, and irrelevant. Because this dataset was prepared for sentiment analysis of topics that are not related to flu, we used all the tweets in this dataset, with the exception of the ones with irrelevant labels as flu-unrelated tweets.

Application Dataset

For validation purposes, we prepared an application dataset by collecting a set of Twitter posts for the first 20 weeks of the year 2018 within the boundary box of the state of Connecticut as a location filter using its associated longitude and latitude. The data were collected from Twitter SNS using a crawler that works with the Twitter API to stream tweets. The crawler is designed to filter the tweets based on keywords that are directly related to flu and verified by health care professionals. The list contains 11 flu-related keywords: fever, headache, sick, respiratory virus, ache, stuffy nose, dehydration, flu, influenza, contagious, and cough. Due to some technical issues, we were able to collect few twitter documents for the 10th week. Therefore, we did not include the period of the 10th week in our experiments. The total number of tweets over the 19 weeks was 8,440,670.

Centers for Disease Control and Prevention Influenza-Like Illness Network Data

ILI weekly rate produced by the CDC ILINet was used as a gold standard for comparison. The official ILI rates consider outpatients with symptoms of influenza who have visited any location of ILINet-participated health care providers around the United States. The data were obtained from the official CDC website [26].

Data of Hospital Emergency Department Syndromic Surveillance System

These data consist of the number of patients who have visited any location of the emergency departments (EDs) of the

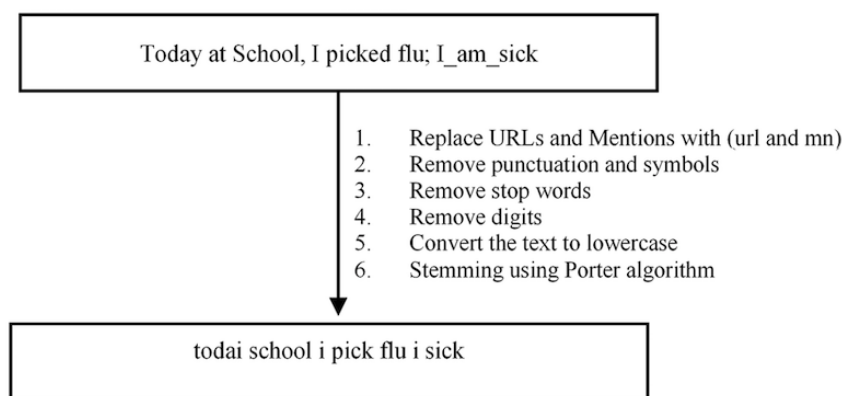
hospitals in Connecticut. Data of Hospital Emergency Department Syndromic Surveillance (HEDSS) generates daily reports about the daily patient visits based on the information received from the EDs. The generated reports include a percentage of patient visits for influenza [27]. These data are used to train the linear regression model for the final flu rate prediction for the state of Connecticut.

Preprocessing

During data preprocessing, stop-words, punctuations, and symbols were removed before the training and testing processes using the NLP toolkit (NLTK) [28]. Stop words such as “the” or “are” are very frequent and may lead to inaccurate classification results if used as features. The preprocessing also includes stemming that is used to reduce words to their roots. There are many stemming algorithms available for use. For this study, the stemming algorithm employed was Porter stemming. It is one of the most commonly used stemming algorithms. It is a rule-based algorithm with 5 steps that is designed based on the idea that English suffixes are made of smaller and simpler ones. A suffix is removed if a rule in the 5 steps passes the conditions and is then accepted [29]. Figure 2 shows the overall preprocessing steps.

URLs, hashtags, and mentions (MN) in the tweets were kept in the corpus. They can be used as features for classification. URLs were replaced with the keyword (url), and MN were replaced with the keyword (mn) to be used as one feature for classification.

Figure 2. Text preprocessing.



Feature Extraction

A maximum classification accuracy can be achieved by selecting the best set of features. Therefore, feature selection is a crucial process in any classification problem. In text classification, the set of features is a subset of words (n -gram) that can be used to distinguish between different classes. The selected words should provide useful information to be used for classification purposes. Thus, it is important to consider different techniques to convert the text in a way that can be processed to gain the required information. In this work, we considered additional features to enhance the classification accuracy. The additional features are

sentiment based features, stylometric features, and flu-related keyword features (Algorithm 1 in Multimedia Appendix 1).

Textual Features

The default features in text classification are the terms and words that make up the document/text. Text classifiers are trained and tested using n -gram features, as basic features, by breaking down the documents/texts into single words (uni-grams), terms composed of 2 words (bi-grams), and terms composed of 3 words (tri-grams) and/or more. A basic technique in text classification is to count n -gram features including the uninformative ones that may yield inaccurate results. Therefore, it is important to use smarter techniques. One of these techniques

is the word/term weighting technique, which weighs the count for every word/term in the text. There are different techniques of word weighting, which include Boolean weighting, term frequency weighting (TF), inverse document frequency weighting (IDF), and TF-IDF. Among the 4 types of word weighting techniques, only the IDF and TF-IDF techniques consider the importance of a word/term in the entire corpus instead of the importance of the word/term in only a document. It has been shown in [22] that TF-IDF is more accurate than IDF. Therefore, in this study, we used TF-IDF to weigh the n -gram features for the conventional machine learning classifiers.

TF-IDF value is obtained by multiplying the TF value by the value of IDF (Equation 1 in Figure 3). TF is the ratio between

the term t with frequency n_t in a given document d and the total numbers of terms n in the document d (Equation 2 in Figure 3). IDF is the inverse of the number of documents that has the term t at least once. It is calculated using Equation 3 in Figure 3, which is the ratio between the frequency N_d of the documents d that have term t , and the total number N of documents d in the analyzed corpus.

For the FT classifier, the representations of textual features of a document are averaged and weighted to be fed to the classifier. For word ordering, FT uses only partial information about the order by using bag of n -grams instead of BoW with the full information of the word ordering [30].

Figure 3. Term frequency-inverse document frequency calculations. TF: term frequency; IDF: inverse document frequency.

$$TF - IDF (t, d) = TF(t, d) \times IDF(t) \quad (1)$$

$$TF(t, d) = \frac{n_t}{n} \quad (2)$$

$$IDF(t) = \frac{N_d}{N} \quad (3)$$

Stylometric Features

Stylometric features of twitter posts include retweets, MN, and URL links. These features were kept in the corpus to be used for classification. URL links and MN to others were preprocessed by replacing them to url and mn keywords.

Topic-Related Keywords-Based Features

It is common to use seed words in text classification. For example, in sentiment analysis, a list of words, including nice and good, is used for positive sentiment and another list of words, including bad and poor, can be used for negative sentiment. In this study, a set of flu-related keywords/terms were used as a set of features for flu-related tweets. The list includes some important influenza-related keywords, symptoms, and treatments. The list of the keywords is kept in an array, and then each tweet is compared against these keywords to keep track of their occurrences.

Sentiment-Based Features

Sentiment analysis is the process of extracting the sentiment of a text using contextual polarity. It is commonly used in classifying reviews of different products in the Internet such as the sentiment of movies. In this study, we used TextBlob library to assign a sentiment to each tweet [31]. TextBlob is a Python library that is used to analyze textual data. On the basis of the polarity score of a tweet, a sentiment value is assigned to the text: positive or negative.

Classification Model Building: Training and Testing

For the sake of accuracy and efficiency, various classifiers are evaluated, including FT and 6 conventional machine learning algorithms [32,33].

FastText

FT is a text classifier. It has been shown that FT produces accurate classification results that are comparable with the results produced by deep neural network classifiers. In addition, it has been shown that the processes of FT training and classification are very fast using a standard computer with a multicore processor. An FT model can be trained using billions of words in just a few minutes, and it can classify about 500,000 sentences in less than a minute [30].

FT utilizes several techniques to enhance the efficiency. It is a linear-based model, scaled to very large data and large output space using a rank constraint and a fast loss approximation. It uses a hierarchal softmax function for a faster search. In addition, only partial information about the word order is used for prediction. Furthermore, it uses the technique of hashing for textual feature mapping [30].

Conventional Machine Learning Classifiers

For training and testing, several supervised classification methods were evaluated to determine one with better classification accuracy [33]. The evaluated conventional classifiers include Random Forest, Naïve Bayes, SVM, C 4.5 Decision Tree, K-nearest neighbors (KNN), and AdaBoost. The preprocessed labeled dataset was used to train and test the model of different classifiers using 10-fold cross validation as the experimental setting. The 10-fold cross validation is a method to validate the studied/built model by iterating through the labeled data 10 times with different subsets of training and testing for each iteration.

Mapping

For weekly rates, a MapReduce (MR) approach was used to process the large dataset of tweets. MR consists of 2 main functions: Map and Reduce. The Map function takes an input as a pair (week number and post), groups all the posts associated with the same week number, and generates intermediate pairs to be passed to the Reduce function. The Reduce function merges all the pairs with the same week number after processing the associated values such as counting or summing them up [34].

Weekly Flu Rate Estimation Based on Regression

To predict the influenza rate at a certain week, we used a regression-based estimator. The proposed flu rate estimator has been evaluated using different regression models. In addition to the linear regression model, 3 different regression techniques were evaluated to determine the one with better estimation accuracy.

A regression model should be trained (fitted) using available data of flu rates, such as the data obtained from FluNearYou [35]—a Web application that uses weekly surveys to collect health status of individuals or the data of flu emergency visits obtained from HEDSS. For this study, we used the data of

Figure 4. Proposed linear regression model.

$$F_w = \beta + \alpha_1 \frac{1}{2016 - 1998} \sum_{y=1998}^{2016} F_w^y + \alpha_2 T_w \quad (4)$$

Other Regression Models

In addition to our proposed linear regression model, 3 different regression techniques were evaluated to determine the technique with better estimation accuracy. The evaluated techniques are polynomial regression, logistic regression, and support vector regression. The measure of Pearson correlation (r) is used to find the most accurate model to be used for the final weekly flu rate estimation.

Results

Classification Results

The results show that the proposed model improves the performance of flu post classifications using a combination of the additional features. The performance results of the used classifiers are shown in Table 1 using the precision, recall and F -measure metrics. The Random Forest method achieved the highest accuracy results, with an F -measure of 90.1%. In addition, we used the receiver operating characteristic (ROC) metric to evaluate the used classifiers. ROC is a curve with points that represent the pair of TP rate (sensitivity) and false

HEDSS for regression model training, where the average ILI rates of previous years and rates of flu-related tweets obtained from the classification results are passed to the regression model as predictors. The regression model was then tested and validated using CDC ILINet data.

Linear Regression Model

Linear regression is used when the dependent variable (response) is continuous and the independent variables (predictors) are either continuous or discrete, and the relationship between the dependent and independent variable(s) is linear. The linear regression indicates that the rate in the change of the mean of the response value is constant with respect to the value of the predictor(s). Therefore, the relationship is represented by an equation of a line [36].

Using the proposed predictors that include a combination of the rate of flu tweets and the average ILI rate of the same week number of past years (from 1998 to 2016), our proposed linear regression model has the following form (Figure 4), where indicates the flu rate at week w , β is the intercept which is the mean value of when all predictors are (0), values represent the regression coefficients, is the actual rate of flu incidents in week w of year y , and is the rate of flu tweets in week w .

positive rate (specificity). A perfect curve is the one that passes through the upper left corner representing 100% sensitivity and 100% specificity. Thus, the closer the curve is to that corner, the better the accuracy is [37]. As shown in Figure 5, Random Forest appears to be the best classifier. The high accuracy results demonstrate the efficiency and effectiveness of the extracted features.

Moreover, the performance results of FT with different sets of features, is presented in Figure 6. The overall accuracy using the F -measure metric ranges between 86.47% and 89.9%. This demonstrates the efficiency of the FT classifier. The highest classification accuracy is achieved by using the 5-gram features, together with all the proposed additional features (F -measure=89.9%) in only 21.53 seconds for training and testing using 10-fold cross validation on a standard computer (2.6 GHz Intel Core *i7* processor, and 16 GB RAM). It has been shown that FT can produce, in a short time, accurate results that are comparable with the results produced by the state-of-the-art deep neural network classifiers [30]. The high accuracy, together with the efficiency of FT make it an optimal classifier for flu disease surveillance models/systems with very large data. Therefore, FT will be used for our further analysis.

Table 1. Performance of classifiers.

Classifier name	Precision	Recall	<i>F</i> -measure
C4.5 Decision Tree	0.876	0.85	0.873
Random Forest	0.905	0.902	0.901
Support vector machine	0.883	0.883	0.883
Naïve Bayes	0.846	0.826	0.824
AdaBoost	0.867	0.864	0.864
K-nearest neighbors	0.874	0.872	0.872
FastText	0.899	0.899	0.899

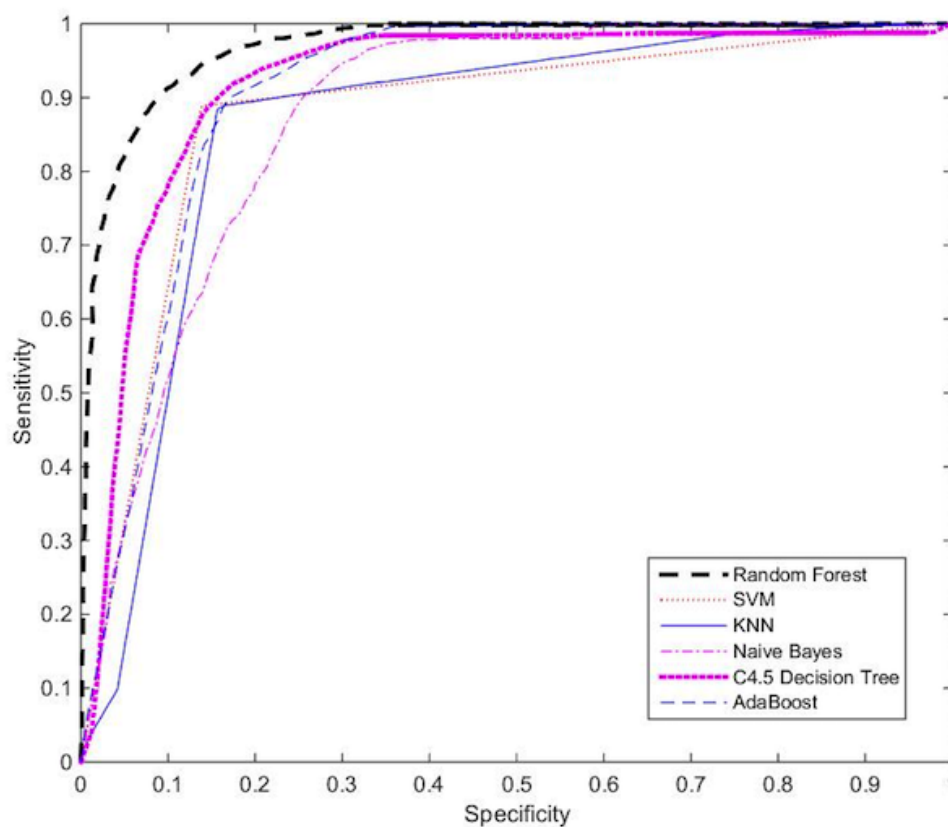
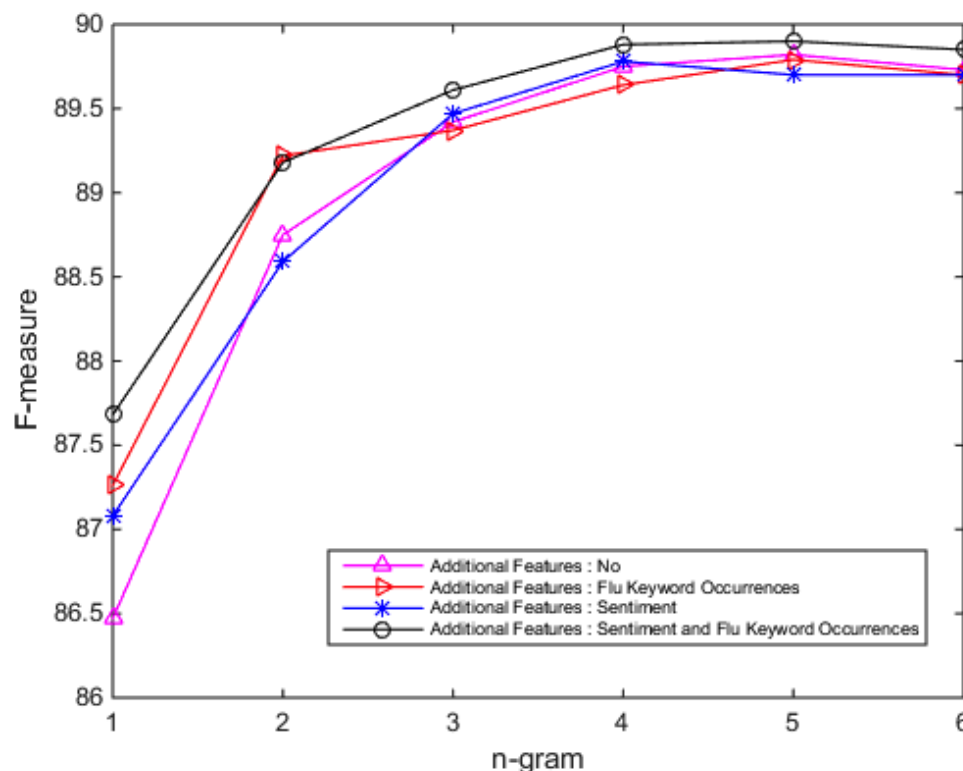
Figure 5. Performance comparison using receiver operating characteristic. SVM: support vector machine; KNN: K-nearest neighbors.

Figure 6. FastText performance using different sets of features.

Many studies have used the available data from Twitter to build faster influenza surveillance systems [17]. All the reviewed studies use conventional machine learning methods to distinguish between flu-relevant and flu-irrelevant posts for further analysis. A summary of the performance results of previous works, which include tweet classification for

Twitter-based flu surveillance systems, is shown in Table 2. The metrics are reported as percentages. The evaluation of flu tweet classification using the *F*-measure shows that the proposed framework using FT, together with the extracted features, achieved high accuracy with *F*-measure value of 89.9%.

Table 2. Summary of the reviewed flu posts classifiers (flu-relevant/flu-irrelevant).

Study	Classifier name	Precision	Recall	<i>F</i> -measure	Note
Broniatowski et al [18], Lamb et al [19]	SVM ^a and Logistic Regression	67	87	75.62	Multilevel classification
Santos and Matos [21]	Naïve Bayes	N/A ^b	N/A	83	— ^c
Aramaki et al [20]	SVM	N/A	N/A	75.6	—
Cui et al [22]	SVM	87.49	92.28	89.68	—
Byrd et al [23]	Naïve Bayes	N/A	N/A	N/A	70% accuracy
Proposed framework	Random Forest	90.5	90.2	90.1	—
Proposed framework	FastText	89.9	89.9	89.9	—

^aSVM: support vector machine.

^bN/A: not applicable.

^cNot available.

Weekly Flu Rate Estimation Results

The framework was evaluated by applying the trained FT model on the application data, which includes over 8,400,000 tweets, for classification. Then, the classification results together with the historical CDC data were passed on to the proposed regression-based estimator as predictors to obtain weekly flu-rates. The results of the flu estimator show a highly correlated output to the gold standard data (CDC). The estimator was evaluated using several regression models. Every model

was fitted using the data of flu emergency visits obtained from HEDSS. Then, it was tested on CDC ILINet data from January 1, 2018 to May 19, 2018.

The performance results of the proposed flu rate estimator based on different regression models are shown in Table 3. The table presents the accuracy results using the Pearson correlation measure *r*.

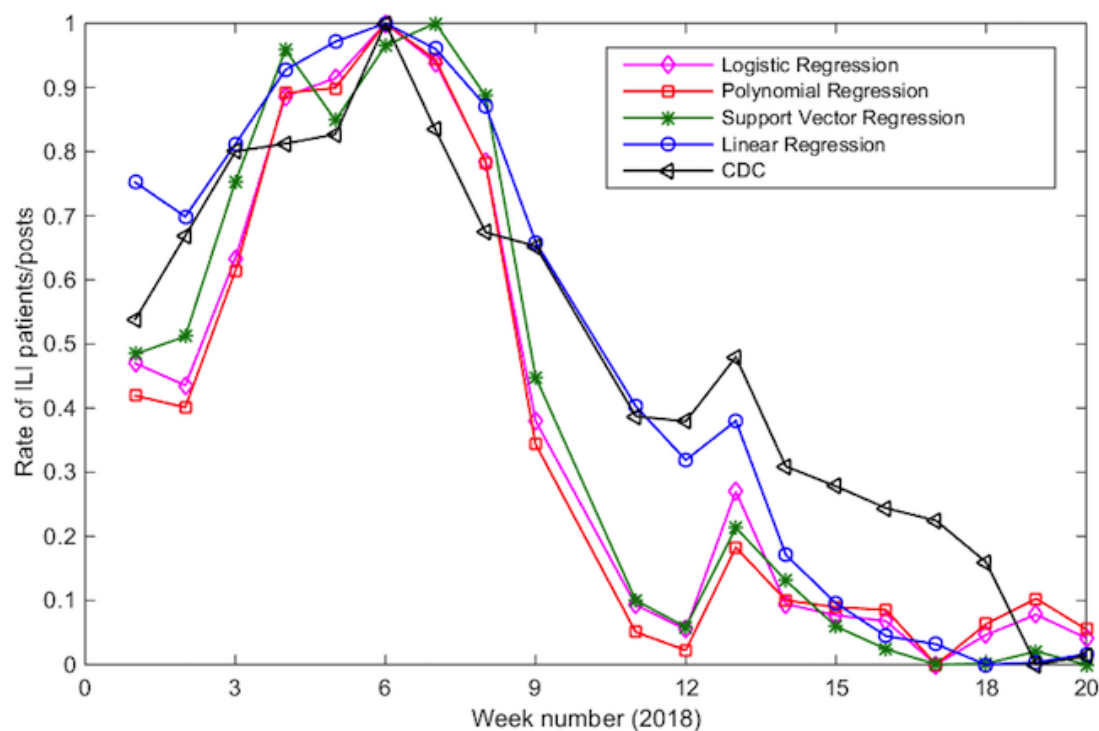
The linear regression-based estimator achieved the highest accuracy results, with a Pearson correlation of 96.2%. Figure 7 also shows that linear regression is the most correlated model with the ground truth (CDC). It shows the normalized rate of ILI patients obtained from the CDC and the normalized rate of

ILI Twitter posts obtained from the output of our proposed solution during the period of January through May of 2018 for the state of Connecticut. The rate values of the proposed framework and ILINet are normalized to a common scale for comparison.

Table 3. Performance of flu rate estimator using different regression models.

Regression model	<i>r</i> value
Polynomial regression	0.895
Logistic regression	0.917
Support vector regression	0.930
Linear regression	0.962

Figure 7. Correlation between the proposed framework and CDC influenza-like illness rate using different regression models. CDC: Centers for Disease Control and Prevention; ILI: influenza like illness.



Discussion

FastText Versus Conventional Machine Learning Classifiers

To build a classification model with better accuracy and efficiency, FT and several supervised classification methods using the proposed additional features were evaluated. In addition to FT, the evaluated classifiers are Random Forest, Naïve Bayes, SVM, C4.5 decision tree, KNN, and AdaBoost. The preprocessed labeled dataset was used to train and test models of the different classifiers with the TF-IDF-based *n*-gram features and the proposed additional ones, which are presented in the Feature Extraction Section.

Computational Complexity

The experiments show that FT produces accurate classification results in only 21.53 seconds for training and testing using

10-fold cross validation on a standard computer (2.6 GHz Intel Core *i7* processor, and 16 GB RAM). FT is an efficient linear-based model. It uses a hierarchical softmax function that reduces the computational complexity to become logarithmic $O(\log n)$, leading to faster classification training and testing [30]. For word ordering, only partial information about the order is used by using a bag of *n*-grams instead of a BoW with the full information of the word ordering. For more efficiency, the bag of *n*-grams are mapped using hashing techniques [30]. On the other hand, the experiments show that Random Forest, which is the most accurate conventional classifier in our experiment with F-measure value of 90.1, requires a longer time (39 min and 26 seconds) for training and testing using the experimental settings. The worst time complexity of Random Forest is quadratic for training $O(n^2 \log n)$ and linear for prediction $O(n)$ [38]. This, together with the experimental results, demonstrates the efficiency and the accuracy of FT classifier. FT is an optimal classifier to detect new outbreaks with new signs and symptoms

published in posts of SNSs. Therefore, FT has been adopted for further analysis in our proposed framework.

FastText as a Flu Post Classification Module

For a better FT model, we evaluated 28 different feature settings using FT, with the parameter values of learning rate of 0.8 and epoch of 8, to determine the best feature set. Initially, the model was trained and tested using 1 setting of n -gram features ($n=1-6$), which are tokens of (n) words including the stylometric features. Then, different settings of the additional features are combined with the tweet text for training and testing using n -grams ($n=1-6$). The settings include a combination of text and sentiment features, a combination of the text and keyword occurrence features, and a combination of all additional features (text + sentiment + hasKeyword):

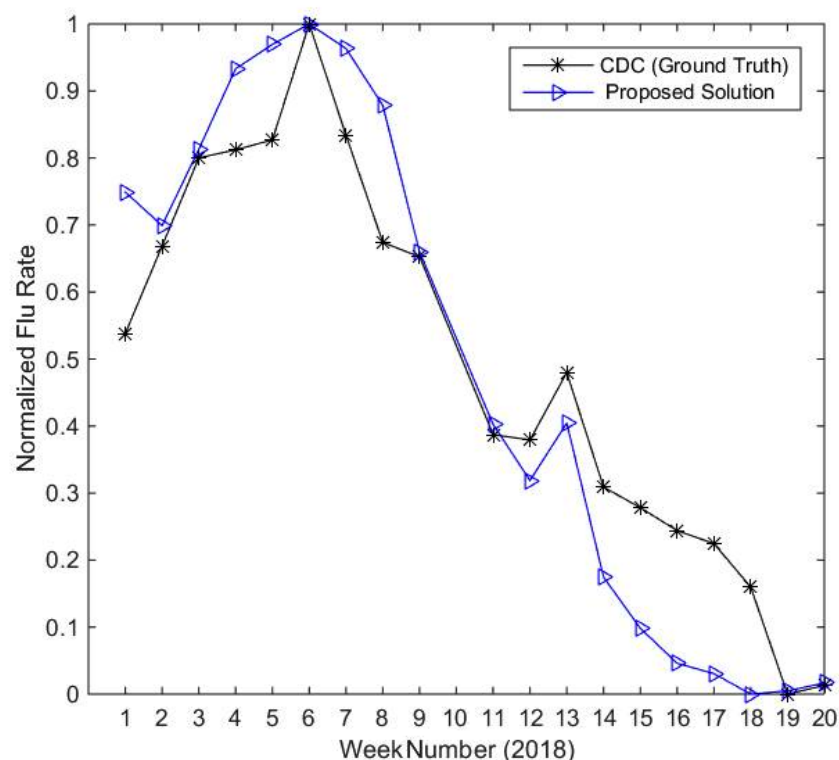
```
__label__<related/unrelated>TEXT _sent_<neg/pos>
_hasKeywrд_<yes/no>
```

With a standard computer (2.6 GHz Intel Core *i7* processor, and 16 GB RAM), the preprocessed labeled dataset was used to train and test the models using 10-fold cross validation as well.

Linear Regression as a Weekly Flu Rate Estimation Module

In addition to the efficiency of linear regression, the experimental results, as shown in Figure 8, demonstrate the model accuracy and confirm the linear relationship between the rates of weekly flu (dependent variable) and flu-related tweets (independent variable). Therefore, the linear regression model is used for the weekly flu rate estimation module.

Figure 8. Correlation between the proposed framework and CDC influenza-like illness rate. CDC: Centers for Disease Control and Prevention.



Statistical Power Analysis

Power analysis has been performed to justify and ensure the appropriateness of the number of instances that are used for this study. Experimental results show that the accuracy of flu tweet classification using FT with the proposed additional features outperform FT with only textual features. Therefore, power analysis is also used to prove this hypothesis, which is stated as an alternative hypothesis H_a , whereas the null hypothesis H_0 is the hypothesis where there is no change in the accuracy using proposed features with respect to only textual features. With the power analysis, a statistical test rejects the null hypothesis when it is false. With this, one can conclude that there is a difference between the accuracies (better accuracy) using additional features and can confirm our alternative hypothesis H_a . If the null hypothesis is not rejected, then the alternative hypothesis should be rejected. The opposing hypotheses for our

work can be stated as shown in Figure 9, where μ_{proposed} is the accuracy average of FT using the proposed additional features, and μ_{textual} is the accuracy average of FT using only textual features for flu tweet classification.

To determine the required sample size n , 4 parameters/factors must be known or estimated, which are as follows:

- α : significance level (1% or 5%)
- p : desired power of the test (80%)
- σ : population SD
- d : effect size (the difference between the 2 groups)

The values of the first 2 parameters are generally fixed. The parameter of significance level α is usually set to either .05 or .01 and is the probability of rejecting the null hypothesis when it is true. The power parameter p is the probability that the effect will be detected and is usually set to either 0.8 or 0.9. On the

other hand, the last 2 parameters are problem dependent. For our analysis, the last 2 parameters are estimated based on our previous experiments. Thus, the values of all the 4 required parameters are stated below:

- $\alpha=5\%$
- $p=80\%$
- $\sigma=0.27$
- $d=0.012$

Using these parameters together with the z -test model to obtain z -scores, the sample size n can be computed by using Equation 7 (Figure 10).

Given the estimated values of the required parameters, we will obtain the computations and values shown in Figure 11.

Using the obtained sample size n and the significance level α , the below parameters (in Figure 12) can be computed to apply the z -test and then make a decision on accepting or rejecting our alternative hypothesis.

Because the obtained value of the z -test (18) is higher than the critical value ($18 > 1.96$), the observed difference is significant and shows that the additional features enhance the accuracy of FT to classify flu tweets. In other words, results of the z -test show that the null hypothesis (H_0) should be rejected, and the sample set of 7941 tweets is sufficient to prove that FT with the proposed additional features is more accurate than FT with only textual features for flu tweet classification. Our experiments included over 10,000 tweets, which is more than enough to prove the hypothesis claims.

Figure 9. The 2 opposing hypotheses.

$$H_0: \mu_{proposed} = \mu_{textual} \quad (5)$$

$$H_a: \mu_{proposed} > \mu_{textual} \quad (6)$$

Figure 10. Sample size formulation.

$$Sample\ size\ (n) = 2 \times \left(\sigma \times \frac{z_{1-\frac{\alpha}{2}} + z_p}{d} \right)^2 \quad (7)$$

Figure 11. Sample size computation results.

$$Sample\ size\ (n) = 2 \times \left(0.27 \times \frac{z_{1-\frac{0.05}{2}} + z_{0.8}}{0.012} \right)^2$$

$$Sample\ size\ (n) = 2 \times \left(0.27 \times \frac{1.959 + 0.8416}{0.012} \right)^2$$

$$Sample\ size\ (n) = 7941$$

Figure 12. Z-test computation and analysis to determine whether to accept or reject the alternative hypothesis.

$$\text{Mean } (\hat{x}) = \frac{\sum x}{n} = \frac{7292}{7941} = 0.918 \quad (8)$$

$$\text{Variance } (\sigma^2) = \frac{\sum(x - \hat{x})^2}{n} = 0.075 \quad (9)$$

$$\text{Standard Deviation } (\sigma) = \sqrt{\sigma^2} = 0.27 \quad (10)$$

$$\text{Critical } z = z_{1-\frac{\alpha}{2}} = 1.96 \quad (11)$$

$$\text{Standard Error}(Sx) = \frac{\sigma}{\sqrt{n}} = 0.003 \quad (12)$$

$$\text{Lower limit} = \hat{x} - \text{Critical } z \times Sx = 0.912 \quad (13)$$

$$\text{Upper limit} = \hat{x} + \text{Critical } z \times Sx = 0.923 \quad (14)$$

$$\text{Null Hypothesis } (H_0): \mu_{\text{proposed}} = \mu_{\text{textual}} = 0.864 \quad (15)$$

$$Z_{\text{test}}(Z) = \frac{\hat{x} - \mu_{\text{textual}}}{Sx} = 18 \quad (16)$$

Performance Metrics

In this section, we present the used performance metrics. The performance of the classifiers are evaluated using different metrics presented in Figure 13, which are as follows: accuracy (Equation 17), precision (Equation 18), recall (Equation 19), and *F*-measure (Equation 20). These metrics are used to provide a better overview of the model performance. The accuracy measure by itself is not a perfect measure if the dataset is not balanced. Precision and recall are better measures in the case of imbalanced datasets. The selected metrics can be computed using true positive (*TP*), true negative (*TN*), false positive (*FP*), and false negative (*FN*) measures, where *TP* refers to the rate of correctly classified instances as positive, *TN* refers to the rate of correctly classified instances as negative, *FP* refers to the rate of incorrectly classified instances as positive, and *FN* refers to the rate of incorrectly classified instances as negative. In this work, we mainly use *F*-measure as a performance metric for evaluation and comparison. *F*-measure is a weighted average

of 2 different performance metrics: precision and recall. Its value ranges between 0 (worst) and 1 (best).

In addition, the performance of flu rate estimation is evaluated using Pearson correlation. It is a metric that evaluates the correlation between 2 datasets using the symbol *r* that ranges between (1) and (− 1): the value of *r*= 1 when both datasets exactly match and the value of *r*= 0 when there is no correlation between the 2 datasets. An available ground truth is usually used to evaluate the quality of the results of the proposed methods and frameworks. For this study, we used the recent CDC weekly reports as the ground truth to be compared with the proposed solution. Let be the observed value of the ground truth (CDC ILINet data), *x_i* be the predicted weekly flu rate value, and *y* and *x* be the average values of *y_i* and *x_i*, respectively. Using these notations, the Pearson correlation value *r* is defined as shown in Equation 21, illustrated in Figure 14 [39].

Figure 13. Performance metrics used to evaluate the proposed work. FN: false negative; FP: false positive; TP: true positive.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (17)$$

$$Precision = \frac{TP}{TP + FP} \quad (18)$$

$$Recall = \frac{TP}{TP + FN} \quad (19)$$

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (20)$$

Figure 14. Pearson correlation value computation.

$$r = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (21)$$

As shown in [Table 4](#) and depicted in [Figure 8](#), the results show a strong correlation (96.29% Pearson correlation) between the output of the proposed framework and the CDC reports. This correlation percentage shows that our proposed solution provides accurate results on par with the best results in our survey, while being more efficient (faster). In addition, we believe that this

is the first work that uses Twitter postings for flu trend predictions in the state of Connecticut with strong correlated results. To the best of our knowledge, this is also the first work that shows a Twitter-based solution for flu prediction using recent data that were collected in the year of 2018.

Table 4. Summary of the reviewed studies with reported Pearson correlation.

Study	Time frame	Location	r value
Broniatowski et al [18]	September 2012-May 2013	United States	0.93
Aramaki et al [20]	November 2008-June 2010	Japan	0.89
Santos and Matos [21]	March 2010-February 2012	Portugal	0.89
Lamb et al [19]	May 2009-October 2010	United States	0.9897
Cui et al [22]	September 2013-December 2013	China	— ^a
Byrd et al [23]	October 2015-November 2015	Ottawa	—
Proposed framework	January 2018-May 2018	Connecticut, United States	0.9629

^aNot applicable.

Conclusions

For disease surveillance models, gathering related information about diseases and then issuing necessary reports at an early stage is crucial for outbreak prevention. Data of microblogging sites, such as Twitter, have become popular enough to be used as triggers for different event prediction such as disease outbreaks. Recently, many studies have used these data to build effective epidemic prediction models such as flu outbreak prediction. It has been observed in the literature that most of the models use conventional machine learning methods to filter

and distinguish between the flu-relevant and irrelevant posts for further analysis. In this study, we introduced a framework based on FT, a state-of-the-art text classifier that uses the features of sentiment analysis and flu keyword occurrences for faster classification. Thereafter, a combination of the classified Twitter documents and historical CDC data was passed to a linear regression-based module for weekly flu rate predictions. The results demonstrated the efficiency and the accuracy of the proposed framework. The final predicted flu trend using Twitter documents showed a strong Pearson correlation of 96.29 % with the ground truth data of CDC for the first few months of 2018.

Authors' Contributions

The manuscript was written by AA and MF. AA and MF conceived the ideas. AA performed the literature survey, simulations, and analysis under the supervision of MF. Both authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional feature extraction algorithm.

[[PDF File \(Adobe PDF File\), 68KB - publichealth_v5i2e12383_app1.pdf](#)]

References

1. 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 Apr;15(4):e85 [[FREE Full text](#)] [doi: [10.2196/jmir.1933](#)] [Medline: [23615206](#)]
2. Corley CD, Cook DJ, Mikler AR, Singh KP. Text and structural data mining of influenza mentions in web and social media. *Int J Environ Res Public Health* 2010 Feb;7(2):596-615 [[FREE Full text](#)] [doi: [10.3390/ijerph7020596](#)] [Medline: [20616993](#)]
3. Zhang LW, Zhu DH. Research of technical development trend and hot points based on text mining. : IEEE; 2010 Dec Presented at: 2010 2nd International Conference on Information Engineering and Computer Science (ICIECS); December 25-26, 2010; Wuhan, China p. 1-5.
4. Polgreen PM, Chen Y, Pennock DM, Nelson FD. Using internet searches for influenza surveillance. *Clin Infect Dis* 2008 Dec 1;47(11):1443-1448. [doi: [10.1086/593098](#)] [Medline: [18954267](#)]
5. Goel S, Hofman JM, Lahaie S, Pennock DM, Watts DJ. Predicting consumer behavior with web search. *Proc Natl Acad Sci U S A* 2010 Oct 12;107(41):17486-17490 [[FREE Full text](#)] [doi: [10.1073/pnas.1005962107](#)] [Medline: [20876140](#)]
6. Scharnow M, Vogelgesang J. Measuring the public agenda using search engine queries. *Int J Public Opin Res* 2011 Mar 1;23(1):104-113. [doi: [10.1093/ijpor/edq048](#)]
7. Dugas AF, Hsieh YH, Levin SR, Pines JM, Mareiniss DP, Mohareb A, et al. Google Flu Trends: correlation with emergency department influenza rates and crowding metrics. *Clin Infect Dis* 2012 Feb 15;54(4):463-469 [[FREE Full text](#)] [doi: [10.1093/cid/cir883](#)] [Medline: [22230244](#)]
8. Morrison JL, Breitling R, Higham DJ, Gilbert DR. GeneRank: using search engine technology for the analysis of microarray experiments. *BMC Bioinformatics* 2005 Sep 21;6:233 [[FREE Full text](#)] [doi: [10.1186/1471-2105-6-233](#)] [Medline: [16176585](#)]
9. 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](#)]
10. Suh B, Hong L, Pirolli P, Chi E. Want to be retweeted? Large scale analytics on factors impacting retweet in twitter network. 2010 Aug Presented at: Second IEEE international conference on Social computing; August 20-22, 2010; Minneapolis, MN, USA p. 177-184. [doi: [10.1109/SocialCom.2010.33](#)]
11. Wakamiya S, Morita M, Kano Y, Ohkuma T, Aramaki E. Tweet classification toward Twitter-based disease surveillance: new data, methods, and evaluations. *J Med Internet Res* 2019 Feb 20;21(2):e12783 [[FREE Full text](#)] [doi: [10.2196/12783](#)] [Medline: [30785407](#)]
12. Kagashe I, Yan Z, Suheryani I. Enhancing seasonal influenza surveillance: topic analysis of widely used medicinal drugs using Twitter data. *J Med Internet Res* 2017 Dec 12;19(9):e315 [[FREE Full text](#)] [doi: [10.2196/jmir.7393](#)] [Medline: [28899847](#)]
13. Woo H, Cho Y, Shim E, Lee J, Lee C, Kim S. Estimating influenza outbreaks using both search engine query data and social media data in South Korea. *J Med Internet Res* 2016 Dec 4;18(7):e177 [[FREE Full text](#)] [doi: [10.2196/jmir.4955](#)] [Medline: [27377323](#)]
14. Xue H, Bai Y, Hu H, Liang H. Regional level influenza study based on Twitter and machine learning method. *PLoS One* 2019;14(4):e0215600 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0215600](#)] [Medline: [31013324](#)]
15. Davidson MW, Haim DA, Radin JM. Using networks to combine. *Sci Rep* 2015 Jan 29;5:8154 [[FREE Full text](#)] [doi: [10.1038/srep08154](#)] [Medline: [25634021](#)]
16. Culotta A. Towards detecting influenza epidemics by analyzing Twitter messages. In: Proceedings of the First Workshop on Social Media Analytics. 2010 Jul Presented at: SOMA'10; July 25-28, 2010; Washington DC p. 115-122. [doi: [10.1145/1964858.1964874](#)]
17. Alessa A, Faezipour M. A review of influenza detection and prediction through social networking sites. *Theor Biol Med Model* 2018 Dec 1;15(1):2 [[FREE Full text](#)] [doi: [10.1186/s12976-017-0074-5](#)] [Medline: [29386017](#)]

18. Broniatowski DA, Paul MJ, Dredze M. National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic. *PLoS One* 2013;8(12):e83672 [FREE Full text] [doi: [10.1371/journal.pone.0083672](https://doi.org/10.1371/journal.pone.0083672)] [Medline: [24349542](https://pubmed.ncbi.nlm.nih.gov/24349542/)]
19. Lamb A, Paul M, Dredze M. Separating fact from fear: Tracking flu infections on twitter. Stroudsburg, PA: Association for Computational Linguistics; 2013 Jun Presented at: Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; June 9-14, 2013; Atlanta, GA p. 789-795 URL: <https://www.aclweb.org/anthology/N13-1097>
20. Aramaki E, Maskawa S, Morita M. Twitter catches the flu: detecting influenza epidemics using twitter. In: Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA: Association for Computational Linguistics; 2011 Jul Presented at: EMNLP'11; July 27-31, 2011; Edinburgh, United Kingdom p. 1568-1576 URL: <https://www.aclweb.org/anthology/D11-1145>
21. Santos JC, Matos S. Analysing Twitter and web queries for flu trend prediction. *Theor Biol Med Model* 2014 May 7;11(Suppl 1):S6 [FREE Full text] [doi: [10.1186/1742-4682-11-S1-S6](https://doi.org/10.1186/1742-4682-11-S1-S6)] [Medline: [25077431](https://pubmed.ncbi.nlm.nih.gov/25077431/)]
22. Cui X, Yang N, Wang Z, Hu C, Zhu W, Li H, et al. Chinese social media analysis for disease surveillance. *Pers Ubiquit Comput* 2015 Sep 11;19(7):1125-1132. [doi: [10.1007/s00779-015-0877-5](https://doi.org/10.1007/s00779-015-0877-5)]
23. Byrd K, Mansurov A, Baysal O. Mining twitter data for influenza detection and surveillance. : IEEE; 2016 May Presented at: 2016 IEEE/ACM International Workshop on Software Engineering in Healthcare Systems (SEHS); May 14-15, 2016; Austin, TX.
24. Wirth R, Hipp J. CRISP-DM: Towards a standard process model for data mining. Blackpool, United Kingdom: Practical Application Company; 2000 Apr Presented at: Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining; April 11-13, 2000; Manchester, UK p. 29-39.
25. Sanders NJ. Sanders Analytics. Sanders twitter sentiment corpus URL: <http://www.sananalytics.com/lab/twitter-sentiment> [WebCite Cache ID [6joVh9V4R](https://www.webcitation.org/6joVh9V4R)]
26. Centers for Disease Control and Prevention. FluView URL: <https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html> [WebCite Cache ID [72oxwpUzu](https://www.webcitation.org/72oxwpUzu)]
27. Dembek ZF, Carley K, Hadler J. Guidelines for constructing a statewide hospital syndromic surveillance network. *MMWR Suppl* 2005 Aug 26;54:21-24. [Medline: [16177688](https://pubmed.ncbi.nlm.nih.gov/16177688/)]
28. Loper E, Bird S. NLTK: The Natural Language Toolkit. Stroudsburg, PA: Association for Computational Linguistics; 2002 Jul 07 Presented at: ETMTNLP '02 Proceedings of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics - Volume 1; July 7, 2002; Philadelphia, PA URL: <https://www.aclweb.org/anthology/W02-0109>
29. Singh J, Gupta V. A systematic review of text stemming techniques. *Artif Intell Rev* 2016 Aug 1;48(2):157-217. [doi: [10.1007/s10462-016-9498-2](https://doi.org/10.1007/s10462-016-9498-2)]
30. Joulin A, Grave E, Bojanowski P, Mikolov T. Bag of tricks for efficient text classification. 2017 Apr Presented at: 15th Conference of the European Chapter of the Association for Computational Linguistics; April 3-7, 2017; Valencia, Spain p. 427-431.
31. Loria S, Keen P, Honnibal M, Yankovsky R, Karesh D, Dempsey E. Textblob: simplified text processing. *Secondary TextBlob: Simplified Text Processing*; 2014. URL: <https://textblob.readthedocs.io/en/dev/> [accessed 2019-06-07] [WebCite Cache ID [78wtnJ9VQ](https://www.webcitation.org/78wtnJ9VQ)]
32. Alessa A, Faezipour M. Tweet Classification Using Sentiment Analysis Features and TF-IDF Weighting for Improved Flu Trend Detection. In: *Machine Learning and Data Mining in Pattern Recognition: 14th International Conference, MLDM 2018 Proceedings, Part I*: Springer; 2018 Jul 08 Presented at: International Conference on Machine Learning and Data Mining in Pattern Recognition; July 15-19, 2018; New York, NY p. 174-186.
33. Alessa A, Faezipour M, Alhassan Z. Text classification of flu-related tweets using fasttext with sentiment and keyword features. : IEEE; 2018 Jun Presented at: 2018 IEEE International Conference on Healthcare Informatics (ICHI); June 4-7, 2018; New York, NY p. 366-367.
34. Al Essa A, Faezipour M. Mapreduce and spark-based analytic framework using social media data for earlier flu outbreak detection. In: *Advances in Data Mining, Applications and Theoretical Aspects*. New York, NY: Springer International Publishing; 2017:246-257.
35. Chunara R, Aman S, Smolinski M, Brownstein J. Flu near you: an online self-reported influenza surveillance system in the us. *Online J Public Health Inform* 2013;5(1) [FREE Full text] [doi: [10.5210/ojphi.v5i1.4456](https://doi.org/10.5210/ojphi.v5i1.4456)]
36. Rawlings J, Pantula S, Dickey D. *Applied Regression Analysis: A Research Tool*. New York, NY: Springer; 2001.
37. Zweig M, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 1993 Apr;39(4):561-577 [FREE Full text] [Medline: [8472349](https://pubmed.ncbi.nlm.nih.gov/8472349/)]
38. Louppe G. arXiv. Wallonia, Belgium: University of Liege; 2014. Understanding random forests: From theory to practice URL: <https://arxiv.org/pdf/1407.7502> [accessed 2019-06-07] [WebCite Cache ID [78wuTYlkk](https://www.webcitation.org/78wuTYlkk)]
39. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining search, social media, and traditional data sources to improve influenza surveillance. *PLoS Comput Biol* 2015 Oct;11(10):e1004513 [FREE Full text] [doi: [10.1371/journal.pcbi.1004513](https://doi.org/10.1371/journal.pcbi.1004513)] [Medline: [26513245](https://pubmed.ncbi.nlm.nih.gov/26513245/)]

Abbreviations

CDC: Centers for Disease Control and Prevention
ED: emergency departments
FN: false negative
FP: false positive
FT: FastText
HEDSS: Data of Hospital Emergency Department Syndromic Surveillance
IDF: inverse document frequency
ILINet: Influenza Like Illness Surveillance Network
KNN: K-nearest neighbors
MI: mutual information
MN: mentions
MR: MapReduce
NLP: natural language processing
NLTK: natural language processing toolkit
ROC: receiver operating characteristic
SNSs: social networking sites
SVM: support vector machine
TF: term frequency
TF-IDF: term frequency-inverse document frequency
TN: true negative
TP: true positive

Edited by G Eysenbach; submitted 07.10.18; peer-reviewed by J Bian, X Wang; comments to author 04.04.19; revised version received 22.05.19; accepted 25.05.19; published 23.06.19.

Please cite as:

Alessa A, Faezipour M

Preliminary Flu Outbreak Prediction Using Twitter Posts Classification and Linear Regression With Historical Centers for Disease Control and Prevention Reports: Prediction Framework Study

JMIR Public Health Surveill 2019;5(2):e12383

URL: <http://publichealth.jmir.org/2019/2/e12383/>

doi: [10.2196/12383](https://doi.org/10.2196/12383)

PMID:

©Ali Alessa, Miad Faezipour. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 23.06.2019. 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

“Where There’s Smoke, There’s Fire”: A Content Analysis of Print and Web-Based News Media Reporting of the Philip Morris–Funded Foundation for a Smoke-Free World

Christina Watts^{1,2}, BHLth; Becky Freeman^{1,3}, PhD

¹School of Public Health, Faculty of Medicine and Health, The University of Sydney, Camperdown, Australia

²Cancer Council New South Wales, Woolloomooloo, Australia

³Charles Perkins Centre, The University of Sydney, Camperdown, Australia

Corresponding Author:

Christina Watts, BHLth

Cancer Council New South Wales

153 Dowling Street

Woolloomooloo,

Australia

Phone: 61 400709347

Email: cmcb8303@uni.sydney.edu.au

Abstract

Background: In September 2017, the Foundation for a Smoke-Free World (FSFW), a not-for-profit organization with a core purpose “to accelerate global efforts to reduce deaths and harm from smoking” was launched. However, the legitimacy of the FSFW’s vision has been questioned by experts in tobacco control because of the organization’s only funding partner, Philip Morris International (PMI).

Objective: This study aimed to examine the response to the FSFW in Web-based and print news media to understand how the FSFW and its funding partner, PMI, were framed.

Methods: News articles published within a 6-month period after the FSFW was announced were downloaded via Google News and Factiva and coded for topic, framing argument, slant, mention of tobacco control policies, and direct quotes or position statements.

Results: A total of 124 news articles were analyzed. The news coverage of the FSFW was framed by 6 key arguments. Over half of the news articles presented a framing argument in opposition to the FSFW (64/124, 51.6%). A further 20.2% (25/124) of articles framed the FSFW positively and 28.2% of articles (35/124) presented a neutral debate with no primary slant. The FSFW was presented as not credible because of the funding link to PMI in 29.0% (36/124) of articles and as a tactic to mislead and undermine effective tobacco control measures in 11.3% of articles (14/124). However, 12.9% of articles (16/124) argued that the FSFW or PMI is part of the solution to reducing the impact of tobacco use. Evidence-based tobacco control policies were mentioned positively in 66.9% (83/124) of news articles and 9.6% (12/124) of articles presented tobacco control policies negatively.

Conclusions: The Web-based and print news media reporting of the formation of the FSFW and its mission and vision has primarily been framed by doubt, skepticism, and disapproval.

(*JMIR Public Health Surveill* 2019;5(2):e14067) doi:[10.2196/14067](https://doi.org/10.2196/14067)

KEYWORDS

tobacco industry; mass media; smoking, nontobacco products

Introduction

Background

Tobacco smoking contributes, in epidemic proportions, to the global burden of disease, causing the death of approximately 7 million people each year [1]. In response to this, efforts to curb

tobacco use have shown great success through the implementation of effective, evidence-based tobacco control policies. The establishment of the World Health Organization (WHO) Framework Convention on Tobacco Control (FCTC) in 2003 was a critical milestone for global tobacco control, when 168 countries signed the convention, signifying widespread international cooperation to accelerate the implementation of

effective tobacco control measures globally [2]. As of January 2019, there are 181 parties to the convention, making the WHO FCTC not only 1 of the most widely adopted United Nations treaties [2] but also 1 of the most impactful because of its role in implementing and advancing tobacco control legislation and programs [3]. However, despite the success of the WHO FCTC, the global tobacco epidemic remains problematic, and the progress toward the implementation of effective tobacco control policies has been slow in many countries, particularly in low- and middle-income countries where 80% of the world's smokers live [1].

On September 13, 2017, the former head of the WHO's tobacco free initiative, Derek Yach, announced the establishment of the *Foundation for a Smoke-Free World* (FSFW), a not-for-profit organization with the mission to "accelerate global efforts to reduce deaths and harm from smoking, with the ultimate goal of eliminating smoking worldwide" [4]. In a commentary written by Derek Yach, he argues that there has been "complacency" in the adoption of WHO FCTC policies, and progress to reduce smoking is "lagging." He calls for an increased "ambition" in tobacco control and proposes that harm reduction from smoking is the solution needed to achieve this change [5]. However, in setting out on the journey to achieve this mission, the FSFW has accepted nearly US \$1 billion over 12 years from Philip Morris International (PMI) [5]. PMI has a long and documented history of opposing evidence-based tobacco control policies, and recent investigations have shown that PMI is continuing to promote its flagship cigarette brands to youth and young adults through advertising via social media influencers on Instagram and Facebook [6], in advertising campaigns such as "Be Marlboro" [7], and by handing out free cigarettes to young people at parties [8]. In 2016, PMI also lost a legal suit that PMI launched against the Government of Uruguay over the implementation of large graphic health warnings and a ban on misleading packaging [9]. These are just a small illustration of PMI's most public actions to oppose tobacco control policies, which misalign with the FSFW's stated ultimate goal of eliminating smoking worldwide.

In response, experts working in tobacco control and public health have largely questioned the legitimacy of the FSFW and Derek Yach's motives and whether the FSFW has any role to play in the efforts to sustain and increase the reduction in tobacco smoking globally. In a statement following the announcement of FSFW, the WHO stated it will not partner with the FSFW and consequently called on governments and the public health community to reject any partnership with the FSFW [10]. In January 2018, the deans from 17 public health schools across the United States also released a statement announcing that they will not accept funding from the FSFW, prompting other universities, associations, and organizations to follow suit [11-13]. Tobacco control experts argued that the move by PMI to fund such an organization aligns with their longstanding agenda to support research for the purposes of promoting a positive public image and to regain influence within

the public policy setting, whereas also publicly distracting the media from evidence-based tobacco control policies [14].

Although the tobacco control and public health community response to the launch of the FSFW has been primarily negative, how the FSFW is positioned in the news media has the potential to be heavily influenced and manipulated by both FSFW and PMI. The news media has the power to influence the general public's knowledge and attitudes regarding health issues [15] and can even shape the health policy agenda [16]. The ways in which news events and issues are reported can influence how the target audience thinks, feels, and decides, and facts can be slanted to a particular side with tone, charisma, and rhetoric [17]. As PMI's involvement with the FSFW aligns with their repeated efforts to rebuild a positive public reputation [18-20], it is likely that both PMI and FSFW have an interest in how the FSFW and affiliated research and programs are portrayed in the news media. At the heart of PMI's objective to rebrand itself as a company dedicated to a "smoke-free future," is an effort to shift public sentiment and ultimately, shape and influence the tobacco policy agenda [21]. It is, therefore, critical to monitor and examine how the global news media have responded to the launch of FSFW.

Objectives

In this study, we examined how the global news media responded to the FSFW during a 6-month period following the announcement of the FSFW, to understand how the media framed the FSFW and its funding partner, PMI. We also documented any mention of proven, evidence-based tobacco control policies in the news reporting. All news actors were captured to build a picture of which key informants and messages are featured in the news coverage, as well as the influence of PMI and FSFW's media releases in shaping the development of news articles.

Methods

Search Strategy

Google News and Dow Jones Factiva were used to search for Web-based and print news articles published within a 6-month period after the FSFW was publicly announced. The time period of data collection spanned from September 13, 2017 to March 13, 2018. Separate searches were conducted on both Google News and Factiva using the search terms "Foundation for a Smoke-Free/Smokefree World" and "Smoke-free/Smokefree" with "Foundation." A further search was conducted on Factiva with the search term "Smoke-free/Smokefree" and PMI selected as the company reference. Searches conducted on Factiva were limited to the following Factiva source categories: newspapers, magazines and journals, and major news and business. Searches conducted on Google News specifically looked for print news articles and Web-based news websites (eg, Reuters, news.com.au, and The Guardian) but excluded blogs and letters.

Table 1. Coding variables.

Coding variable	Description
Topic	Overall, what is the news article about? (1 topic per article coded)
Framing argument	What is the argument presented in relation to the FSFW ^a ? Framing argument was determined by identifying the argument that was presented most frequently within the articles
Slant	Is the article presenting the FSFW as a positive or negative initiative? Or is the article neutral toward the FSFW? Positive: defined by a framing argument that is in favor of the FSFW. Negative: defined by a framing argument that is opposed to the FSFW. Neutral: defined by no framing argument, but rather, neutral debate is presented
Mention of tobacco control policies	What tobacco control policies or initiatives are highlighted or mentioned in the article?
Direct quotes or position statements	Who is quoted or paraphrased in the article with a position or opinion on the FSFW?

^aFSFW: Foundation for a Smoke-Free World.

The news articles that appeared in the search results were scanned for relevance. Any news article that did not make at least 1 reference to the FSFW was not downloaded for inclusion in the study. News articles obtained through Google News were downloaded using the Web-browser extension, NCapture, and were then imported into NVivo; a qualitative data analysis program from QSR International. News articles obtained through Factiva were downloaded and imported into NVivo. Exact duplicates of news articles were removed; however, news articles that were similar, but not exact replications, were retained.

Content Analysis

All news articles were read in full by the first author, CW. News articles were coded and analyzed for topic, framing argument, overall slant, mention of tobacco control policies, and direct quotes or position statements (Table 1). Coding categories were developed iteratively following a similar methodological approach taken by Smith and Wakefield [22]. Each article was coded to 1 of 8 different topics, which represented the primary message of the article. Articles were also content analyzed and coded for the argument that framed the article, which was determined by identifying the argument that was presented most frequently within the article. Each of the framing arguments were categorized as either being in favor of, against, or neutral toward the FSFW. Articles were also content analyzed for references to tobacco control policies or initiatives and were coded in 11 categories. Direct quotes and position statements about the FSFW from individuals, associations, companies, or organizations were coded in 14 categories.

Finally, all news articles were also cross-checked against media releases from PMI and the FSFW that were issued within the data collection period. Media releases were sourced directly from the FSFW [23] and PMI websites [24], where all public media releases are published. Articles were coded as to whether the media releases were copied verbatim or were altered. This was determined by identifying articles published within 1 week of the media release date and systematically auditing articles for key quotations from the media releases. Articles that were written with the aid of a press release, but were not copied verbatim, were also coded for whether the article followed the same slant as the press release or not.

A random sample of 5 articles were selected and coded by an additional 4 coders, including the second author, using coding guidelines, to test inter-rater reliability. The inter-rater reliability was 93.5% (116/124).

Results

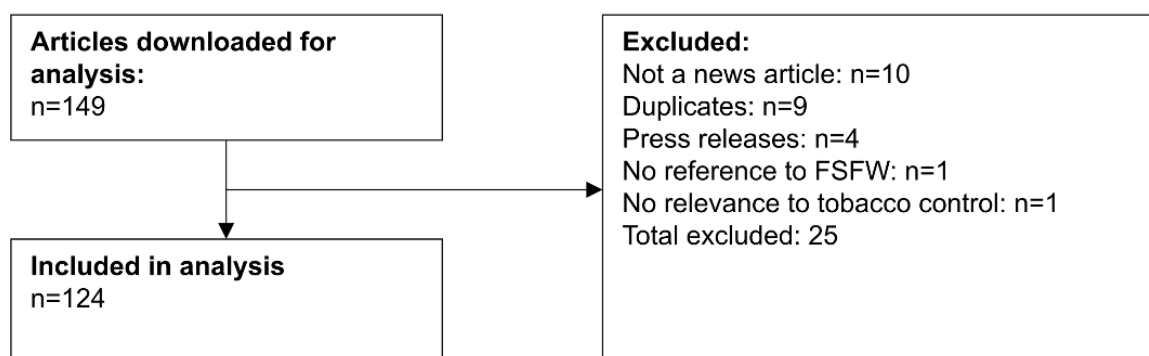
Overview

A total of 149 news articles were downloaded from Google News and Factiva. A total of 9 news articles were exact duplicates and were removed from the sample. Upon reading each article in full, a further 4 articles were found to be media releases, 1 article did not make any reference to the FSFW, 1 had no relevance to tobacco control, and 10 articles were news bulletins or less than a paragraph in length, and so were also excluded from the study (Figure 1). A total of 124 news articles were included in the final content analysis. A total of 4 of these articles were published interviews with André Calantzopoulos (PMI chief executive officer), Derek Yach (FSFW president), Michael Bloomberg (WHO global ambassador for noncommunicable diseases and Bloomberg Philanthropies founder), and Allan Erickson (team leader of National Tobacco Reform Initiative).

Topics and Overall Slant

Articles that reported on PMI wanting a smoke-free future or PMI's strategy to achieve a smoke-free future was the most frequent article topic, with 24.2% (30/124) of all articles coded under this topic. A further 23.4% (29/124) of articles were about individuals or entities refusing funding or any affiliation with the FSFW, and 21.8% (27/124) of articles explored and questioned whether the vision of the FSFW is plausible or not.

When cross-tabulated with the overall slant of each article, whether that be positive, negative, or neutral toward the FSFW, it was found that among articles with an overall positive slant (25/124, 20.2%), articles written about PMI wanting a smoke-free future were most frequent (13/124, 10.5%). Among articles with an overall negative slant toward FSFW (64/124, 51.6%), articles written about individuals or entities refusing funding or any affiliation with the FSFW were most frequent (21/124, 16.9%), followed by articles that explored and questioned the plausibility of the vision of the FSFW (14/124, 11.3%). The results of the article topics by slant are summarized in Table 2.

Figure 1. Flowchart for identification of articles to be included in analysis. FSFW: Foundation for a Smoke-Free World.**Table 2.** Article themes by overall slant.

Topic	Brief description	Slant, n (%)			
		Positive toward the FSFW ^a	Negative toward the FSFW	Neutral toward the FSFW	Total
Bloomberg STOP Foundation announced	Articles that report on the announcement or establishment of the STOP Foundation, funded by Bloomberg Philanthropies.	0 (0.0)	7 (5.6)	1 (0.8)	8 (6.5)
FSFW announced	Articles that report on the announcement or establishment of the FSFW.	1 (0.8)	0 (0.0)	4 (3.2)	5 (4.0)
New grants, research, and programs from the FSFW	Articles that report on any new grants, research projects, or programs that are funded from the FSFW.	4 (3.2)	2 (1.6)	0 (0.0)	6 (4.8)
PMI ^b wants a smoke-free world	Articles that report on PMI wanting a smoke-free future or world and its future investment strategy to achieve this.	13 (10.5)	8 (6.5)	9 (7.3)	30 (24.2)
Reduced-risk products (RRP)	Articles that report on the regulation of RRP or the need for new avenues in tobacco control and the potential of RRP fulfilling this need.	3 (2.4)	0 (0.0)	4 (3.2)	7 (5.6)
Refusing funding or affiliation with the FSFW	Articles about individuals or entities refusing funding or any affiliation with the FSFW.	1 (0.8)	21 (16.9)	7 (5.6)	29 (23.4)
The plausibility of the vision of the FSFW	Articles that explore/question whether the vision of the FSFW is plausible or not. Articles within this category generally explore perspectives on the FSFW specifically to uncover whether it can be trusted or not.	3 (2.4)	14 (11.3)	10 (8.1)	27 (21.8)
Tobacco industry interference	Articles about how the tobacco industry is interfering or the ways in which the tobacco industry continues to interfere in the development and implementation of effective tobacco control policies.	0 (0.0)	12 (9.7)	0 (0.0)	12 (9.7)
Total		25 (20.2)	64 (51.6)	35 (28.2)	— ^c

^aFSFW: Foundation for a Smoke-Free World.

^bPMI: Philip Morris International.

^cNot applicable.

Table 3. Framing arguments presented.

Framing argument	Articles by frame, n (%)
Argument in support of FSFW^a	25 (20.2)
PMI ^b is funding the FSFW in an effort to ensure sustained profitability of the tobacco industry.	9 (7.3)
The FSFW or PMI is part of the solution to reducing the impact of tobacco use.	16 (12.9)
Argument in opposition to FSFW	64 (51.6)
FSFW is not credible simply because of funding link to PMI	36 (29.0)
PMI is funding the FSFW to disingenuously makeover their image	10 (8.1)
The FSFW is a tactic to ensure sustained profitability of the tobacco industry.	4 (3.2)
The FSFW is a tactic to mislead and undermine effective tobacco control measures	14 (11.3)
Balanced debate	35 (28.2)

^aFSFW: Foundation for a Smoke-Free World.

^bPMI: Philip Morris International.

Framing Arguments

The news coverage of the FSFW was framed by 6 key arguments, as can be seen in [Table 3](#). Over half of the news articles presented an overall framing argument that was classified as against the FSFW (64/124, 51.6%), whereas arguments in favor of the FSFW made up 20.2% (25/124). A further 28.2% of articles (35/124) presented a neutral debate with no primary framing argument. The most frequent framing argument presented was that the FSFW is not credible because of the funding link to PMI (36/124, 29.0%). Statements questioning the credibility of the FSFW included:

It is the height of hypocrisy for PMI to proclaim that it is helping solve the tobacco problem while it aggressively markets cigarettes. [25]

Philip Morris isn't the solution to the tobacco problem. Philip Morris is the cause. [26]

Articles also argued that the FSFW was established as a tactic to mislead and undermine effective tobacco control measures (14/124, 11.3%), to deceitfully ensure sustained profitability of the tobacco industry (4/124, 3.2%), and that PMI is funding the FSFW to disingenuously makeover their corporate image (10/124, 8.1%). Articles that presented arguments in favor of the FSFW most frequently argued that the FSFW or PMI is part

of the solution to reducing the impact of tobacco use (16/124, 12.9%). A further 7.3% of articles (9/124) argued that PMI's decision to fund the FSFW is a logical and genuine effort to ensure the industry sustains profitability in the future through investment in reduced-risk products. For example:

Yet Philip Morris International seems sincere in wanting to have people stop smoking combustible cigarettes—to go smoke-free—and switch over to e-cigs as an alternative. [27]

Tobacco Control Policies and Measures Mentioned

Evidence-based tobacco control policies were commonly referenced or mentioned within the news articles analyzed. The WHO FCTC was mentioned most frequently across all articles (31/124, 25.0%), with the majority of articles presenting the WHO FCTC in a positive or neutral light (14/124, 11.3% and 15/124, 12.1% of the total number of articles, respectively). Tobacco excise was also frequently mentioned in articles (26/124, 20.8%), along with advertising and promotion bans (22/124, 17.7%) and graphic health warnings (21/124, 16.9%). Overall, 66.9% (83/124) of all articles mentioned tobacco control policies or measures positively, whereas only 9.6% (12/124) of articles presented tobacco control policies or measures negatively. A full list of tobacco control policies and measures referenced in the articles can be seen in [Table 4](#).

Table 4. Tobacco control policies and measures referenced in news articles (% of total number of articles).

Tobacco control policies mentioned ^a	Positive ^b , n (%)	Neutral ^c , n (%)	Negative ^d , n (%)	Total, n (%)
WHO FCTC ^e	14 (11.3)	15 (12.1)	2 (1.6)	31 (25.0)
Tobacco excise	17 (13.7)	7 (5.6)	2 (1.6)	26 (20.8)
Advertising and promotion bans	17 (13.6)	4 (3.2)	1 (0.8)	22 (17.7)
Graphic health warnings	16 (12.9)	4 (3.2)	1 (0.8)	21 (16.9)
Smoke-free environment policies	7 (5.6)	5 (4.0)	2 (1.6)	14 (11.3)
Plain packaging	5 (4.0)	6 (4.8)	0 (0.0)	11 (8.9)
Low-nicotine cigarettes	0 (0.0)	10 (8.1)	0 (0.0)	10 (8.1)
Smoking cessation support	4 (3.2)	2 (1.6)	3 (2.4)	9 (7.2)
Point of sale restrictions	2 (1.6)	2 (1.6)	0 (0.0)	4 (3.2)
Mass media campaigns	0 (0.0)	1 (0.8)	1 (0.8)	2 (1.6)
Single cigarette sales ban	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.8)
Total	83 (66.9)	56 (46.5)	12 (9.6)	

^aIf multiple policies were mentioned within 1 article, they were coded separately so the total exceeds the number of articles.

^bPolicy mentioned is framed as being effective, proven, or evidence-based.

^cPolicy mentioned with no statement of effectiveness or whether or not the policy is proven or evidence-based.

^dPolicy mentioned is framed as being ineffective or not having adequate effect on reducing smoking.

^eWHO FCTC: World Health Organization Framework Convention on Tobacco Control.

Quotes or Position Statements

Articles were coded for quotes and position statements found within the articles. Quotes from tobacco companies were most frequently paraphrased in all articles, with 43.5% (54/124) of all articles having at least 1 quote or statement from a tobacco company. Although the majority of these were quotes from PMI (49/124, 39.2%), which described the FSFW and PMI's vision for a smoke-free world as positive, a small fraction of quotes was from other tobacco companies (6/124, 4.8%):

One senior executive at a rival manufacturer praised PMI for scoring a "huge PR coup", but claimed that the Swiss-based company was not doing anything markedly different from its competitors. "It has said it will go smoke-free "as soon as possible", although it has not put a time frame on achieving it," he said. "I think you'll see PMI selling Marlboro cigarettes for many years to come. [28] [Unidentified tobacco company executive]

Both the FSFW and public health groups, associations, and organizations were each quoted or had their position statement paraphrased in approximately 40.3% (50/124) of all articles (38.7% and 40.0%, respectively). Derek Yach was the primary spokesperson for the FSFW, with 38 of the total of 48 articles that quoted or paraphrased the FSFW, specifically quoting Yach. A total of 1 article was a published interview with Yach. The next most frequent organization or individual quoted or paraphrased was the United Nations, including the WHO, with 37.6% (47/124) of all articles including statements or quotes from the WHO. All statements attributed to the WHO were opposed to the formation and funding of FSFW. Health academics were also quoted in 27.2% (34/124) of the articles;

however, they were slightly more varied in their position on the FSFW. Of the 23 health academics quoted, 3 were in favor of the FSFW, 18 were against, and 2 held neutral positions. Examples of positive and negative quotes or statements from health academics include:

Another anti-smoking expert and Adjunct Professor at the Centre for Health Law, Policy & Ethics at University of Ottawa, David Sweanor, is on the same page as Yach. He is inviting everyone at the Tobacco or Health Conference to consider a different, and perhaps more effective approach to the usual "let's wrestle tobacco companies to the ground" kind of reasoning. [29]

The idea of taking money that's from the tobacco industry is just antithetical to everything we do, Karen Emmons, Dean for academic affairs at Harvard's public health school. [30]

Media Releases

News articles are frequently generated as a direct result of media releases, which can be reproduced verbatim as a published news article or rewritten with quotes or content from the media release paraphrased with at times, a different overall slant or angle. During the data collection period, there were 4 media releases from the FSFW and 1 from PMI that were included in at least 1 article. A total of 3 news articles in total published the media release verbatim with no changes, 3 used quotes from the media releases but kept the same slant and angle as the media release, and 10 articles had quotes from the media releases, though the overall slant and angle of the article was different or conflicting with that of the media release (Table 5).

Table 5. Article reliance on media releases from Philip Morris International (PMI) and Foundation for a Smoke-Free World (FSFW).

Media release title	Organization	Media release reproduced verbatim (n=3), n	Quote from media release, different angle (n=10), n	Quote from media release, same angle (n=3), n
Global foundation launches to accelerate an end to smoking	FSFW	0	1	1
Letter to the World Health Organization	FSFW	0	5	1
Public invited to shape Foundation for a Smoke-Free World's US \$1 billion research agenda	FSFW	1	1	0
Accelerating an end to smoking	FSFW	1	0	0
Philip Morris International announces support for the establishment of the Foundation for a Smoke-Free World	PMI	1	3	1

Discussion

Principal Findings

The global news media have reported the announcement of the FSFW; however, the portrayal of the foundation and its mission and vision has primarily been framed by doubt, skepticism, and disapproval. A majority of articles were slanted negatively toward the FSFW, and the most frequent framing argument presented was that the FSFW was not credible because of the funding link to PMI. Other arguments against the FSFW were that the foundation was established as a tactic to mislead and undermine effective tobacco control measures, to deceitfully ensure sustained profitability of the tobacco industry, and to disingenuously makeover PMI's corporate image. Encouragingly, at least 1 evidence-based tobacco control measure was mentioned or referenced positively within a majority of the news articles analyzed.

Despite attempts by both PMI and the FSFW to steer media coverage in a positive direction, the tobacco control community's views opposing this new venture dominated the news reporting. This is a positive finding for public health advocates as it is indicative that tobacco control policies and perspectives are generally accepted as reputable, newsworthy, and fundamental in the efforts to reducing the burden of smoking. News media have repeatedly shown to be influential in communicating health messages to the public [31-34], and news editors are fundamental in determining what stories are published and how these issues are framed. Given that 66.9% of all articles (83/124) made mention of evidence-based tobacco control policies positively and just 9.6% of articles (12/124) referred to evidence-based tobacco control policies negatively, it appears that the decades of effort to raise public support for tobacco control policies has influenced news editors' and journalists' perceptions of tobacco issues and solutions and overall acceptance of public health viewpoints. This aligns with previous research, which found that newspaper editors' largely expressed support for tobacco control measures and objectives in tobacco-related editorials [22]. Given the high dependency of politicians and key decision makers on the media and how issues are framed, there is therefore potential for a widespread rejection of the FSFW among governments, policy makers, and the public if news reporting continues to align with the viewpoints of the tobacco control community [35].

The ways in which issues are reported on in the media can also be heavily impacted by the content of media releases. Corporate strategies for engagement with the media are seen to be essential to shape a company's reputation and build respectability and credibility [35]. However, inaccuracies or embellishments in original media releases can easily be reproduced as fact in news media articles [36,37]. As PMI has attempted to rebuild its public reputation through public relations and social media strategies [18-20], it was essential to examine whether the media releases from PMI and the FSFW resulted in additional news articles and whether or not these news articles were reproduced verbatim or published with the same or a conflicting angle. The 4 media releases from the FSFW and 1 from PMI did not result in a large response in Web-based and print news media, with only 3 articles reproduced verbatim from the media release. Although 10 articles had used the messaging from the media releases, the resulting articles were not framed in support of FSFW.

In what was clearly an attempt to secure positive news coverage, we noted that an Australian journalist traveled to Japan as a guest of PMI. This was revealed in a concluding disclosure statement within the news article. The article written by this journalist presented a balanced debate about PMI's involvement in reduced-risk products, citing quotes from reputable evidence-based organizations such as the WHO and Cancer Council Australia, alongside quotes from the manager of PMI in Japan [38]. Another journalist wrote of his experience going to the PMI headquarters in Switzerland and interviewing PMI's chief executive, André Calantzopoulos, which may indicate that the trip was paid for by PMI [39]. The subsequent news article that was published by this journalist argued that PMI is part of the solution to reducing tobacco related disease. It is possible that other journalists received similar benefits without public disclosure, which may have had an impact upon the way the way in which PMI and the FSFW was framed in the final news articles.

Although the findings of this study broadly indicate positive results for the public health and tobacco control agenda, the FSFW and PMI were still able to deliver their key messages and have key media spokesperson quotes within a significant proportion of all news articles (44.0% and 38.7%, respectively). The most frequent topic presented in the news articles was PMI's key public relations message that it is working for a smoke-free future, irrespective of the overall slant of the article.

The presentation of some supportive views from health academics in the media may also create confusion and a perception of “in-fighting” and disunity among experts in the field.

Given PMI’s deceptive history and false attempts to publicly rebrand itself as a socially responsible company committed to a smoke-free future, it is essential that the tobacco control community unite to reinforce why the FSFW cannot be considered a reputable research-focused not-for-profit organization. The FSFW lacks independence from PMI, as revealed from a review of the FSFW governing documents showing that PMI’s funding is contingent on the FSFW working toward its stated goals, which are aligned with PMI’s agenda. The FSFW also lacks transparency and accountability, and board members, who were appointed by the FSFW president, are paid US \$50,000 per year, which is far beyond the standard of other nongovernment organizations [40]. In recent months, Derek Yach wrote an open letter on behalf of the FSFW inviting the WHO executive board to work with the FSFW [41], which subsequently resulted in the Global Centre for Good Governance in Tobacco Control coordinating a response from 120 organizations and 182 individuals calling on the WHO to reject the FSFW’s bid to partner together to reduce the impact of smoking [42]. Such coordinated efforts are key to minimizing the influence of FSFW and PMI in tobacco control research and policy making into the future. It is also critical that the news media is continually monitored to ensure that the tobacco industry and associated organizations, such as the FSFW, are

not successfully undermining public health efforts to reduce tobacco use. This is critically important as the FSFW begins to roll out funding and grants to public health researchers, as has happened in New Zealand with the creation of the *Centre of Research Excellence on Indigenous Sovereignty & Smoking* [43].

Limitations

News articles were not able to be analyzed for readership numbers, nor for the number of times each article was shared on social media; therefore, the overall reach of the articles was not determined. The study also focused on Web-based and print news articles only and excluded other Web-based channels that are used to obtain and share information, such as social media. As the data were obtained through Google News and Factiva, there is also no guarantee that all Web-based and print news articles were included in the analysis.

Conclusions

In the 6 months following the announcement of the FSFW, the global news media primarily portrayed the FSFW and PMI with doubt, skepticism, and disapproval, despite efforts to shape the development of news articles through corporate media releases and media spokespeople. Although this is a positive finding for the public health agenda, it is essential that the tobacco control community continue to provide a unified voice in the media on this issue to minimize the influence of the FSFW and PMI in tobacco control research and policy development.

Acknowledgments

The authors would like to thank the staff at Cancer Council New South Wales, Tobacco Control Unit for coding a sample of news articles to test inter-rater reliability.

Conflicts of Interest

None declared.

References

1. World Health Organization. 2018 Mar 09. Tobacco Fact Sheet URL: <https://www.who.int/en/news-room/fact-sheets/detail/tobacco> [accessed 2019-03-18] [WebCite Cache ID 76xB1eb39]
2. World Health Organization. Parties to the WHO Framework Convention on Tobacco Control URL: <https://www.who.int/fctc/cop/en/> [accessed 2019-03-18] [WebCite Cache ID 76xB7cPDE]
3. Chung-Hall J, Craig L, Gravely S, Sansone N, Fong GT. World Health Organization. 2016. Impact of the WHO Framework Convention on Tobacco Control on the implementation and effectiveness of tobacco control measures: a global evidence review URL: <https://www.who.int/fctc/cop/WHO-FCTC-ITC-Global-Evidence-Review.pdf> [accessed 2019-03-18] [WebCite Cache ID 76xIk0Rhz]
4. Foundation for a Smoke-Free World. Global foundation launches to accelerate an end to smoking URL: <https://www.smokefreeworld.org/newsroom/global-foundation-launches-accelerate-end-smoking> [accessed 2019-03-18] [WebCite Cache ID 76xBUpIIF]
5. Yach D. Foundation for a Smoke-Free World. *Lancet* 2017 Oct 14;390(10104):1807-1810. [doi: [10.1016/S0140-6736\(17\)32602-8](https://doi.org/10.1016/S0140-6736(17)32602-8)] [Medline: 29047446]
6. TakeApart Campaign. Where there's smoke URL: <https://www.takeapart.org/wherethersmoke/> [accessed 2019-03-18] [WebCite Cache ID 76xBdsGXN]
7. Boseley S. The Guardian. 2014 Mar 12. Marlboro marketing campaign aimed at young people, anti-tobacco report says URL: <https://www.theguardian.com/society/2014/mar/12/marlboro-marketing-campaign-aimed-young-people-cigarettes-report> [accessed 2018-10-12] [WebCite Cache ID 76xJ1BAIf]

8. Kalra A, Bansal P, Lasseter T, Wilson D. Reuters. 2017 Jul 18. Philip Morris takes aim at young people in India, and health officials are fuming URL: <https://www.reuters.com/investigates/special-report/pmi-india/> [accessed 2019-03-18] [WebCite Cache ID 76xJHldap]
9. University of Bath, Tobacco Tactics. 2017. Philip Morris vs the Government of Uruguay URL: http://www.tobaccotactics.org/index.php/Philip_Morris_vs_the_Government_of_Uruguay [accessed 2019-03-18] [WebCite Cache ID 76xJU5CF6]
10. World Health Organization. 2017. WHO Statement on Philip Morris funded Foundation for a Smoke-Free World URL: <https://www.who.int/news-room/detail/28-09-2017-who-statement-on-philip-morris-funded-foundation-for-a-smoke-free-world> [accessed 2019-03-18] [WebCite Cache ID 76xJrMXjh]
11. Thomas DP, Hefler M, Bonevski B, Calma T, Carapetis J, Chamberlain C, et al. Australian researchers oppose funding from the Foundation for a Smoke-Free World. *Aust N Z J Public Health* 2018 Dec;42(6):506-507. [doi: [10.1111/1753-6405.12861](https://doi.org/10.1111/1753-6405.12861)] [Medline: [30512245](https://pubmed.ncbi.nlm.nih.gov/30512245/)]
12. Young K, Samet JM, Guilarte TR, Williams MA, Giles WH, Halverson PK, et al. Harvard School of Public Health. 2018. Statement on the Foundation for a Smoke-Free World URL: <https://www.hsph.harvard.edu/deans-office/2018/01/26/smoke-free-world-letter/> [accessed 2019-03-18] [WebCite Cache ID 76xKgAuYq]
13. Council of Academic Public Health Institutions Australasia. Statement on Foundation for a Smoke Free World URL: <http://caphia.com.au/statement-on-foundation-for-a-smoke-free-world/> [accessed 2019-03-18] [WebCite Cache ID 76xL8VGLQ]
14. Daube M, Moodie R, McKee M. Towards a smoke-free world? Philip Morris International's new Foundation is not credible. *Lancet* 2017 Oct 14;390(10104):1722-1724. [doi: [10.1016/S0140-6736\(17\)32561-8](https://doi.org/10.1016/S0140-6736(17)32561-8)] [Medline: [29047432](https://pubmed.ncbi.nlm.nih.gov/29047432/)]
15. Hayes M, Ross IE, Gasher M, Gutstein D, Dunn JR, Hackett RA. Telling stories: news media, health literacy and public policy in Canada. *Soc Sci Med* 2007 May;64(9):1842-1852. [doi: [10.1016/j.socscimed.2007.01.015](https://doi.org/10.1016/j.socscimed.2007.01.015)] [Medline: [17337317](https://pubmed.ncbi.nlm.nih.gov/17337317/)]
16. Baillie RK. Determining the effects of media portrayals of alcohol: going beyond short term influence. *Alcohol Alcohol* 1996 May;31(3):235-242. [doi: [10.1093/oxfordjournals.alcalc.a008142](https://doi.org/10.1093/oxfordjournals.alcalc.a008142)] [Medline: [8844028](https://pubmed.ncbi.nlm.nih.gov/8844028/)]
17. Entman RM. Framing bias: media in the distribution of power. *J Commun* 2007;57(1):163-173. [doi: [10.1111/j.1460-2466.2006.00336.x](https://doi.org/10.1111/j.1460-2466.2006.00336.x)]
18. McDaniel P, Malone R. Creating the "desired mindset": Philip Morris's efforts to improve its corporate image among women. *Women Health* 2009;49(5):441-474 [FREE Full text] [doi: [10.1080/03630240903238800](https://doi.org/10.1080/03630240903238800)] [Medline: [19851947](https://pubmed.ncbi.nlm.nih.gov/19851947/)]
19. Smith EA, Malone RE. Altria means tobacco: Philip Morris's identity crisis. *Am J Public Health* 2003 Apr;93(4):553-556. [doi: [10.2105/ajph.93.4.553](https://doi.org/10.2105/ajph.93.4.553)] [Medline: [12660196](https://pubmed.ncbi.nlm.nih.gov/12660196/)]
20. Watts C, Hefler M, Freeman B. 'We have a rich heritage and, we believe, a bright future': how transnational tobacco companies are using Twitter to oppose policy and shape their public identity. *Tob Control* 2019 Dec;28(2):227-232. [doi: [10.1136/tobaccocontrol-2017-054188](https://doi.org/10.1136/tobaccocontrol-2017-054188)] [Medline: [29666168](https://pubmed.ncbi.nlm.nih.gov/29666168/)]
21. Philip Morris International. Designing a Smoke-Free Future URL: <https://www.pmi.com/who-we-are/designing-a-smoke-free-future> [accessed 2019-03-18] [WebCite Cache ID 76xLcMuzr]
22. Smith KC, Wakefield M. Textual analysis of tobacco editorials: how are key media gatekeepers framing the issues? *Am J Health Promot* 2005;19(5):361-368. [doi: [10.4278/0890-1171-19.5.361](https://doi.org/10.4278/0890-1171-19.5.361)] [Medline: [15895539](https://pubmed.ncbi.nlm.nih.gov/15895539/)]
23. SmokeFreeWorld.org Newsroom. Foundation for a Smoke-Free World URL: <https://www.smokefreeworld.org/newsroom> [accessed 2019-04-29] [WebCite Cache ID 77zMNfo48]
24. Philip Morris International. Press Releases URL: <https://www.pmi.com/media-center/press-releases> [accessed 2019-04-29] [WebCite Cache ID 77zMXD6a6]
25. Myers M, Koval R. Fortune. 2018 Jan 5. Commentary: Philip Morris Says It Wants to Quit Cigarettes. But It's Just Blowing Smoke URL: <http://fortune.com/2018/01/05/philip-morris-international-quit-smoking-cigarettes-tobacco/> [accessed 2019-03-18] [WebCite Cache ID 76xLzKafi]
26. Hurt R. Post Bulletin. 2017 Nov 27. Hurt: Don't be fooled by Big Tobacco smokescreen URL: https://www.postbulletin.com/opinion/columnists/hurt-don-t-be-fooled-by-big-tobacco-smokescreen/article_0a3c98c8-42a0-540d-8c48-cd5743024679.html [accessed 2019-03-01] [WebCite Cache ID 76xMBNg6z]
27. Walsh D. Still in the habit of generating cash. Scotland: The Times; Oct 26, 2017.
28. Chambers S. Business Live. 2018 Jan 31. Black cloud hangs over substitutes for cigarettes URL: <https://www.businesslive.co.za/bd/life/2018-01-31-black-cloud-hangs-over-substitutes-for-cigarettes/> [accessed 2019-05-21] [WebCite Cache ID 78Wpsj0Nr]
29. Caruana D. Bloomberg Funds Agency to Monitor Tobacco Industry. 2018 Mar 12. URL: <https://www.vapingpost.com/2018/03/12/bloomberg-funds-agency-to-monitor-tobacco-industry/> [accessed 2019-05-21] [WebCite Cache ID 78WqDxqMH]
30. Binkley C. Business Insider. 2018 Jan 25. 17 universities oppose anti-smoking group with tobacco ties URL: <https://www.businessinsider.com/ap-17-universities-oppose-anti-smoking-group-with-tobacco-ties-2018-1/?r=AU&IR=T> [accessed 2019-03-01] [WebCite Cache ID 76xMVgifN]
31. Krewski D, Lemyre L, Turner M, Lee J, Dallaire C, Bouchard L, et al. Public perception of population health risks in Canada: health hazards and sources of information. *Hum Ecol Risk Assess* 2006 Aug;12(4):626-644. [doi: [10.1080/10807030600561832](https://doi.org/10.1080/10807030600561832)]

32. Malone RE, Boyd E, Bero LA. Science in the news: journalists' constructions of passive smoking as a social problem. *Soc Stud Sci* 2016 Jun 29;30(5):713-735. [doi: [10.1177/030631200030005003](https://doi.org/10.1177/030631200030005003)]
33. Scheufele DA, Tewksbury D. Framing, agenda setting, and priming: the evolution of three media effects models. *J Commun* 2007;57(1):9-20. [doi: [10.1111/j.0021-9916.2007.00326.x](https://doi.org/10.1111/j.0021-9916.2007.00326.x)]
34. Chapman S, Haynes A, Derrick G, Sturk H, Hall W, St George A. Reaching "an audience that you would never dream of speaking to": influential public health researchers' views on the role of news media in influencing policy and public understanding. *J Health Commun* 2014;19(2):260-273. [doi: [10.1080/10810730.2013.811327](https://doi.org/10.1080/10810730.2013.811327)] [Medline: [24156565](https://pubmed.ncbi.nlm.nih.gov/24156565/)]
35. Einwiller S, Carroll C, Korn K. Under what conditions do the news media influence corporate reputation? The roles of media dependency and need for orientation. *Corp Reputation Rev* 2010 Jan 20;12(4):299-315. [doi: [10.1057/crr.2009.28](https://doi.org/10.1057/crr.2009.28)]
36. Chau J, McGill B, Freeman B, Bonfiglioli C, Bauman A. Overselling sit-stand desks: news coverage of workplace sitting guidelines. *Health Commun* 2018 Dec;33(12):1475-1481. [doi: [10.1080/10410236.2017.1359034](https://doi.org/10.1080/10410236.2017.1359034)] [Medline: [28850250](https://pubmed.ncbi.nlm.nih.gov/28850250/)]
37. MacKenzie R, Freeman B. Second-hand smoke in cars: how did the "23 times more toxic" myth turn into fact? *Can Med Assoc J* 2010 May 18;182(8):796-799 [FREE Full text] [doi: [10.1503/cmaj.090993](https://doi.org/10.1503/cmaj.090993)] [Medline: [20385738](https://pubmed.ncbi.nlm.nih.gov/20385738/)]
38. Young M. news.com.au. 2018 Jan 07. Tobacco giant Philip Morris publishes advertisement declaring it wants to 'give up smoking'. This is not new URL: <https://www.news.com.au/lifestyle/health/health-problems/tobacco-giant-philip-morris-publishes-advertisement-declaring-it-wants-to-give-up-smoking-this-is-not-new/news-story/6abbf96c26eed5e8d2b3317537fc454b> [accessed 2019-03-18] [WebCite Cache ID 76xMj7Agt]
39. The Economist. 2017 Dec 19. The tobacco paradox: Combustible cigarettes kill millions a year. Can Big Tobacco save them? URL: <https://www.economist.com/business/2017/12/19/combustible-cigarettes-kill-millions-a-year-can-big-tobacco-save-them> [accessed 2018-11-28] [WebCite Cache ID 76xMtLriL]
40. van der Eijk Y, Bero LA, Malone RE. Philip Morris International-funded 'Foundation for a Smoke-Free World': analysing its claims of independence. *Tob Control* 2018 Sep 21:4278. [doi: [10.1136/tobaccocontrol-2018-054278](https://doi.org/10.1136/tobaccocontrol-2018-054278)] [Medline: [30242044](https://pubmed.ncbi.nlm.nih.gov/30242044/)]
41. Yach D. SmokeFreeWorld.org. 2019 Jan 24. An open letter to the who executive board URL: <https://www.smokefreeworld.org/blog/open-letter-who-executive-board> [accessed 2019-03-18] [WebCite Cache ID 76xN5RJB6]
42. Global Centre for Good Governance in Tobacco Control. 2019 Jan 28. An open letter to the Director General and Executive Board of the World Health Organization URL: <https://ggtc.world/2019/01/28/an-open-letter-to-the-director-general-and-executive-board-of-the-world-health-organization/> [accessed 2019-03-01] [WebCite Cache ID 76xNJ1mpO]
43. Foundation for a Smoke-Free World. Centers of Excellence URL: <https://www.smokefreeworld.org/centers-grants/centers-excellence> [accessed 2019-03-18] [WebCite Cache ID 76xNU2q5w]

Abbreviations

FCTC: Framework Convention on Tobacco Control

FSFW: Foundation for a Smoke-Free World

PMI: Philip Morris International

WHO: World Health Organization

Edited by G Eysenbach; submitted 20.03.19; peer-reviewed by K Evans-Reeves, R Malone; comments to author 17.04.19; revised version received 06.05.19; accepted 07.05.19; published 06.06.19.

Please cite as:

Watts C, Freeman B

"Where There's Smoke, There's Fire": A Content Analysis of Print and Web-Based News Media Reporting of the Philip Morris-Funded Foundation for a Smoke-Free World

JMIR Public Health Surveill 2019;5(2):e14067

URL: <http://publichealth.jmir.org/2019/2/e14067/>

doi: [10.2196/14067](https://doi.org/10.2196/14067)

PMID: [31172959](https://pubmed.ncbi.nlm.nih.gov/31172959/)

©Christina Watts, Becky Freeman. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 06.06.2019. 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 for Sexual Relationships: Mixed-Methods Study on Digital Flirting and Online Dating Among Adolescent Youth and Young Adults

James Lykens¹, MA; Molly Pilloton², MPH; Cara Silva², MPH; Emma Schlam², BA; Kate Wilburn², BS; Emma Pence², BA

¹Center for Research and Education on Gender and Sexuality, San Francisco State University, San Francisco, CA, United States

²Youth Tech Health, Oakland, CA, United States

Corresponding Author:

James Lykens, MA
Center for Research and Education on Gender and Sexuality
San Francisco State University
835 Market Street
5th Floor, Suite 525
San Francisco, CA,
United States
Phone: 1 415 817 4520
Email: james.lykens@etr.org

Related Article:

This is a corrected version. See correction statement: <http://publichealth.jmir.org/2019/2/e14815/>

Abstract

Background: According to a 2015 report from the Pew Research Center, nearly 24% of teens go online almost constantly and 92% of teens are accessing the internet daily; consequently, a large part of adolescent romantic exploration has moved online, where young people are turning to the Web for romantic relationship-building and sexual experience. This digital change in romantic behaviors among youth has implications for public health and sexual health programs, but little is known about the ways in which young people use online spaces for sexual exploration. An examination of youth sexual health and relationships online and the implications for adolescent health programs has yet to be fully explored.

Objective: Although studies have documented increasing rates of sexually transmitted infections and HIV among young people, many programs continue to neglect online spaces as avenues for understanding sexual exploration. Little is known about the online sexual health practices of young people, including digital flirting and online dating. This study explores the current behaviors and opinions of youth throughout online sexual exploration, relationship-building, and online dating, further providing insights into youth behavior for intervention opportunities.

Methods: From January through December 2016, an exploratory study titled TECHsex used a mixed-methods approach to document information-seeking behaviors and sexual health building behaviors of youth online in the United States. Data from a national quantitative survey of 1500 youth and 12 qualitative focus groups (66 youth) were triangulated to understand the experiences and desires of young people as they navigate their sexual relationships through social media, online chatting, and online dating.

Results: Young people are using the internet to begin sexual relationships with others, including dating, online flirting, and hooking up. Despite the fact that dating sites have explicit rules against minor use, under 18 youth are using these products regardless in order to make friends and begin romantic relationships, albeit at a lower rate than their older peers (19.0% [64/336] vs 37.8% [440/1163], respectively). Nearly 70% of youth who have used online dating sites met up with someone in person (44.78% [30/67] under 18 vs 74.0% [324/438] over 18). Focus group respondents provided further context into online sexual exploration; many learned of sex through pornography, online dating profiles, or through flirting on social media. Social media played an important role in vetting potential partners and beginning romantic relationships. Youth also reported using online dating and flirting despite fears of violence or catfishing, in which online profiles are used to deceive others.

Conclusions: Youth are turning to online spaces to build sexual relationships, particularly in areas where access to peers is limited. Although online dating site use is somewhat high, more youth turn to social media for online dating. Sexual relationship-building included online flirting and online dating websites and/or apps. These findings have implications for future sexual health programs interested in improving the sexual health outcomes of young people. Researchers may be neglecting to include social media as potential sources of youth hookup culture and dating. We implore researchers and organizations to consider the relationships young people have with technology in order to more strategically use these platforms to create successful and youth-centered programs to improve sexual health outcomes.

(*JMIR Public Health Surveill* 2019;5(2):e10695) doi:[10.2196/10695](https://doi.org/10.2196/10695)

KEYWORDS

youth; sexual health; online dating; adolescent health

Introduction

Recent increases in personal technology ownership among youth [1] combined with the developmental stage of sexual exploration and identity development [2] have changed the way young people prefer to build relationships and undergo sexual exploration. Websites and dating apps are popular spaces in which to develop sexual relationships and locate sexual health information [3-4], experiment with sexual play [2], discover pornography, and begin dating [5]. Previous studies have examined the role of online sexual exploration among adolescents as a time of identity construction and expression [6-7]; a decade ago, early chat rooms and social media platforms provided ideal ways for young people to connect and construct identities. However, due to constantly changing online tools for social connection and dating, few studies have recently explored the new ways in which adolescents use and describe their online use for sexual exploration and relationship-building. In addition, few studies have focused specifically on the use of online dating websites among youth younger than 18 years; instead, most studies have either focused on the victimization of minors online and the moral panics around youth sexuality and new media use or have relied on a monolithic description of youth that fails to consider developmental differences among ages in regard to sexual health research [8].

An exploration into the use of online dating and social media for sexual exploration among youth is timely considering the increased rates of sexually transmitted infections and HIV among youth today [9]. As these trends continue to increase, it is vital for researchers and program developers to consider the unique interactions that young people have with sexual health and relationships online in order to create successful programs, particularly due to the rapidly changing landscape of online dating and social media platforms. Further, a historical failure to focus on minor youth (younger than 18 years) differences and relying on traditional approaches to understanding sexual and romantic behaviors of young people may be insufficient to fully capture the sexual health implications of online sexual exploration and relationship development.

While many projects have found success bringing sexual health programs online [10-12], and recent studies have begun to incorporate a more complicated definition of youth [13], an examination of the unique relationships young people have with their sexual exploration and relationships online has yet to be fully explored. In particular, few studies have focused on the

amalgamation of online dating and social media [14] or the level of trust and engagement that youth have invested in online dating and flirting, particularly in the context of public health. Exploratory studies are needed to remain up to date on youth trends and behaviors in order to locate opportunities for health interventions and future research.

In response, the TECHsex national research project examined the online sexual exploration habits of young people in order to paint a broader picture of online sexual relationships, including online dating, trusted websites, age differences, behaviors, and online flirting. The implications of these data demonstrate that future programs should consider the varied experiences, levels of trust, age, and behaviors of youth prior to developing sexual health programs if they wish to yield greater health outcomes among young people.

Methods

In this exploratory study, we used a mixed-methods approach to capture a broad understanding of sexual exploration and relationships online among young people. The study received human protections approval from the Quorum Institutional Review Board.

Quantitative Phase

A national self-report survey, hosted on the online survey software Qualtrics, was conducted from September 2015 through July 2016. A total of 1500 youth ages 13 to 24 years (average age 19.70 years) responded; 22.40% (336/1500) of respondents were younger than 18 years.

Survey questions were determined by a comprehensive literature review of recent reports and articles on youth technology use, positive youth development [15], sexual exploration, and sexual health in the United States. The survey asked detailed demographic and behavioral questions on technology use, sexual behavior online, romantic relationship behavior online, and trust in dating and flirting online. Data reported here are a subset of our complete TECHsex report [16], which includes additional data on other youth technology trends. Data detailed here are self-reported responses to closed-ended survey questions. We analyzed these data using descriptive statistics including cross-tabulations and frequencies.

Qualitative Phase

In addition to the online quantitative survey, semistructured qualitative focus groups were conducted to contextualize

quantitative findings. Findings from the quantitative survey directly influenced the creation of the semistructured interview guide for focus groups. Focus groups were conducted in the south, west, midwest, and northeast United States in the following cities: Berkeley, CA; Oakland, CA; Tunica, MS; Birmingham, AL; Newark, NJ; New Orleans, LA; and Chicago, IL. Study sites were chosen for regional diversity and documented sexual health needs [17-18]. Focus groups lasted approximately 90 minutes. All participants were required to sign a consent form (if minors, participants were required to have consent forms signed by parents or guardians). Efforts were made to increase the diversity of respondents by working with local youth-serving organizations to recruit at least half of the respondents as youth of color and/or lesbian, gay, bisexual, transgender, and queer youth. Two focus groups per site were conducted and stratified by age, one for participants younger than 18 years and the other for those over 18 years.

Focus groups were audio recorded, transcribed, and de-identified for confidentiality purposes. Using a thematic analysis approach [19], the lead authors identified major themes across focus groups, including sexual health information online, online dating, and digital flirting. Using the Cohen kappa statistic, interrater reliability was established between two coders with an overall agreement percentage of 97.45% (range 73.90% to 100% agreement across all codes) [20]. Quotes from focus groups are included to contextualize quantitative data from the national survey.

Recruitment

Participants for the national survey were recruited via community partners with flyers and postcards, as well as through panel companies (organizations that enroll interested users based on reported demographics and provide small incentives for survey completion). Participants received US \$5 in digital currency for survey completion. Focus group participants were recruited through community partners at each site using traditional methods of recruitment including flyers, postcards,

and social media posts in youth-centric locations including school campuses, clinics, and public areas. Focus group participants were compensated with a US \$25 Visa gift card for their participation.

Quantitative National Survey

A sample of 1500 youth across the United States, ages 13 to 24 years (average age 19.7 years), completed the online survey. All four regions as defined by the US Census Bureau [21] were represented: west (414/1500, 27.60%), midwest (362/1500, 24.13%), south (410/1500, 27.33%), and northeast (314/1500, 20.93%). A total of 22.40% (336/1500) of respondents were younger than age 18 years. The race and ethnicity breakdown of the sample was 57.60% (864/1500) white, 20.13% (302/1500) black/African American, 5.33% (70/1500) American Indian or Alaska Native, 9.33% (140/1500) Asian, 1.46% (22/1500) Native Hawaiian or other Pacific Islander, and 11.40% (171/1500) another race not listed. Nearly 24% (356/1500, 23.73%) of the sample identified as Hispanic or Latino. The gender breakdown was 62.00% (930/1500) women, 33.67% (505/1500) men, and 4.30% (65/1500) transgender-spectrum (consisting of individuals who identify as a different gender than their sex assigned at birth including transgender woman, transgender man, and genderqueer). Further information on employment status and education are reported in [Table 1](#).

Qualitative Focus Groups

A total of 66 youth participated in the focus groups. The demographic breakdown by race was 81% (53/66) black/African American, 3% (2/66) Asian, 2% (1/66) American Indian or Alaska Native, 2% (1/66) Native Hawaiian or other Pacific Islander, 5% (3/66) white, and 2% (1/66) another race not listed. Approximately 20% (13/66) of the sample identified as Hispanic or Latino. The gender breakdown was 52% (34/66) women, 42% (28/66) men, and 6% (4/66) transgender-spectrum. Further information on employment status and education are reported in [Table 1](#).

Table 1. Sample demographics.

Characteristic	Quantitative survey (n=1500), n (%)	Qualitative focus group (n=66), n (%)
Gender		
Male	495 (33.0)	27 (40.9)
Female	930 (62.0)	34 (51.5)
Transgender-spectrum	75 (5.0)	5 (7.6)
Age		
Minor (younger than 18 years)	336 (22.4)	28 (42.4)
Adult (18 years and older)	1164 (77.6)	38 (57.6)
Highest level of education		
Junior high or middle school	72 (4.8)	7 (10.6)
Some high school	223 (14.8)	24 (36.4)
High school diploma	444 (29.6)	14 (21.2)
Technical school	39 (2.6)	0 (0)
Some college/university	472 (31.5)	11 (16.7)
Undergraduate degree	148 (9.9)	4 (6.1)
Graduate degree	102 (6.8)	6 (9.1)
Employment		
Unemployed	247 (16.4)	25 (38.9)
Full-time student	585 (39.0)	20 (30.3)
Part-time student (less than 30 hours)	79 (5.2)	7 (10.6)
Part-time student (more than 30 hours)	84 (5.6)	4 (6.1)
Part-time	107 (7.1)	4 (6.1)
Full-time	256 (17.1)	3 (4.5)
Contract, freelance, or temporary employee	8 (0.5)	0 (0)
Self-employed	26 (1.7)	0 (0)
Homemaker (at home without children)	28 (1.8)	0 (0)
Stay-at-home parent (at home with children)	42 (2.8)	1 (1.5)
Other	38 (2.5)	2 (3.0)

Results

Using Online Dating and Online Flirting for Sexual Exploration

Current perceptions of sexual exploration and online romantic relationship formation are complicated among youth. While online romantic relationships and dating were certainly present, youth across all ages held a pervasive fear of dating violence and cyber abuse; despite this fear, many used online dating sites as a place to find and begin romantic relationships. Participants often were forced to reconcile these warring feelings in order to continue exploring sexual relationships online and accessing online dating spaces. For many younger adolescents (younger than age 18 years) using social media as an alternative space for online dating allows them to bypass feelings of fear due to the heightened level of trust in these platforms. The following results focus on these two themes, including benefits and fear of technology for sexual exploration and relationships. Both

themes are addressed below, supported by triangulated data from the national survey and regional focus groups.

Benefits of Technology for Sexual Exploration and Relationships

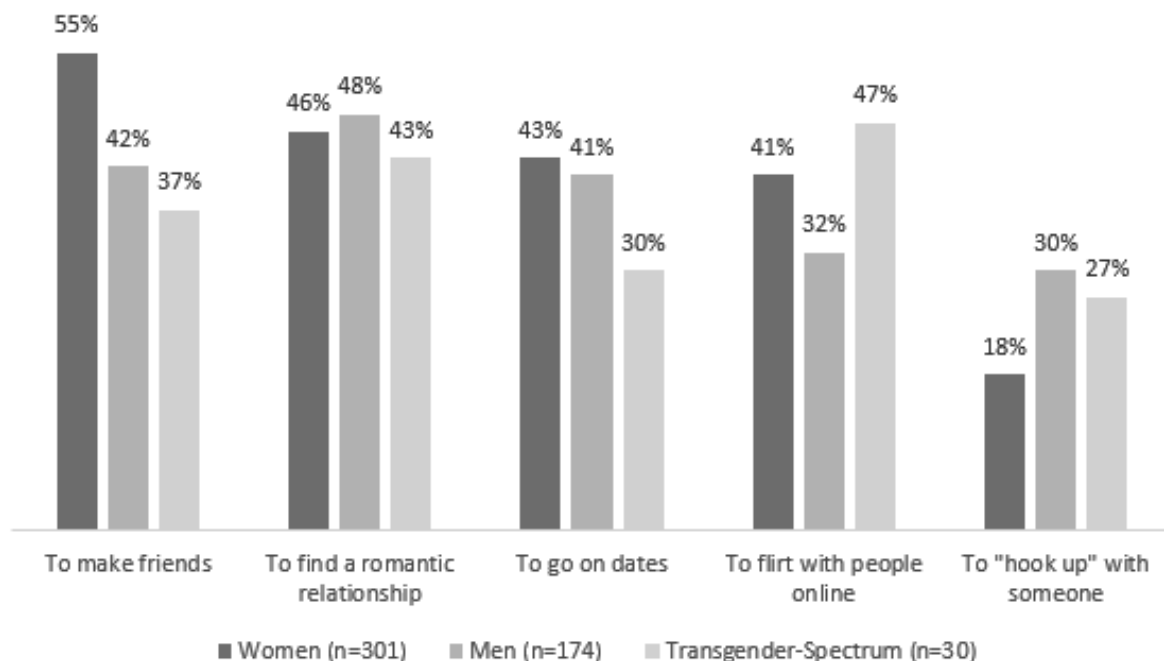
Quantitative Survey

Youth are forming romantic relationships online and using dating apps frequently. About 34.00% (510/1500) of survey respondents reported that they use dating sites/apps or have in the past. Despite the fact that these dating sites have explicit rules against minor use, under 18 youth are using these products in order to make friends and begin romantic relationships, albeit at a lower rate than their older peers (19.0% [64/336] vs 37.80% [440/1163], respectively). Nearly 80.0% (404/505) of youth who have used online dating sites met up with someone in person (45% [30/67] under 18 vs 74.0% [324/438] over 18). The most popular dating platforms were Tinder, OkCupid, and Match, with Tinder most popular among both minors (27/67, 40%) and over 18 youth (227/438, 51.8%). Of those who

reported current or past use of dating apps, under 18 youth reported that they primarily used them to make friends online (41/67, 61%) compared with over 18 youth primarily using them for finding romantic relationships (214/438, 48.9%) (see Figure 1 for further breakdown by gender). About a third (58/175, 33.1%) of men who use or have used dating sites reported using them to hook up (ie, to first meet online and then decide on a location for a casual sexual encounter) with someone. In contrast, only 17.9% (54/302) of women and 27% (8/30) of transgender-spectrum youth reported using dating sites to hook up with someone. This relatively popular use of online dating has implications for health interventions; some current sexual health education programs that use dating apps to recruit and disseminate information have shown initial promise, particularly with sexual and gender minority youth [16].

Although the majority of youth do not use online dating websites (990/1500, 66.00%), youth are still building romantic relationships online by flirting with others through social media. Digital flirting often takes the form of comments, private messaging, innuendo emojis (ie, the eggplant or water squirt emoji to mimic a penis or ejaculation), or liking someone's photos on social media. Nearly 41.00% (615/1500) of respondents said that they used social media to flirt with others. When asked how they flirted with others online, women were most likely to send messages to flirt with someone (280/346, 80.9%), men were most likely to like someone's photos (191/239, 79.9%), and transgender-spectrum youth were most likely to message someone (33/39, 85%).

Figure 1. Reasons for using a dating app by proportion and gender.



Qualitative Focus Groups

Within focus groups, online flirting was cited as the entry point for hookups and dating. Instagram, Facebook, and Snapchat were cited as places where youth could find others who have similar interests and begin conversations with potential partners. Youth explained that they flirt online to start a conversation with someone they are interested in. Usually youth flirt online with their extended network and can discover other people of interest by looking through the friend lists of people they already know, which could potentially lead to a sexual encounter. Because digital flirting often takes place over social media and behind a screen, it also allows youth to more confidently approach someone. As one over 18 youth in Oakland, CA, reports, online dating is a way to find romantic partners anywhere.

It's pretty easy to find people that don't live near you. So if you wanna find somebody without online dating you have to like go and actually find them. With online

dating you can like search everywhere. [Oakland, CA, participant; over 18 years]

Although the majority of focus group participants had experience with digital flirting or had flirted with others online, online dating sites (eg, Tinder or Grindr) elicited a more mixed response, particularly among youth under 18. Online dating sites were perceived as doors to hookup culture or a place to trick or "catfish" someone (ie, when someone has a false online dating profile with the intention of deceiving others). In rare cases, some youth did believe that you could develop romantic and long-term relationships online. However, reactions to online dating were typically negative and couched within general mistrust of using technology to meet someone.

I know like I have friends who do that, they just meet up with dudes [from Tinder] and like get them to buy bottles of stuff in exchange for like sex. And they will like sleep at their house and stuff like that for nights at a time and it's just crazy. [Oakland, CA, participant; under 18 years]

Another under 18 participant in Tunica, MS, shared fears of catfishing when using online dating.

Online dating, it's very easy for you to lie. And they can just find where you live. Just show up where you're at. [Tunica, MS, participant; under 18 years]

Despite the overall negative reaction to online dating, participants did share some positive experiences with online dating. Some focus group participants were adamant about the positives of online dating, citing it as another place to potentially find the one. Older focus group participants were very aware of some of the dangers but also understood that online dating works for some people and can result in happy relationships. Participants who identified within the lesbian, gay, bisexual, and transgender-spectrum (LGBT) community were more likely to think highly of online dating as a place to find like-minded peers.

Most participants who decried online dating for fear of false information or threats to personal security reported similar online dating behaviors through social media flirting, where they often located potential partners and began romantic and sexual relationships with those they met online. For these participants, the vetted experience of social media through connected peers eased the fear of potential violence, catfishing, or abuse.

Discussion

Principal Findings

Results indicate that online spaces often serve as primary avenues to begin romantic relationships and foster sexual identities through online flirting and dating, regardless of minor or adult status. By building these relationships online, young people are fully involved in the many facets of online sexual exploration including developing first sexual relationships and accessing online hookup culture. For many, this is a positive experience, particularly for LGBT respondents and some rurally located participants who would otherwise not be able to connect with peers. For others, online dating was interpreted as dangerous and to be avoided, despite reporting the same online activities through social media to begin romantic relationships. Due to this fear, many young people have migrated to social media as a safer online dating option, possibly due to the increased familiarity and visibility of such platforms. This is particularly true for under 18 focus group participants, who were far more likely to consider dating those they met through social media than online dating.

Limitations

The nature of our nonrepresentative sample inhibits us from being able to generalize our findings to a larger youth population. However, the large sample size provides initial steps in understanding the way in which young people use and trust the internet and social media for romantic relationship-building and sexual exploration. In addition, because the quantitative

survey was hosted online, it may be that some youth without access to personal internet-enabled devices were left out of the study. Future studies should consider representative samples, particularly among gender identity and age, as well as how to include young people who may not be reached due to location or access, including rural youth. In addition, due to the exploratory nature of this study, this paper does not provide data on predictive behavior of sexual exploration and relationship-building online among youth.

Conclusion

Results from this study indicate that online dating platforms and social media are both promising spaces for health interventions and information dissemination. Researchers and organizations should consider the unique relationships between young people, sexual health and relationships, and trusted online spaces to create tailored and effective online health programs. These complicated relationships illustrate that the internet and social media are ideal platforms for program implementation, but researchers must be conscious of the varying amounts of trust that youth place in these spaces. In particular, these data highlight the need for dating-based interventions even among minor youth, who may be left out of such programs due to the misconception that they are not using online dating for sexual exploration and relationship formation because of legal restraints. In addition, online dating is a useful platform for interventions, but researchers must not exclude dating and hookup practices on social media, which are widely used and more trusted among youth.

In addition, the interest and participation in online dating and digital flirting may provide a window into the sexual health and behaviors of young people online and could provide opportunities for enhancing healthy relationship formation. One 2017 study explored the nature of healthy online dating dynamics, which could be an important new field of sexual health education for young people [22]. Other studies have highlighted the intersection of online dating and cyberbullying [23], further stressing the need for contemporary relationship education in order to enhance sexual and mental wellness among youth.

In sum, young people are clearly using online spaces for building romantic relationships, even when they reject traditional forms of online dating. It should be further noted that this data does not provide a condemnation of these practices, but simply demonstrates the complicated relationship that youth have with online sexual exploration and relationship development. In fact, these online practices can provide many benefits as well as risks, and are vital considerations for future health programs. These findings could have implications for sexual health and risk for future programs aimed at enhancing youth sexual health. A nuanced approach that considers age difference among youth and the amount of trust in the dating platform will provide a more tailored approach to reducing negative health outcomes.

Acknowledgments

The authors would like to acknowledge the time and effort of our community partners, including Tunica Teens in Action, REALM Charter School, Hetrick-Martin Institute, University of Alabama Children's Hospital, Illinois Caucus for Adolescent Health, Berkeley Youth Alternatives, Louisiana Public Health Institute, and Vodafone Americas. Additionally, the authors would like to acknowledge our funding partner, the David and Lucile Packard Foundation.

Conflicts of Interest

None declared.

References

1. Madden M, Lenhart A, Duggan M, Cortesi S, Gasser U. Teens and technology 2013. Washington: Pew Internet and American Life Project URL: https://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2013/PIP_TeensandTechnology2013.pdf [accessed 2019-04-15] [WebCite Cache ID 77ePNhwK6]
2. Pujazon-Zazik MA, Manasse SM, Orrell-Valente JK. Adolescents' self-presentation on a teen dating web site: a risk-content analysis. *J Adolesc Health* 2012 May;50(5):517-520. [doi: [10.1016/j.jadohealth.2011.11.015](https://doi.org/10.1016/j.jadohealth.2011.11.015)] [Medline: [22525117](https://pubmed.ncbi.nlm.nih.gov/22525117/)]
3. Mitchell KJ, Ybarra ML, Korchmaros JD, Kosciw JG. Accessing sexual health information online: use, motivations and consequences for youth with different sexual orientations. *Health Educ Res* 2014 Feb;29(1):147-157 [FREE Full text] [doi: [10.1093/her/cyt071](https://doi.org/10.1093/her/cyt071)] [Medline: [23861481](https://pubmed.ncbi.nlm.nih.gov/23861481/)]
4. Buzi RS, Smith PB, Barrera C. Talk with Tiff: teen's inquiries to a sexual health website. *J Sex Marital Ther* 2015;41(2):126-133. [doi: [10.1080/0092623X.2013.857375](https://doi.org/10.1080/0092623X.2013.857375)] [Medline: [24354340](https://pubmed.ncbi.nlm.nih.gov/24354340/)]
5. Dolcini MM, Warren J, Towner SL, Catania JA, Harper GW. Information age: do urban African American youth find sexual health information online? *Sex Res Social Policy* 2015 Jun 01;12(2):110-114 [FREE Full text] [doi: [10.1007/s13178-014-0174-5](https://doi.org/10.1007/s13178-014-0174-5)] [Medline: [26052264](https://pubmed.ncbi.nlm.nih.gov/26052264/)]
6. Subrahmanyam K, Smahel D, Greenfield P. Connecting developmental constructions to the internet: identity presentation and sexual exploration in online teen chat rooms. *Dev Psychol* 2006 May;42(3):395-406. [doi: [10.1037/0012-1649.42.3.395](https://doi.org/10.1037/0012-1649.42.3.395)] [Medline: [16756432](https://pubmed.ncbi.nlm.nih.gov/16756432/)]
7. Valkenburg PM, Peter J. Online communication among adolescents: an integrated model of its attraction, opportunities, and risks. *J Adolesc Health* 2011 Feb;48(2):121-127. [doi: [10.1016/j.jadohealth.2010.08.020](https://doi.org/10.1016/j.jadohealth.2010.08.020)] [Medline: [21257109](https://pubmed.ncbi.nlm.nih.gov/21257109/)]
8. Pascoe C. Resource and risk: youth sexuality and new media use. *Sex Res Soc Policy* 2011 Mar 12;8(1):5-17. [doi: [10.1007/s13178-011-0042-5](https://doi.org/10.1007/s13178-011-0042-5)]
9. Allison S, Bauermeister JA, Bull S, Lightfoot M, Mustanski B, Shegog R, et al. The intersection of youth, technology, and new media with sexual health: moving the research agenda forward. *J Adolesc Health* 2012 Sep;51(3):207-212 [FREE Full text] [doi: [10.1016/j.jadohealth.2012.06.012](https://doi.org/10.1016/j.jadohealth.2012.06.012)] [Medline: [22921129](https://pubmed.ncbi.nlm.nih.gov/22921129/)]
10. Mustanski B, Greene GJ, Ryan D, Whitton SW. Feasibility, acceptability, and initial efficacy of an online sexual health promotion program for LGBT youth: the Queer Sex Ed intervention. *J Sex Res* 2015;52(2):220-230. [doi: [10.1080/00224499.2013.867924](https://doi.org/10.1080/00224499.2013.867924)] [Medline: [24588408](https://pubmed.ncbi.nlm.nih.gov/24588408/)]
11. Levesque DA, Johnson JL, Welch CA, Prochaska JM, Paiva AL. Teen dating violence prevention: cluster-randomized trial of teen choices, an online, stage-based program for healthy, nonviolent relationships. *Psychol Violence* 2016 Jul;6(3):421-432 [FREE Full text] [doi: [10.1037/vio0000049](https://doi.org/10.1037/vio0000049)] [Medline: [27482470](https://pubmed.ncbi.nlm.nih.gov/27482470/)]
12. Sheoran B, Braun RA, Gaarde JP, Levine DK. The hookup: collaborative evaluation of a youth sexual health program using text messaging technology. *JMIR Mhealth Uhealth* 2014;2(4):e51 [FREE Full text] [doi: [10.2196/mhealth.3583](https://doi.org/10.2196/mhealth.3583)] [Medline: [25367444](https://pubmed.ncbi.nlm.nih.gov/25367444/)]
13. Gkatzidou V, Hone K, Sutcliffe L, Gibbs J, Sadiq ST, Szczepura A, et al. User interface design for mobile-based sexual health interventions for young people: design recommendations from a qualitative study on an online Chlamydia clinical care pathway. *BMC Med Inform Decis Mak* 2015 Aug 26;15:72 [FREE Full text] [doi: [10.1186/s12911-015-0197-8](https://doi.org/10.1186/s12911-015-0197-8)] [Medline: [26307056](https://pubmed.ncbi.nlm.nih.gov/26307056/)]
14. Van Ouytsel J, Van Gool E, Walrave M, Ponnet K, Peeters E. Exploring the role of social networking sites within adolescent romantic relationships and dating experiences. *Comput Hum Behav* 2016 Feb;55:76-86. [doi: [10.1016/j.chb.2015.08.042](https://doi.org/10.1016/j.chb.2015.08.042)]
15. Lerner R. The positive youth development perspective: theoretical and empirical bases of a strengths-based approach to adolescent development. In: Lopez SJ, Snyder CD, editors. *The Oxford Handbook of Positive Psychology*. 2nd Edition. New York: Oxford University Press; 2012.
16. Lykens J, Pilloton M, Silva C, Schlamm E, Sheoran B. TECHsex: Youth Sexuality and Health Online.: YTH URL: <http://yth.org/projects/ythtechsex/> [accessed 2018-04-04] [WebCite Cache ID 6yR9WoPWe]
17. Foster PP, Thomas M, Lewis D. Reverse migration, the black church and sexual health: implications for building HIV/AIDS prevention capacity in the Deep South. *AIMS Public Health* 2016;3(2):242-254 [FREE Full text] [doi: [10.3934/publichealth.2016.2.242](https://doi.org/10.3934/publichealth.2016.2.242)] [Medline: [29546159](https://pubmed.ncbi.nlm.nih.gov/29546159/)]
18. Smith MU, DiClemente RJ. Students Together Against Negative Decisions (STAND): a peer educator training curriculum for sexual risk reduction in the rural South. *Prev Med* 2000 Jun;30(6):441-449. [Medline: [10901486](https://pubmed.ncbi.nlm.nih.gov/10901486/)]

19. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
20. McHugh M. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-282 [FREE Full text] [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]
21. Geographic terms and concepts: Census divisions and Census regions.: United States Census Bureau URL: https://www.census.gov/geo/reference/gtc/gtc_census_divreg.html [accessed 2019-01-13] [WebCite Cache ID 75OyLrsW7]
22. Howard D, Debnam K, Strausser A. "I'm a stalker and proud of it": adolescent girls' perceptions of the mixed utilities associated with internet and social networking use in their dating relationships. *Youth Soc* 2017:1-20.
23. Yahner J, Dank M, Zweig JM, Lachman P. The co-occurrence of physical and cyber dating violence and bullying among teens. *J Interpers Violence* 2015 Apr;30(7):1079-1089. [doi: [10.1177/0886260514540324](https://doi.org/10.1177/0886260514540324)] [Medline: [25038223](https://pubmed.ncbi.nlm.nih.gov/25038223/)]

Abbreviations

LGBT: lesbian, gay, bisexual, and transgender-spectrum

Edited by G Eysenbach; submitted 04.04.18; peer-reviewed by J Gaarde, B Hawkins, Z Ma; comments to author 17.08.18; revised version received 13.01.19; accepted 17.02.19; published 16.05.19.

Please cite as:

Lykens J, Pilloton M, Silva C, Schlamm E, Wilburn K, Pence E

Google for Sexual Relationships: Mixed-Methods Study on Digital Flirting and Online Dating Among Adolescent Youth and Young Adults

JMIR Public Health Surveill 2019;5(2):e10695

URL: <http://publichealth.jmir.org/2019/2/e10695/>

doi: [10.2196/10695](https://doi.org/10.2196/10695)

PMID: [31099335](https://pubmed.ncbi.nlm.nih.gov/31099335/)

©James Lykens, Molly Pilloton, Cara Silva, Emma Schlamm, Kate Wilburn, Emma Pence. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 16.05.2019. 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

Tools for the Diagnosis of Herpes Simplex Virus 1/2: Systematic Review of Studies Published Between 2012 and 2018

Zeeshaan Arshad^{1*}, BSc; Abrar Alturkistani^{2*}, BSc, MPH; David Brindley³, MEng, MSc, DPhil, FRSA; Ching Lam³, MEng; Kimberley Foley², BSc, BEd, MSc, PhD; Edward Meinert³, MA, MSc, MBA, MPA, PhD, CEng FBCS, EUR ING

¹Department of Medicine, University of Cambridge, Cambridge, United Kingdom

²Global Digital Health Unit, Imperial College London, London, United Kingdom

³Healthcare Translation Research Group, Department of Paediatrics, University of Oxford, Oxford, United Kingdom

* these authors contributed equally

Corresponding Author:

Edward Meinert, MA, MSc, MBA, MPA, PhD, CEng FBCS, EUR ING

Healthcare Translation Research Group

Department of Paediatrics

University of Oxford

Level 2, Children's Hospital

John Radcliffe Hospital

Oxford, OX3 9DU

United Kingdom

Phone: 44 7824446808

Email: edward.meinert@paediatrics.ox.ac.uk

Abstract

Background: Herpes simplex virus (HSV)-1 and HSV-2 are common infections affecting the global population, with HSV-1 estimated to affect 67% of the global population. HSV can have rare but severe manifestations, such as encephalitis and neonatal herpes, necessitating the use of reliable and accurate diagnostic tools for the detection of the viruses. Currently used HSV diagnostic tools require highly specialized skills and availability of a laboratory setting but may lack sensitivity. The numerous recently developed HSV diagnostic tools need to be identified and compared in a systematic way to make the best decision about which diagnostic tool to use. The diagnosis of HSV is essential for prompt treatment with antivirals. To select the best test for a patient, knowledge of the performance and limitations of each test is critical.

Objective: This systematic review has summarized recent studies evaluating HSV-1 and HSV-2 diagnostic tools.

Methods: Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines, selection criteria, data extraction, and data analysis were determined before the commencement of the study. Studies assessing the specificity/sensitivity of HSV-1 or HSV-2 diagnostic tools published between 2012 and 2018 were included. Quality assessment of included studies was performed using the quality assessment of diagnostic accuracy studies (QUADAS-2) tool.

Results: Searches of the PubMed database yielded 264 studies; 11 studies included 11 molecular assays, and 8 studies included 19 different serological assays for the detection of HSV-1, HSV-2, or both. A greater proportion of molecular assay-based tools are being developed by commercial entities. Studies that tested molecular assays mostly focused on cutaneous and mucosal HSV infections (n=13); 2 studies focused on ocular disease, whereas only 1 study focused on the central nervous system manifestations. The Simplexa HSV 1 & 2 Direct is currently the only Food and Drug Administration-approved device for use on cerebrospinal fluid. No tools focused on prenatal screening. We also present performance metrics of tests for benchmarking of future technology. Most of the included studies had a high risk of bias rating in half of the QUADAS-2 tool risk of bias domains.

Conclusions: The use of serologic tests to diagnose genital lesions is inappropriate because positive results may be due to chronic infection, whereas negative results may overlook recent infection. The incidence of acute infections is rising. As these infections present the greatest risk to fetuses, work needs to be done to prevent vertical transfer. Prenatal screening for primary infection and subsequent medical intervention will assist in lowering the rate of neonatal herpes. In conclusion, HSV diagnosis is moving away from culture-based methods to serology-based or polymerase chain reaction-based methods. Sensitive, rapid, and efficient HSV diagnostic tools should be adopted for the prevention of acute infections and neonatal herpes.

KEYWORDS

diagnostic techniques and procedures; herpes simplex; diagnosis

Introduction

Herpes simplex virus (HSV)-1 and HSV-2 are DNA viruses that belong to Alphaherpesvirinae, a subfamily of the Herpesviridae family [1]. They are the causative agents in a wide range of human diseases that include oral and genital mucocutaneous lesions and some rare but life-threatening conditions such as fulminate encephalitis [2].

HSV is a very common condition, with HSV-1 and HSV-2 infection estimated to affect 67% and 11% of the global population, respectively [2]. The prevalence of HSV can be even higher in low- and middle-income countries or among certain patient subpopulations. For instance, HSV-2 prevalence in sub-Saharan countries is estimated to be as high as 53.7% among individuals aged 15 to 25 years [3]. The prevalence of HSV-2 can be used as a biomarker for the HIV epidemic because of its high association among patients with an HSV-2 infection [4].

One of the most severe manifestations of HSV is encephalitis, which can have a mortality rate up to 97% [5]. Encephalitis is a rare manifestation, which could result in about 4 to 5 cases/million population/year in developed countries [6,7]. HSV infections are very common among women of reproductive age, which can increase the risk of virus transmission from the mother to the child during birth, resulting in neonatal herpes [4]. Neonatal herpes is another severe manifestation of the disease and can cause long-term health complications requiring appropriate and reliable identification of the disease.

In addition to the prevalence and potential complications, lesions caused by HSV-1 and/or HSV-2 are nonspecific and have variable presentations. Chronic carriage necessitates the use of different laboratory testing methods appropriate to each case [8]. Currently, viral culture is the most commonly used method of diagnosis, and it is considered the gold standard method [8]. This method is limited, however, by the need of a laboratory setting, aseptic technique, and variable accuracy dependent on disease stage [8]. It also takes up to 7 days to get the viral culture results, during which time the infection may have developed further. The viral culture will have a lowered sensitivity after the first 48 hours of appearance of the symptoms and is best if administered as soon as the symptoms appear [9]. Although antigens specific to HSV-1 and HSV-2 can also be readily detected by direct immunofluorescence assay, these methods have been found to lack sensitivity in nonsymptomatic patients [8].

Other methods for HSV diagnosis include serological tests, which use blood sample to check for the disease *antibodies* [9]. These tests can take between 1 day and 3 weeks [9]. Serological tests are suitable to detect the subtype of the HSV virus (HSV-1 or HSV-2) and can detect asymptomatic patients [10]. Serology-based tests can also be used to confirm a clinical diagnosis of HSV because it is a more reliable method than

clinical diagnosis [10]. It has been argued that serology-based tests are more *direct* and *economical* compared with viral cultures [10].

More modern methods used for detection of infection include techniques that detect viral DNA such as polymerase chain reaction (PCR). These methods use molecular-based assays and are widely used for the detection of infectious diseases [11]. Such methods have been found to be more sensitive and rapid and require less stringent conditions in terms of collection data [11,12]. Type-specific diagnosis is also possible using antigen detection techniques such as enzyme immunoassays targeting the glycoprotein G antigen [8].

A thorough knowledge of the performance and limitations of available tests is critical to select the most appropriate test for a patient as well as inform clinical decision making and the development of future technologies. Previous systematic reviews regarding HSV diagnostic tools reviewed only commercially available tests used to diagnose only one type of the virus (HSV-2) and only reviewed tests used in sub-Saharan Africa [13]. Another review looked at HSV diagnostic tools but limited its population only to pregnant women and neonates [14]. This systematic review summarizes recent studies evaluating HSV-1 and HSV-2 diagnostic tools. We will evaluate the study characteristics, the performance of the various tests included in the studies, and finally discuss their limitations and strengths. We will then suggest areas with unmet medical needs.

Methods

Selection criteria, data extraction, and methods of analysis were decided before the study, in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [15].

Eligibility Criteria

We selected only original articles published between 2012 and 2018. We used the Population, Intervention, Comparator, Outcome (PICO) framework to highlight the review's inclusion and exclusion criteria (Textbox 1). The population included both animal and human populations of all sex and age groups. Patients with multimorbidities were also included, for example, patients with hepatitis and/or HIV because of the possibility that HSV coexists with diseases such as HIV and hepatitis in some patient populations [16,17]. The intervention under study included diagnostic tools to detect all types of HSV-1 or HSV-2 infections including oral, genital, ocular, or central nervous system (CNS)-related infections. The elements in the PICO framework were determined as a result of a preliminary search in PubMed before the formulation of the review's eligibility criteria. To help compare studies and their outcomes, we only included studies that reported sensitivity and/or specificity as performance indicators of the diagnostic tools [18]. We excluded book chapters, reviews and reports, and articles not in the English language. To capture the most up to date evidence, we

restricted the search to the years 2012 to 2018 ([Multimedia Appendix 1](#)).

Search Strategy and Sources of Information

The complete search strategy used to identify studies is detailed in [Multimedia Appendix 1](#). PubMed (MEDLINE) was searched from January 1, 2012 to December 31, 2018. We only searched the PubMed database because most of the studies concerning diagnostic tools are indexed in MEDLINE [19]. We also searched the bibliographies of relevant articles and reviews that we identified in our initial scoping searches.

Study Selection

Studies were selected first by removing duplicates from the initially identified studies, then manually screening the titles and abstracts of the studies against the eligibility criteria, and finally performing full-text reading of eligible studies. ZA and AA completed study screenings. Study records were managed using Mendeley Desktop (v1.16.3) software.

Data Extraction

Data extraction sheets were developed before data extraction. Data extraction was completed separately by the study authors ZA and AA to increase accuracy. Items collected in the data extraction sheets were identified and finalized through iterative discussions with the study authors.

The data extraction sheets collected information on the following:

1. General study information including country of the study, HSV subtype, sample size, study design (retrospective or prospective), the name of the diagnostic test under study, as well as the comparison tools.
2. Methodological details including sample collection method, transportation/storage method, study aims, and conclusions.
3. After identification of the diagnostic test under study, the tests were categorized as follows: serology-based HSV detection assays, molecular assays for detection of HSV, or culture-based methods. Tables for each method were then created that summarized the performance of each test. Although sensitivity and specificity were our primary performance metrics, agreement (total, positive, and negative), kappa, area under receiver operator curve, negative predictive value, and positive predictive value were also recorded. Cost and time to run the test were also extracted. In addition, the Food and Drug Administration and European Medicines Agency status of each included diagnostic tool were ascertained through search of the respective organizations websites [20,21]. For studies that reported analytical performance and clinical performance, only the clinical performance results were reported in this review.

Risk of Bias in Studies

The risk of bias of included studies was assessed using the QUADAS-2 tool [22].

Textbox 1. Review inclusion and exclusion criteria reported using the Population, Intervention, Comparator, Outcome framework.

Population	<ul style="list-style-type: none"> • All animal and human populations, including individuals of all ages and both sexes such as neonates, infants, children, adolescents, and pregnant women. • Individuals with multimorbidity will also be included such as patient coinfecting with herpes simplex virus (HSV) and HIV, or HSV and hepatitis C.
Intervention	<ul style="list-style-type: none"> • All diagnostic tests for HSV-1 or HSV-2 used for any type of the infection such as oral, anogenital, ocular, or central nervous system infections. • Diagnostic tests that detect other viruses in addition to HSV-1 and HSV-2 will not be included.
Comparator	<ul style="list-style-type: none"> • Not specified.
Outcome	<ul style="list-style-type: none"> • Sensitivity and/or specificity as performance measured of the diagnostic test.
Dates	<ul style="list-style-type: none"> • Database searches will be limited to the dates April 17, 2012 to December 31, 2018.

Results

Search Results and Study Inclusion

Following PRISMA guidelines [23], we identified 264 records from the PubMed search and an additional 30 through the bibliography search. There were no duplicates in the records identified; hence, all 294 record titles and abstracts were

screened for eligibility; 235 articles were excluded based on the eligibility criteria. Full-text reading of the 59 screened articles resulted in a total of 19 full-text studies for inclusion in this systematic review ([Multimedia Appendix 2](#)). We screened the abstracts of 264 articles for relevance and found 59 records that could not be excluded based purely on reviewing the abstracts and titles and these were assessed as full texts. We subsequently excluded 40 articles if, after deeper examination, we found that

there was no performance metric measured (n=8), the article did not include a diagnostic tool (n=9), study described viruses other than HSV-1 or HSV-2 (n=11), sensitivity and/or specificity were not measured (n=8), or the article was a review (n=4).

Study Characteristics

Summaries of the 19 included studies can be found in [Tables 1 and 2](#). The majority of studies were prospective in nature (n=12; fewer studies (n=7) relied on archived samples ([Table 1](#)). The number of studies over the last 6 years has grown, reaching its highest in 2017: from 5 studies in 2012 to 11 in 2017 ([Table 1](#)). This growth declined in 2018, with only 1 study found to be focused on HSV-1 or 2 diagnostic tools. Although the search was completed on December 31, 2018, there may be other 2018 studies that were yet to be published on the day of the search. Sample sizes within studies ranged from 60 in Al - Shobaili et al [[24](#)] and Loughman et al [[25](#)] to 3408 in Patel et al. [[26](#)] with a median sample size being of 179 patients ([Table 1](#)). Studies investigated diagnostic tools for HSV-1/2 (n=10), HSV-2 only (n=7), or HSV-1 only (n=2), and in 1 study, the viral subtype could not be ascertained [[27](#)], so it was labelled as HSV-1/2. In addition, 58% of tests were molecular assays, whereas the remainder used serology-based HSV detection methods. No study used culture-based method as their primary method; studies did use them as a comparison method ([Table 1](#)). In terms of disease state, 2 studies investigated patients with ocular disease [[27,28](#)], 1 for CNS infection [[29](#)], and the remainder for patients with suspected mucosal and/or cutaneous lesions ([Table 1](#)). Given the heterogeneous nature of the studies included within this review in terms of sample type, collection, and patient demographics, it was not possible to draw statistical inferences or perform a meta-analysis.

The methodological characteristics of included studies are presented in [Table 2](#). Basic demographic information was reported in 13 studies. In studies that reported participant gender (n=11), between 30% and 92% of study participants were male. Participants' age ranged from 5 months to 85 years ([Table 2](#)).

Diagnostic Test Performance

Molecular Assays

Molecular assays for detection of HSV are summarized in [Table 3](#). A total of 11 diagnostic tools were investigated within the 11 studies (including 1 study that used both molecular and serological methods in conjunction [[27](#)], out of which 6 have received FDA approval; [Table 3](#)). Some FDA-approved tools had an approval for only some components of the test or certain uses of the test only. For example, the Viper HSV-Qx assay [[34](#)] and the BD ProbeTec HSV-Qx (HSVQx) system [[40](#)] were only approved for anogenital lesions, Luminex ARIES HSV-1 & 2 Assay [[35](#)] was approved for cutaneous or mucocutaneous lesions only; 8 study samples were collected using mucosal or cutaneous swabs ([Table 3](#)); in the remaining 3 studies, 1 collected samples through corneal scrapings or conjunctival swabs, 1 collected CSF, and 1 tears (for ocular manifestations; [Multimedia Appendix 3](#)).

The performance metrics of the various molecular tests are summarized for comparison with future technologies summarized in [Figures 1 and 2](#). One study [[35](#)] reported positive percent agreement and negative percent agreement instead of sensitivity and specificity, which is equivalent to the sensitivity and specificity measures in the absence of a standard [[43](#)]. Sensitivities for molecular methods ranged from 76.9% (27) to 100% (for a number of tools) ([Figures 1 and 2](#)). The study by Shoji et al [[27](#)] was the single study that used both a serological and molecular technique in conjunction using tears as a sample material. The study reported the lowest specificity of 82.6% using their ELIZA and PCR techniques. The second worst sensitivity was indicated by a real-time singleplex PCR tested in an academic center by Barrado et al [[28](#)]. Here, a sensitivity of 77.8% using corneal scrapings was found [[28](#)]. All other values were above 90% ([Multimedia Appendix 3](#); [Figures 1 and 2](#)). Gitman et al [[31](#)] aimed to compare cell culture, DFA, and a laboratory-developed real-time TaqMan PCR (LDT HSV PCR) for the detection of HSV in dermal, genital, ocular, mouth, or other swab samples [[31](#)]. Conventional culture was found to have a sensitivity and specificity of 87.9 (0.768-0.943) and 99.1 (0.945-1.000), respectively. The difference when using direct immunofluorescence, TaqMan PCR, or Simplexa Direct PCR did not improve these metrics [[31](#)].

Table 1. Summary of included studies. Table summarizes study location, herpes subtype, sample size, study design (retrospective or prospective), and the name of the diagnostic test under study as well as the tools it will be compared with.

Study ID	Country	Study design	HSV ^a -1/2	Sample size	Disease	Test	Test type	Comparison
Al - Shobaili et al [24]	Saudi Arabia	Prospective	HSV-2	60: 35 male and 25 female	Genital ulcer disease (genital)	Herpe Select Express Rapid Test (Focus Technologies Inc, Cypress, CA)	Immunochromatographic (serology-based HSV detection assay)	HerpeSelect 2 IgGELISA (Focus Technologies Inc, Cypress, CA); Kalon HSV-2 IgG ^b ELISA ^c assay (Kalon Biological Ltd, Guilford, United Kingdom); MAb-EIA ^d
Barrado et al [28]	Spain	Prospective	HSV-1	188	Suspicion of herpetic epithelial keratitis and nonherpetic corneal ulcer (ocular)	Real-time single-plex PCR ^e	Molecular assays for detection of HSV	Cell Culture- csPCR, csCC, cswPCR, csWCC
Burton et al [30]	United States	Retrospective	HSV-2	84	Veterans with hepatitis C and HSV-2 coinfection	Focus HerpeSelect Hsv-2 IgG (Focus Technologies Inc, Cypress, CA)	Serology-based HSV detection assay	Biokit HSV-2 rapid assay (Biokit United States, Lexington, MA)
Gitman et al [31]	United States	Prospective	HSV-1, HSV-2	171	Dermal, genital, ocular, mouth	Simplexa HSV 1 & 2 PCR (Focus Diagnostics, Cypress, CA)	Molecular assays for the detection of HSV	Cell culture (Diagnostic Hybrids, Athens, OH); Cytospin-enhanced DFA ^f ; Real-time Taq-Man PCR (LDT ^g HSV PCR)
Granato et al [32]	United States	Prospective	HSV-1, HSV-2	1351	Cutaneous and mucocutaneous herpes infection (cutaneous)	AmpliVue HSV-1 & 2 assay (Quidel, San Diego, CA)	Molecular assays for detection of HSV	ELVIS HSV ID and D3 Typing System (Quidel DHI, Athens, OH)
Hobbs et al [33]	Kenya	Retrospective	HSV-2	198	Not specified	HerpeSelect 2 (Focus Diagnostics, Cypress, CA); Kelon ELISA (Kalon Biological, Guildford, United Kingdom)	Serology-based HSV detection assay	Same tests but with serum sample
Lang et al [34]	Canada	Prospective	HSV-1, HSV-2	276	Anogenital, oral	Viper HSV-Qx assay (BD Molecular Diagnostics)	Molecular assays for detection of HSV	LightCycler 2.0 platform (HSV-LC) (Roche Diagnostics, Basel, Switzerland)
Lee et al [35]	Singapore	Retrospective	HSV-1, HSV-2	117	Not specified	Luminex ARIES HSV-1 & 2 assay (Luminex Corp, Austen, TX)	Molecular assays for detection of HSV	Forster resonance energy transfer-based PCR assay and FTD Neuro 9 assay (Fast Track Diagnostics, Junglinster, Luxembourg)

Study ID	Country	Study design	HSV ^a -1/2	Sample size	Disease	Test	Test type	Comparison
Liermann et al [36]	Germany	Retrospective	HSV-1, HSV-2	263	Not specified	Orgentec ELISA Anti-HSV-2 IgM (Mainz, Germany); Orgentec ELISA Anti-HSV-1 IgM (Mainz, Germany); Orgentec ELISA anti-HSV-1/2 IgM (Mainz, Germany); Serion ELISA classic HSV-2 IgM IgM (Würzburg, Germany); Serion ELISA classic HSV-1 IgM (Würzburg, Germany); Serion ELISA classic HSV-1 + 2 IgM IgM (Würzburg, Germany); Orgentec ELISA Anti-HSV-2 IgG (Mainz, Germany); Orgentec ELISA anti-HSV-1 IgG (Mainz, Germany); Serion ELISA classic HSV-2 IgG IgM (Würzburg, Germany); Serion ELISA classic HSV-1 IgG IgM (Würzburg, Germany); Serion ELISA classic HSV-1 + 2 IgG IgM (Würzburg, Germany)	Serology-based HSV detection assay	HerpeSelect 1 ELISA IgG (Focus Diagnostics, Cypress, CA); HerpeSelect 2 ELISA IgG (Focus Diagnostics, Cypress, CA); Immunoblot assay recomLine HSV-1 & 2 IgG (Mikrogen, Neuried, Germany)
Liu et al [37]	China	Prospective	HSV-2	318	Not specified	gG _{321580His} ELISA test	Serology-based HSV detection assay	HerpeSelect 2 ELISA IgG (Focus Diagnostics, Cypress, CA)
Loughman et al [25]	France	Retrospective	HSV-2	60: 42 female and 18 male	Not specified	HSV-2 biochip test using reagents from the HerpeSelect® 2 IgG kit (Focus Diagnostics, CA; product EL0920G)	Serology-based HSV-2 immunoassay with biochip	LIAISON HSV-2 type specific IgG chemiluminescent immunoassay test (DiaSorin, MN)
Binnicker et al [29]	United States	Prospective	HSV-1, HSV-2	100	Central nervous system infection due to herpes simplex virus (CNS)	Simplexa HSV 1 & 2 Direct (Focus Diagnostics, Cypress, CA)	Molecular assays for detection of HSV	Roche HSV-1/2 ASR (Roche Diagnostics, Indianapolis)
Miller et al [38]	United States	Prospective	HSV-1, HSV-2	179	Genital herpes	IDbox HSV-1/2 assay (GenturaDx)	Molecular assays for detection of HSV	Culture; MultiCode HSV PCR
Parra-Sánchez et al [39]	Spain	Prospective	HSV-1, HSV-2	283	Various mucosal and cutaneous lesions, CSF, urinary	HSV OligoGen kit (Operon-Immuno&Molecular Diagnostics, Zaragoza, Spain)	Molecular assays for detection of HSV	Roche LightCycler HSV ₆ Qual Kit assay (Roche Diagnostics, Basel, Switzerland)

Study ID	Country	Study design	HSV ^a -1/2	Sample size	Disease	Test	Test type	Comparison
Patel et al [26]	South Africa and Zambia	Retrospective	HSV-2	3408	HSV-2/HIV-1 coinfecting individuals and HIV-uninfected heterosexual partners	Kalon HerpeSimplex virus type 2 IgG ELISA (Kalon Biological Ltd, Surrey, United Kingdom)	Serology-based HSV detection assay	Washington HSV western blot
Van Der Pol et al [40]	United States	Prospective	HSV-1, HSV-2	508	Anogenital infections (anogenital)	BD ProbeTec HSV-Qx (HSVQx) system (BD Diagnostics, Sparks, MD)	Molecular assays for detection of HSV	Elvis culture system (San Diego); Laboratory-developed PCR assay
Shevlin and Morrow [41]	United States	Retrospective	HSV-2	100	Not specified	Uni-Gold HSV-2 rapid (Trinity Biotech, Ireland)	Serology-based HSV detection assay	Washington HSV western blot
Shoji et al [27]	Japan	Prospective	Viral subtype could not be ascertained	59 + 23 healthy volunteers	Herpes simplex keratitis (ocular)	HSV DNA (by PCR) and HSV specific sIgA antibody levels (ELIZA) in tears	Serology-based HSV detection assay and molecular assays for detection of HSV	— ^h Control patients?
Tong et al [42]	United States	Prospective	HSV-1, HSV-2	176	Genital and oral lesions (oral and genital)	IsoGlow HSV typing assay (BioHelix Corp, Beverly, MA)	Molecular assays for detection of HSV	IsoAmp HSV (BioHelix Corp, Beverly, MA); ELVIS Shell Vial Assay (Diagnostic Hybrids, Athens, OH)

^aHSV: herpes simplex virus.

^bIgG: immunoglobulin G.

^cELISA: enzyme-linked immunosorbent assay.

^dMAb-EIA: monoclonal antibody enzyme immunoassay.

^ePCR: polymerase chain reaction.

^fDFA: direct fluorescent antibody.

^gLDT: laboratory developed test.

^hNot applicable.

Table 2. Methodological characteristics of included studies, summarizing study aims, conclusions, and demographic data.

Study ID	Patient demographic data	Aims/rationale	Conclusions
Al - Shobaili et al [24]	58% male; mean age 37 (SD 13.6) years	To evaluate a point-of-care test (HerpeSelect Express Rapid Test) for more rapid turnaround of results in a nonlaboratory setting.	The HerpeSelect Express Rapid Test has adequate sensitivity and specificity for confirming HSV ^a -2 infection in patients with genitourinary disease. The test is good at diagnosing high-risk individuals.
Barrado et al [28]	52.2% male; mean age 56.9 (SD 19.7) years	Assess if conjunctival swab samples were equivalent to corneal scrapings to diagnose herpetic keratitis.	Conjunctival swabs may serve as supplemental method for the diagnosis of typical HK ^b despite limited sensitivity when the collection of corneal scrapings would not be feasible. PCR ^c must be considered the gold standard for diagnosis of typical HK.
Burton et al [30]	92% male; 67% African American; 33% white; 97% heterosexual	To assess type-specific tests for HSV-2 in patients with chronic hepatitis C infection	In veterans with chronic hepatitis C infection, HerpeSelect 1 HSV-2 index values between 1.1 and 2.89 should be confirmed with an alternate test for HSV-2 infection. Only effective within certain range.
Gitman et al [31]	26.4% male; 148 adult; 23 pediatric	Compare the performance, time to result, and cost of the Simplexa HSV 1 & 2 Direct PCR with those of conventional	Simplexa HSV 1 & 2 Direct PCR is expensive but required the least training, had the lowest hands-on time and fastest assay time (75 min vs 3 hours by LDT ^d HSV PCR), detected most positives, considered an internal control, and provided the HSV type.
Granato et al [32]	— ^e	To evaluate the performance of Amplivue HSV-1 + 2 assay compared with ELVIS HSV IDD3 in detecting HSV-1-2 from clinical specimens.	The results of this study show that the AmpliVue HSV-1/2 assay was more sensitive than ELVIS culture for detecting HSV-1 and HSV-2 in a wide variety of cutaneous and mucocutaneous specimens
Hobbs et al [33]	198 adult male and female; 18 samples used for Kalon assay; 178 samples used for the Focus assay	To evaluate and validate the use of Focus and Kalon ELISAs ^f in detecting HSV-2 antibodies from dried blood sample as part of an HIV prevention trial.	The use of dried blood spots can reduce time and provide an effective way to test for HSV-2 in resource-limited settings. The study found that the use of dried blood spots with the Kalon assay did not perform well and that the use of dried blood spots with the Focus assay resulted in comparable sensitivity and specificity to the use of serum sample.
Lang et al [34]	—	Compare HSV-Q and real-time HSV PCR in terms of accuracy and cost. Device approved for anogenital infections; this study investigates its use on lesions from other anatomical locations.	The HSV-Qx assay was found to be highly sensitive and accurate. Gray zone may be required for specimens with values falling between 50 and 800 maximum relative fluorescence units. HSV-Qx has a lower cost per specimen (\$22) compared with that of HSV-LC (\$34). Samples lying near the positivity cut off should be retested.
Lee et al [35]	—	To evaluate the analytical and clinical performance of the Luminex ARIES assay in detecting HSV-1 and HSV-2 compared with an approved assay for the HSV testing for the diagnosis of meningitis/encephalitis.	The analytical performance results of the assay showed that the assay had lower sensitivity than the comparator. On the other hand, the clinical performance results of the assay showed that it had comparable performance to the routinely used assay. Compared with 2 other routinely used assays, the ARIES assay required the shortest amount of hands-on and assay time and was the least labor-intensive. The study concluded that Luminex ARIES assay can be used for successful detection of HSV-1 and HSV-2 for skin and genital infections, meningitis, and encephalitis.
Liermann et al [36]	Healthy children aged between 5 months and 3 years, healthy voluntary blood donors aged 18 to 65 years, or hospitalized patients aged between 14 and 70 years	Evaluate the HSV type-common and type-specific IgG ^g and IgM ^h enzyme-linked immunosorbent assays for the diagnosis of acute and latent HSV infections.	HSV type-common IgM ELISAs can be useful to confirm acute newly acquired HSV infections; the use of single-type IgM ELISAs on the basis of whole-virus antigen is dispensable.
Liu et al [37]	64% male; mean age 35.9 (SD 4.52) years for males; mean age 30.7 (SD 4.65) years for females	Need for a convenient, high-quality, rapid, and inexpensive domestic serodiagnostic kit that differentiates between HSV-1 and HSV-2.	The study indicates that gG _{321-580H} has a high diagnostic potential for HSV-2 virus serodiagnosis in humans.

Study ID	Patient demographic data	Aims/rationale	Conclusions
Loughman et al [25]	30% male; median age of 40.5 (range: 15-85)	To validate the HSV-2 biochip assay as a point-of-care test to be able to use it in resource-limited settings.	Membrane-touch biochip requires some improvements such as expanded calibration before it can be used as a rapid diagnosis tool. The tool was validated and performs comparable to other standard diagnostic methods. The study concluded that the assay could work potentially as a rapid, point-of-care test for HSV-2.
Binnicker et al [29]	—	Investigates a more rapid, cheaper method of detecting HSV in CSF ⁱ fluid that only uses a small amount of fluid. This makes it suitable for use in neonates.	The Simplexa HSV-1/2 assays demonstrated a combined sensitivity and specificity of 96.2% and 97.9%, respectively. Assay does not require nucleic acid extraction. Results are available in 60 min. The Simplexa assay requires only 50 mL of CSF.
Miller et al [38]	—	New system designed to maximize sensitivity. This study assesses whether this is achieved and is fully automated.	Assay shows acceptable clinical performance characteristics and demonstrates promise for further development of this fully automated platform for detection of pathogen nucleic acid in clinical laboratories
Parra-Sánchez et al [39]	42.2% male; median age 30.5 (Q1-Q3 24.0-39.0) years	Evaluate a new assay, the HSV OligoGen Kit for the detection of HSV in several types of clinical samples	Detection of HSV-1 and HSV-2 in clinical samples by HOK ^j was not significantly different from detection by LCHK assay ($P \geq .8$, t test). Statistical data obtained in this study confirm the usefulness and reliable results of this new assay from a variety of specimens.
Patel et al [26]	51% male; median age (IQR) of 32.0 (27-39); HIV-positive: 50%	Evaluate an onsite HSV tool in a population of HIV infected and noninfected patients and to evaluate the precision of the Kalon HSV-2 IgG ELISA.	In populations with optimal diagnostic accuracy, Kalon is a reliable stand-alone method for on-site HSV-2 IgG antibody detection. Kalon can be utilized in resource-limited settings, enhancing the feasibility to monitor the epidemic and assess intervention efforts.
Van Der Pol et al [40]	34% male; median age: 25 years (range: 17-71)	Results of a multicenter study evaluating the performance of a recently FDA ^k -approved, commercially available, type-specific nucleic acid amplification test that allows type-specific HSV discrimination.	Assay performs as well as the other assays on a fully automated system that provides results within a few hours rather than many days. No differences in test performance based on gender, clinic type, location of the lesion, or type of lesion were observed.
Shevlin and Morrow [41]	35% male (estimate); Pediatric sera (age <18 years) were not included	Evaluation of a point-of-care and inexpensive device to detect anti-HSV antibody in serum or blood.	UHR is a reliable, low-cost alternative to other point-of-care HSV-2 diagnostic tests. Showed both sensitivity and specificity in a small group of adults.
Shoji et al [27]	51% male	Assesses whether the combined utilization of IgA antibody (HSV-sIgA) levels in tears and real-time PCR can improve diagnostic ability.	The combination of laboratory detection of HSV DNA by real-time PCR and of HSV-sIgA by ELISA using tear samples enables higher reliability in diagnosing the subgroups of HSK ^l .
Tong et al [42]	—	Substantial market need for a low-cost, point-of-care HSV typing assay. Such a device is assessed within this study.	Both formats of the IsoGlow HSV typing assay had sensitivities comparable to that of the FDA-cleared IsoAmp HSV test and specificity for the 2 types of HSV comparable with that of ELVIS HSV turnaround time of around 1 hour.

^aHSV: herpes simplex virus.

^bHK: Herpes Keratitis.

^cPCR: polymerase chain reaction.

^dLDT: laboratory developed test.

^eNot applicable.

^fELISA: enzyme-linked immunosorbent assay.

^gIgG: immunoglobulin G.

^hIgM: immunoglobulin M.

ⁱCSF: cerebrospinal fluid.

^jHOK: HSV OligoGen kit.

^kFDA: Food and Drug Administration.

^lHSK: Herpes Simplex Keratitis.

Table 3. Molecular assays for the detection of herpes simplex virus, summarizing the different molecular assays studied within the 10 included studies. Information includes regulatory status, collection/storage/transport method, and performance.

Test	Study ID	Manufacturer	FDA ^a status	EMA ^b status	Sample type	Collection method
Real-time singleplex PCR ^c	Barrado et al [28]	University Hospital Madrid, Spain (academic)	No	No	Corneal scrapings and conjunctival swabs	Corneal scrapings by a platinum Kimura spatula using as viral transport medium UTM (Universal transport medium, Copan Diagnostic, Inc). Conjunctival swabs were obtained by polyester swabs.
Simplexa HSV 1 & 2 Direct	Gitman et al [31]	Focus Diagnostics, Cypress, CA	Yes	Unclear	Dermal, ocular, mouth	Swabs were collected from local clinics, doctors' offices, and inpatient wards in 3 mL of M4 viral transport medium (Remel, Lenexa, KS)
Simplexa HSV 1 & 2 Direct	Binnicker et al [29]	Focus Diagnostics, Cypress, CA	Yes	Unclear	CSF ^d	Unknown how CSF collected.
AmpliVue HSV-1 & 2 assay	Granato et al [32]	Quidel, San Diego, CA	Yes	Unclear	Cutaneous and mucocutaneous swabs	All specimens were collected on swabs, transported to the laboratory in viral transport medium
Viper HSV-Qx assay	Lang et al [34]	BD Diagnostics, Sparks, MD	Yes (only for anogenital lesions)	Unclear	Anogenital and oral swabs	— ^e
Luminex ARIES HSV-1 & 2 assay	Lee et al [35]	Luminex Corp, Austen, TX	Yes (only for cutaneous or mucocutaneous lesion samples)	Unclear	Lesion swab samples	Lesion swab samples were collected in clinic and deidentified for the study
BD ProbeTec HSV-Qx (HSVQx) system	Van Der Pol et al [40]	BD Diagnostics, Sparks, MD	Yes (only for anogenital lesions)	—	Anogenital swabs	Sampled using a polyester swab, which was placed directly into universal viral transport medium (Becton, Dickinson [BD], Sparks, MD), and then using a second, foam-tipped swab, which was placed directly into the BDQx liquid wet-swab transport medium (Qx WS).
IDbox HSV-1/2 assay	Miller et al [38]	GenturaDx, Hayward, CA	No	No	Genital swabs	Swabs collected in 3 mL of universal transport medium, transported to labs, aliquoted samples stored in -70°C
HSV OligoGen kit	Parra-Sánchez et al [39]	Operon-Immuno&Molecular Diagnostics, Zaragoza, Spain	No	No	Various swabs; 110 ulcer specimens, 48 urine, 48 endocervical, 43 CSFs, 4 urethral, and 3 pharyngeal swabs	Cobas PCR media urine sample kit or cobas PCR media female swab sample kit (Roche Diagnostics GmbH, Mannheim, Germany); 110 ulcer specimens, 48 urine, 48 endocervical, 43 CSF, 4 urethral, and 3 pharyngeal swabs
HSV DNA (by PCR) and HSV specific sIgA antibody levels (ELIZA)	Shoji et al [27]	School of Medicine, Tokyo, Japan (academic)	No	No	Tears	Using Schirmer strips (Schirmer Tear Production Measuring Strips; Showa- Yakuhihako, Tokyo, Japan)
IsoGlow HSV typing assay	Tong et al [42]	BioHelix Corp, Beverly, MA	No	Unclear	Swabs	Clinical swabs suspended in vial transport medium were collected from the Cleveland Clinic (Cleveland, OH). The samples were shipped on ice for overnight delivery and were aliquoted on receipt. Some were placed at -80°C for long-term storage, and some were placed at -20°C for short-term storage or near-term testing.

- ^aFDA: Food and Drug Administration.
- ^bEMA: European Medicines Agency.
- ^cPCR: polymerase chain reaction.
- ^dCSF: cerebrospinal fluid.
- ^eNot applicable.

Figure 1. Sensitivity and specificity of molecular-based diagnostics. Because of the heterogeneous nature of studies included in this systematic review, it is not possible to draw strict conclusions with regard to performance metrics. This figure, therefore, aims solely to act as a means of visually displaying the sensitivity and specificity results of each study. HSV: herpes simplex viruses.

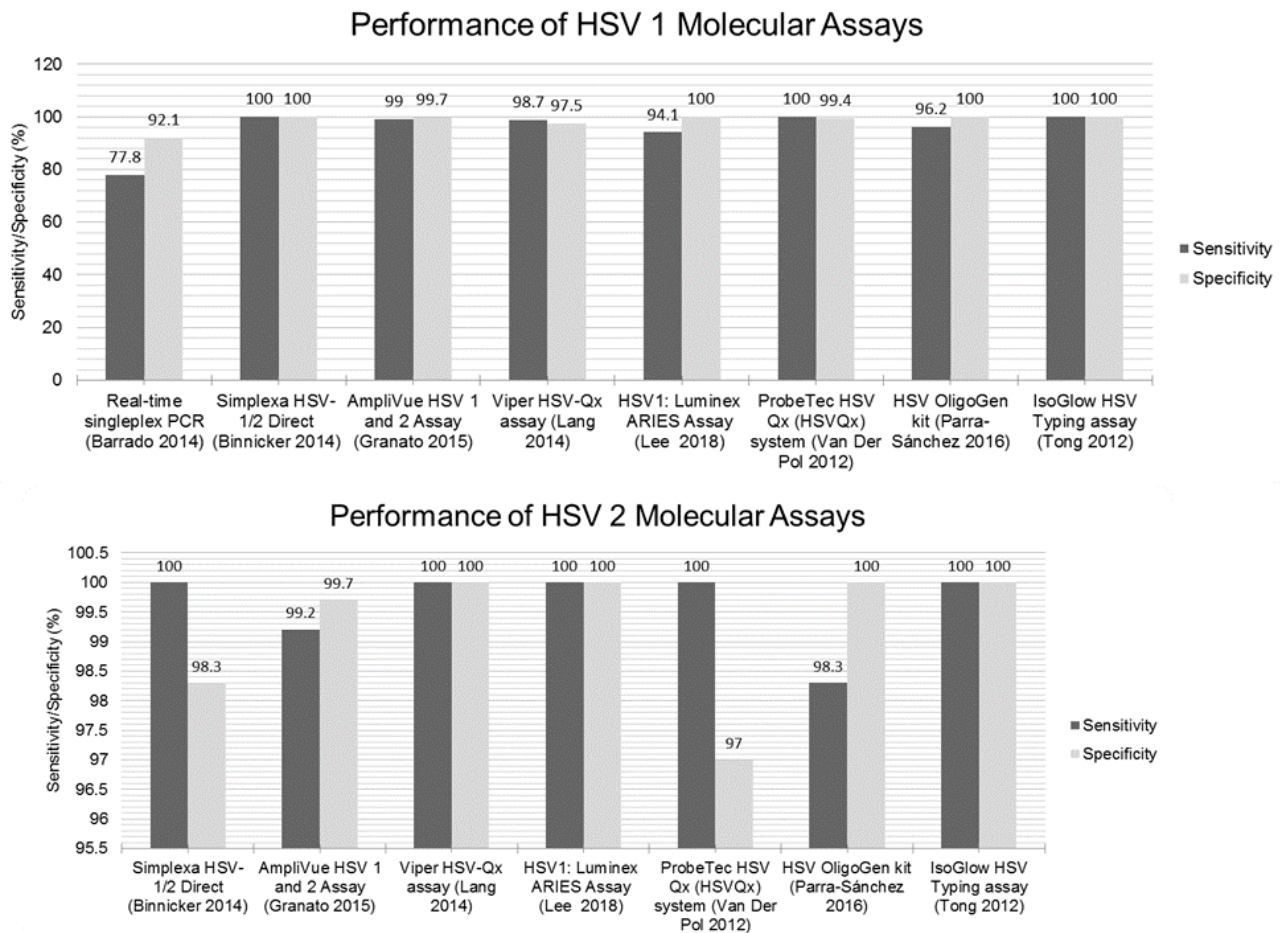
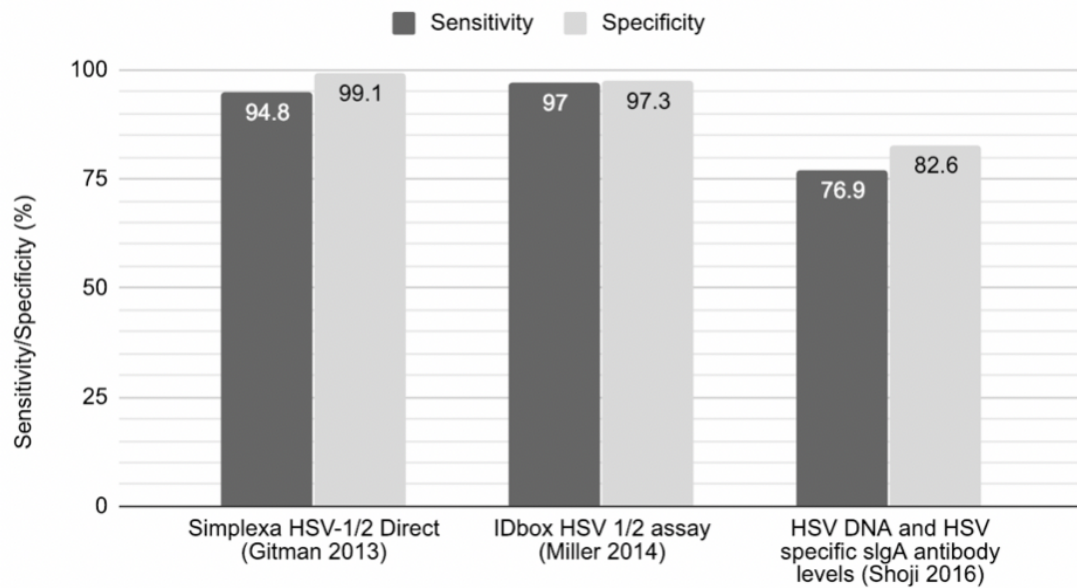


Figure 2. Sensitivity and specificity of molecular-based diagnostics. Because of the heterogeneous nature of studies included in this systematic review, it is not possible to draw strict conclusions with regard to performance metrics. This figure, therefore, aims solely to act as a means of visually displaying the sensitivity and specificity results of each study. HSV: herpes simplex viruses.



Serological Assays

Serological assays for the detection of HSV are summarized in [Multimedia Appendix 3](#). A total of 19 serological diagnostic tools were investigated within 8 studies (plus 1 study [27] that used both molecular and serological methods in conjunction—detailed in [Table 3](#)); 4 of the serological diagnostic tools have received FDA approval ([Multimedia Appendix 3](#)) of which HerpeSelect 2 [33] was approved for use with serum only, and the test used by Loughman et al [25] only used reagents from the FDA-approved HerpeSelect 2 IgG kit. All of the studies in this category used serum samples, and only 1 study also tested the use of dried blood spots instead [33]. The performance metrics of the various tests are summarized for comparison to future technologies ([Multimedia Appendix 3](#) and [Figure 3](#)). Sensitivity for serological tools ranged from 39% in

hepatitis-coinfected individuals using the Herpeselect Hsv-2 IgG [30] tool to 100% (for a number of tools; [Figure 4](#)). All others were equal to or above 80% ([Multimedia Appendix 3](#) and [Figure 3](#)). The study by Liermann et al [36] investigated the detection of primary infection and concluded that its HSV IgM serology tool should not be used to make decisions for antiviral treatment in HSV.

Cost

Gitman et al [31] carried out the most comprehensive cost analysis, comparing the performance, time to result, and cost of the Simplexa HSV 1 & 2 Direct PCR with those of conventional cell culture, DFA, and a laboratory-developed real-time TaqMan PCR (LDT HSV PCR). The results of this are displayed in [Table 4](#) (table was taken directly from the corresponding article [31]).

Figure 3. Sensitivity and specificity of molecular-based diagnostics. Because of the heterogeneous nature of studies included in this systematic review, it is not possible to draw strict conclusions with regard to performance metrics. This figure, therefore, aims solely to act as a means of visually displaying the sensitivity and specificity results of each study. HSV: herpes simplex viruses.

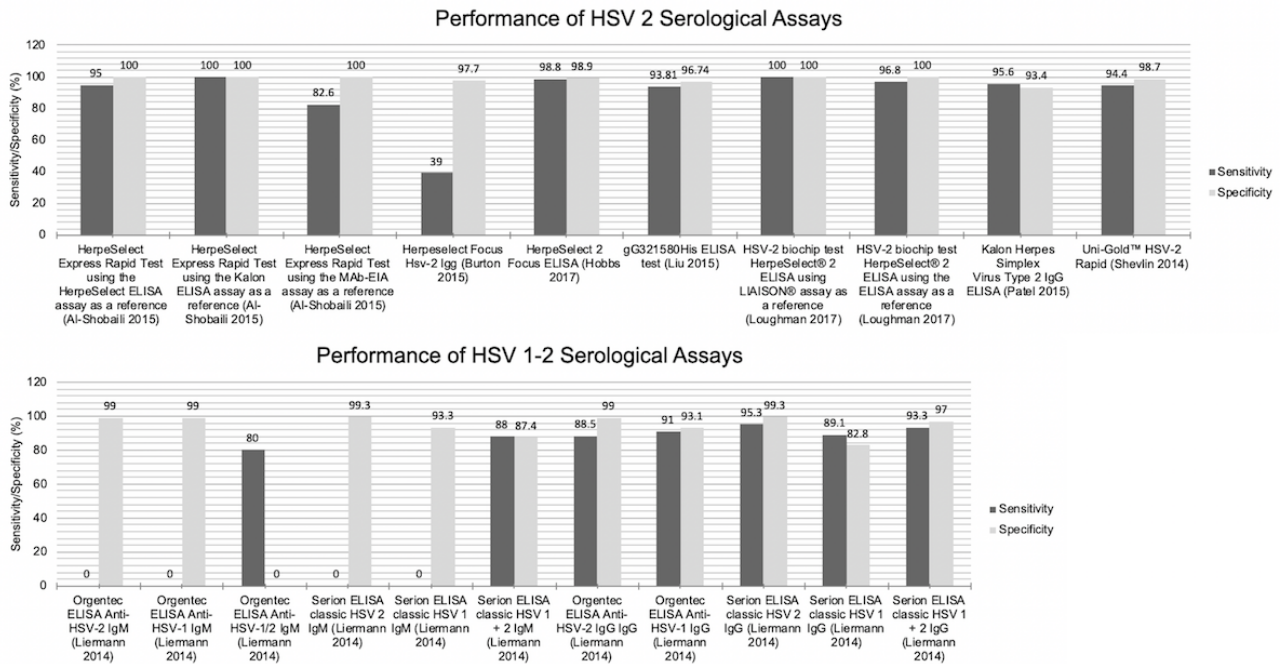


Figure 4. Graphical display for the Quality Assessment of Diagnostic Accuracy Studies-2 tool results. QUADAS: quality assessment of diagnostic accuracy studies.

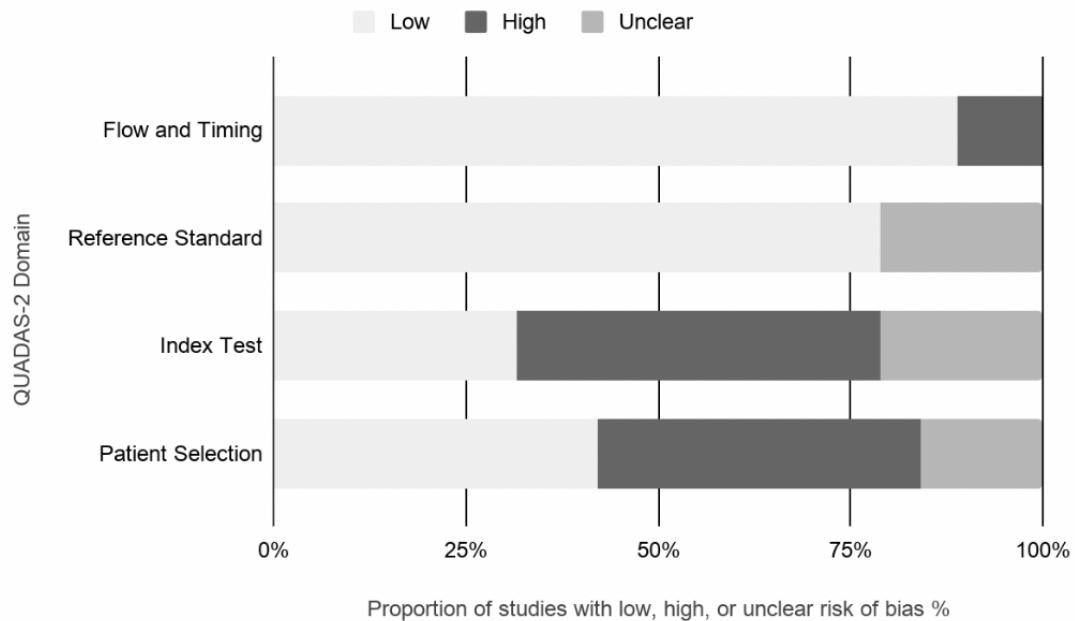


Table 4. Comparison of estimated costs per reportable result and expertise required. Table adapted from Gitman et al [31].

Test	Assay time	Frequency of testing	Cost (\$)			Expertise
			Materials and reagents	Labor	Total	
Cytospin-DFA ^a	90 min	On demand	5.14	8.78	13.92	Mid-high
Conventional culture	1-7 days	Examined once a day	5.4	11.94	17.94	Mid-high
Simplexa Direct PCR	75 min	Once a day	39.54	2.69	42.23	Mid
LDT ^b HSV ^c PCR ^d with extraction	3 h	Once a day	19.96	5.07	25.03	High

^aDFA: direct fluorescent antibody.

^bLDT: laboratory developed test.

^cHSV: herpes simplex virus.

^dPCR: polymerase chain reaction.

Gitman et al [31] concluded that the Simplexa HSV 1 & 2 Direct PCR was the most expensive but required the least training of the assays used, had the lowest hands-on time, and fastest assay time (75 min, vs 3 hours by LDT HSV PCR), and provided the HSV type. The Simplexa Direct PCR is an FDA-approved device (Table 3), and such could be used to benchmark developing technologies.

The cost-to-run per specimen was also compared between the Viper platform and real-time PRC in the study by Lang et al [34]. Compared with the HSV-LC system, the Viper instrument was found to be \$12 cheaper to run per specimen (\$22 vs \$34). In addition, the platform is fully automated [34]. Although they did not carry out a formal cost analysis, the study by Binnicker et al [29] compared time with run for real-time PCR assay and the commercially available Simplexa HSV 1 & 2 Direct (Focus Diagnostics, Cypress, CA) using CSF samples. In addition, Lee et al [35] has reported that the automated ARIES assay takes only 2 hours/12 samples and has the fastest turnaround time compared with other in-house assays. In addition, to avoiding the need for nucleic acid extraction, the Simplexa devices run time was quoted as 60 min in comparison to 4 hours by conventional methods. A separate tool, the HSV-Qx assay, was found to produce results in a few hours compared with many days for conventional culture techniques [40]. Tong et al [42] also demonstrated a turnaround time of 1 hour using their portable device, although it was not fully automated (requires hands on time of 5 min).

Sample Type and Patient Subpopulations

Binnicker et al [29] found that the Simplexa HSV-1/2 assays demonstrated a combined HSV-1 and HSV-2 sensitivity and

specificity of 96.2% and 97.9%, respectively. Critically, the results are available in 60 min, and the test only requires 50 µL of CSF [29].

Risk of Bias in Individual Studies

Most of the studies received high risk of bias in the patient selection and in the index test domains of the QUADAS tool (Table 5; Figure 4). High risk of bias in the patient selection domain was mostly because of the lack of randomization in recruiting patients and the use of case-control design in selecting participants and/or recruited patients with *confirmed diagnosis*. Only few studies included patients with suspected disease and no pretested samples [40,41] or recruited a random sample [26]. High risk of bias in the index test category was mostly because of the index test results being interpreted with the knowledge of the reference test results; most of the studies used the reference test first, and after knowing the results, performed the index test. Because all of the studies used the gold standard or an approved and validated reference tests, the reference standard mostly received low risk of bias (Figure 4). Studies that received high bias rating in the flow and timing domain used different reference standards for different patient groups and did not include all patients or samples in the analysis [27,36]. Studies that received low risk of bias rating in general recruited a random sample [26], blinded study conductors to the results of the reference and index tests [26,29,36], deidentified previously collected samples [41], or conducted reference and index tests simultaneously [34]. All of the studies received low risk of bias in the applicability concerns domains of the QUADAS tool because of the reviews' inclusion criteria being inclusive of a wide range of patients, index tests, and references standards (Table 5).

Table 5. Tabular presentation for the Quality Assessment of Diagnostic Accuracy Studies tool results. Table format adapted from Whiting et al [22].

Study	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Al-Shobaili et al [24]	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Low risk	Low risk
Barrado et al [28]	Low risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
Burton et al [30]	High risk	Unclear risk	Unclear risk	Low risk	Low risk	Low risk	Low risk
Gitman et al [31]	Low risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
Granato et al [32]	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Hobbs et al [33]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
Lang et al [34]	Unclear risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Lee et al [35]	Low risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
Liermann et al [36]	Unclear risk	Low risk	Low risk	High risk	Low risk	Low risk	Low risk
Liu et al [37]	High risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	Low risk
Loughman et al [25]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
Binnicker et al [29]	High risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Miller et al [38]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
Parra-Sánchez et al [39]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
Patel et al [26]	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Van Der Pol et al [40]	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Low risk	Low risk
Shevlin and Morrow [41]	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Shoji et al [27]	High risk	High risk	Unclear risk	High risk	Low risk	Low risk	Low risk
Tong et al [42]	Unclear risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk

Discussion

Overview

To select the most appropriate test for a patient, a thorough knowledge of the performance and limitations of available tests is critical to inform clinical decision making and the development of future technologies that may penetrate innovation gaps. This systematic review aimed to summarize recent study articles evaluating HSV-1 and HSV-2 diagnostic tools. We evaluated the study characteristics and the performance of the various tests included in the studies; here, we conclude by discussing the various strengths and limitations of various tools and how they may relate to the development of future diagnostics.

Performance

It is evident that further work needs to be done with regard to tools that work better with different sample types. It also appears that tools developed in academic centers perform worse. This may be because they are at an earlier stage of the developmental pathways and thus, not nearing approval or commercialization. Overall, tests were more specific than sensitive, and it appears that serological tests perform worse.

Culture- and laboratory-based methods have become the gold standard choice for the diagnosis of HSV-1 and HSV-2 infections. However, translational effort is focused on developing alternate techniques based on serological and

molecular assays as evidenced by the absence of novel culture-based methods in this review.

Specimen quality and adequate transport and handling to maintain infectivity are essential while using culture methods. Because of the enveloped nature of the HSV-1 and HSV-2 virus, specimens collected using a swab must be transferred to suitable viral transport media with the addition of antimicrobials. Using standard culture methods, results are usually available within 5 days; although in some cases, the result may take up to 2 weeks. Such issues are largely mitigated when using molecular methods. The use of an antigen detection system such as Herpcheck (DuPont, Wilmington, DE) can decrease turnaround time [44]. Shell vial culture also has a decreased average turnaround time, with high sensitivity achieved within 24 hours [45]. In Tong et al [42], the IsoGlow HSV Typing assay (Nucleic acid amplification tests—based technique) was compared with the ELVIS Shell Vial Assay where it displayed similar performance, while having the advantages that it was more rapid, portable, and does not require laboratory equipment. Another limitation of traditional culture methods is that viral subtyping requires the additional step of using HSV-1 and HSV-2 monoclonal antibodies [46,47]. Similarly Lee et al [35], found that the ARIES assay required the least amount of time when compared with 2 other routinely used assays and the least technical skills and knowledge [35]. The enzyme-linked virus-inducible system (ELVIS; Diagnostic Hybrids, Inc, USA) enables faster turnaround than standard culture or shell vial

methods, with similar performance characteristics [45]. High sensitivities using culture-based methods are hindered by highly variable specimen quality, dependency on stage of the lesion, as well as primary versus recurrent infection [47].

Turnaround Time and Cost

In HSV infections, especially when it results in CNS-related infections, rapid turnaround is essential for treatment [48,49]. Rapid turnaround is important for providing the treatment to patient as early as possible. In the case of HSV, which has no cure, antivirals are the only treatment available to manage the symptoms of the disease [50]. Antivirals are recommended to be administered within the first 48 to 72 hours of the appearance of symptoms for best results [51], which is a very short time interval that requires a rapid turnaround time in case the provider has to wait for the diagnostic test results before prescribing the medication. Even more important is the throughput time (ie, how many tests can be run using the system within certain amount of time) of a diagnostic test. In most cases, a confirmed diagnostic is necessary to provide antiviral treatment for HSV since the prolonged use of antivirals can lead to antiviral resistance [50]. According to Lee et al [35], the ARIES assay's throughput time was reported as 12 samples in 2 hours. This throughput time is rapid compared with 3 to 7 days of processing time needed for viral culture [52].

Automation increases turnaround time, improve reproducibility of the test, and reduce the chance of human error [53]. Although not quantified, capital would be saved because of the automated nature and more rapid turnaround of such devices. Automated diagnostic tests may only require minimal technical knowledge reducing the likelihood or *economically wasteful misclassification* because of limited knowledge of the lab technician [53]. Rapid turnaround time is important for most infectious or viral diseases. For example, for respiratory viruses, diagnostic test should be provided at the point-of-care (possible with results ready in less than 1.6 hours) to ensure the best outcomes for the patient [54].

Sample Type and Patient Subpopulations

The HIV status of an individual does not appear to affect the performance of HSV ELISAs [55]. This is an important point because of the high rate of coinfection and increased risk of transmission and acquisition of HIV by HSV-positive individuals [56,57]. Among other common sexually transmitted diseases, only coinfection with *Neisseria gonorrhoeae* has been shown to significantly reduce HSV immunoassay's performance [57]. However, the specificity of immunoassays can differ based on population demographics based on the findings of a meta-analysis conducted on studies from Africa [13].

While HSV can rarely be cultured from the CSF, molecular assays can yield results in a matter of hours, which can be important when dealing with possible HSV encephalitis or meningitis. Currently, the only device approved for use for HSV virus detection in CSF is the Simplexa HSV 1 & 2 Direct (Focus Diagnostics, Cypress, CA) (Table 4). The latter is important for the use of the test in the neonatal population. Neonates immune system has a limited ability in fighting off infections, hence resulting in mortality rates exceeding 80% from HSV infection

[58]. A PCR-based HSV detection system from CSF is critical for this patient population. Given this information, it is concerning that greatest incidence of HSV infections occurs in women of reproductive age increasing the risk of maternal transmission of the virus to the fetus or neonate [59,60]. This is a major public health concern in both developed and developing countries: a 3-year study in Canada found a neonatal HSV incidence of 5.9 per 100,000 live births and a case fatality rate of 15.5% [61]. Prevention is challenging given the asymptomatic nature of the disease in its latent phase [61], and current guidelines recommend delivery by cesarean section. This is because 85% of vertical transmissions occur during the vaginal delivery where the fetus comes into contact with infected genital secretions [61]. A 33-year study conducted in the New York, USA, concluded that deaths from HSV have increased over the time period (incidence 0.82 deaths per 100,000 live births), despite reductions in deaths from HIV and syphilis [62]. They stated this is due to an increasing number of pregnant women having no immunity to HSV type 1 putting them at increased risk of contracting the disease during pregnancy [62]. Interestingly, neonatal HSV-1 infection in developing countries is rare with the vast majority of infections being caused by HSV-2, which carries a worse prognosis. In addition, acute infection is associated with a higher transmission rate in part because of the lack of maternal to fetal antibody transfer in the third trimester [60]. A test that can accurately diagnose acute (not just latent infection) is urgently required, as current methods may not detect early infection. The study by Liermann et al [36] was the only study to investigate detection of primary infection, and concluded that its HSV IgM serology tool should not be used to make decisions for antiviral treatment in HSV [36]. Thus, no diagnostic tools included in this systematic review have addressed this unmet medical need.

The use of serologic tests to diagnose suspected genital lesions are inappropriate because positive results may be due to chronic infection, whereas negative results may overlook recent infection.

Risk of Bias of the Included Studies

A common source of bias in diagnostic tool studies occurs when studies recruit patients based on their disease status. This often leads to overestimation of the results and may exaggerate diagnostic accuracy [22]. All of the studies included in this review used sample populations that had a higher than average prevalence of HSV infection. As a result, these sensitivities may be overestimated.

Limitations

The studies included within this systematic review employed heterogeneous methods while also using different patient populations. For this reason, a statistical comparison of results was not conducted, and instead, the results displayed in a qualitative manner to benchmark future technology. Although a PRISMA-compliant search was searching the PubMed database, it is possible that some studies may have been inadvertently missed.

Conclusions

The landscape of diagnostic tools for HSV-1 and HSV-2 infections is rapidly moving away from laboratory-based and culture methods that have long been considered the gold standard technique. A majority of tools study cutaneous and mucosal HSV infections (n=13); 2 tests focused on ocular disease, whereas a single one on CNS manifestations. No diagnostic tools included in our systematic review are currently

suitable for use as prenatal tools, however. The incidence of acute infections is rising, and because these infections present the greatest risk to unborn fetuses, further work needs to be done to develop diagnostic tools to detect primary infection in expectant mothers to prevent vertical transfer. This will assist in lowering the rate of neonatal herpes, which can be a life-threatening condition. We believe this can only be achieved through prenatal screening for primary infection and subsequent medical intervention.

Acknowledgments

The authors would like to thank the librarians at Imperial College London for their assistance in the search strategy. This work was funded by the European Institute of Innovation and Technology Health (grant 18654).

Conflicts of Interest

DB is a stockholder in Translation Ventures Ltd (Charlbury, UK), IP Asset Ventures Ltd, and Biolacuna Ltd (Oxford, UK), companies that, among other services, provide biomanufacturing, regulatory, and financial advice to pharmaceutical clients. DB is also subject to the CFA Institute's codes, standards, and guidelines, so he must stress that this piece is provided for academic interest only and must not be construed in any way as an investment recommendation. Additionally, at time of publication, DB and the organizations with which he is affiliated may or may not have agreed to and/or pending funding commitments from the organizations named herein.

Multimedia Appendix 1

Search terms.

[[DOCX File, 14KB - publichealth_v5i2e14216_app1.docx](#)]

Multimedia Appendix 2

Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2009 flow diagram.

[[DOC File, 59KB - publichealth_v5i2e14216_app2.doc](#)]

Multimedia Appendix 3

Serological assays for detection of herpes simplex virus and summary of the different serological assays studied within the 6 included studies.

[[DOCX File, 23KB - publichealth_v5i2e14216_app3.docx](#)]

References

1. Davison A. Herpesviruses: General Features. In: Encyclopedia of Virology (Third Edition). Oxford: Academic Press; 2008.
2. World Health Organization. 2017. Herpes simplex virus URL: <https://www.who.int/news-room/fact-sheets/detail/herpes-simplex-virus> [accessed 2019-05-07] [WebCite Cache ID 78Bz0QSTM]
3. Torrone EA, Morrison CS, Chen P, Kwok C, Francis SC, Hayes RJ, STIMA Working Group. Prevalence of sexually transmitted infections and bacterial vaginosis among women in sub-Saharan Africa: an individual participant data meta-analysis of 18 HIV prevention studies. PLoS Med 2018 Dec;15(2):e1002511 [FREE Full text] [doi: [10.1371/journal.pmed.1002511](https://doi.org/10.1371/journal.pmed.1002511)] [Medline: [29485986](https://pubmed.ncbi.nlm.nih.gov/29485986/)]
4. Kouyoumjian SP, Heijnen M, Chaabna K, Mumtaz GR, Omori R, Vickerman P, et al. Global population-level association between herpes simplex virus 2 prevalence and HIV prevalence. AIDS 2018 Jun 19;32(10):1343-1352 [FREE Full text] [doi: [10.1097/QAD.0000000000001828](https://doi.org/10.1097/QAD.0000000000001828)] [Medline: [29794495](https://pubmed.ncbi.nlm.nih.gov/29794495/)]
5. Bradshaw MJ, Venkatesan A. Herpes simplex virus-1 encephalitis in adults: pathophysiology, diagnosis, and management. Neurotherapeutics 2016 Dec;13(3):493-508 [FREE Full text] [doi: [10.1007/s13311-016-0433-7](https://doi.org/10.1007/s13311-016-0433-7)] [Medline: [27106239](https://pubmed.ncbi.nlm.nih.gov/27106239/)]
6. Jørgensen LK, Dalgaard LS, Østergaard LJ, Nørgaard M, Mogensen TH. Incidence and mortality of herpes simplex encephalitis in Denmark: a nationwide registry-based cohort study. J Infect 2017 Dec;74(1):42-49. [doi: [10.1016/j.jinf.2016.09.004](https://doi.org/10.1016/j.jinf.2016.09.004)] [Medline: [27717782](https://pubmed.ncbi.nlm.nih.gov/27717782/)]
7. Granerod J, Cousens S, Davies NW, Crowcroft NS, Thomas SL. New estimates of incidence of encephalitis in England. Emerg Infect Dis 2013;19(9) [FREE Full text] [doi: [10.3201/eid1909.130064](https://doi.org/10.3201/eid1909.130064)] [Medline: [23969035](https://pubmed.ncbi.nlm.nih.gov/23969035/)]

8. LeGoff J, Péré H, Bélec L. Diagnosis of genital herpes simplex virus infection in the clinical laboratory. *Virology* 2014 May 12;11:83 [FREE Full text] [doi: [10.1186/1743-422X-11-83](https://doi.org/10.1186/1743-422X-11-83)] [Medline: [24885431](https://pubmed.ncbi.nlm.nih.gov/24885431/)]
9. WebMD. 2005. Herpes Test: What You Should Know URL: <https://www.webmd.com/genital-herpes/herpes-tests-what-you-should-know> [accessed 2019-05-07] [WebCite Cache ID 78BzH2JUC]
10. Wald A, Ashley-Morrow R. Serological testing for herpes simplex virus (HSV)-1 and HSV-2 infection. *Clin Infect Dis* 2002 Oct 15;35(Suppl 2):S173-S182. [doi: [10.1086/342104](https://doi.org/10.1086/342104)] [Medline: [12353203](https://pubmed.ncbi.nlm.nih.gov/12353203/)]
11. Burd EM. Validation of laboratory-developed molecular assays for infectious diseases. *Clin Microbiol Rev* 2010 Jul;23(3):550-576 [FREE Full text] [doi: [10.1128/CMR.00074-09](https://doi.org/10.1128/CMR.00074-09)] [Medline: [20610823](https://pubmed.ncbi.nlm.nih.gov/20610823/)]
12. Wald A, Huang M, Carrell D, Selke S, Corey L. Polymerase chain reaction for detection of herpes simplex virus (HSV) DNA on mucosal surfaces: comparison with HSV isolation in cell culture. *J Infect Dis* 2003 Nov 01;188(9):1345-1351. [doi: [10.1086/379043](https://doi.org/10.1086/379043)] [Medline: [14593592](https://pubmed.ncbi.nlm.nih.gov/14593592/)]
13. Biraro S, Mayaud P, Morrow RA, Grosskurth H, Weiss HA. Performance of commercial herpes simplex virus type-2 antibody tests using serum samples from Sub-Saharan Africa: a systematic review and meta-analysis. *Sex Transm Dis* 2011 Feb;38(2):140-147. [doi: [10.1097/OLQ.0b013e3181f0bafb](https://doi.org/10.1097/OLQ.0b013e3181f0bafb)] [Medline: [20706175](https://pubmed.ncbi.nlm.nih.gov/20706175/)]
14. Anzivino E, Fioriti D, Mischitelli M, Bellizzi A, Barucca V, Chiarini F, et al. Herpes simplex virus infection in pregnancy and in neonate: status of art of epidemiology, diagnosis, therapy and prevention. *Virology* 2009 Apr 06;6:40 [FREE Full text] [doi: [10.1186/1743-422X-6-40](https://doi.org/10.1186/1743-422X-6-40)] [Medline: [19348670](https://pubmed.ncbi.nlm.nih.gov/19348670/)]
15. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009 Oct;62(10):e1-34 [FREE Full text] [doi: [10.1016/j.jclinepi.2009.06.006](https://doi.org/10.1016/j.jclinepi.2009.06.006)] [Medline: [19631507](https://pubmed.ncbi.nlm.nih.gov/19631507/)]
16. Burton MJ, Penman A, Adams-McAlpine G, Overman TL, Hook EW. Prevalence and characteristics of herpes simplex virus type-2 coinfection in veterans with hepatitis C. *Am J Med Sci* 2012 Dec;344(6):436-440. [doi: [10.1097/MAJ.0b013e31824585bd](https://doi.org/10.1097/MAJ.0b013e31824585bd)] [Medline: [22395713](https://pubmed.ncbi.nlm.nih.gov/22395713/)]
17. Looker, Elmes JA, Gottlieb SL, Schiffer JT, Vickerman P, Turner KM, et al. Effect of HSV-2 infection on subsequent HIV acquisition: an updated systematic review and meta-analysis. *Lancet Infect Dis* 2017 Dec;17(12):1303-1316 [FREE Full text] [doi: [10.1016/S1473-3099\(17\)30405-X](https://doi.org/10.1016/S1473-3099(17)30405-X)] [Medline: [28843576](https://pubmed.ncbi.nlm.nih.gov/28843576/)]
18. Altman DG, Bland JM. Diagnostic tests. 1: sensitivity and specificity. *Br Med J* 1994 Jun 11;308(6943):1552. [Medline: [8019315](https://pubmed.ncbi.nlm.nih.gov/8019315/)]
19. Mckibbon KA, Marks S. Searching for the best evidence. Part 2: searching CINAHL and Medline. *Evid Based Nurs* 1998 Oct 01;1(4):105-107. [doi: [10.1136/ebn.1.4.105](https://doi.org/10.1136/ebn.1.4.105)]
20. US Food & Drug Administration. URL: <https://www.fda.gov/> [accessed 2019-05-07] [WebCite Cache ID 78Bzxz9P]
21. European Medicines Agency. URL: <https://www.ema.europa.eu/en> [accessed 2019-05-07] [WebCite Cache ID 78C09rGxw]
22. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011 Oct 18;155(8):529-536. [doi: [10.7326/0003-4819-155-8-201110180-00009](https://doi.org/10.7326/0003-4819-155-8-201110180-00009)] [Medline: [22007046](https://pubmed.ncbi.nlm.nih.gov/22007046/)]
23. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)] [Medline: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)]
24. Al-Shobaili H, Hassanein KM, Mostafa MS, Al Duways AS. Evaluation of the HerpeSelect Express rapid test in the detection of herpes simplex virus type 2 antibodies in patients with genital ulcer disease. *J Clin Lab Anal* 2015 Jan;29(1):43-46. [doi: [10.1002/jcla.21725](https://doi.org/10.1002/jcla.21725)] [Medline: [24687953](https://pubmed.ncbi.nlm.nih.gov/24687953/)]
25. Loughman T, Singh B, Seddon B, Noone P, Santhosh P. Validation of a membrane touch biosensor for the qualitative detection of IgG class antibodies to herpes simplex virus type 2. *Analyst* 2017 Jul 24;142(15):2725-2734. [doi: [10.1039/c7an00666g](https://doi.org/10.1039/c7an00666g)] [Medline: [28708188](https://pubmed.ncbi.nlm.nih.gov/28708188/)]
26. Patel EU, Manucci J, Kahle EM, Lingappa JR, Morrow RA, Piwowar-Manning E, et al. Precision of the Kalon Herpes Simplex Virus Type 2 IgG ELISA: an international inter-laboratory assessment. *BMC Infect Dis* 2015 Dec 30;15:398 [FREE Full text] [doi: [10.1186/s12879-015-1130-6](https://doi.org/10.1186/s12879-015-1130-6)] [Medline: [26423888](https://pubmed.ncbi.nlm.nih.gov/26423888/)]
27. Shoji J, Sakimoto T, Inada N, Kamei Y, Matsubara M, Takamura E, et al. A diagnostic method for herpes simplex keratitis by simultaneous measurement of viral DNA and virus-specific secretory IgA in tears: an evaluation. *Jpn J Ophthalmol* 2016 Jul;60(4):294-301. [doi: [10.1007/s10384-016-0448-y](https://doi.org/10.1007/s10384-016-0448-y)] [Medline: [27126382](https://pubmed.ncbi.nlm.nih.gov/27126382/)]
28. Barrado L, Suarez MJ, Pérez-Blázquez E, Otero JR, Folgueira MD. Could polymerase chain reaction tests on conjunctival swabs be useful to diagnose herpetic keratitis? *Enferm Infecc Microbiol Clin* 2014 Jan;32(1):28-30. [doi: [10.1016/j.eimc.2013.07.005](https://doi.org/10.1016/j.eimc.2013.07.005)] [Medline: [24060450](https://pubmed.ncbi.nlm.nih.gov/24060450/)]
29. Binnicker MJ, Espy MJ, Irish CL. Rapid and direct detection of herpes simplex virus in cerebrospinal fluid by use of a commercial real-time PCR assay. *J Clin Microbiol* 2014 Dec;52(12):4361-4362 [FREE Full text] [doi: [10.1128/JCM.02623-14](https://doi.org/10.1128/JCM.02623-14)] [Medline: [25274992](https://pubmed.ncbi.nlm.nih.gov/25274992/)]
30. Burton M, Van Wagoner NJ, Sunesara I, Penman A, Swiatlo E, Hook EW. Evaluating the performance of the focus HerpeSelect® HSV-2 IgG in veterans with chronic hepatitis C infection. *J Med Virol* 2015 Aug;87(8):1377-1381. [doi: [10.1002/jmv.24195](https://doi.org/10.1002/jmv.24195)] [Medline: [25908332](https://pubmed.ncbi.nlm.nih.gov/25908332/)]

31. Gitman MR, Ferguson D, Landry ML. Comparison of Simplexa HSV 1 & 2 PCR with culture, immunofluorescence, and laboratory-developed TaqMan PCR for detection of herpes simplex virus in swab specimens. *J Clin Microbiol* 2013 Nov;51(11):3765-3769 [FREE Full text] [doi: [10.1128/JCM.01413-13](https://doi.org/10.1128/JCM.01413-13)] [Medline: [24006008](https://pubmed.ncbi.nlm.nih.gov/24006008/)]
32. Granato PA, Alkins BR, Yen-Lieberman B, Greene WH, Connolly J, Buchan BW, et al. Comparative evaluation of AmpliVue HSV 1+2 assay with ELVIS culture for detecting herpes simplex virus 1 (HSV-1) and HSV-2 in clinical specimens. *J Clin Microbiol* 2015 Dec;53(12):3922-3925 [FREE Full text] [doi: [10.1128/JCM.01905-15](https://doi.org/10.1128/JCM.01905-15)] [Medline: [26468497](https://pubmed.ncbi.nlm.nih.gov/26468497/)]
33. Hobbs MM, Mwanyumba SW, Luseno WK, Hartman S, Halpern CT, Hallfors DD, et al. Evaluation of herpes simplex virus type 2 serological tests for use with dried blood spots in Kenya. *Sex Transm Dis* 2017 Dec;44(2):101-103 [FREE Full text] [doi: [10.1097/OLQ.0000000000000557](https://doi.org/10.1097/OLQ.0000000000000557)] [Medline: [28081046](https://pubmed.ncbi.nlm.nih.gov/28081046/)]
34. Lang AL, Roberts C, Mazzulli T, Hachette TF, LeBlanc JJ. Detection and differentiation of herpes simplex viruses by use of the viper platform: advantages, limitations, and concerns. *J Clin Microbiol* 2014 Jun;52(6):2186-2188 [FREE Full text] [doi: [10.1128/JCM.03636-13](https://doi.org/10.1128/JCM.03636-13)] [Medline: [24696023](https://pubmed.ncbi.nlm.nih.gov/24696023/)]
35. Lee CK, Chai CN, Capinpin SM, Ang A, Ng SY, Lee PL, et al. Evaluation of the Luminex ARIES HSV 1&2 Assay and comparison with the FTD Neuro 9 and in-house real-time PCR assays for detecting herpes simplex viruses. *Ann Lab Med* 2018 Sep;38(5):440-445 [FREE Full text] [doi: [10.3343/alm.2018.38.5.440](https://doi.org/10.3343/alm.2018.38.5.440)] [Medline: [29797814](https://pubmed.ncbi.nlm.nih.gov/29797814/)]
36. Liermann K, Schäfler A, Henke A, Sauerbrei A. Evaluation of commercial herpes simplex virus IgG and IgM enzyme immunoassays. *J Virol Methods* 2014 Apr;199:29-34. [doi: [10.1016/j.jviromet.2014.01.001](https://doi.org/10.1016/j.jviromet.2014.01.001)] [Medline: [24418499](https://pubmed.ncbi.nlm.nih.gov/24418499/)]
37. Liu T, Liu JF, Yu H, Si GJ, Hu J, Li J. Production of a fragment of glycoprotein G of herpes simplex virus type 2 and evaluation of its diagnostic potential. *Singapore Med J* 2015 Jun;56(6):346-352 [FREE Full text] [doi: [10.11622/smedj.2014197](https://doi.org/10.11622/smedj.2014197)] [Medline: [25532518](https://pubmed.ncbi.nlm.nih.gov/25532518/)]
38. Miller S, Samayoa E, Post L, Wright C, McKinley G, Wood M, et al. Development and clinical evaluation of a novel fully automated qualitative PCR assay for the diagnosis of anogenital herpes simplex virus infection. *Diagn Microbiol Infect Dis* 2014 Oct;80(2):102-106. [doi: [10.1016/j.diagmicrobio.2014.06.013](https://doi.org/10.1016/j.diagmicrobio.2014.06.013)] [Medline: [25128985](https://pubmed.ncbi.nlm.nih.gov/25128985/)]
39. Parra-Sánchez M, Marcuello López A, García-Rey S, Zakariya-Yousef Breval I, Bernal Martínez S, Pueyo Rodríguez I, et al. Performance of the HSV OligoGen kit for the diagnosis of herpes simplex virus type 1 and 2. *Diagn Microbiol Infect Dis* 2016 Jul;85(3):315-317. [doi: [10.1016/j.diagmicrobio.2016.04.019](https://doi.org/10.1016/j.diagmicrobio.2016.04.019)] [Medline: [27185644](https://pubmed.ncbi.nlm.nih.gov/27185644/)]
40. Van Der Pol B, Warren T, Taylor SN, Martens M, Jerome KR, Mena L, et al. Type-specific identification of anogenital herpes simplex virus infections by use of a commercially available nucleic acid amplification test. *J Clin Microbiol* 2012 Nov;50(11):3466-3471 [FREE Full text] [doi: [10.1128/JCM.01685-12](https://doi.org/10.1128/JCM.01685-12)] [Medline: [22875892](https://pubmed.ncbi.nlm.nih.gov/22875892/)]
41. Shevlin E, Morrow RA. Comparative performance of the Uni-Gold™ HSV-2 Rapid: a point-of-care HSV-2 diagnostic test in unselected sera from a reference laboratory. *J Clin Virol* 2014 Nov;61(3):378-381 [FREE Full text] [doi: [10.1016/j.jcv.2014.08.012](https://doi.org/10.1016/j.jcv.2014.08.012)] [Medline: [25200648](https://pubmed.ncbi.nlm.nih.gov/25200648/)]
42. Tong Y, McCarthy K, Kong H, Lemieux B. Development and comparison of a rapid isothermal nucleic acid amplification test for typing of herpes simplex virus types 1 and 2 on a portable fluorescence detector. *J Mol Diagn* 2012 Nov;14(6):569-576 [FREE Full text] [doi: [10.1016/j.jmoldx.2012.05.005](https://doi.org/10.1016/j.jmoldx.2012.05.005)] [Medline: [22951487](https://pubmed.ncbi.nlm.nih.gov/22951487/)]
43. US Food & Drug Administration. 2007. Statistical guidance on reporting results from studies evaluating diagnostic tests - Guidance for industry and FDA staff URL: <https://www.fda.gov/RegulatoryInformation/Guidances/ucm071148.htm> [accessed 2019-05-07] [WebCite Cache ID 78C0PNg1O]
44. Espy MJ, Uhl JR, Mitchell PS, Thorvilson JN, Svien KA, Wold AD, et al. Diagnosis of herpes simplex virus infections in the clinical laboratory by LightCycler PCR. *J Clin Microbiol* 2000 Feb;38(2):795-799 [FREE Full text] [Medline: [10655387](https://pubmed.ncbi.nlm.nih.gov/10655387/)]
45. Proffitt MR, Schindler SA. Rapid detection of HSV with an enzyme-linked virus inducible system (ELVIS) employing a genetically modified cell line. *Clin Diagn Virol* 1995 Aug;4(2):175-182. [Medline: [15566838](https://pubmed.ncbi.nlm.nih.gov/15566838/)]
46. Langenberg A, Zbanyszek R, Dragavon J, Ashley R, Corey L. Comparison of diploid fibroblast and rabbit kidney tissue cultures and a diploid fibroblast microtiter plate system for the isolation of herpes simplex virus. *J Clin Microbiol* 1988 Sep;26(9):1772-1774 [FREE Full text] [Medline: [2846647](https://pubmed.ncbi.nlm.nih.gov/2846647/)]
47. Nerurkar LS, Jacob AJ, Madden DL, Sever JL. Detection of genital herpes simplex infections by a tissue culture-fluorescent-antibody technique with biotin-avidin. *J Clin Microbiol* 1983 Jan;17(1):149-154 [FREE Full text] [Medline: [6298272](https://pubmed.ncbi.nlm.nih.gov/6298272/)]
48. Valenstein P. Laboratory turnaround time. *Am J Clin Pathol* 1996 Jun;105(6):676-688. [Medline: [8659441](https://pubmed.ncbi.nlm.nih.gov/8659441/)]
49. Chaudhuri A, Kennedy PG. Diagnosis and treatment of viral encephalitis. *Postgrad Med J* 2002 Oct;78(924):575-583 [FREE Full text] [Medline: [12415078](https://pubmed.ncbi.nlm.nih.gov/12415078/)]
50. Poole CL, James SH. Antiviral therapies for herpesviruses: current agents and new directions. *Clin Ther* 2018 Aug;40(8):1282-1298. [doi: [10.1016/j.clinthera.2018.07.006](https://doi.org/10.1016/j.clinthera.2018.07.006)] [Medline: [30104016](https://pubmed.ncbi.nlm.nih.gov/30104016/)]
51. British Medical Journal Best Practice. 2018. Herpes simplex infection: Treatment algorithm URL: <https://bestpractice.bmj.com/topics/en-gb/53> [accessed 2019-05-07] [WebCite Cache ID 78C0cmHRW]
52. Glass N, Nelson H, Huffman L. Screening for genital herpes simplex: Brief update for the U.S. Preventive Services Task Force. *Genital Herpes Infection: Serologic Screening* 2016:2018 [FREE Full text]
53. Jorgenson J. Automation in clinical microbiology. Boca Raton, FL: CRC Press; 1987.

54. Brendish NJ, Malachira AK, Beard KR, Ewings S, Clark TW. Impact of turnaround time on outcome with point-of-care testing for respiratory viruses: a analysis from a randomised controlled trial. *Eur Respir J* 2018 Dec;52(2). [doi: [10.1183/13993003.00555-2018](https://doi.org/10.1183/13993003.00555-2018)] [Medline: [29946003](https://pubmed.ncbi.nlm.nih.gov/29946003/)]
55. Lingappa J, Nakku-Joloba E, Magaret A, Friedrich D, Dragavon J, Kambugu F, et al. Sensitivity and specificity of herpes simplex virus-2 serological assays among HIV-infected and uninfected urban Ugandans. *Int J STD AIDS* 2010 Sep;21(9):611-616 [FREE Full text] [doi: [10.1258/ijsa.2009.008477](https://doi.org/10.1258/ijsa.2009.008477)] [Medline: [21097732](https://pubmed.ncbi.nlm.nih.gov/21097732/)]
56. Corey L, Wald A, Celum CL, Quinn TC. The effects of herpes simplex virus-2 on HIV-1 acquisition and transmission: a review of two overlapping epidemics. *J Acquir Immune Defic Syndr* 2004 Apr 15;35(5):435-445. [Medline: [15021308](https://pubmed.ncbi.nlm.nih.gov/15021308/)]
57. Summerton J, Riedesel M, Laeyendecker O, Gaydos C, Maldeis NE, Hardick A, et al. Effect of sexually transmitted disease (STD) coinfections on performance of three commercially available immunosorbent assays used for detection of herpes simplex virus type 2-specific antibody in men attending Baltimore, Maryland, STD clinics. *Clin Vaccine Immunol* 2007 Dec;14(12):1545-1549 [FREE Full text] [doi: [10.1128/CVI.00120-07](https://doi.org/10.1128/CVI.00120-07)] [Medline: [17913866](https://pubmed.ncbi.nlm.nih.gov/17913866/)]
58. Whitley R, Gnann W. *Mucocutaneous manifestations of viral diseases: An illustrated guide to diagnosis and management*. Boca Raton, FL: CRC Press; 2016.
59. Sauerbrei A, Wutzler P. Herpes simplex and varicella-zoster virus infections during pregnancy: current concepts of prevention, diagnosis and therapy. Part 2: Varicella-zoster virus infections. *Med Microbiol Immunol* 2007 Jun;196(2):95-102. [doi: [10.1007/s00430-006-0032-z](https://doi.org/10.1007/s00430-006-0032-z)] [Medline: [17180380](https://pubmed.ncbi.nlm.nih.gov/17180380/)]
60. Brown ZA, Wald A, Morrow RA, Selke S, Zeh J, Corey L. Effect of serologic status and cesarean delivery on transmission rates of herpes simplex virus from mother to infant. *J Am Med Assoc* 2003 Jan 08;289(2):203-209. [Medline: [12517231](https://pubmed.ncbi.nlm.nih.gov/12517231/)]
61. Kropp RY, Wong T, Cormier L, Ringrose A, Burton S, Embree JE, et al. Neonatal herpes simplex virus infections in Canada: results of a 3-year national prospective study. *Pediatrics* 2006 Jun;117(6):1955-1962. [doi: [10.1542/peds.2005-1778](https://doi.org/10.1542/peds.2005-1778)] [Medline: [16740836](https://pubmed.ncbi.nlm.nih.gov/16740836/)]
62. Sampath A, Maduro G, Schillinger JA. Infant deaths due to herpes simplex virus, congenital syphilis, and HIV in New York City. *Pediatrics* 2016 Dec;137(4) [FREE Full text] [doi: [10.1542/peds.2015-2387](https://doi.org/10.1542/peds.2015-2387)] [Medline: [26933212](https://pubmed.ncbi.nlm.nih.gov/26933212/)]

Abbreviations

CNS: central nervous system

CSF: cerebrospinal fluid

DFA: direct fluorescent antibody

ELISA: enzyme-linked immunosorbent assay

FDA: Food and Drug Administration

HSV: herpes simplex viruses

PICO: Population, Intervention, Comparator, Outcome

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

QUADAS: quality assessment of diagnostic accuracy studies

Edited by G Eysenbach; submitted 31.03.19; peer-reviewed by S Dunbar; comments to author 23.04.19; revised version received 26.04.19; accepted 26.04.19; published 23.05.19.

Please cite as:

Arshad Z, Alturkistani A, Brindley D, Lam C, Foley K, Meinert E

Tools for the Diagnosis of Herpes Simplex Virus 1/2: Systematic Review of Studies Published Between 2012 and 2018

JMIR Public Health Surveill 2019;5(2):e14216

URL: <http://publichealth.jmir.org/2019/2/e14216/>

doi: [10.2196/14216](https://doi.org/10.2196/14216)

PMID: [31124465](https://pubmed.ncbi.nlm.nih.gov/31124465/)

©Zeeshan Arshad, Abrar Alturkistani, David Brindley, Ching Lam, Kimberley Foley, Edward Meinert. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 23.05.2019. 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

#PrEP4Love: An Evaluation of a Sex-Positive HIV Prevention Campaign

Jessica M Dehlin^{1*}, LCSW; Ryan Stillwagon^{2*}, MA; Jim Pickett³; Lance Keene⁴, MSW; John A Schneider¹, MD, MPH

¹Department of Medicine, University of Chicago, Chicago, IL, United States

²Department of Sociology, University of British Columbia, Vancouver, BC, Canada

³AIDS Foundation of Chicago, Chicago, IL, United States

⁴Social Service Administration, University of Chicago, Chicago, IL, United States

*these authors contributed equally

Corresponding Author:

Jessica M Dehlin, LCSW

Department of Medicine

University of Chicago

1525 E 55th St, Suite 314

Chicago, IL, 60615

United States

Phone: 1 7737024377

Email: jdehlin@medicine.bsd.uchicago.edu

Abstract

Background: Pre-exposure prophylaxis (PrEP) is an effective but underutilized method for preventing HIV transmission in communities vulnerable to HIV. Public health campaigns aimed at increasing PrEP awareness and access have less evaluation data.

Objective: The aim of this study was to evaluate Chicago's PrEP campaign, PrEP4Love (P4L), a campaign that uses health equity and sex-positivity approaches for information dissemination.

Methods: P4L launched in February 2016 and remains an active campaign to date. The analysis period for this paper was from the launch date in February 2016 through May 15, 2016. Our analysis reviews the Web-based reach of the campaign through views on social media platforms (Facebook and Instagram), *smart ads*, or ads served to individuals across a variety of Web platforms based on their demographics and browsing history, and P4L website clicks.

Results: In total, 40,913,560 unique views were generated across various social media platforms. A total of 24,548 users clicked on P4L ads and 32,223,987 views were received from *smart ads*. The 3 most clicked on ads were *STD Signs & Symptoms—More Information on STD Symptoms, HIV & AIDS Prevention*, and *HIV Prevention Medication*. An additional 6,970,127 views were gained through Facebook and another 1,719,446 views through Instagram. There was an average of 182 clicks per day on the P4L website.

Conclusions: This is the first study investigating public responses to a health equity and sex-positive social marketing campaign for PrEP. Overall, the campaign reached millions of individuals. More studies of PrEP social marketing are needed to evaluate the relationship of targeted public health campaigns on stigma and to guide future PrEP promotion strategies.

(*JMIR Public Health Surveill* 2019;5(2):e12822) doi:[10.2196/12822](https://doi.org/10.2196/12822)

KEYWORDS

pre-exposure prophylaxis; public health; health equity; health promotion

Introduction

Background

Chronic illnesses such as HIV or AIDS require renewed public health responses centered around health equity [1] to reduce growing disparities between different racial communities in the United States. In absolute terms, the rate of new HIV infections in the United States has remained stable from 2012 to 2016 among the general population and there were 38,739 diagnoses 2017 [2]. In relative terms, the rates of new HIV infection are on the rise [3] and disproportionately affect gay, black, and Latinx men, where 1 in 2 black, gay men and 1 in 4 gay Latinx men are projected to acquire HIV during their lifetimes [4]. A health equity approach considers how the confluence of upstream, structural oppressions [5], and downstream, individual outcomes, merge to produce widening health disparities among certain populations in the United States. Such populations in Chicago include black, gay and bisexual men and other men who have sex with men (BMSM), transgender women of color, and cisgender women of color [6] who are highly impacted by the HIV epidemic [7]. We used an evidence-based, health equity approach that aims to increase PrEP awareness and adoption among individuals most at risk of HIV infection in Chicago [8].

Pre-Exposure Prophylaxis as Prevention

Pre-exposure Prophylaxis (PrEP; Tenofovir-Emtricitabine, sold under the brand name Truvada) is an effective biomedical intervention that can decrease the risk of HIV infection by up to 99% [9]. Although PrEP uptake in the United States has increased over the years when the Food and Drug Administration (FDA) approved it in 2012 [10], to date, PrEP adoption among sexual and gender minority communities (eg, racial, gender, and sexual) remains limited [8], most likely because of unaddressed social determinants such as lesbian, gay, bisexual, transgender and queer (LGBTQ) stigma [11], intergenerational poverty, inadequate housing, and high rates of unemployment and incarceration—all of which impact health [12,13]. Furthermore, Gilead, a pharmaceutical company and the only company with FDA approval to manufacture and market PrEP, previously stated that they did not view PrEP as a commercial opportunity and had not participated in commercial advertising [14].

Grassroots Pre-Exposure Prophylaxis Promotion

PrEP activism amid pharmaceutical and societal indifference is one of many *health social movements* [15] that have thrived through organization and agitation. The aim of PrEP activism in the United States has been to increase knowledge about the rising rates of new HIV infections among black and Latinx individuals, prioritize the issue of health access and treatment among individuals regardless of their social position or financial resources, and give voice to the populations most impacted by the epidemic today in nonstigmatizing, empowering, and relatable ways [16]. Previous public health campaigns have covered topics such as the impact of air pollution on the African American community [17], the health risks of tobacco and vaping among young people [18,19], and teen pregnancy prevention [20]. Within the field of HIV prevention, campaigns

have increased condom use among teens [21] and routine HIV testing among women [22].

Where Gilead was absent, the PrEP health social movement was active. Across the globe nonprofit and public health organizations filled the pharmaceutical marketing gap through mass media campaigns designed to increase PrEP awareness, acceptability, and access in the United States [23–26] and around the world [27–30]. PrEP advertising thus emerged as a localized grassroots phenomenon among LGBTQ health and public health organizations in various US and global cities [31]. One of the benefits of PrEP marketing without Gilead was the freedom that various LGBTQ health and public organizations had in tailoring its message to the communities most impacted within a given region. Without consulting local community members affected by the HIV epidemic, interventions crafted by pharmaceutical and state-led coalitions far removed from these communities risk being irrelevant or offensive [32–34]. An example of this occurred in 2010, when the Illinois Department of Public Health launched a campaign that stitched together the faces of 4 different men of color with the catchphrase, “He’s the one that could infect you” [35]. This fear-based campaign directly linked an entire community of black men as the primary *transmitters* of HIV—essentially vectors of disease—and stigmatized men of color as unhealthy and untrustworthy [36]. Conversely, subsequent campaigns have employed sex-affirming and culturally relevant messaging to appeal to communities impacted by HIV. In early 2016, AIDS Project Los Angeles launched *It Feels Good*, a PrEP campaign that paired its slogan with HIV prevention tips, such as “It feels good protecting myself against HIV with PrEP; It feels good discussing PrEP with my doctor; [and] It feels good telling my friends that PrEP helps protect them against HIV” [37]. The campaign tapped into concerns voiced by the LGBTQ community within Los Angeles about PrEP and leveraged them into ways to promote PrEP as an option for HIV prevention.

PrEP4Love: #ContractHeat, #SpreadTingle, #CatchDesire #TransmitLove

In Chicago, the PrEP4Love (P4L) campaign was launched in early 2016 across the Chicago Transit Authority (CTA) through a Web-based presence with campaign ads, a campaign website and through P4L-sponsored pop-up events [31]. Although advertising PrEP as an HIV prevention tool, the campaign prioritized intimacy and desire, framing consensual sexual activities as healthy and pleasurable [38]. The sex-positive approach of P4L is a means of working toward health equity by affirming pleasure and desire in gender and sexual minority communities; the prevention campaign is uniquely risk reductive *and* sex-positive. This approach celebrates non-normative gender expressions and sexualities that are often stigmatized and marginalized. Furthermore, it builds trust between local communities and their providers, while validating diverse relationships and reinserting and encouraging sexual pleasure as a part of HIV prevention [39]. Addressing stigma and building community are the 2 unique contextual equity factors that the P4L campaign addressed. As Gilead has started using advertising in 2018, they have also adopted a sex-positive approach assumed to be inspired by campaigns such as P4L [40].

This paper aimed to address the intersection of health equity, public health, and social media marketing in the context of a public health campaign that offers PrEP as an HIV prevention strategy for underserved gender and sexual minority populations. This is done by evaluating if the P4L campaign met its primary goal of increasing PrEP awareness among black, gay, and bisexual men and other men who have sex with men and black transwomen in Chicago.

Methods

Campaign Development

The P4L campaign was produced by the Chicago PrEP Working Group, now significantly expanded and called the Illinois PrEP Working Group (IPWG). Initially, this group kick-started the P4L campaign through US \$350K raised from a challenge grant from the Alphawood Foundation (US \$250K) and private donations from individual Chicagoans (US \$100K) [31]. IPWG secured pro bono creative support and volunteer hours in the summer of 2015 from a collection of Publicis Groupe advertising agencies with shops in Chicago, including Leo Burnett, Starcom, Razorfish, and Spark. Private funding and perspectives leveraged by community organizations familiar with sexual subcultures heavily affected by HIV and sexually transmitted infections enabled the IPWG to challenge and redesign the narrative on HIV prevention to one that takes seriously feelings of desire, sexual pleasure, and sexual intimacy shared between sexual actors without the use of fear as a tactic within these relationships. The sex-positive message was driven by research: A 2015 study noted one of the strongest predictors for adopting PrEP among gay men and other men who have sex with men (MSM) was the desire to increase intimacy between themselves

and their partner [41], a finding that recently challenged assumptions built into some former high-profile fear-based HIV prevention campaigns [32]. Concurrently, the IPWG also sought to use the campaign as a vehicle to address health inequities in Chicago [31].

The IPWG developed culturally competent marketing materials, utilizing cross-organizational collaborations to reimagine the ways public health messaging reached vulnerable populations within the city. Developing P4L's sex-positive HIV prevention marketing campaigns took the combined efforts of 60 creative professionals over 10 months, clocking over 1200 hours, to produce 8 final photographs and 50 plus unique files across 15 media formats [31]. The P4L campaign paired fear-based epidemiological language commonly used to describe HIV infection (eg, catch, transmit, contract, and spread) with words describing positive aspects of sex and intimacy (desire, love, heat, and tingle) [42]. Figures 1 and 2 showcase the visual advertisements, where words were painted on the bodies of 2 individuals in each other's embrace. What makes P4L notable is its impressive digital and physical scale and reach, its ability to ally with various media marketing firms, and its capacity to tailor PrEP messaging to specific gender and sexual marginalized communities in Chicago [29]. The final photographs are simple artistic portraits of community members posed in intimate positions that represent the 3 populations most affected by new HIV infections: black, gay men and other MSM, transgender women of color, and black cisgender heterosexual women (Figures 1 to 3). Campaign content was shared with 3 focus groups of 7 to 10 participants representing each of these populations and a fourth group mixed with a proportion of the 3 target populations to assess attitudes on the campaign and to gain feedback to refine the creative content.

Figure 1. PrEP4Love advertising photos, 2016.



Figure 2. PrEP4Love campaign at a Chicago Transit Authority bus stop, Chicago, Illinois.

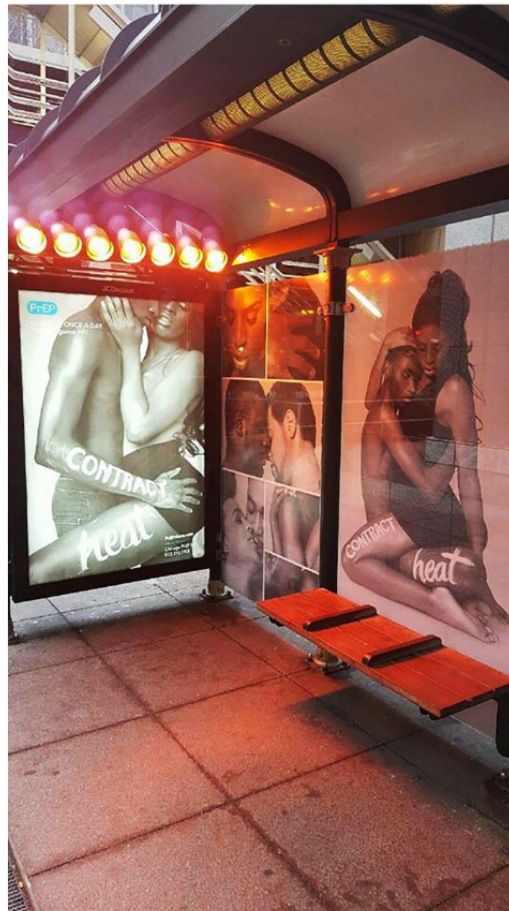
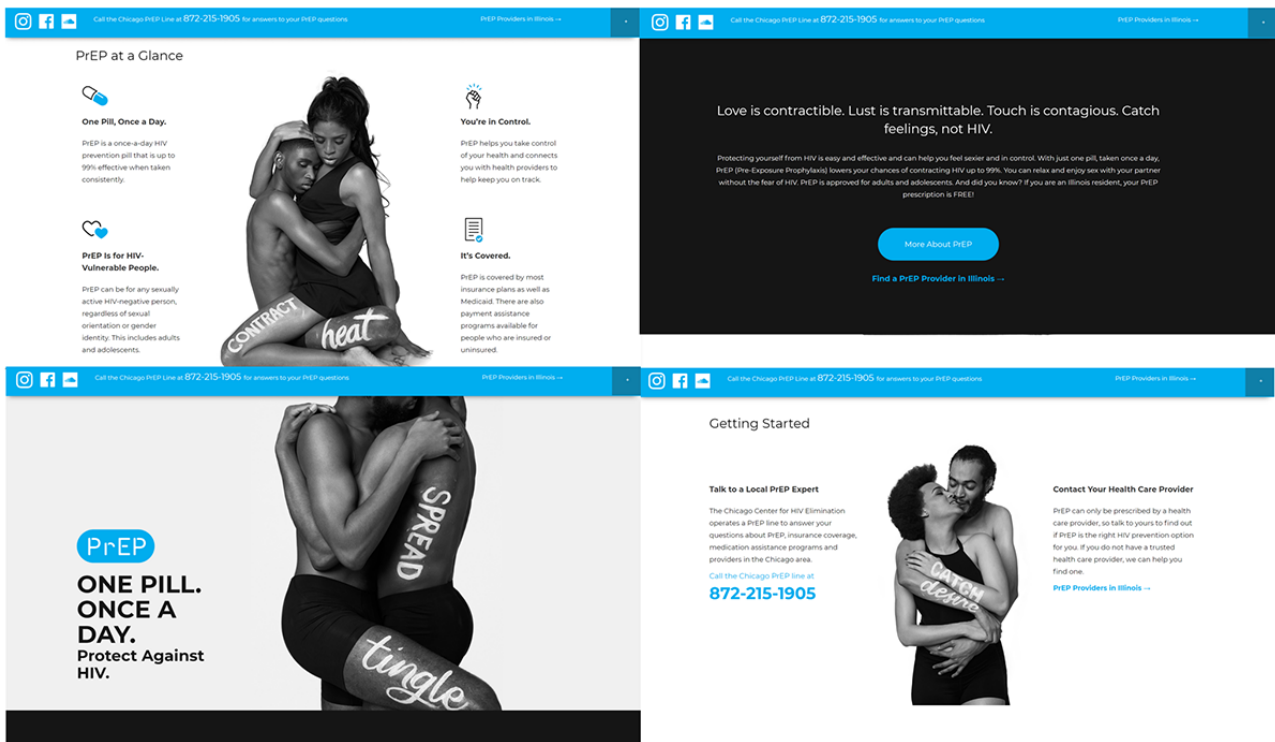


Figure 3. The PrEP4Love website: 4 different snapshots.



Implementation

P4L Web-based advertising launched in February 2016 (Facebook, Instagram, and *smart ads*, defined as ad views that are acquired from a cohort of specific keyword Google searches) and across the CTA, including buses, train stations, platforms and cars, and gay bars using posters and drink coasters throughout Chicago. The placement of these advertisements was based on 3 specific train lines of Chicago (Orange, Green, and Red) that run through primarily black neighborhoods on the south (Green line and Red line) and west sides (Orange line) of Chicago. The red line also reaches the far north side and the *Chicago gay enclave* of *boystown* [43,44]. As P4L aimed to reach the black, gay and trans communities and because of Chicago's racial community segregation, the campaign stakeholders decided to place advertisements along these routes to increase their reach among these populations.

The campaign advertisements directed viewers to the P4L website [45]. The P4L website (Figure 3) directs visitors to the Chicago PrEPLine, a warm-line that provides tailored information about PrEP and links visitors to PrEP services. The Chicago PrEPLine is staffed by employees of the University of Chicago who connect callers to PrEP care. This includes answering questions about PrEP, scheduling their PrEP appointments, and providing transportation services to get to scheduled appointments. The P4L website also guides visitors to a *live* list of PrEP providers in Illinois as well as the *PrEP at a glance* tool, which includes information on PrEP and insurance coverage, who PrEP is for, and how often you would need to take PrEP for it to be effective. The website also has a list of frequently asked questions and upcoming P4L pop-up events. The backdrop of the website is the P4L campaign images (shown in Figure 3).

Data Collection

We collected the first round of user behavior data from 4 different sources: (1) the CTA website, used to report information on Chicago's commuter traffic data and physical P4L advertisements; (2) Google Analytics, used to report information on the P4L Web-based ad campaign and the P4L website data; (3) the Chicago PrEPLine, used to report outcomes of those seeking PrEP from the campaign; and (4) the P4L pop-ups [46], used to continue the conversation of P4L and initiate and drive community-determined health care.

Chicago's commuter traffic data were acquired through the CTA website [47]. We used the following Google Analytics metrics to acquire the remainder of our results: *impressions*, or data on P4L ad views, ad clicks (including clicks per day), *smart ad* views, ad views acquired from social media broken down between Facebook and Instagram, and lastly P4L website Web-based sessions, which includes time spent on the website, demographics of those who spent time on the website, and navigation statistics on the website. The impressions, or ad views, were the key performance indicator, as increasing PrEP awareness was the primary outcome for the campaign. To attribute website traffic to a campaign, tracking codes may be used; however, tracking codes were not used within the P4L campaign. Nevertheless, because this was a grassroots campaign effort that was only advertised on the CTA system, at local bars

and clubs, and through the Web-based ads P4L allotted, we may be able to assume that website traffic was mostly attributed to the P4L campaign.

PrEPLine data are collected in the data collection tool, REDCap [48], by the staff members of the University of Chicago. Typically, a caller will call the PrEPLine number, and the PrEPLine staff assigned *on call* will answer the phone (or call back within 24 hours if missed) and answer any questions that the caller has about PrEP. The staff will also aim to link the caller to a PrEP provider if the caller indicates interest. All data, including demographics, scheduling information, and outcomes, are recorded in REDCap. PrEPLine data are cleaned monthly.

There were 2 live P4L pop-up events [46]: *Crazy Sexy Cool* that focused on young, gay, black men held on February 8, 2016, and *Formation* that focused on black and Latinx cis and transgender women held on March 9, 2016, included in our analysis. Both pop-ups were local, community-led gatherings, accompanied by live music, performers, educational material with discussion, and a P4L photo booth. There were no formal data collected at these events outside of the number of events themselves, but there was promotion of the events on the P4L Instagram and Facebook pages, as well as sponsorship for the event by P4L.

Data Analysis

The analysis period was February through May 15, 2016. The analysis included descriptive statistics for impressions or data on P4L ad views, ad clicks (including clicks per day), *smart ad* views, ad views acquired from social media broken down between Facebook and Instagram, and lastly P4L website Web sessions. The data we received were a direct output from Google Analytics, which is tracked through the Urchin Tracking Module system (UTM). UTMs are the parameters appended to a URL that indicate how site data should be attributed per page session. These data were then sent to our team where the descriptive analyses were conducted.

The analysis methods for the PrEPLine data include running a REDCap report of those who have called the PrEPLine within the data analysis period of their demographics and outcomes.

There were no analysis methods for the pop-up events because no formal data were collected at the events.

Results

Physical Locations

Advertising with the CTA generated an estimated 7 million views during the month of February 2016 based on commuter flow data. P4L advertising was plastered in the interiors of 800 trains and buses for 1 month and was represented on 100 train platforms. In addition, P4L was featured in a specially branded, heated, bus shelter (emphasizing the *contract heat* message; Figure 3) in a heavily trafficked, downtown area of Chicago at the crossroads of numerous bus and rail lines. Branded beverage coasters were deployed at 16 bars across Chicago for 2 months. During this time, 7 of the 16 bars also featured print advertising in restrooms [31].

PrEP4Love Web-based Ad Campaign

Views

The digital campaign coincided with the CTA launch on February 12, 2016, and lasted 2 months, finishing its first round on April 12, 2016. A total of 40,913,560 unique views were generated across various social media platforms (Table 1). Unique views indicate that the website was visited on about 41

million separate occasions, and individuals may have visited the website multiple times. A total of 32,223,987 views were received from *smart ads*, advertising attached to word searches on search engines such as Google; the 3 most clicked on ads were *STD Signs & Symptoms—More Information on STD Symptoms, HIV & AIDS Prevention*, and *HIV Prevention Medication*. An additional 6,970,127 views were gained through Facebook and another 1,719,446 views through Instagram.

Table 1. PrEP4Love ad views.

Type of views	Views, n
Smart ad views	32,223,987
Facebook ad views	6,970,127
Instagram ad views	1,719,446
Total views	40,913,560

Clicks

A total of 24,548 users clicked on P4L ads when the ads appeared on their desktops or media devices. There was an average of 182 clicks per day on the P4L website. *Click-through rate* or CTR or the number of clicks divided by impressions or views, is used to determine the rate of clicks, or interest, an ad is receiving. Industry standards for a social campaign that uses social media platforms such as Facebook or Instagram, hold a benchmark standard CTR range of 0.5% to 0.9%. P4L's CTR was 0.0006, or 0.06%, (24,548/40,913,560); therefore when rating P4L as a social campaign, it is below industry standards.

PrEP4Love Website

Data analysis from the P4L website started on February 1, 2016, and is ongoing through the AIDS Foundation of Chicago and the University of Chicago. As of May 15, 2016, there were 30,881 Web-based sessions among 28,485 users. Unique users accounted for 83.00% (25,632/30,881) of these sessions and 16.99% (5249/30,881) of sessions were returning users. The average time spent on the site was 55 seconds per user (range=20 seconds to 2 min 58 seconds), a rate that had steadily risen since the site launched and hovered over 1 min during the final 4 weeks of analysis. User demographics demonstrated more males (53.00%, 15,098/28,485) visiting the site compared with females (46.99%, 13,387/28,485). Interestingly, there was a higher proportion of 25- to 34-year-olds (34.00%, 9685/28,485) than 18- to 24-year-olds (18.00%, 5128/28,485) who visited the site. User behavior data revealed patterns on how users interacted with the P4L website. In total, 6.04% (1721/28,485) of Web users clicked on a link that takes them to more information

about Chicago-based PrEP providers. In addition, users clicked into the *PrEP clinics and provider* link 2441 times and spent an average of 3 min on the page. *What is PrEP* has received the most user clicks (n=1732), followed by *What about condoms* (n=1556), *Side effects* (n=1459), and *Is PrEP for me* (n=1308). There may be a relationship to further investigate between campaign funding and website activity: because the advertising has stopped, the number of clicks across the site has dropped (data not shown).

Chicago PrEPLine

An additional way of tracking those who were referred for PrEP care through the influence of P4L is to review the data from PrEPLine during the analysis period. P4L partnered with PrEPLine, a program that serves as a liaison between potential PrEP initiators and PrEP providers by linking those interested in PrEP care through phone calls, text messages, Facebook messages, and in-person encounters. PrEPLine's warm-line number was disseminated through all P4L physical and digital marketing campaigns. PrEPLine logged 83 unduplicated encounters with participants during the analysis period. The majority of PrEPLine callers identified as African American or black, young, cisgender, gay or same gender loving men (as shown in Table 2).

During the analysis period, 51% (42/83) indicated being ready for PrEP, 25% (21/83) scheduled an appointment, and 18% (15/83) initiated PrEP. Of the 15 individuals who initiated PrEP, more than half (53%, 8/15) named hearing about PrEP through the P4L campaign. In the 83 encounters as a whole, 34% (28/83) named hearing about PrEP or PrEPLine through the P4L campaign.

Table 2. Sociodemographic characteristics by PrEPLine Outcome (n=83).

Demographics	Total ^a (N=83), n (% ^b)	PrEP ^c questions and concerns addressed (n=29), n (%)	Successfully initiated PrEP (n=15), n (%)	Lost to follow up and no longer interested in PrEP (n=39), n (%)
Age (years)				
<20	6 (12)	2 (18)	0 (0)	4 (17)
20-24	14 (29)	2 (18)	3 (20)	9 (39)
25-29	13 (27)	4 (36)	6 (40)	3 (13)
30-34	4 (8)	0 (0)	2 (13)	2 (9)
35+	12 (24)	3 (27)	4 (27)	5 (22)
Gender				
Cis male	42 (81)	9 (64)	13 (87)	20 (87)
Cis female	9 (17)	5 (36)	1 (7)	3 (13)
Other or not listed	1 (2)	0 (0)	1 (7)	0 (0)
Race (check all that apply)				
Black	44 (77)	12 (70)	13 (72)	19 (86)
Latin(x)	4 (7)	1 (6)	2 (11)	1 (5)
White	4 (7)	1 (6)	2 (11)	1 (5)
Don't know/other/not listed/declined	5 (9)	3 (18)	1 (5)	1 (5)
Sexual orientation				
Gay or same gender loving	19 (50)	3 (50)	8 (53)	8 (47)
Bisexual	5 (13)	0 (0)	2 (13)	3 (18)
Heterosexual	14 (37)	3 (50)	5 (33)	6 (35)

^aMissing data not included in table.

^bMay not sum to 100% because of rounding off.

^cPrEP: pre-exposure prophylaxis.

Discussion

Major Themes

I'm just letting you know that I am sitting next to a couple at the State [and] Lake bus stop, who discussed #PrEP and praised it. And then started making out while leaning up against your picture... right next to me. My night is made and you are changing the world.
[Feedback from a heated bus shelter, February 2016]

Our analysis provides 3 major themes that contribute to equity-focused health research. First, we met our main objective of the campaign which was to disseminate awareness and information regarding a new biomedical HIV prevention method targeting individuals who are most at risk of contracting HIV, which was positively received by the public using a health equity approach. This was demonstrated by the reach of nearly 41 million ad views and the high volume of website engagement. P4L user behavior data collected from the website shows us that the first wave of marketing had a positive aggregate impact on PrEP awareness within Chicago through the percentage of returning site users and the increasing amount of time users spent on the website. Furthermore, the PrEP provider page was the most popular link clicked by users, where users spent an average of 3 min per visit. This could indicate that users were

seeking to connect with providers for PrEP adoption. Data reviewed from the PrEPLine further support evidence that the campaign demand has moved beyond Web-based features to in-person calls made to obtain additional PrEP-related information and PrEP initiation. Moreover, the demographics of the PrEPLine callers indicate the P4L campaign is reaching young BMSM as this was how the majority of the PrEPLine callers identified.

Second, our analysis highlights that a health social movement can successfully side-step the increasing power of commercial and market interests in shaping public health interventions in terms of reach, measured by impressions, or ad views. Medical sociologist, Peter Conrad, argues that the increasing power of private entities, such as pharmaceutical and insurance companies, has crowded out other social forces central to health care decision making (eg, physicians and local governments) and play a major role in determining how medical information and resources are currently distributed within the United States [49]. Gilead, the sole marketer of PrEP, waited 6 years to market Truvada as a biomedical prevention option, as designated by the FDA in 2012. Activists, nonprofit community organizations, researchers, and public health departments therefore banded together to increase overall PrEP awareness around the world and to make PrEP available and accessible to populations most

vulnerable to new HIV infection. The project benefited from being collaborative, collective, community-designed, and community-owned, with hundreds of members of the Chicago PrEP Working Group (now the IPWG with about 350 members).

Finally, although we successfully met our primary goal of PrEP awareness, the campaign's CTR was below industry standards. This can happen for many reasons: the creativity of the ad did not resonate with the population that was viewing it, the ad was not reaching the intended population, the budget was not big enough to keep the *smart ads* and other campaign components running long enough, and so forth. However, based on the data from the ad views, website, and PrEPLine outcomes, we have demonstrated that grassroots campaigns, such as P4L, can be implemented successfully and reach millions of people to increase PrEP awareness using sex-positive and health equity approaches. In addition to the successes monitored through the website and ads, other research was also conducted through Northwestern University on the P4L campaign between June 2017 and April 2018 [50]. The RADAR Study, with a cohort of 700 people, those who had seen a P4L ad in Chicago (75.9%, 531.3/700) were nearly 3 times as likely to talk to their physicians about PrEP, 2 times as likely to initiate the conversation about PrEP with their physician, and nearly 2 times as likely to initiate PrEP within the last 6 months [50]. These findings contribute to the knowledge we have of the PrEP outcomes of those who have seen the P4L ads and demonstrate not only the reach of the campaign but the direct impact on PrEP uptake.

Campaign Feasibility and Evaluation

Private funding allowed for artistic freedom in the campaign design and underlines the importance of pleasure and intimacy in HIV prevention. Sociologists such as Kane Race note the remarkable void of pleasure-based HIV prevention campaigns when it comes to harm reduction sexual health campaigns [39]. We demonstrate the feasibility of reincluding pleasure in harm reduction sexual health campaigns. Though fear-based campaigns can be effective in changing health behaviors in the short term [51], they are the subject of intense ethical debate and possibly less effective long-term than positive or gain-based public health messaging [32,33,35].

As indicated early on, the budget for the initial launch of the campaign was US \$350,000. There is a direct relationship between funding and user engagements. Site activity dropped and, therefore, so did PrEP awareness reach, when sponsored advertising stopped in mid-April 2016 (data not shown). Mobile technology traffic, particularly, fell dramatically after sponsorship ended, accounting for 20.00% (8,182,712/40,913,560) of all site traffic. To have a lasting impact and a robust Web presence, public health campaigns require persistent funding to maintain various digital modalities with which users regularly engage.

This campaign is specific to Chicago in several ways: Chicago is an urban, liberal area that has the support and community to launch a sex-positive HIV prevention campaign that displays embracing partnerships of gender and sexual minorities. It should be mentioned that PrEPLine did receive a small percentage of calls that provided racist and homophobic

feedback relating to the campaign and there were some instances of reported vandalism on the physical ads. This feedback, however, was not overwhelming and was used as an opportunity to educate the callers if the callers were willing to engage. There is also a commuter system in place that is able to double as an advertisement launch pad for the campaign, which was strategically placed based on the racial segregation within the city of Chicago that is mapped out by the train lines. Finally, PrEP care in Chicago is readily available and accessible. This campaign is feasible to recreate in similar urban areas but will have other conditions to consider when planning the campaign launch in other settings, for example, in the southern states or in rural areas. Some potential barriers may include access to PrEP (cost, insurance coverage, transportation, and distance), stigma, and campaign funding. However, the potential solutions to these barriers would be for primary care physicians to educate themselves and each other on PrEP for HIV prevention as a way to minimize transportation, distance and stigma as a barrier. Another way to mitigate stigma would be to offer peer change agent strategies to those who interact with the campaign for their own benefit or interest, but ultimately providing them with strategies to disseminate PrEP information to others in their network who would benefit from PrEP for HIV prevention [52]. These strategies can be offered on the campaign website and could even be integrated as part of the *PrEPLine's*, or campaign-directed warm-line's, discussion of *how to talk to the callers' networks about PrEP*. Regarding the financial burden of PrEP, advocacy around Medicaid coverage for Truvada as PrEP, Gilead-funded payment programs that cover the cost of Truvada as PrEP, and working with eligible health care organizations who are registered to collaborate with the 340B Drug Discount Program, the US federal government program used by pharmacies to lower the cost of drugs significantly, [53] are all ways to provide low-cost or no-cost Truvada as PrEP for uninsured or underinsured patients.

Other important components that any city or state can implement from P4L include something similar to the PrEPLine or the information hub for those who want more information on PrEP. This information line can develop resources and systems to support those who are looking to get linked to PrEP care. The collaborative manner in how the campaign was created was also essential in its success. Groups that are constructing a sex-positive public health campaign should equally partner with community members so that the campaign is culturally competent, inclusive, and affirming. Finally, an evaluation process is important and not always conducted with public health campaigns. Yet, the evaluation component is essential for understanding whether the campaign is effectively delivering the expected outcomes. If these recommendations are implemented, there may be higher rates of success in future health campaigns that recreate a similar campaign as P4L.

Limitations

Limitations exist in the approach of P4L as well. Although using sex-positive messaging in the campaign is intended to be empowering and was well-received by target audiences, it can also provoke stigmatizing responses. The IPWG engaged sex-positivity in its advertising campaign as a way to reintroduce the discourse of sexual pleasure into medical conversations on

HIV risk and prevention. Normative understandings of gender expression and sexualities continue to impact sex-positive HIV prevention work and will subsequently shape future public health campaigns' reach and constitution. Another limitation is that we do not possess more detailed demographic data about those who were counted among the unique views, clicks, and word searches. Even though transportation advertisements were intentionally placed in the pockets of Chicago in areas populated by people of color and therefore those most affected by HIV, it is difficult to conclude whether we effectively reached gay, black men, other black MSM, black ciswomen, and transgender women of color during our analysis period. However, Northwestern University's data, collected between June 2017 and April 2018, do have demographic information as the demographic and P4L data were attached to their research project, RADAR [50]. This cohort (n=700) that reported having 75.9% (531.3/700) who had seen a P4L ad in Chicago, was made up of 31.7% (222/700) Hispanic or Latinx, 30.3% (212/700) white, 27% (189/700) black or African American, and 11% (77/700) other. Participants in RADAR's cohort identified as 91.6% (641/700) cisgender male, 4.3% (30/700) transgender female, and 4.1% (29/700) other [50]. The age make up was 16 to 31 years and 69.6% (487/700) identified as gay [50], 17.7% (124/700) identified as bisexual, 7% (49/700) identified as queer, 2.7% (19/700) identified as heterosexual or straight, 1.1% (8/700) identified as unsure or questioning, and 1.9 (12/700) % identified as other [50]. From these data, we can see that the campaign did reach young, BMSM, which aligns with the aim of the P4L campaign. Another limitation would be that we did not have tracking codes within the P4L website. Although we can assume that the traffic is attributed to the promotion from the campaign, we cannot be fully confident in that statement. We also did not meet industry standards for the CTR and would want to analyze this further and test in the future. With regard to PrEPLine, although we did collect and report on PrEPLine demographic data that underline PrEPLine and P4L reaching young BMSM, the total number was low and there were missing data. This was because although PrEPLine staff attempted to collect demographic information from all callers, there was less success in completing the demographics portion of the database with those who were simply calling for PrEP information instead of those who were ready to schedule a PrEP appointment. Finally, although pop-up events [46] spurred P4L activity in important ways across Chicago, the campaign did not use any data collection tools during these early on events during the analysis period. Since our analysis period, there have been 2 additional live pop-up events: *Black Joy* created for cis and transgender women held on October 4, 2016, and *Black Tea* designed for young, gay, black men on May 5, 2017.

Future Research

Future research should gather demographic information either through increased advertising for PrEPLine or short, digital surveys when each user visits the website. Future research could also use such data to inform predictive factors of those who spent the most time on the website and those who had the most clicks, and we could potentially link individuals who completed the Web-based demographic survey to PrEPLine and PrEP outcomes. More specifically, it is essential to collect data on

those who visit the P4L website to better understand how the Web-based components impact their PrEP awareness, linkage, and uptake. It is also important for public health campaigns, such as P4L to invite more young people and cisgender and transgender, black women to join the movement so as to ensure their health care experiences are culturally competent and sexually affirming. As there were no individuals who were younger than 20 years, and low numbers of cisgender and transgender, black women successfully linked to PrEP care through the PrEPLine in this analysis, more attention should be brought to these groups. This may include more activity on the P4L social media sites, but also potentially hosting youth-centered, female-centered, and transgender and gender nonconforming pop-up events. This would be a great opportunity to acquire in-person data collection that was not collected at the first 4 pop-ups (*Crazy Sexy Cool*, *Formation*, *Black Joy*, and *Black Tea*). There were also 3 focus groups that were used to inform the campaign and analyzing data collected during those focus groups may be a substantial contribution to the sex-positive public health campaign research field as well.

P4L is the first campaign of its kind in terms of reach and scale. Similar grassroots PrEP campaigns have appeared across the world in response to new understandings of HIV prevention. Since the campaign, there have been several initiatives that have been launched that continue to promote sex-positivity and *spread tingle*. New initiatives with European partners have helped shape the P4L campaign in France [54], for example, where they have adopted the P4L campaign, recreated the photos, and made videos for their website that demonstrates the same messaging as Chicago's campaign. P4L has also been a local success with support offered from the Chicago Department of Public Health, which launched the second wave of the campaign in the fall of 2017 across the city's transit system. AIDS Project of the Ozarks, a community-based grassroots organization housed in Springfield, Missouri, is launching the original P4L campaign in southwest Missouri. The only adaptation is the addition of their local PrEP linkage and support information. P4L is a work in progress and, because of that, there is no definitive timeline of when the campaign will end. P4L is not *done*, even during times without paid advertising, it lives through its live pop-up events and ongoing social media posts and continues to track engagement and reach in the community. The next steps for the P4L campaign are to launch a Latinx-specific component. This paper demonstrates the success and limitations of the first wave of a data-driven, sex-positive sexual health campaign that fostered respectful community engagement. Through the success of this campaign we also underscore a couple of take-away points. First, implementing P4L required a substantial amount of start-up financial capital. Second, the importance of readily available and accessible PrEP resources to which the campaign can refer inquiring and interested viewers. Finally, the success of this campaign was brought about, in part, through the relationships brokered by Jim Pickett with supportive marketing and advertising agencies. These resources, together, were important in the launch and success of the sex-positive P4L campaign and therefore central to the dissemination of PrEP health information, PrEP awareness, and accessibility of PrEP in Chicago.

Acknowledgments

The authors would like to sincerely thank Allison Joson who provided insight and expertise of the marketing methods and analytics component of this study, which greatly improved the quality of this paper.

Conflicts of Interest

None declared.

References

1. Braveman P. What are health disparities and health equity? We need to be clear. *Public Health Rep* 2014;129(Suppl 2):5-8 [FREE Full text] [doi: [10.1177/00333549141291S203](https://doi.org/10.1177/00333549141291S203)] [Medline: [24385658](https://pubmed.ncbi.nlm.nih.gov/24385658/)]
2. Centers for Disease control and Prevention. 2016. HIV Basics: Basic Statistics URL: <http://www.cdc.gov/hiv/basics/statistics.html> [accessed 2019-05-29] [WebCite Cache ID 78jbopEHM]
3. Rubin MS, Colen CG, Link BG. Examination of inequalities in HIV/AIDS mortality in the United States from a fundamental cause perspective. *Am J Public Health* 2010 Jun;100(6):1053-1059. [doi: [10.2105/AJPH.2009.170241](https://doi.org/10.2105/AJPH.2009.170241)] [Medline: [20403885](https://pubmed.ncbi.nlm.nih.gov/20403885/)]
4. Centers for Disease Control and Prevention. 2016 Feb 23. Half of black gay men and a quarter of Latino gay men projected to be diagnosed within their lifetime URL: <https://www.cdc.gov/nchhstp/newsroom/2016/croi-press-release-risk.html> [accessed 2019-05-28] [WebCite Cache ID 78je6gyRd]
5. Williams DR, Sternthal M. Understanding racial-ethnic disparities in health: sociological contributions. *J Health Soc Behav* 2010;51 Suppl:S15-S27 [FREE Full text] [doi: [10.1177/0022146510383838](https://doi.org/10.1177/0022146510383838)] [Medline: [20943580](https://pubmed.ncbi.nlm.nih.gov/20943580/)]
6. Sullivan PS, Salazar L, Buchbinder S, Sanchez TH. Estimating the proportion of HIV transmissions from main sex partners among men who have sex with men in five US cities. *AIDS* 2009 Jun 1;23(9):1153-1162. [doi: [10.1097/QAD.0b013e32832baa34](https://doi.org/10.1097/QAD.0b013e32832baa34)] [Medline: [19417579](https://pubmed.ncbi.nlm.nih.gov/19417579/)]
7. Chicago Department of Public Health. 2015. HIV/STI Surveillance Report, Chicago: December 2015 URL: https://www.chicago.gov/content/dam/city/depts/cdph/HIV_STI/HIV_STI_Surveillance_Report_2015.pdf [accessed 2019-05-29] [WebCite Cache ID 78jey7PNQ]
8. Smith DK. HIV prevention pill is not reaching most who could potentially benefit- especially African Americans and Latinos. 2018 Presented at: Conference on Retroviruses and Opportunistic Infections; March 4-7, 2018; Boston, MA URL: https://www.cdc.gov/nchhstp/newsroom/images/2018/hiv/PrEP_highres.jpg
9. Anderson PL, Glidden DV, Liu A, Buchbinder S, Lama JR, Guanira JV, et al. Emtricitabine-tenofovir concentrations and pre-exposure prophylaxis efficacy in men who have sex with men. *Sci Transl Med* 2012 Sep 12;4(151):151ra125 [FREE Full text] [doi: [10.1126/scitranslmed.3004006](https://doi.org/10.1126/scitranslmed.3004006)] [Medline: [22972843](https://pubmed.ncbi.nlm.nih.gov/22972843/)]
10. Bush S, Magnuson D, Rawlings MK, Hawkins T, McCallister S, Giler RM. Racial Characteristics of FTC/TDF for Pre-exposure Prophylaxis (PrEP) Users in the US #2651. 2016 Presented at: ASM Microbe; June 16-20, 2016; Boston, MA URL: http://www.natap.org/2016/HIV/062216_02.htm
11. Collins PH. Prisons for Our Bodies, Closets for Our Minds, in *Black Sexual Politics*. New York, New York: Routledge; 2004:87-116.
12. Wheeler DP, Lucas J, Wilton L, Nelson LE, Hucks-Ortiz C, Watson CC, et al. Building effective multilevel HIV prevention partnerships with black men who have sex with men: experience from HPTN 073, a pre-exposure prophylaxis study in three US cities. *J Int AIDS Soc* 2018;21(S7):e25180. [Medline: [30334600](https://pubmed.ncbi.nlm.nih.gov/30334600/)]
13. Schneider JA, Bouris A, Smith DK. Race and the public health impact potential of pre-exposure prophylaxis in the United States. *J Acquir Immune Defic Syndr* 2015 Sep 01;70(1):e30-e32. [doi: [10.1097/QAI.0000000000000670](https://doi.org/10.1097/QAI.0000000000000670)] [Medline: [25950209](https://pubmed.ncbi.nlm.nih.gov/25950209/)]
14. Staton T. Fierce Pharma. 2015. Here's a switcheroo: HIV activists target Gilead for holding back on Truvada marketing URL: <https://www.fiercepharma.com/sales-and-marketing/here-s-a-switcheroo-hiv-activists-target-gilead-for-holding-back-on-truvada> [accessed 2019-05-29] [WebCite Cache ID 78juFPVm2]
15. Brown P, Zavestoski S, McCormick S, Mayer B, Morello-Frosch R, Gasior Altman R. Embodied health movements: new approaches to social movements in health. *Sociol Health Illn* 2004 Jan;26(1):50-80 [FREE Full text] [doi: [10.1111/j.1467-9566.2004.00378.x](https://doi.org/10.1111/j.1467-9566.2004.00378.x)] [Medline: [15027990](https://pubmed.ncbi.nlm.nih.gov/15027990/)]
16. Randolph W, Viswanath K. Lessons learned from public health mass media campaigns: marketing health in a crowded media world. *Annu Rev Public Health* 2004;25:419-437. [doi: [10.1146/annurev.publhealth.25.101802.123046](https://doi.org/10.1146/annurev.publhealth.25.101802.123046)] [Medline: [15015928](https://pubmed.ncbi.nlm.nih.gov/15015928/)]
17. Karnik R. Sierra Club. 2015 Sep 03. Environmental, public health, and civil rights groups unveil new environmental justice campaign: Ads highlight disproportionate impacts of smog pollution on communities of color, call on EPA for strong standards URL: <https://content.sierraclub.org/press-releases/2015/09/>

- [environmental-public-health-and-civil-rights-groups-unveil-new-environmental](#) [accessed 2019-06-04] [WebCite Cache ID 78sKnuQO7]
18. Cision. 2013 Nov 25. Marketing shop LimeGreen partners with Chicago Department of Public Health to launch flavored tobacco awareness campaign URL: <https://www.prnewswire.com/news-releases/marketing-shop-limegreen-partners-with-chicago-department-of-public-health-to-launch-flavored-tobacco-awareness-campaign-233312041.html> [accessed 2019-05-29] [WebCite Cache ID 78jv2GYIC]
 19. Gurciullo B. Chicago Tribune. 2015 Dec 23. Public health campaign takes aim at vaping among city youth URL: <https://www.chicagotribune.com/news/breaking/ct-vaping-campaign-met-20151223-story.html> [accessed 2019-05-29] [WebCite Cache ID 78jvEq2iZ]
 20. City of Chicago. 2013 May 14. Provocative new campaign sparks citywide conversations on teen parenthood URL: https://www.chicago.gov/city/en/depts/cdph/provdrs/healthychicago/news/2013/may/provocative_new_campaignsparkscitywideconversationsonteenparenth.html [accessed 2019-05-29] [WebCite Cache ID 78jvTzSjW]
 21. CBS Chicago. 2015 Dec 01. Chicago youth help launch URL: <https://chicago.cbslocal.com/2015/12/01/chicago-youth-help-launch-chicago-wears-condoms-campaign/> [accessed 2019-05-29] [WebCite Cache ID 78jw51Uom]
 22. Stein E. City of Chicago. 2012 Mar 09. City Launches "Take Charge. Take the Test." HIV Prevention Campaign for African-American Women URL: https://www.chicago.gov/city/en/depts/cdph/provdrs/health_services/news/2012/mar/city_launches_campaigntoencouragehighriskwomentogetestforhiv.html [accessed 2019-05-29] [WebCite Cache ID 78jwVWRoq]
 23. The Stigma Project. 2016. URL: http://www.thestigmaproject.org.usrfiles.com/html/88e4f8_06f07071bee5da2981e9f3f7735bf2a5.html [accessed 2019-05-29] [WebCite Cache ID 78jwZuZG]
 24. How To Get PrEP. 2016 Apr 21. URL: <http://www.howtogetprep.com/> [accessed 2019-05-29] [WebCite Cache ID 78jx2TOKB]
 25. Harlem United. 2015 Jun 30. Swallow This URL: <http://www.harlemunited.org/swallow-this-launch/>
 26. PrEP Facts. 2016. PrEP URL: <https://prepfacts.org/> [accessed 2019-05-29] [WebCite Cache ID 78jxmJBzT]
 27. Star Observer. 2015 Sep 18. Melbourne Poster Saying Gay Men Can "F*** Raw" on PrEP Follow Report Revealing Record Figures of Condomless Sex URL: <http://www.starobserver.com.au/news/national-news/melbourne-posters-saying-gay-men-can-f-raw-on-prep-come-after-report-reveals-record-figures-of-unprotected-sex/140834> [accessed 2019-05-29] [WebCite Cache ID 78jvEREaF]
 28. Apoyo Positivo. 2015 Nov 17. #placajealVIH 2015 URL: <http://apoyopositivo.org/campanas-apoyo/placajealvih-2015/>
 29. APCOM. 2015 Jul 08. PrEPARING ASIA: A new direction for HIV prevention among MSM in Asia URL: <http://www.apcom.org/2015/07/08/preparing-asia-a-new-direction-for-hiv-prevention-among-msm-in-asia/>
 30. Basta TB, Reece M, Wilson MG. The transtheoretical model and exercise among individuals living with HIV. *Am J Health Behav* 2008;32(4):356-367. [doi: [10.5555/ajhb.2008.32.4.356](https://doi.org/10.5555/ajhb.2008.32.4.356)] [Medline: [18092896](https://pubmed.ncbi.nlm.nih.gov/18092896/)]
 31. Pickett J. #TransmitLove across Chicago: PrEP4Love Campaign Update. In: AIDS Foundation of Chicago. 2016 Presented at: Chicago PrEP Working Group; June 6, 2016; Chicago, IL.
 32. Weinstein M. Business Wire. 2015 Jun 18. Correcting and Replacing Graphic AHF President Calls Attention to URL: <https://www.businesswire.com/news/home/20150616006589/en/CORRECTING%20and-REPLACING%20GRAPHIC-AHF-President-Calls-Attention-%E2%80%9CThe> [accessed 2019-05-29] [WebCite Cache ID 78jzm4kgb]
 33. Duran D. Huffington Post. 2012 Nov 12. Truvada whores? URL: https://www.huffpost.com/entry/truvada-whores_b_2113588 [accessed 2019-05-29] [WebCite Cache ID 78k01K65o]
 34. Mamo L, Epstein S. The pharmaceuticalization of sexual risk: vaccine development and the new politics of cancer prevention. *Soc Sci Med* 2014 Jan;101:155-165. [doi: [10.1016/j.socscimed.2013.11.028](https://doi.org/10.1016/j.socscimed.2013.11.028)] [Medline: [24560236](https://pubmed.ncbi.nlm.nih.gov/24560236/)]
 35. Trice DT. Chicago Tribune. 2010 May 13. HIV prevention campaign ends after outcry URL: <https://www.chicagotribune.com/news/ct-xpm-2010-05-13-ct-met-trice-aids-campaign-20100513-story.html> [accessed 2019-05-29] [WebCite Cache ID 78k0HiA9I]
 36. Bailey MM. No tea, no shade: New writings in black queer studies. Durham, NC: Duke University Press; 2016.
 37. HIV Equal. 2016 Feb 09. APLA Reveals "It Feels Good" PrEP Campaign URL: <http://www.hivequal.org/hiv-equal-online/apla-reveals-it-feels-good-prep-campaign> [accessed 2019-05-29] [WebCite Cache ID 78k0Z1Mmi]
 38. Ohkura M, Bonenfant N, Upton MD. A sex-positive approach to healthcare, and Truvada as HIV Pre-Exposure Prophylaxis (PrEP). 2018 Presented at: Larner College of Medicine Faculty Publications; January 8, 2018; Burlington, VT URL: <https://scholarworks.uvm.edu/comfac/3/>
 39. Race K. The use of pleasure in harm reduction: perspectives from the history of sexuality. *Int J Drug Policy* 2008;19(5):417-423. [doi: [10.1016/j.drugpo.2007.08.008](https://doi.org/10.1016/j.drugpo.2007.08.008)] [Medline: [17904347](https://pubmed.ncbi.nlm.nih.gov/17904347/)]
 40. NBC News. 2018 May 30. With TV Ads, Drugmaker Gilead Brings HIV Prevention Drug Out of the Closet URL: <https://www.nbcnews.com/feature/nbc-out/tv-ads-drugmaker-gilead-brings-hiv-prevention-drug-out-of-the-closet-n878721> [accessed 2019-05-29] [WebCite Cache ID 78k161Gwd]

41. Gamarel KE, Golub SA. Intimacy motivations and pre-exposure prophylaxis (PrEP) adoption intentions among HIV-negative men who have sex with men (MSM) in romantic relationships. *Ann Behav Med* 2015 Apr;49(2):177-186 [FREE Full text] [doi: [10.1007/s12160-014-9646-3](https://doi.org/10.1007/s12160-014-9646-3)] [Medline: [25124457](https://pubmed.ncbi.nlm.nih.gov/25124457/)]
42. AIDS Foundation of Chicago. 2016 Feb 01. Citywide campaign transmits love shows the sexier side of HIV prevention URL: <https://www.aidschicago.org/page/news/all-news/citywide-campaign-transmits-love-and-shows-the-sexier-side-of-hiv-prevention> [accessed 2019-05-29] [WebCite Cache ID [78k1LZYVu](https://www.webcitation.org/78k1LZYVu)]
43. Ghaziani A. *There Goes the Gayborhood*. Princeton, NJ: Princeton University Press; 2014.
44. Orne J. *Boystown: Sex & Community in Chicago*. Chicago & London: The University of Chicago Press; 2017.
45. PrEP4Love. One Pill. Once a Day. Protect Against HIV. URL: <https://prep4love.com/> [accessed 2019-05-31]
46. Ghaziani A, Stillwagon R. Queer pop-ups. *Contexts* 2018;17(1):78-80. [doi: [10.1177/1536504218766544](https://doi.org/10.1177/1536504218766544)]
47. Chicago Transit Authority. 2016. URL: <https://www.transitchicago.com/> [accessed 2019-05-29] [WebCite Cache ID [78k29VCeJ](https://www.webcitation.org/78k29VCeJ)]
48. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
49. Conrad P. The shifting engines of medicalization. *J Health Soc Behav* 2005 Mar;46(1):3-14. [doi: [10.1177/002214650504600102](https://doi.org/10.1177/002214650504600102)] [Medline: [15869117](https://pubmed.ncbi.nlm.nih.gov/15869117/)]
50. Phillips II G, Lindeman PT, Raman AB, Hayford CS, Mustanski B. Influence of PrEP4Love Campaign on PrEP Uptake Among YMSM in Chicago. 2019 Presented at: Conference on Retroviruses and Opportunistic Infections; March 4-7, 2018; Seattle, WA URL: <https://www.croiconference.org/sessions/influence-prep4love-campaign-prep-uptake-among-ymsm-chicago>
51. Bhatia R, Koch A, Gutierrez E, Prachand N. PrEP uptake in a community sample of MSM in Chicago. 2015 Presented at: American Public Health Association; October 31-November 4, 2015; Chicago, IL URL: <https://apha.confex.com/apha/143am/webprogram/Paper336951.html>
52. Young LE, Schumm P, Alon L, Bouris A, Ferreira M, Hill B, et al. PrEP Chicago: a randomized controlled peer change agent intervention to promote the adoption of pre-exposure prophylaxis for HIV prevention among young black men who have sex with men. *Clinical Trials* 2018;15(1):44-52. [Medline: [28862483](https://pubmed.ncbi.nlm.nih.gov/28862483/)]
53. 340b Health. 2019. Overview of the 340b drug pricing program URL: <https://www.340bhealth.org/members/340b-program/overview/>
54. AIDES. 2018 Jul 12. La PrEP Dossier URL: <https://www.aides.org/campagne/prep4love> [accessed 2019-05-29] [WebCite Cache ID [78k3pfvzR](https://www.webcitation.org/78k3pfvzR)]

Abbreviations

BMSM: black, gay and bisexual men and other men who have sex with men

CTA: Chicago Transit Authority

CTR: click-through rate

FDA: Food and Drug Administration

IPWG: Illinois PrEP working group

LGBTQ: lesbian, gay, bisexual, transgender and queer

MSM: men who have sex with men

P4L: PrEP4Love

PrEP: pre-exposure prophylaxis

UTM: urchin tracking module

Edited by G Eysenbach; submitted 14.11.18; peer-reviewed by V McMahan, V Patel; comments to author 17.01.19; revised version received 13.03.19; accepted 10.05.19; published 17.06.19.

Please cite as:

Dehlin JM, Stillwagon R, Pickett J, Keene L, Schneider JA

#PrEP4Love: An Evaluation of a Sex-Positive HIV Prevention Campaign

JMIR Public Health Surveill 2019;5(2):e12822

URL: <http://publichealth.jmir.org/2019/2/e12822/>

doi: [10.2196/12822](https://doi.org/10.2196/12822)

PMID: [31210141](https://pubmed.ncbi.nlm.nih.gov/31210141/)

©Jessica M Dehlin, Ryan Stillwagon, Jim Pickett, Lance Keene, John A Schneider. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 17.06.2019. 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

Factors Associated With Willingness to Use Pre-Exposure Prophylaxis in Brazil, Mexico, and Peru: Web-Based Survey Among Men Who Have Sex With Men

Thiago Silva Torres^{1*}, PhD; Kelika A Konda^{2*}, PhD; E Hamid Vega-Ramirez^{3,4*}, MD, MSc; Oliver A Elorreaga², MSc; Dulce Diaz-Sosa^{3,4}, PsyD; Brenda Hoagland¹, MD, PhD; Steven Diaz⁵, MD; Cristina Pimenta⁶, PhD; Marcos Benedetti¹, MSc; Hugo Lopez-Gatell⁷, PhD; Rebeca Robles-Garcia⁴, PhD; Beatriz Grinsztejn¹, MD, PhD; Carlos Caceres², MD, PhD; Valdilea G Veloso¹, MD, PhD; ImPrEP Study Group^{1,2,3}

¹Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz (INI/Fiocruz), Rio de Janeiro, Brazil

²Centro de Investigación Interdisciplinaria en Sexualidad Sida y Sociedad, UPCH, Lima, Peru

³Condesa & Condesa-Iztapalapa Specialized Clinics, Mexico City, Mexico

⁴National Institute of Psychiatry Ramon de la Fuente Muñiz, Mexico City, Mexico

⁵Center for Prevention and Comprehensive Healthcare for HIV/AIDS of Mexico City, Mexico City, Mexico

⁶Brazilian Ministry of Health, Brasília, Brazil

⁷National Institute of Public Health, Mexico City, Mexico

*these authors contributed equally

Corresponding Author:

Thiago Silva Torres, PhD

Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz (INI/Fiocruz)

Av Brasil 4365 Manguinhos

Rio de Janeiro, 21040-900

Brazil

Phone: 55 38659623

Fax: 55 38659679

Email: thiago.torres@ini.fiocruz.br

Abstract

Background: HIV disproportionately affects key populations including men who have sex with men (MSM). HIV prevalence among MSM varies from 17% in Brazil and Mexico to 13% in Peru, whereas it is below 0.5% for the general population in each country. Pre-exposure prophylaxis (PrEP) with emtricitabine/tenofovir is being implemented in the context of combination HIV prevention. Reports on willingness to use PrEP among MSM have started to emerge over the last few years. Previously reported factors associated with willingness to use PrEP include awareness, higher sexual risk behavior, and previous sexually transmitted infection.

Objective: This study aimed to evaluate the factors associated with willingness to use daily oral PrEP among MSM in 3 Latin American, middle-income countries (Brazil, Mexico, and Peru).

Methods: This Web-based, cross-sectional survey was advertised in 2 gay social network apps (Grindr and Hornet) used by MSM to find sexual partners and on Facebook during 2 months in 2018. Inclusion criteria were being 18 years or older, cisgender men, and HIV-negative by self-report. Eligible individuals answered questions on demographics, behavior, and PrEP (awareness, willingness to use, barriers, and facilitators). Multivariable logistic regression modeling was performed to assess the factors associated with willingness to use daily oral PrEP in each country.

Results: From a total sample of 43,687 individuals, 44.54% of MSM (19,457/43,687) were eligible and completed the Web-based survey—Brazil: 58.42% (11,367/19,457), Mexico: 30.50% (5934/19,457), and Peru: 11.08% (2156/19,457); median age was 28 years (interquartile range: 24–34), and almost half lived in large urban cities. Most participants were recruited on Grindr (69%, 13,349/19,457). Almost 20% (3862/19,352) had never tested for HIV, and condomless receptive anal sex was reported by 40% (7755/19,326) in the previous 6 months. Whereas 67.51% (13,110/19,376) would be eligible for PrEP, only 9.80% (1858/18,959) of participants had high HIV risk perception. PrEP awareness was reported by 64.92% (12,592/19,396); this was lower in Peru (46.60%, 1002/2156). Overall, willingness to use PrEP was reported by 64.23% (12,498/19,457); it was highest in Mexico (70%,

4158/5934) and lowest in Peru (58%, 1241/2156). In multivariate regression models adjusted for age, schooling, and income in each country, willingness to use PrEP was positively associated with PrEP awareness and PrEP facilitators (eg, free PrEP and HIV testing) and negatively associated with behavioral (eg, concerned by daily pill regimen) and belief barriers (eg, sexual partners may expect condomless sex).

Conclusions: In this first cross-country, Web-based survey in Latin America, willingness to use PrEP was found to be high and directly related to PrEP awareness. Interventions to increase awareness and PrEP knowledge about safety and efficacy are crucial to increase PrEP demand. This study provides important information to support the implementation of PrEP in Brazil, Mexico, and Peru.

(*JMIR Public Health Surveill* 2019;5(2):e13771) doi:[10.2196/13771](https://doi.org/10.2196/13771)

KEYWORDS

pre-exposure prophylaxis; men who have sex with men; prevention; Latin America; surveys and questionnaires

Introduction

Background

HIV continues to be a major health problem worldwide. The Latin American region has the fourth largest number of individuals living with HIV (1.8 million accumulated cases) and is fifth with regard to new HIV infections (96,000). About 90% of new HIV infections in 2016 in Latin America occurred in 10 countries, including Brazil (49%), Mexico (13%), and Peru (4%) [1]. In this region, HIV disproportionately affects key populations, and these are primarily gay, bisexual, and other men who have sex with men (MSM) and transgender women [2]. HIV prevalence among MSM varies from 17% in Brazil [3] and Mexico [4] to 13% in Peru [5], whereas it is below 0.5% for the general population in each country [6-8]. This continuing burden highlights the need for a more energetic, integrated, and strategic focus on combination HIV prevention among this population.

Pre-exposure prophylaxis (PrEP) with daily oral emtricitabine/tenofovir (FTC/TDF) is an effective biomedical prevention strategy to reduce HIV acquisition among MSM [9-12]. PrEP guidelines were first published in 2012 by the US Centers for Disease Control and Prevention [13], followed by the World Health Organization (WHO) in 2014 [14], and in 2017, by the Brazilian Ministry of Health (MoH) [15]. PrEP is being implemented by public health services in Brazil [16], and it is being considered for implementation in Mexico and Peru, even though neither country has issued guidance with regard to PrEP.

Reports on PrEP awareness, willingness to use, and acceptability among MSM have started to emerge over the last few years. Few MSM were aware of PrEP (<50%) in 11 studies in low- and middle-income countries [17]. Willingness to use PrEP among MSM varies by country and time, ranging from 32% to 92% [18], suggesting a diversity of factors associated with willingness to use PrEP. Previously reported factors associated with willingness to use PrEP included the following: awareness, low cost of PrEP, higher sexual risk, previous sexually transmitted infection (STI), and unwillingness to use condoms [19]. However, these factors may vary according to the characteristics of the region or country and may change rapidly. Recent reports of awareness and willingness to use PrEP among MSM are available for Brazil [20-22], but no information was

captured after the availability of PrEP in the public health system. Conversely, no recent reports are available for either Mexico or Peru.

The Implementation PrEP Project (ImPrEP) is the first transnational project in Latin America aiming to generate evidence on the feasibility, acceptability, and cost-effectiveness of PrEP among key populations (MSM and transgender people) specific to the cultural contexts and health systems in Brazil, Mexico, and Peru. Within ImPrEP, several studies are being conducted, including a PrEP demonstration study with 7500 MSM and transgender people in these countries. The results of ImPrEP [23] will permit stakeholders to evaluate and incorporate PrEP as part of combination HIV prevention within their countries. The data presented herein are from a formative survey conducted for ImPrEP, with the aim to better understand the characteristics, sexual behavior, and the knowledge and opinions regarding PrEP among MSM.

Objective

This study evaluated the factors associated with willingness to use PrEP among MSM in Brazil, Mexico, and Peru.

Methods

Study Design

This was a cross-sectional, Web-based survey for MSM from 3 countries in Latin America: Brazil, Mexico, and Peru. Individuals who met the eligibility criteria (age ≥ 18 years, cisgender men, and HIV uninfected by self-report) and provided informed consent were directed to the Web-based questionnaire.

The questionnaire was conducted using SurveyGizmo in Brazil and SurveyMonkey in Mexico and Peru. The project was advertised on 2 geosocial networking apps for sexual encounters among men (Hornet and Grindr) and Facebook social media. On Facebook, advertisements focused toward male gender, country (Brazil, Mexico, or Peru), and related interests, applying keywords frequently used by gay and bisexual population, for instance, gay pride, gay community, and homophobia. Hornet users received 2 inbox messages with a link to the survey, and Grindr users received pop-up advertisements, 1 per week for 2 months. The advertisements were launched in March for Brazil and Peru and in June for Mexico, all in 2018. The questionnaire remained open for 2 months in each country. The informed consent stated clearly that the survey was targeting MSM.

Another survey targeting transgender people within ImPrEP is still ongoing. No incentives were provided for answering the survey.

Variables

Sociodemographics

Age was categorized in the following 3 brackets: 18 to 24, 25 to 34, and ≥ 35 years. Race was dichotomized as white versus nonwhite (black, mixed-race, Asian, or indigenous) for Brazil and Peru; in Mexico, this question was dichotomized as indigenous versus nonindigenous. Definition of monthly income varied in each country (see Table 1 footnote). Schooling was dichotomized by those who had less than or complete secondary education versus any postsecondary education.

HIV Testing and HIV Risk Perception in the Next Year

Individuals were asked when they were last tested for HIV. The options were as follows: last 3 months, last 6 months, last year, more than 1 year, and never tested. We dichotomized this question in never versus at least once in a lifetime for the multivariate analyses. HIV risk perception was assessed with the question *What is your chance of getting HIV in the next year?* with possible options grouped into 3 categories for analysis: low (none/low risk), middle (some risk), and high (high risk/certainty of infection) [22].

Sexual Behavior and Preliminary Eligibility for Pre-Exposure Prophylaxis (Adapted From the World Health Organization Risk Criteria for Pre-Exposure Prophylaxis)

Sexual behavior was assessed in the last 6 months with the following: number of male sexual partners (0, 1-5, 6-10, and ≥ 10); condomless receptive anal sex, condomless insertive anal sex, sex with HIV-positive male partner, sex with a male partner of unknown HIV status, sex under the influence of alcohol, and chemsex or sex under the influence of excitatory drugs (all dichotomized yes/no).

We evaluated preliminary eligibility for PrEP considering the ImPrEP demonstration study's risk-related inclusion criteria, which were adapted from the WHO recommendations for PrEP use (all in the past 6 months): reporting unprotected sex, having an HIV-positive male partner, exchanging sex for money, or having an STI [24]. Individuals were then dichotomized (yes/no) as preliminarily eligible for PrEP based on risk if they fulfilled any of the listed criteria.

Substance Use

Binge drinking [25] was evaluated as follows: *In the last six months, did you drink five or more drinks in a couple of hours?* A dose was defined as 1 can of beer (300 mL) or 1 glass of wine (120 mL) or 1 shot of distilled alcohol (30 mL of ex. cachaça, vodka, whisky, tequila, mezcal, or pisco). Substance use in the past 6 months included the following: tobacco, marijuana, hash, stimulants (cocaine, crack, amphetamines, and 4-hydroxybutanoic acid), hallucinogens (lysergic acid diethylamide, ketamine), *poppers* (alkyl nitrites), and erectile dysfunction drugs.

Awareness, Willingness to Use, Barriers, and Facilitators of Daily Oral Pre-Exposure Prophylaxis

Awareness was assessed as follows: *Have you ever heard of PrEP?* and dichotomized yes/no. A brief explanation of PrEP was provided after respondents had answered the PrEP awareness question. PrEP use was categorized as follows: *never, current, and past*. We asked respondents to rank 3 PrEP regimens from 1-most preferred to 3-least preferred: daily oral (1 FTC/TDF pill/day), event-driven (2 pills before intercourse and 1 pill 24 hours and 48 hours after), and injectable (injection every 2 months).

Willingness is understood as an intentional behavior based on 2 principles: (1) recognition of the behavioral objective (ie, taking PrEP) and (2) strategies implemented to achieve this objective. The second principle can be *experiential processes*, which refer to cognitive strategies (eg, believing that PrEP will protect against HIV) or *behavioral processes*, which are strategies to produce and maintain the behavior (eg, setting an alarm as a reminder to take PrEP) [26]. We used a 5-point Likert scale to assess willingness to use PrEP, anticipated risk compensation, and if participants would use PrEP provided by the MoH at no cost. Willingness to use PrEP was defined as responding *highly likely* to the statement *I would use one daily pill for PrEP*. Anticipated risk compensation was defined as responding *highly likely* or *likely* to the following statement: *I would not use a condom if I used PrEP*. PrEP if provided by the MoH was defined as responding *highly likely* to the statement *I would use PrEP if part of a MoH program*. The inclusion of the *likely* category only for anticipated risk compensation was to be more conservative with this as a potentially negative outcome of PrEP use.

Barriers and facilitators to daily oral PrEP were accessed using a 5-point Likert scale (*very important* to *not important*) [27]. For descriptive statistics, barriers were defined as responding *very important* or *important*, and facilitators were defined as responding *very important*. The inclusion of *important* category only for barriers was to more broadly identify relevant obstacles of willingness to use PrEP. Barriers were grouped into the following: informational (eg, concerned by short-term or long-term side effects, that antiretroviral therapy would not work if they became HIV infected while on PrEP, and by lack of 100% protection against HIV), behavioral (eg, concerned by daily pill regimen, talking to a doctor about sex, and having quarterly HIV/STI testing), and beliefs (eg, concerned that taking PrEP implies being at risk of HIV acquisition, sexual partners may expect condomless anal sex, people may assume they are HIV-positive, and people may ask why they are taking pills). For facilitators, we assessed the following: free PrEP, access to free HIV testing, access to other free exams (eg, HIV/STI testing), access to personal PrEP counseling, adherence support from apps, support, and counseling about sex life. Using these groupings, we performed confirmatory factor analysis and calculated the Cronbach alpha to verify if these items could be grouped (each Cronbach alpha was $>.70$).

Percentage of willingness to use daily oral prep is also given per region. For Brazil, region was categorized according to geopolitical regions: North (7 states), Northeast (9 states),

Central-West (3 states and Federal District), South (3 states), and Southeast (2 states). São Paulo and Rio de Janeiro are part of Southeast Brazil, but they were maintained separated in the analysis because of the high number of responders from these 2 states. For Peru, regions are grouped according to their geographical characteristics and political division: Lima (Lima city and Callao), Coast (Lima region and other coastal cities), Sierra (cities of the northern, central, and southern highlands), and Jungle. For Mexico, the regions were Northwest (Baja California, Baja California Sur, Chihuahua, Durango, Sinaloa, and Sonora), Northeast (Coahuila, Nuevo León, and Tamaulipas), West (Nayarit, Jalisco, Colima, and Michoacán), East (Puebla, Veracruz, Tlaxcala, and Hidalgo), North Center (Aguascalientes, Guanajuato, San Luis Potosí, Zacatecas, and Querétaro), South Central (Morelos, State of Mexico, and Mexico City), and South (Guerrero, Oaxaca, Chiapas, Tabasco, Campeche, Quintana Roo, and Yucatán). The last 2 regions (Southwest and Southeast), given their sociodemographic similarities, were included as a single region.

Ethical Issues

An ethical review board in each country approved the study: in Brazil, INI Evandro Chagas-FIOCRUZ institutional review board (#CAAE 82021918.0.0000.5262); in Mexico, the research ethics committee of the National Institute of Psychiatry *Ramón de la Fuente Muñiz* (#CEI/C/038/2018); and in Peru, Universidad Peruana Cayetano Heredia Ethical Committee for Research with Human Subjects (#101460). All study participants provided their informed consent electronically before initiating the Web-based survey. No identification of participants was collected.

Statistical Analysis

Participant characteristics, awareness, willingness to use, barriers, and facilitators of PrEP were described by their frequencies, and chi-square tests were used to compare the variables per country (Brazil, Mexico, and Peru). Some survey questions included options such as *I do not want to answer* or *I do not know* to aid participant comfort. These answers were considered missing for data analysis. We compared the main outcome (willingness to use PrEP) by participant characteristics using chi-square tests for each country. Multivariable logistic regression was performed calculating adjusted odds ratios to explore factors associated with willingness to use PrEP for each country. Age, income, and schooling (and race in Brazil) were defined a priori as confounders and were kept in the final model for each country, irrespective of statistical significance. Models were created using a backwards stepwise modeling approach; variables with a bivariate *P* value of less than .01 were included in the initial model and subsequently excluded if their *P* value was greater than .05. Exclusions were done variable by variable,

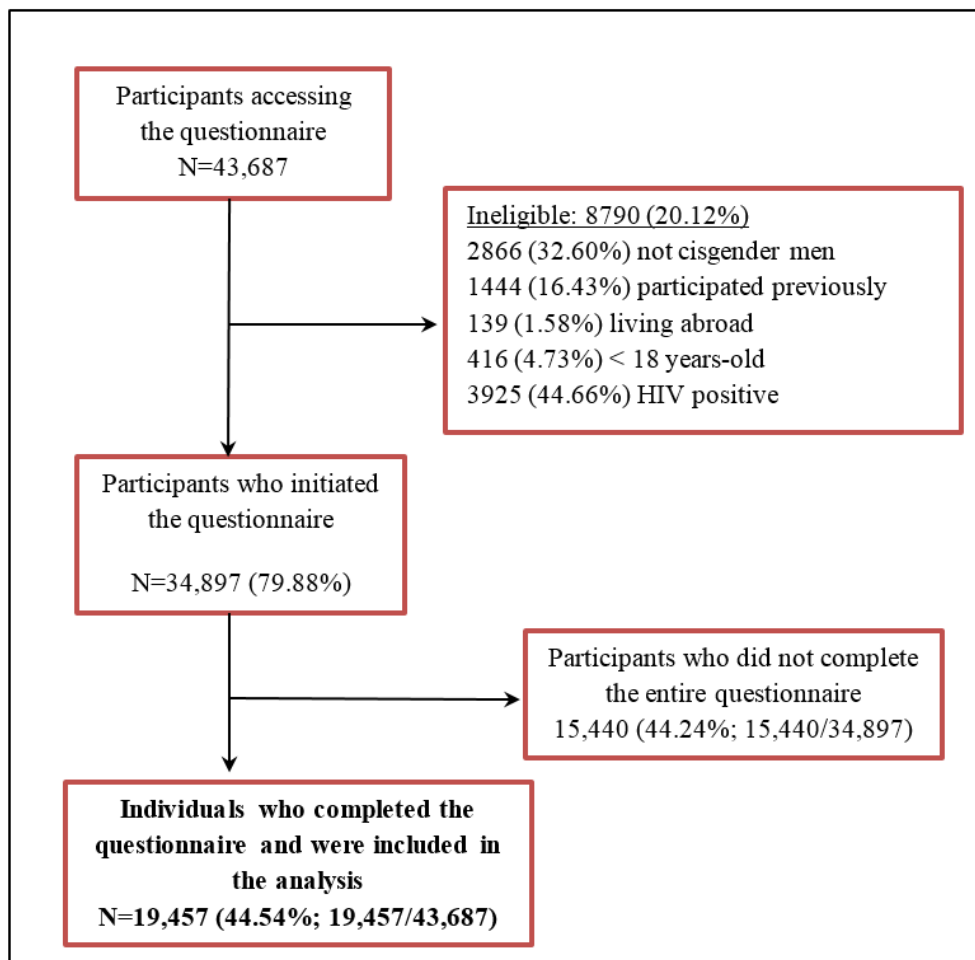
starting by excluding the variable with the highest *P* value and then rerunning the model to repeat the process. The final multivariable models included variables that remained significant (threshold of $P < .05$) after the backwards stepwise process and the a priori defined confounders. All analyses were performed using STATA version 14 (College Station, TX).

Results

During the study, a total of 43,687 individuals provided their informed consent; 20.12% (8790/43,687) were ineligible (Figure 1). Of the 34,897/43,687 (79.88%) eligible individuals, 19,457/43,687 (44.54%) completed the questionnaire and were included in this analysis. Among them, 11,367/19,457 (58.42%) were from Brazil, 5934/19,457 (30.50%) were from Mexico, and 2156/19,457 (11.08%) were from Peru. Almost half of respondents (8981/19,457, 46.17%) lived in the largest urban centers of these countries: São Paulo, Brazil (3198/19,457, 16.44%), Mexico City, Mexico (2618/19,457, 13.46%), Rio de Janeiro, Brazil (1589/19,457, 8.17%), and Lima, Peru (1576/19,457, 8.10%).

Table 1 shows the characteristics of the participants who completed the questionnaire in each country. Their median age was 28 years (interquartile range: 24-34). The proportion of young individuals (18-24 years) was high in Peru (41.23%, 889/2156) and similar in Brazil (28.35%; 3222/11,367) and Mexico (29.76%; 1766/5934). Almost half of respondents from Brazil were white (52.82%, 6004/11,366), whereas the majority from Peru were nonwhite (80.42%, 1672/2079). In Brazil and Peru, fewer individuals were in the high-income category (13.47%, 1531/11,367 and 14.09%; 272/1930, respectively), whereas in Mexico more than a quarter (26.46%, 1419/5363) were in this category. Most participants had at least some postsecondary education in all 3 countries, but this proportion was lower in Brazil (61.14%, 6888/11,266).

Most participants were sexually attracted only to men (89.20%, 17,306/19,401), and the proportion of those attracted to both men and women was highest in Peru (17.13%, 367/2142), followed by Mexico (10.94%, 647/5914) and then Brazil (7.12%, 808/11,348). Among those who were only sexually attracted to women, almost half (45.79%, 125/273) had sex with men in the previous 6 months. Most participants did not have a steady partner (73.57%, 14,195/19,294). Most respondents were recruited on Grindr (68.61%, 13,349/19,457), and this was true for Brazil (67.06%, 7623/11,367) and Mexico (80.91%, 4801/5934), whereas in Peru, the plurality was recruited via Facebook (48.10%, 1037/19,457). Peru has the highest proportion of individuals who never used apps for sexual encounters (19.20%, 414/2156), whereas most Brazilians used them daily (54.05%, 6142/11,363).

Figure 1. Study flowchart. Brazil, Mexico and Peru, 2018.

Overall, 67.51% (13,110/19,419) of respondents would be eligible for PrEP, whereas only 9.80% (1858/18,959) of participants had high HIV risk perception, which was lowest in Brazil (8.94%, 986/11,029). Almost 20% (3862/19,352) had never tested for HIV, and this was highest in Peru (24.47%, 524/2141). A total of 45.66% (8836/19,352) MSM had tested for HIV in the previous 3 or 6 months, and this was highest in Brazil (48.73%, 5505/11,297). Condomless receptive anal sex was reported by 40.13% (7755/19,326) overall, being higher in Peru (44.48%, 954/2145) than in Mexico (40.04%, 2368/5914) and Brazil (39.34%, 4433/11,268). Even with the relatively high proportion of postexposure prophylaxis (PEP) awareness in Brazil (66.07%, 7510/11,367) and Mexico (55.09%, 3269/5934), ever use of PEP was very low in these countries, with only slightly more than 10% (1452/11,326) in Brazil.

Most MSM had heard about PrEP (64.92%, 12,592/19,396; Table 2); however, this proportion was much lower in Peru 46.60% (1002/2150), compared with over 60% in Brazil (68.82, 7794/11,325) and Mexico (64.11; 3796/5921). The majority had never used PrEP (96.19%, 18,685/19,425). Overall, willingness to use PrEP was reported by 64.23%

(12498/19,457); this was highest in Mexico (70.07%, 4158/5934), followed by Brazil (62.45%, 7099/11,367) and Peru (57.56%, 1241/2156). Anticipated risk compensation if taking PrEP was reported by 21.83% (4248/19,457) overall. Injectable PrEP was the preferred PrEP regimen among a plurality of the Brazilians (44.61%, 5071/11,368) and Peruvians (41.00%, 884/2156), whereas daily oral PrEP was preferred among the Mexicans (37.92%, 2260/5960).

All information barriers to PrEP were considered *important* or *very important* by a majority of respondents in each country. There was less importance assigned to the behavioral and belief barriers to PrEP, the barrier *taking pills everyday* was considered *important* or *very important* by 47.90% (9320/19,457) overall, though this was less of a concern in Mexico 42.08% (2497/5934). The majority of respondents in Mexico and Peru were concerned that taking PrEP indicates being at risk of HIV infection and that their partners might expect condomless anal sex. Among the assessed facilitators to PrEP, only 2 were not considered *very important* by the majority of the participants: support for taking daily PrEP from apps or messages and counseling about sex.

Table 1. Characteristics of the individuals who completed the questionnaire in Brazil, Mexico, and Peru (2018).

Characteristics	Brazil (N=11,367; 58.42%)	Mexico (N=5,934; 30.50%)	Peru (N=2,156; 11.08%)	Total (N=19,457)	P value ^a
Age (years; n=19,456), median (interquartile range)	29 (24-35)	28 (24-34)	26 (22-31)	28 (24-34)	N/A^b
18-24	3222 (28.35)	1766 (29.76)	889 (41.23)	5877 (30.21)	<.001
25-35	5364 (47.19)	2991 (50.40)	970 (44.99)	9325 (47.93)	N/A
≥36	2780 (24.46)	1177 (19.83)	297 (13.78)	4254 (21.86)	N/A
Race (nonwhite ^c ; n=13,446), n (%)	5363 (47.18)	N/A	1672 (80.42)	7035 (52.32)	<.001
Monthly income^d (n=18,660), n (%)	N/A	N/A	N/A	N/A	<.001
Low	5136 (45.18)	1544 (28.79)	710 (36.79)	7390 (39.60)	—
Middle	4700 (41.35)	2400 (44.75)	948 (49.12)	7087 (37.98)	N/A
High	1531 (13.47)	1419 (26.46)	272 (14.09)	4183 (22.42)	N/A
Schooling (≤ secondary education; n=19,300), n (%)	4378 (38.86)	1395 (23.60)	461 (21.69)	6234 (32.30)	<.001
Sexual attraction (n=19,401), n (%)	N/A	N/A	N/A	N/A	<.001
Men	10,386 (91.54)	5212 (88.14)	1708 (79.74)	17,306 (89.20)	N/A
Women	152 (1.34)	54 (0.91)	67 (3.13)	273 (1.41)	N/A
Men/women	808 (7.12)	647 (10.94)	367 (17.13)	1822 (9.39)	N/A
Steady partner (yes; n=19,294), n (%)	2916 (25.83)	1576 (26.77)	607 (28.67)	5099 (26.43)	.02
Recruitment (n=19,457), n (%)	N/A	N/A	N/A	N/A	<.001
Grindr	7623 (67.06)	4801 (80.91)	925 (42.90)	13,349 (68.61)	N/A
Hornet	2489 (21.90)	569 (9.59)	20 (0.93)	3078 (15.82)	N/A
Facebook	780 (6.86)	230 (3.88)	1037 (48.10)	2047 (10.52)	N/A
Other	475 (4.18)	334 (5.63)	174 (8.07)	983 (5.05)	N/A
Use of apps for sexual encounters (n=19,454), n (%)	N/A	N/A	N/A	N/A	<.001
Never	835 (7.35)	432 (7.28)	414 (19.20)	1681 (8.64)	N/A
Sometimes	4387 (38.60)	3336 (56.22)	1033 (47.91)	8756 (45.01)	N/A
Daily	6142 (54.05)	2166 (36.50)	709 (32.88)	9017 (46.35)	N/A
Last HIV testing (n=19,352), n (%)	N/A	N/A	N/A	N/A	<.001
Previous 3 months	3516 (31.12)	1428 (21.15)	551 (25.74)	5495 (28.39)	N/A
Previous 6 months	1989 (17.60)	1031 (17.44)	321 (14.99)	3341 (17.26)	N/A
Previous 12 months	1835 (16.24)	843 (14.26)	277 (12.94)	2955 (15.27)	N/A
More than 12 months	1972 (17.45)	1259 (21.29)	468 (21.86)	3699 (19.11)	N/A
Never	1986 (17.58)	1352 (22.86)	524 (24.47)	3862 (19.96)	N/A
HIV risk perception^e (n=18,959), n (%)	N/A	N/A	N/A	N/A	<.001
Low	7746 (70.22)	3341 (57.41)	1218 (57.78)	12,305 (64.90)	N/A
Middle	2299 (20.84)	1845 (31.70)	652 (30.93)	4796 (25.30)	N/A
High	986 (8.94)	634 (10.89)	238 (11.29)	1858 (9.80)	N/A
Preliminary eligibility for PrEP ^f (n=19,419), n (%)	7938 (70.03)	3668 (61.85)	1504 (69.82)	13,110 (67.51)	<.001
Sexual behavior					
Number of male sexual partners^g (n=19,376), n (%)	N/A	N/A	N/A	N/A	<.001
None	873 (7.68)	1401 (23.82)	345 (16.2)	2619 (13.52)	N/A
1-5	6036 (53.11)	2872 (48.84)	1145 (53.78)	10,053 (51.88)	N/A

Characteristics	Brazil (N=11,367; 58.42%)	Mexico (N=5,934; 30.50%)	Peru N=2,156; 11.08%)	Total (N=19,457)	P value ^a
6-10	1958 (17.23)	821 (13.96)	290 (13.62)	3069 (15.84)	N/A
>10	2499 (21.99)	787 (13.38)	349 (16.39)	3635 (18.76)	N/A
Condomless receptive anal sex ^g (yes; n=19,326), n (%)	4433 (39.34)	2368 (40.04)	954 (44.48)	7755 (40.13)	<.001
Condomless insertive anal sex ^g (yes; n=19,322), n (%)	5106 (45.31)	2511 (42.45)	960 (44.90)	8577 (44.39)	<.001
Sex with HIV+ male partner ^g (yes; n=19,270), n (%)	1217 (10.82)	712 (12.08)	269 (12.64)	2198 (11.41)	.002
Sex with unknown HIV status male partner ^g (yes; n=19,133), n (%)	3911 (35.27)	2956 (49.98)	1108 (52.02)	7975 (41.68)	<.001
Sex under alcohol use ^g (yes; n=19,414), n (%)	4190 (36.97)	2089 (35.23)	778 (36.15)	7057 (36.35)	.08
Chemsex ^g (yes; n=19,401), n (%)	1993 (17.60)	948 (16.00)	243 (11.29)	3184 (16.41)	<.001
Transactional sex ^g (yes; n=19,456), n (%)	616 (5.42)	271 (4.57)	155 (7.19)	1042 (5.36)	<.001
STI diagnoses ^{fg,h} (yes; n=18,875), n (%)	1476 (13.25)	372 (6.52)	197 (9.70)	2045 (10.83)	<.001
PEP ⁱ awareness (yes; n=19,457), n (%)	7510 (66.07)	3269 (55.09)	778 (36.09)	11,557 (59.40)	<.001
PEP use (yes; n=19,393), n (%)	1452 (12.82)	310 (5.24)	81 (3.76)	1843 (9.50)	<.001
Binge drinking ^g (yes; n=19,390), n (%)	7774 (68.39)	4094 (69.43)	1518 (71.40)	13,386 (69.04)	<.001
Substance use^g (n=18,906), n (%)	N/A	N/A	N/A	N/A	N/A
Tobacco	2905 (25.56)	1981 (35.87)	539 (26.72)	5425 (28.69)	<.001
Marijuana or hash	2907 (25.57)	1369 (24.79)	471 (23.35)	4747 (25.11)	.09
Stimulants ^j	1622 (14.27)	545 (9.87)	130 (6.45)	2297 (12.15)	<.001
Hallucinogens ^k	994 (8.74)	248 (4.49)	28 (1.39)	1270 (6.72)	<.001
Poppers	735 (6.47)	1233 (22.33)	195 (9.67)	2163 (11.44)	<.001
Erectile dysfunction drug ^g	1054 (9.27)	605 (10.96)	90 (4.46)	1749 (9.25)	<.001

^aChi-square test.

^bN/A: not applicable.

^cBlack, Asian, Native American, and Mix race; this question was not available for Mexican respondents.

^dFor Brazil, we considered the number of minimum wages in the family monthly income: low ≤ 3 , middle 4-10, high >10 (monthly minimum wage in 2018 was 954 BRL=US \$250, currency from June 2018). For Peru, we considered individual monthly income, categorized by number of minimum salaries: low ≤ 3 , middle 4-10, high >10 (monthly minimum wage in 2018 was 850 PEN=US \$265). For Mexico, we considered individual monthly income, categorized by number of minimum salaries: low: from no income to <3 , middle 3-4, high ≥ 5 (monthly minimum wage in 2018 was 2686 MXN=US \$141).

^eIn the next year.

^fAdapted from the WHO criteria for pre-exposure prophylaxis, which included the following: unprotected anal sex with a male or trans partner, sex with an HIV-positive partner, sex work, or STI diagnosis; all in the past 6 months.

^gDuring the previous 6 months.

^hSyphilis, gonorrhea, or rectal chlamydia.

ⁱPEP: postexposure prophylaxis.

^jCocaine, crack, ecstasy, and GHB (4-hydroxybutanoic acid).

^kSolvents, lysergic acid diethylamide, and ketamine.

Table 2. Awareness, willingness to use, barriers, and facilitators to daily oral pre-exposure prophylaxis in Brazil, Mexico, and Peru (2018).

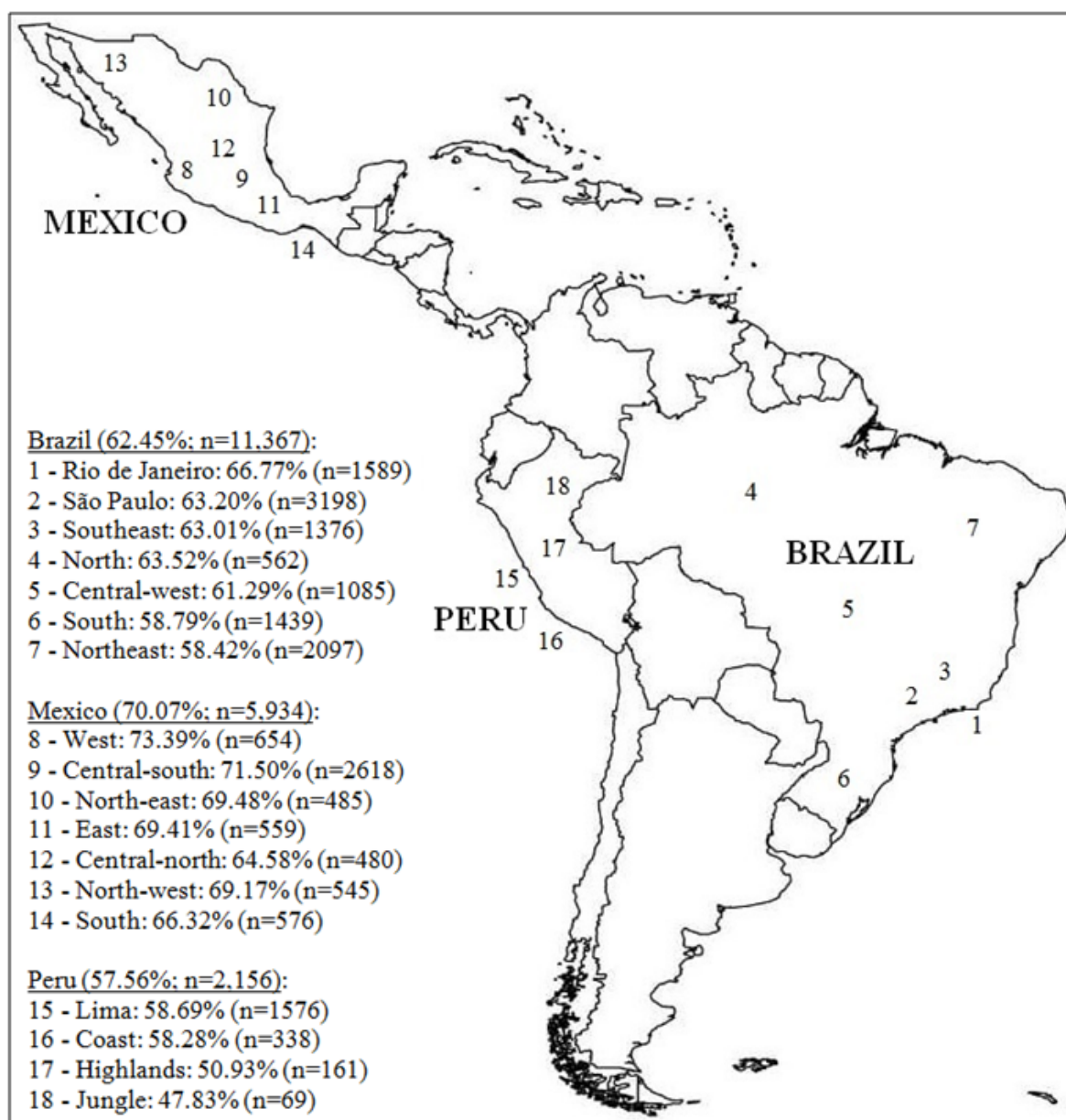
Variables	Brazil (N=11,367; 58.42%), n (%)	Mexico (N=5934; 30.50%), n (%)	Peru (N=2156; 11.08%), n (%)	Total (N=19,457), n (%)	P value ^a
PrEP ^b awareness (yes; n=19,396 ^c)	7794 (68.82)	3796 (64.11)	1002 (46.60)	12,592 (64.92)	<.001
Willingness to use PrEP (yes)	7099 (62.45)	4158 (70.07)	1241 (57.56)	12,498 (64.23)	<.001
PrEP use (n=19,425)					<.001
Never	10,875 (95.67)	5735 (96.97)	2075 (96.78)	18,685 (96.19)	N/A ^d
Current	266 (2.34)	87 (1.47)	42 (1.96)	395 (2.03)	N/A
Past	226 (1.99)	92 (1.56)	27 (1.26)	345 (1.78)	N/A
Anticipated risk compensation (yes)	2325 (20.45)	1380 (23.26)	543 (25.19)	4248 (21.83)	<.001
PrEP if part of Ministry of Health program (yes)	5814 (51.15)	4113 (69.31)	1165 (54.04)	11,092 (57.01)	<.001
Preferred PrEP regimen					<.001
Daily oral	3802 (33.45)	2260 (37.92)	719 (33.35)	6771 (34.80)	N/A
Event-driven	2494 (21.94)	1491 (25.13)	553 (25.65)	4538 (23.32)	N/A
Injectable	5071 (44.61)	2193 (36.96)	884 (41.00)	8148 (41.88)	N/A
Barriers to daily oral PrEP^e					N/A
Information					
Afraid of short-term side effects	6805 (59.87)	4126 (69.53)	1614 (74.86)	12,545 (64.48)	<.001
Afraid of long-term side effects	7601 (66.87)	4601 (77.54)	1721 (79.82)	13,923 (71.56)	<.001
Afraid that antiretroviral therapy would not work if infected	6522 (57.38)	5031 (84.78)	1810 (83.95)	13,363 (68.68)	<.001
Afraid of not being 100% protected against HIV	7831 (68.89)	4900 (82.57)	1794 (83.21)	14,525 (74.65)	<.001
Behaviors					
Taking pills everyday	5774 (50.80)	2497 (42.08)	1049 (48.65)	9320 (47.90)	<.001
Talking to a doctor about sex life	2670 (23.49)	1897 (31.97)	886 (41.09)	5453 (28.03)	<.001
Having HIV/ sexually transmitted infection testing every 3 months	5025 (44.21)	2615 (44.07)	1120 (51.95)	8760 (45.02)	<.001
Beliefs					
Taking PrEP means that I am at risk of HIV infection	4540 (39.94)	3339 (56.27)	1203 (55.80)	9082 (46.68)	<.001
My partners may expect condomless anal sex	4310 (37.92)	3584 (60.40)	1190 (55.19)	9084 (46.69)	<.001
Afraid that people may think I am HIV+	4303 (37.86)	1966 (33.13)	950 (44.06)	7219 (37.10)	<.001
Afraid that people ask me why I am using PrEP when they see me taking the pills	3709 (32.63)	1663 (28.02)	880 (40.82)	6252 (32.13)	<.001
Facilitators to daily oral PrEP^f					N/A
Free PrEP	8797 (77.39)	4317 (72.75)	1474 (68.37)	14,588 (74.98)	<.001
Access to free HIV test	8573 (75.42)	4554 (76.74)	1567 (72.68)	14,694 (75.52)	.001
Access to other free exams	8052 (70.84)	4285 (72.21)	1499 (69.53)	13,836 (71.11)	.04
Access to personal PrEP counseling	7103 (62.49)	4129 (69.58)	1444 (66.98)	12,676 (65.15)	<.001
Support from apps or messages	5464 (48.07)	2880 (48.53)	1160 (53.80)	9504 (48.85)	<.001
Support and counseling about my sex life	4272 (37.58)	3138 (52.88)	1175 (54.50)	8585 (44.12)	<.001

^aChi-square test.^bPrEP: pre-exposure prophylaxis.^cThe full sample, N=19,457, is included unless otherwise specified.^dN/A: not applicable.

^cBarriers were grouped into PrEP information, beliefs, and behaviors. The percent shown is of participants responding that the barrier was *important* or *very important*.

^fThe percent shown is of participants responding that the facilitator was *very important*.

Figure 2. Willingness to use daily oral pre-exposure prophylaxis per region. Brazil, Mexico and Peru, 2018. Percentage of willingness to use daily oral prep is given per region; n is total number of participants per region or country.



The proportion of individuals willing to use PrEP by characteristics and country are provided in the [Multimedia Appendix 1](#). There were geographic differences in willingness to use PrEP (see [Figure 2](#)). In Brazil, willingness to use PrEP varied across the country, being higher in Rio de Janeiro (66.77%, 1061/1589) and São Paulo (63.20%, 2021/3198; $P<.001$). In Mexico, it was higher in the West (73.39%, 480/654) and Central-South (71.50%, 1872/2618), but lower in Central-North (64.58%, 310/480; $P=.002$). There was no statistically significant difference by region in Peru ($P=.09$).

In each country's final multivariable model, PrEP awareness and PrEP facilitators were positively associated, whereas

behavioral and belief barriers were negatively associated with willingness to use PrEP. HIV risk perception was positively associated with willingness to use PrEP in Brazil and Peru, whereas daily use of apps for sexual encounters was positively associated only in Mexico and Peru. HIV testing at least once in a lifetime was positively associated with willingness to use PrEP only in Brazil. Mexico was the only country where age was significantly associated with willingness to use PrEP, with the older age categories (25+ years) reporting significantly less willingness compared with the youngest category (18-24 years). Informational barriers were negatively associated in Brazil but positively associated with willingness to use PrEP in Peru ([Table 3](#)).

Table 3. Factors associated with willingness to use daily oral pre-exposure prophylaxis in Brazil, Mexico, and Peru (2018).

Variable	Brazil		Mexico		Peru	
	Bivariate models, OR ^a (95% CI)	Multivariate model, AOR ^b (95% CI) ^c	Bivariate models, OR (95% CI)	Multivariate model, AOR (95% CI) ^c	Bivariate models, OR (95% CI)	Multivariate model, AOR (95% CI) ^c
Age (years)						
18-24	Ref. ^d	Ref.	Ref.	Ref.	Ref.	Ref.
25-35	1.11 (1.02-1.21)	0.95 (0.86-1.06)	1.06 (0.93-1.21)	0.84 (0.71-0.98)	1.32 (1.10-1.59)	1.06 (0.85-1.34)
≥36	1.13 (1.01-1.25)	0.97 (0.86-1.10)	0.95 (0.81-1.11)	0.81 (0.66-0.99)	1.47 (1.13-1.93)	1.36 (0.97-1.90)
Color (nonwhite ^e vs white)	0.95 (0.88-1.02)	1.01 (0.93-1.10)	N/A ^f	N/A	0.88 (0.70-1.09)	— ^g
Monthly income^h						
Low	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Middle	1.20 (1.11-1.31)	1.08 (0.99-1.19)	1.05 (0.90-1.23)	0.99 (0.84-1.18)	1.37 (1.12-1.66)	1.28 (1.02-1.63)
High	1.31 (1.16-1.47)	1.27 (1.10-1.46)	1.11 (0.96-1.28)	0.92 (0.78-1.11)	1.51 (1.14-2.01)	1.20 (0.84-1.71)
Schooling (any postsecondary education vs ≤ secondary education)	1.15 (1.07-1.25)	0.99 (0.90-1.08)	1.24 (1.09-1.41)	1.07 (0.92-1.26)	1.19 (0.96-1.46)	0.86 (0.66-1.12)
Steady partner (yes vs no)	1.06 (0.97-1.16)	—	1.14 (1.01-1.30)	—	1.05 (0.87-1.27)	—
HIV testing (at least once lifetime vs never)	1.48 (1.34-1.63)	1.12 (1.00-1.25)	1.42 (1.25-1.61)	—	1.36 (1.12-1.66)	—
Use of apps for sexual encounters						
Never	Ref.	—	Ref.	Ref.	Ref.	Ref.
Sometimes	1.06 (0.91-1.23)	—	1.38 (1.13-1.70)	1.15 (0.90-1.47)	1.32 (1.05-1.66)	1.18 (0.91-1.53)
Daily	1.39 (1.20-1.62)	—	1.84 (1.48-2.28)	1.34 (1.04-1.73)	1.65 (1.29-2.11)	1.34 (1.01-1.78)
HIV risk perceptionⁱ						
Low	Ref.	Ref.	Ref.	—	Ref.	Ref.
Middle	1.38 (1.26-1.53)	1.18 (1.06-1.31)	1.34 (1.18-1.51)	—	1.34 (1.11-1.63)	1.32 (1.06-1.63)
High	1.99 (1.71-2.31)	1.53 (1.30-1.80)	1.75 (1.43-2.14)	—	1.83 (1.36-2.46)	1.58 (1.14-2.19)
Preliminary eligibility for PrEP ^j (yes vs no)	1.62 (1.50-1.76)	1.33 (1.22-1.46)	1.47 (1.32-1.65)	1.29 (1.13-1.47)	1.20 (1.00-1.44)	—
Number of male sexual partners ^k (>5 vs ≤5)	1.55 (1.44-1.68)	1.19 (1.09-1.30)	1.60 (1.40-1.83)	1.18 (1.01-1.38)	1.42 (1.17-1.71)	—
Sex under alcohol use ^k (yes vs no)	1.16 (1.07-1.26)	—	1.17 (1.03-1.33)	—	1.08 (0.89-1.30)	—
Chemsex ^k (yes vs no)	1.24 (1.12-1.38)	—	1.48 (1.25-1.76)	—	1.44 (1.09-1.92)	—
PrEP awareness (yes vs no)	1.98 (1.82-2.14)	1.66 (1.52-1.81)	2.37 (2.11-2.66)	2.05 (1.79-2.34)	2.03 (1.70-2.42)	1.84 (1.51-2.26)
Anticipated risk compensation (yes vs no)	1.48 (1.34-1.63)	1.32 (1.18-1.47)	1.37 (1.19-1.57)	1.37 (1.17-1.59)	1.24 (1.02-1.52)	—
Barriers: Information, mean (SD)	0.94 (0.93-0.95)	0.95 (0.94-0.97)	0.97 (0.96-0.99)	—	1.03 (1.01-1.06)	1.05 (1.01-1.08)
Barriers: Behaviors, mean (SD)	0.93 (0.92-0.94)	0.95 (0.94-0.96)	0.92 (0.90-0.93)	0.94 (0.92-0.95)	0.95 (0.93-0.97)	0.95 (0.92-0.98)
Barriers: Believes, mean (SD)	0.95 (0.94-0.95)	0.97 (0.96-0.98)	0.93 (0.91-0.94)	0.95 (0.93-0.97)	0.95 (0.93-0.97)	0.95 (0.92-0.97)
Facilitators, mean (SD)	1.07 (1.06-1.08)	1.09 (1.08-1.10)	1.07 (1.05-1.08)	1.11 (1.09-1.12)	1.07 (1.05-1.09)	1.11 (1.08-1.14)

^aOR: odds ratio.^bAOR: adjusted odds ratio.^cVariables with $P < .01$ in bivariate models were included in the initial multivariable model. Variables with $P < .05$ were kept in the final multivariable

models, excepted for age, monthly income, and schooling defined a priori for all countries, and race only for Brazil; statistically significant associations at $P < .05$ in italics. Region did not remain in the final multivariate models and bivariate analysis is not shown.

^dRef.: reference.

^eBlack, Asian, Native American, or Mix race.

^fN/A: not applicable.

^gNot statistically significant.

^hFor Brazil, we considered the number of minimum wages in the family monthly income: low ≤ 3 , middle 4-10, high > 10 (monthly minimum wage in 2018 was 954 BRL=US \$250, currency from June 2018). For Peru, we considered individual monthly income, categorized by number of minimum salaries: low ≤ 3 , middle 4-10, high > 10 (monthly minimum wage in 2018 was 850 PEN=US \$265). For Mexico, we considered individual monthly income, categorized by number of minimum salaries: low, from no income to < 3 , middle 3-4, high ≥ 5 (monthly minimum wage in 2018 was 2686 MXN=US \$141).

ⁱIn the next 12 months.

^jAdapted from the WHO criteria for pre-exposure prophylaxis, which included the following: unprotected anal sex with a male or trans partner, sex with an HIV-positive partner, sex work, or STI diagnosis; all in the past 6 months.

^kDuring the previous 6 months.

Discussion

This study shows that MSM from Brazil, Mexico, and Peru are willing to use daily oral PrEP. Willingness to use PrEP was higher in Mexico, followed by Brazil and then Peru. There were similarities in the factors associated with willingness to use PrEP across the 3 countries, and the most important factor was PrEP awareness. It is likely that as PrEP awareness increases in these settings, willingness to use PrEP will also increase. This trend was observed in Brazil [20,21] and in a recent meta-analysis evaluating the factors associated with willingness to use PrEP in low- and middle-income settings [17].

In our models, younger age was significantly associated with willingness to use PrEP in Mexico but not in Brazil and Peru. This is a concern as younger MSM are at high risk of HIV acquisition in Latin America [28] and are experiencing an increase in HIV cases in Brazil [7], Mexico [29], and Peru [30]. The association of income with willingness was different by country; higher income was associated with willingness in Peru and Brazil, whereas there was no association in Mexico. To improve equity in PrEP access, interventions to increase knowledge of PrEP among lower-income and young MSM are essential, and this could be achieved with community-based educators and Web-based advertisements, as internet access is becoming more available in Latin America [31]. Nevertheless, this finding should be interpreted with caution as respondents with lower income may associate their willingness to use PrEP with accessibility to PrEP free of charge within the public health system. Of note, more Mexican MSM reported interest in PrEP as part of a MoH program (free of charge) than Brazilians or Peruvians. In addition, almost 75% of the entire sample thought that PrEP at no cost is *very important* or *important*.

Our regression models captured that informational barriers were significantly associated with willingness to use PrEP in Peru. This unexpected result is likely owing to the limited available information, reflected by the low observed PrEP awareness. In all 3 countries, higher HIV risk perception and higher risky sexual behavior were associated with willingness to use PrEP, reflecting that the individuals to whom PrEP is targeted are those willing to use it. This is corroborated by other studies accessing willingness to use PrEP [21,22,32]. Anticipated risk compensation (the possibility of not using condoms while on PrEP) was associated with willingness in Brazil and Mexico

but not in Peru. Risk compensation is a recurrent concern related to PrEP, and evidence of this was recently observed in other studies [33,34]; conversely, there was no statistically significant increase in condomless receptive anal sex during the PrEP Brazil study [35] or in a study conducted in Thailand [36]. Risk compensation continues to require monitoring and adequate PrEP education.

Notably in Brazil and Peru, the majority of participants would prefer injectable PrEP, and for Mexico, it was the second preferred regimen, just slightly below daily oral PrEP. Event-driven PrEP was third in all 3 countries. Injectable PrEP regimens, currently under study, may be useful and acceptable for MSM who have shown low adherence for oral PrEP regimens. As new PrEP technologies become available, it will be important to continue providing education and collecting information on preferences.

The samples included from each country were distinct; most likely, this was related to the recruitment methods. The fact that a higher proportion of lower-income MSM was included in Brazil could be explained by the increase of internet access in this country, where it is estimated 1 smartphone per person [31]. As expected, most of respondents are from each country's urban large regions where more MSM are concentrated, as well as where access to internet is easier. Brazilian MSM reported more male sexual partners than Mexican and Peruvians but lower rates of steady partner and condomless receptive anal sex. In a recent, large sample of MSM in Brazil, having more sex partners was associated with condomless receptive anal sex among MSM aged ≥ 25 years but not among those aged 18 to 24 years [37]. The higher proportions of STI diagnoses and PEP use in Brazil could be explained by the higher availability and access to exams and PEP free of charge in this country, especially in the Southeast region where most respondents live. Peruvian MSM reported more condomless receptive anal sex, transactional sex, binge drinking, and more had never tested for HIV than in the other countries.

Some study limitations should be considered. First, Web-based studies are not probabilistic sampling strategies, precluding the generalization of the findings. Given the cross-sectional nature of the data, causality and the direction of association may not be inferred. All collected data were self-reported by participants and may be subject to bias. Our data are subject to recall bias

owing to 6-month or 12-month recall periods. There is also a concern about participants taking the survey multiple times. To mitigate this bias, the first question of the survey was as follows: *Are you answering this survey for the first time?* (3.3% of participants answered *no* and were excluded from the study). Finally, we have no ability to assess the veracity of respondents truly being MSM; the analysis included self-identified MSM independent of their sexually activity in the previous 6 months or their self-identified sexual attraction. There were respondents who reported being sexually attracted only to women—though importantly, among this group, 45.8% (125/273) reported having

sex with a man in the past 6 months. We believe the population of interest were reached by and responded to this survey.

Conclusions

In this first cross-country, Web-based survey in Latin America willingness to use PrEP was found to be high and directly related to levels of awareness. Interventions to increase awareness and PrEP knowledge about safety and efficacy are crucial to increase PrEP demand among MSM at high risk who can benefit from this prevention technology. Social media campaigns and campaigns in STI clinics may be valuable to achieve this goal. This study presents important information to support the implementation of PrEP for MSM in Brazil, Mexico, and Peru.

Acknowledgments

This project was made possible, thanks to Unitaid's funding and support. Unitaid accelerates access to innovative health products and lays the foundations for their scale-up by countries and partners. Unitaid is a hosted partnership of the WHO. We would like to thank the MoH of Brazil, Mexico, and Peru.

Authors' Contributions

TST, KAK, and EHVR did the literature search and developed the survey. TST, KAK, EHVR, and OAE analyzed the data and generated the tables and figures. TST, KAK, and EHVR interpreted the data. TST, KAK, EHVR, and DDS drafted the manuscript. BH, SD, CP, MB, HLG, RRG, BG, CC, and VGV revised the manuscript and contributed intellectually.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Proportion of individuals willing to use pre-exposure prophylaxis by characteristics and country.

[DOCX File, 49KB - [publichealth_v5i2e13771_app1.docx](#)]

References

1. UNAIDS. 2017. Ending AIDS: Progress towards the 90-90-90 targets URL: http://www.unaids.org/en/resources/documents/2017/20170720_Global_AIDS_update_2017 [accessed 2019-05-27] [WebCite Cache ID 76Krr7M59]
2. De Boni R, Veloso VG, Grinsztejn B. Epidemiology of HIV in Latin America and the Caribbean. *Curr Opin HIV AIDS* 2014 Mar;9(2):192-198. [doi: [10.1097/COH.0000000000000031](https://doi.org/10.1097/COH.0000000000000031)] [Medline: [24356327](https://pubmed.ncbi.nlm.nih.gov/24356327/)]
3. Kerr L, Kendall C, Guimaraes M, Salani MR, Veras M, Dourado I. HIV prevalence among men who have sex with men in Brazil: results of the 2nd national survey using respondent-driven sampling. *Medicine (Baltimore)* 2018;97(1S Suppl 1):S9-15.
4. Bautista-Arredondo S, Colchero MA, Romero M, Conde-Glez CJ, Sosa-Rubí SG. Is the HIV epidemic stable among MSM in Mexico? HIV prevalence and risk behavior results from a nationally representative survey among men who have sex with men. *PLoS One* 2013;8(9):e72616 [FREE Full text] [doi: [10.1371/journal.pone.0072616](https://doi.org/10.1371/journal.pone.0072616)] [Medline: [24039786](https://pubmed.ncbi.nlm.nih.gov/24039786/)]
5. Clark J, Konda K, Silva-Santisteban A, Peinado J, Lama J, Kusunoki L, et al. Sampling methodologies for epidemiologic surveillance of men who have sex with men and transgender women in Latin America: an empiric comparison of convenience sampling, time space sampling, and respondent driven sampling. *AIDS Behav* 2014 Dec;18(12):2338-2348 [FREE Full text] [doi: [10.1007/s10461-013-0680-0](https://doi.org/10.1007/s10461-013-0680-0)] [Medline: [24362754](https://pubmed.ncbi.nlm.nih.gov/24362754/)]
6. Ormaeche M, Whittombury A, Pun M, Suárez-Ognio L. Hepatitis B virus, syphilis, and HIV seroprevalence in pregnant women and their male partners from six indigenous populations of the Peruvian Amazon Basin, 2007-2008. *Int J Infect Dis* 2012 Oct;16(10):e724-e730 [FREE Full text] [doi: [10.1016/j.ijid.2012.05.1032](https://doi.org/10.1016/j.ijid.2012.05.1032)] [Medline: [22884008](https://pubmed.ncbi.nlm.nih.gov/22884008/)]
7. Department of Surveillance, Prevention and Control of Sexually Transmitted Infections, HIV / AIDS and Viral Hepatitis. *Boletim epidemiológico HIV/Aids* 2018 URL: <http://www.aids.gov.br/pt-br/pub/2018/boletim-epidemiologico-hiv-aids-2018> [accessed 2019-05-27] [WebCite Cache ID 78gXQ0xt9]
8. UNAIDS. 2019. Factsheet Peru 2017 URL: <http://www.unaids.org/es/regionscountries/countries/peru> [accessed 2019-05-29] [WebCite Cache ID 76Krr7M59]

9. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, iPrEx Study Team. 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/)]
10. Grant R, Anderson P, McMahan V, Liu A, Amico K, Mehrotra M, iPrEx study team. Uptake of pre-exposure prophylaxis, sexual practices, and HIV incidence in men and transgender women who have sex with men: a cohort study. *Lancet Infect Dis* 2014 Sep;14(9):820-829 [FREE Full text] [doi: [10.1016/S1473-3099\(14\)70847-3](https://doi.org/10.1016/S1473-3099(14)70847-3)] [Medline: [25065857](https://pubmed.ncbi.nlm.nih.gov/25065857/)]
11. McCormack S, Dunn D, Desai M, Dolling D, Gafos M, Gilson R, et al. Pre-exposure prophylaxis to prevent the acquisition of HIV-1 infection (PROUD): effectiveness results from the pilot phase of a pragmatic open-label randomised trial. *Lancet* 2016 Jan 2;387(10013):53-60 [FREE Full text] [doi: [10.1016/S0140-6736\(15\)00056-2](https://doi.org/10.1016/S0140-6736(15)00056-2)] [Medline: [26364263](https://pubmed.ncbi.nlm.nih.gov/26364263/)]
12. Molina J, Capitant C, Spire B, Pialoux G, Cotte L, Charreau I, et al. On-demand preexposure prophylaxis in men at high risk for HIV-1 infection. *N Engl J Med* 2015 Dec 3;373(23):2237-2246. [doi: [10.1056/nejmoa1506273](https://doi.org/10.1056/nejmoa1506273)]
13. Centers for Disease Control and Prevention. 2014. Preexposure prophylaxis for the prevention of HIV Infection in the United States - clinical practice guideline 30 URL: <https://www.cdc.gov/hiv/pdf/risk/prep/cdc-hiv-prep-guidelines-2017.pdf> [accessed 2019-05-29] [WebCite Cache ID 76KvjBaci]
14. World Health Organization. 2016. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: Recommendations for a public health approach 2nd Edition URL: <https://www.who.int/hiv/pub/arv/arv-2016/en/> [accessed 2019-05-27] [WebCite Cache ID 78gXaBDmM]
15. AIDS.gov. 2017. Protocolo Clínico e Diretrizes terapêuticas para Profilaxia pré-exposição (PrEP) de risco à infecção pelo HIV URL: <http://www.aids.gov.br/pt-br/pub/2017/protocolo-clinico-e-diretrizes-terapeuticas-para-profilaxia-pre-exposicao-prep-de-risco> [accessed 2019-05-27] [WebCite Cache ID 78gXbULLH]
16. World Health Organization. 2017. Brazil begins PrEP roll-out on World AIDS Day URL: <https://www.who.int/hiv/mediacentre/news/brazil-prep/en/> [WebCite Cache ID 76KwxueZT]
17. Yi S, Tuot S, Mwai G, Ngin C, Chhim K, Pal K, et al. Awareness and willingness to use HIV pre-exposure prophylaxis among men who have sex with men in low- and middle-income countries: a systematic review and meta-analysis. *J Int AIDS Soc* 2017 Jun 26;20(1):21580 [FREE Full text] [doi: [10.7448/IAS.20.1.21580](https://doi.org/10.7448/IAS.20.1.21580)] [Medline: [28691439](https://pubmed.ncbi.nlm.nih.gov/28691439/)]
18. Patrick R, Forrest D, Cardenas G, Opoku J, Magnus M, Phillips G, et al. Awareness, willingness, and use of pre-exposure prophylaxis among men who have sex with men in Washington, DC and Miami-Dade County, FL: National HIV Behavioral Surveillance, 2011 and 2014. *J Acquir Immune Defic Syndr* 2017 Dec 1;75(Suppl 3):S375-S382 [FREE Full text] [doi: [10.1097/QAL.0000000000001414](https://doi.org/10.1097/QAL.0000000000001414)] [Medline: [28604442](https://pubmed.ncbi.nlm.nih.gov/28604442/)]
19. Bauermeister J, Meanley S, Pingel E, Soler J, Harper G. PrEP awareness and perceived barriers among single young men who have sex with men. *Curr HIV Res* 2013 Oct;11(7):520-527 [FREE Full text] [Medline: [24476355](https://pubmed.ncbi.nlm.nih.gov/24476355/)]
20. Torres TS, Luz PM, de Boni RB, de Vasconcellos MT, Hoagland B, Garner A, et al. Factors associated with PrEP awareness according to age and willingness to use HIV prevention technologies: the 2017 online survey among MSM in Brazil. *AIDS Care* 2019 May 23;1-10. [doi: [10.1080/09540121.2019.1619665](https://doi.org/10.1080/09540121.2019.1619665)] [Medline: [31122033](https://pubmed.ncbi.nlm.nih.gov/31122033/)]
21. Torres TS, de Boni RB, de Vasconcellos MT, Luz P, Hoagland B, Moreira R, et al. Awareness of prevention strategies and willingness to use preexposure prophylaxis in Brazilian men who have sex with men using apps for sexual encounters: online cross-sectional study. *JMIR Public Health Surveill* 2018 Jan 22;4(1):e11 [FREE Full text] [doi: [10.2196/publichealth.8997](https://doi.org/10.2196/publichealth.8997)] [Medline: [29358160](https://pubmed.ncbi.nlm.nih.gov/29358160/)]
22. Hoagland B, de boni R, Moreira R, Madruga J, Kallas E, Goulart S. Awareness and willingness to use pre-exposure prophylaxis (PrEP) among men who have sex with men and transgender women in Brazil. *AIDS Behav* 2017;21(5):1278-1287. [Medline: [27531461](https://pubmed.ncbi.nlm.nih.gov/27531461/)]
23. Imprep. URL: <http://imprep.org/> [accessed 2019-05-28] [WebCite Cache ID 78hoRAVGi]
24. World Health Organization. 2017. WHO Implementation tool for the pre-exposure prophylaxis (PrEP) of HIV infection URL: <https://www.who.int/hiv/pub/prep/prep-implementation-tool/en/> [WebCite Cache ID 76KucCCie]
25. National Institute on Alcohol Abuse and Alcoholism. 2004. NIAAA council approves definition of binge drinking URL: https://pubs.niaaa.nih.gov/publications/Newsletter/winter2004/Newsletter_Number3.pdf [WebCite Cache ID 76KvJN5Cu]
26. Prochaska JO, DiClemente CC, Velicer WF, Rossi JS. Standardized, individualized, interactive, and personalized self-help programs for smoking cessation. *Health Psychol* 1993 Sep;12(5):399-405. [Medline: [8223364](https://pubmed.ncbi.nlm.nih.gov/8223364/)]
27. Golub SA, Gamarel KE, Rendina HJ, Surace A, Lelutiu-Weinberger CL. From efficacy to effectiveness: facilitators and barriers to PrEP acceptability and motivations for adherence among MSM and transgender women in New York City. *AIDS Patient Care STDS* 2013 Apr;27(4):248-254 [FREE Full text] [doi: [10.1089/apc.2012.0419](https://doi.org/10.1089/apc.2012.0419)] [Medline: [23565928](https://pubmed.ncbi.nlm.nih.gov/23565928/)]
28. UNAIDS. 2016. Prevention gap report URL: http://www.unaids.org/sites/default/files/media_asset/2016-prevention-gap-report_en.pdf [WebCite Cache ID 76Kw6qo3i]
29. UNAIDS. [National report on the monitoring of commitments and expanded goals to end AIDS (GAM Report)] URL: http://www.unaids.org/sites/default/files/country/documents/MEX_2018_countryreport.pdf [WebCite Cache ID 76KwOd65q]
30. LIBROS. 2018. [Final Evaluation Report of the Multisectoral Strategic Plan for HIV / STIs of Peru (2015-2019)] URL: http://www.iessdeh.org/usuario/pem/Propuesta_de_PEM_VIH_2018-2022_para_consulta_a_CONAMUSA-16-19-18.pdf [WebCite Cache ID 76Kwi2STF]

31. Cetic.br. 2016. ICT Households 2015: survey on the use of information and communication technologies in Brazilian Households URL: https://cetic.br/media/docs/publicacoes/2/TIC_Dom_2015_LIVRO_ELETRONICO.pdf [WebCite Cache ID 76KwFRWoR]
32. Hoots BE, Finlayson T, Nerlander L, Paz-Bailey G, National HIV Behavioral Surveillance Study Group. Willingness to take, use of, and indications for pre-exposure prophylaxis among men who have sex with men—20 US cities, 2014. *Clin Infect Dis* 2016 Sep 1;63(5):672-677 [FREE Full text] [doi: [10.1093/cid/ciw367](https://doi.org/10.1093/cid/ciw367)] [Medline: [27282710](https://pubmed.ncbi.nlm.nih.gov/27282710/)]
33. Oldenburg CE, Nunn AS, Montgomery M, Almonte A, Mena L, Patel RR, et al. Behavioral changes following uptake of HIV pre-exposure prophylaxis among men who have sex with men in a clinical setting. *AIDS Behav* 2018 Dec;22(4):1075-1079 [FREE Full text] [doi: [10.1007/s10461-017-1701-1](https://doi.org/10.1007/s10461-017-1701-1)] [Medline: [28150120](https://pubmed.ncbi.nlm.nih.gov/28150120/)]
34. Newcomb M, Moran K, Feinstein B, Forscher E, Mustanski B. Pre-exposure prophylaxis (PrEP) use and condomless anal sex: evidence of risk compensation in a cohort of young men who have sex with men. *J Acquir Immune Defic Syndr* 2018 Dec 1;77(4):358-364 [FREE Full text] [doi: [10.1097/QAI.0000000000001604](https://doi.org/10.1097/QAI.0000000000001604)] [Medline: [29210834](https://pubmed.ncbi.nlm.nih.gov/29210834/)]
35. Grinsztejn B, Hoagland B, Moreira R, Kallas E, Madruga J, Goulart S, PrEP Brasil Study Team. Retention, engagement, and adherence to pre-exposure prophylaxis for men who have sex with men and transgender women in PrEP Brasil: 48 week results of a demonstration study. *Lancet HIV* 2018 Dec;5(3):e136-e145. [doi: [10.1016/S2352-3018\(18\)30008-0](https://doi.org/10.1016/S2352-3018(18)30008-0)] [Medline: [29467098](https://pubmed.ncbi.nlm.nih.gov/29467098/)]
36. Colby D, Crowell T, Sirivichayakul S, Pinyakorn S, Kroon E, Benjapornpong K, Search 010/RV 254 Study Group. Declining trend in transmitted drug resistance detected in a prospective cohort study of acute HIV infection in Bangkok, Thailand. *J Int AIDS Soc* 2016;19(1):20966 [FREE Full text] [doi: [10.7448/IAS.19.1.20966](https://doi.org/10.7448/IAS.19.1.20966)] [Medline: [27802846](https://pubmed.ncbi.nlm.nih.gov/27802846/)]
37. Rocha GM, Guimarães MD, de Brito AM, Dourado I, Veras MA, Magno L, Brazilian HIV/MSM Group. High rates of unprotected receptive anal intercourse and their correlates among young and older MSM in Brazil. *AIDS Behav* 2019 Mar 16. [doi: [10.1007/s10461-019-02459-y](https://doi.org/10.1007/s10461-019-02459-y)] [Medline: [30879210](https://pubmed.ncbi.nlm.nih.gov/30879210/)]

Abbreviations

FTC/TDF: emtricitabine/tenofovir
ImPrEP: implementation PrEP
MoH: Ministry of Health
MSM: men who have sex with men
PEP: postexposure prophylaxis
PrEP: pre-exposure prophylaxis
STI: sexually transmitted infection
WHO: World Health Organization

Edited by G Eysenbach; submitted 21.02.19; peer-reviewed by G Ravasi, A Nunn, S Yi, B Castonguay; comments to author 20.03.19; revised version received 15.05.19; accepted 16.05.19; published 17.06.19.

Please cite as:

Torres TS, Konda KA, Vega-Ramirez EH, Elorreaga OA, Diaz-Sosa D, Hoagland B, Diaz S, Pimenta C, Benedetti M, Lopez-Gatell H, Robles-Garcia R, Grinsztejn B, Caceres C, Veloso VG, ImPrEP Study Group

Factors Associated With Willingness to Use Pre-Exposure Prophylaxis in Brazil, Mexico, and Peru: Web-Based Survey Among Men Who Have Sex With Men

JMIR Public Health Surveill 2019;5(2):e13771

URL: <http://publichealth.jmir.org/2019/2/e13771/>

doi:[10.2196/13771](https://doi.org/10.2196/13771)

PMID:

©Thiago Silva Torres, Kelika A Konda, E Hamid Vega-Ramirez, Oliver A Elorreaga, Dulce Diaz-Sosa, Brenda Hoagland, Steven Diaz, Cristina Pimenta, Marcos Benedetti, Hugo Lopez-Gatell, Rebeca Robles-Garcia, Beatriz Grinsztejn, Carlos Caceres, Valdilea G Veloso, ImPrEP Study Group. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 17.06.2019. 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

Continuous Glucose Monitoring in the Real World Using Photosurveillance of #Dexcom on Instagram: Exploratory Mixed Methods Study

Michelle L Litchman¹, PhD, FNP-BC, FAANP; Sarah E Wawrzynski^{1,2}, BSN, RN; Whitney S Woodruff¹, DNP, RN; Joseph B Arrington³, BS, MBA; Quynh C Nguyen⁴, PhD; Perry M Gee^{1,2}, RN, PhD

¹College of Nursing, University of Utah, Salt Lake City, UT, United States

²Intermountain Healthcare, Salt Lake City, UT, United States

³Department of Biology, University of Utah, Salt Lake City, UT, United States

⁴Department of Epidemiology and Biostatistics, School of Public Health, University of Maryland, College Park, MD, United States

Corresponding Author:

Michelle L Litchman, PhD, FNP-BC, FAANP

College of Nursing

University of Utah

10 2000 East

Salt Lake City, UT, 84112

United States

Phone: 1 8015859612

Email: michelle.litchman@nurs.utah.edu

Abstract

Background: Individuals with diabetes are using social media as a method to share and gather information about their health via the diabetes online community. Infoveillance is one methodological approach to examine health care trends. However, infoveillance, while very effective in identifying many real-world health trends, may miss opportunities that use photographs as primary sources for data. We propose a new methodology, photosurveillance, in which photographs are analyzed to examine real-world trends.

Objective: The purpose of this research is to (1) assess the use of photosurveillance as a research method to examine real-world trends in diabetes and (2) report on real-world use of continuous glucose monitoring (CGM) on Instagram.

Methods: This exploratory mixed methods study examined all photographs posted on Instagram that were identified with the hashtag #dexcom over a 3-month period—December 2016 to February 2017. Photographs were coded by CGM location on the body. Original posts and corresponding comments were textually coded for length of CGM device wear and CGM failure and were analyzed for emerging themes.

Results: A total of 2923 photographs were manually screened; 12.08% (353/2923) depicted a photograph with a CGM site location. The majority (225/353, 63.7%) of the photographs showed a CGM site in an off-label location, while 26.2% (92/353) were in an FDA-approved location (ie, abdomen) and 10.2% (36/353) were in an unidentifiable location. There were no significant differences in the number of likes or comments based on US Food and Drug Administration (FDA) approval. Five themes emerged from the analysis of original posts (N=353) and corresponding comments (N=2364): (1) endorsement of CGM as providing a sense of well-being; (2) reciprocating information, encouragement, and support; (3) reciprocating CGM-related frustrations; (4) life hacks to optimize CGM use; and (5) sharing and learning about off-label CGM activity.

Conclusions: Our results indicate that individuals successfully used CGM in off-label locations, posting photos of these areas with greater frequency than of the abdomen, with no indication of sensor failure. While these photographs only capture a snapshot in time, these posts can be used to inform providers and industry leaders of real-world trends in CGM use. Additionally, there were instances in which sensors were worn beyond the FDA-approved 7-day period; however, they represented the minority in this study.

(*JMIR Public Health Surveill* 2019;5(2):e11024) doi:[10.2196/11024](https://doi.org/10.2196/11024)

KEYWORDS

diabetes; continuous glucose monitoring; off-label use; social media; Instagram; photosurveillance

Introduction

Individuals with diabetes are using social media as a method to share and gather information about their health via the diabetes online community. The diabetes online community is a grassroots collection of stakeholders affected by diabetes, including people with diabetes, caregivers (ie, parents of children with diabetes), health care providers, researchers, and industry who use Internet resources (ie, forums, Facebook, and Twitter) to discuss health-related issues [1-3]. Through shared experiences with virtual peers [2-7], diabetes online community users experience a sense of normalcy. As such, individuals are sharing information about a variety of aspects related to diabetes, including diabetes treatment options such as diabetes technology [2]. One social media site used within the diabetes online community that supports photo sharing is Instagram. If a picture is worth a thousand words, a diabetes-related photo posted on Instagram is worth a million *likes*. Diabetes online community users can optimize diabetes-related conversations on Instagram through the use of hashtags.

One technology that individuals with diabetes use to support self-care is real-time continuous glucose monitoring (CGM). CGM is comprised of a small sensor placed into the subcutaneous tissue, a transmitter, and a receiver or mobile phone that displays glucose levels and trends in real time. One of the two companies that supply CGM monitors for patient use is Dexcom. Currently, the Dexcom CGM system is approved by the US Food and Drug Administration (FDA) to be worn on the abdomen only for adults, and abdomen or buttocks for children. Further, the Dexcom G4 and G5 CGM systems are FDA approved to be worn for 7 days. Through clinical observation and anecdotal reports, patients are wearing CGM devices in off-label sites and extending the wear beyond 7 days. However, we do not know the extent to which CGM users are using off-label site locations or for how long.

Infoveillance is one methodological approach to examine health care trends [8,9]. This approach, which overlays Twitter data and geographic data, has been successful in identifying infectious diseases, such as influenza [10,11] and Zika [12], and suicide in veterans [13,14]. Infoveillance, while very effective in identifying many real-world health trends, may miss opportunities that use photographs as primary sources for data. We propose a new methodology, photosurveillance, in which photographs are analyzed to examine real-world trends. Photographs represent microreports of events of day-to-day life [15], such as diabetes management. Since some CGM users are sharing photographs of their diabetes experiences online, analyzing Instagram using a photosurveillance approach can help observe real-world use of CGM, including off-label activity. Off-label CGM activity may be difficult to track using other data collection methods, such as surveys or interviews, due to response bias. Thus, photosurveillance provides an opportunity to identify how individuals are using CGM in the real world without fear of disapproval from health care providers and researchers due to nonadherence to FDA guidelines. The

purpose of this exploratory research is to (1) assess the use of photosurveillance as a research method to examine real-world trends in diabetes, and (2) report on real-world use of continuous glucose monitoring on Instagram. We anticipate that we will be able to successfully gather information about CGM use using photosurveillance. We hypothesize that there will be more off-label site locations compared to FDA-approved locations and that CGM wear greater than 7 days will be observed.

Methods

Dataset Acquisition and Sampling

This study was acknowledged as exempt by the University of Utah, Salt Lake City, Utah, ethics board. This descriptive study was conducted by hand-searching all photographs (N=2923) posted on Instagram identified with the hashtag #dexcom. A 3-month period was chosen, from December 2016 to February 2017, in order to keep the time-intensive process of hand-searching and coding data manageable, while allowing for an adequate sample size. During the search period, Dexcom was approved for 7-day wear. The hashtag #dexcom was chosen because it was more commonly used (N=46,105) when compared to other CGM-related hashtags—#dexcomg5, N=8350; #dexcomg4, N=6135; #dexcomcgm, N=1105; #enlitesensor, N=956; and #medtronicCGM, N=195. Instagram users include hashtags on their posts for searchability purposes. Additionally, #dexcom was frequently accompanied with other Dexcom-related hashtags, such as those just mentioned, resulting in duplication of posts. Thus, only one hashtag, #dexcom, was used in this analysis. Photographs were included for analysis if they depicted a CGM site and the original post was written in English (N=353). Photographs were excluded if they included a CGM site but were initiated from a company advertising their product (ie, adhesive decal).

Photograph Analysis

Photosurveillance was conducted by categorizing photographs by site location, including FDA-approved sites (ie, in adults, this includes the abdomen only), off-label sites (ie, in adults, this includes the posterior arm, anterior arm, forearm, back, buttocks, thigh, and calf), and *other* sites. The *other* category represented photos of a CGM device on skin without clear body landmarks (ie, hand, foot, or belly button) that identified the exact location. Based on photograph content, we were not able to code for demographic factors, including age, gender, or diabetes type. Therefore, photographs of children and adults were not dichotomized, despite having different FDA approval for site locations. There were no duplicate photos. Two independent researchers (JA and WSW) input data into REDCap, a Web-based data capture program, to organize data [16]. Text from the original posts were analyzed to support data input. For example, in some instances, the photo would state that the CGM device was in a specific location, even if it was minutely visible. There were not instances in which photographs were recoded based on text content. A content analysis approach was utilized in which counts were categorized. Frequencies

were used to describe the differences between FDA-approved and off-label activity.

Post and Comment Analysis

Original posts (N=353) and corresponding comments (N=2364) were analyzed to examine discussions related to CGM activity. Comments were initially read and reread by two investigators (ML and JA) to develop the initial coding schema using an open-code approach. Successful CGM use was determined if the comment provided affirmation that a specific CGM site location worked for them (ie, CGM was reported to be most accurate, comfortable, or last the full length of the sensor, at minimum). CGM failure was categorized by inaccurate readings (ie, CGM readings did not match glucometer readings in a way that was significant to the user), CGM device ripping or falling off, too painful to continue wearing, bleeding impacting use, and unknown reasons. If the post or comment did not affirm success or failure, it was coded as *not applicable*. Post or comment mentions of off-label use of CGM wear more than 7 days were examined. Data were categorized for every 7 days worn beyond FDA approval (ie, 8-14 days, 15-21 days, etc). The frequency of *likes* (ie, “hearts”) and comments were analyzed using *t* tests.

Qualitative Thematic Analysis

Subjective bias by the research team is an inherent possibility in qualitative research. Our team used *bracketing*, or the identification of one’s own bias and preunderstanding of CGM use for people with diabetes. We noted our individual prejudice and discussed this as a team prior to beginning this qualitative content analysis; we then re-evaluated and communicated frequently as a team to be sure our individual bias did not become part of the research findings [17,18]. Additional senior researchers who were experienced with qualitative content analysis independently verified the findings; this process was

included in this study design [18]. Throughout the analysis, the open and axial coding processes were verified frequently with experienced researchers.

Qualitative thematic analysis was conducted by examining each original post (N=353) and their corresponding comments (N=2364). Codes were used to organize similar data in order to identify discussions about FDA-approved and off-label CGM activity [19]. The codes were then systematically applied to all of the data using an open-code approach to capture any data that was not specific to FDA-approved or off-label CGM activity [19,20]. A matrix was created to maintain an audit trail [21]. Themes were developed from the data [22]. Content of data, and not code frequency, was used to assess data saturation [23]. To avoid risk of identification, no direct quotes were used in this manuscript.

Results

Photograph Analysis

Of the 2923 photographs examined during the study period, 12.08% (353/2923) depicted a photograph with a CGM site location. The remaining photos in the original sample (2570/2923, 87.92%) not further analyzed in this study included pictures of CGM trend data on receivers and/or mobile devices; glucometers; CGM supplies (eg, transmitters and sensors); references to food, beverages, or exercise; photos of a person or people; and memes. There were 194 unique users who posted the 353 unique photographs, each user posting 1-4 photographs each. Multiple posts by individuals were linked by an Instagram handle in the dataset. The majority (225/353, 63.7%) of the photographs showed CGM sites in off-label locations, while 26.2% (92/353) were in an FDA-approved location (ie, abdomen) and 10.2% (36/353) were in unidentifiable locations (ie, *unknown* category; see Table 1).

Table 1. Instagram photographs by CGM^a site (N=353).

CGM site location	Frequency, n (%)
Posterior arm	139 (39.4)
Abdomen (FDA ^b approved)	92 (26.1)
Thigh	45 (12.7)
Unknown	36 (10.2)
Forearm	12 (3.4)
Back	10 (2.8)
Anterior arm	10 (2.8)
Calf	7 (2.0)
Buttocks	3 (0.8)

^aCGM: continuous glucose monitoring.

^bFDA: US Food and Drug Administration.

Post and Comment Analysis

Original posts (N=353) captured discussions about successful CGM use and failures. Additionally, some of the comments (N=2364) discussed CGM successes and failures, though not all; Table 2 reports sample sizes that include both the original

post and comments for each site location. Success rates were similar in the abdomen and posterior arm (see Table 2). Inaccuracy concerns were noted more often in the abdomen, calf, and buttocks, although the sample size of the buttocks was relatively low (n=10). There were 40 individual users who noted

wearing their CGM device successfully for more than 7 days (see Figure 1).

Viral Measures

To examine the viral spread of CGM photos and associated comments indicating off-label or FDA-approved site locations,

likes and comments were analyzed. There were no significant differences in the number of likes (FDA approved, N=9152; off-label, N=24,534; $P=.85$) or comments (FDA approved, N=707; off-label, N=1500; $P=.16$) based on FDA approval (see Table 3). However, the number of likes and comments were 3-4 times greater for off-label locations.

Table 2. CGM^a site discussions based on success and failure rates.

CGM site location	Success, n (%)	Inaccurate, n (%)	Ripped or fell off, n (%)	Too painful, n (%)	Bleeding, n (%)	Unknown reason, n (%)	Unable to determine, n (%)
Abdomen (N=178)	137 (77.4)	11 (6.2)	3 (1.7)	6 (3.4)	0 (0)	1 (0.6)	20 (11.2)
Posterior arm (N=273)	218 (79.9)	6 (2.2)	3 (1.1)	0 (0)	0 (0)	2 (0.7)	44 (16.1)
Thigh (N=198)	137 (69.2)	6 (3.0)	2 (1.0)	0 (0)	5 (2.5)	3 (1.5)	45 (22.7)
Back (N=42)	24 (57)	2 (5)	0 (0)	0 (0)	0 (0)	0 (0)	16 (38)
Calf (N=48)	20 (42)	3 (6)	0 (0)	0 (0)	1 (2)	0 (0)	24 (50)
Anterior arm (N=30)	19 (63)	0 (0)	0 (0)	0 (0)	1 (3)	0 (0)	10 (33)
Forearm (N=38)	22 (58)	0 (0)	1 (3)	0 (0)	0 (0)	0 (0)	15 (39)
Buttocks (N=10)	6 (60)	1 (10)	1 (10)	0 (0)	0 (0)	0 (0)	2 (20)

^aCGM: continuous glucose monitoring.

Figure 1. Reported continuous glucose monitoring (CGM) device use for more than 7 days (n=40).

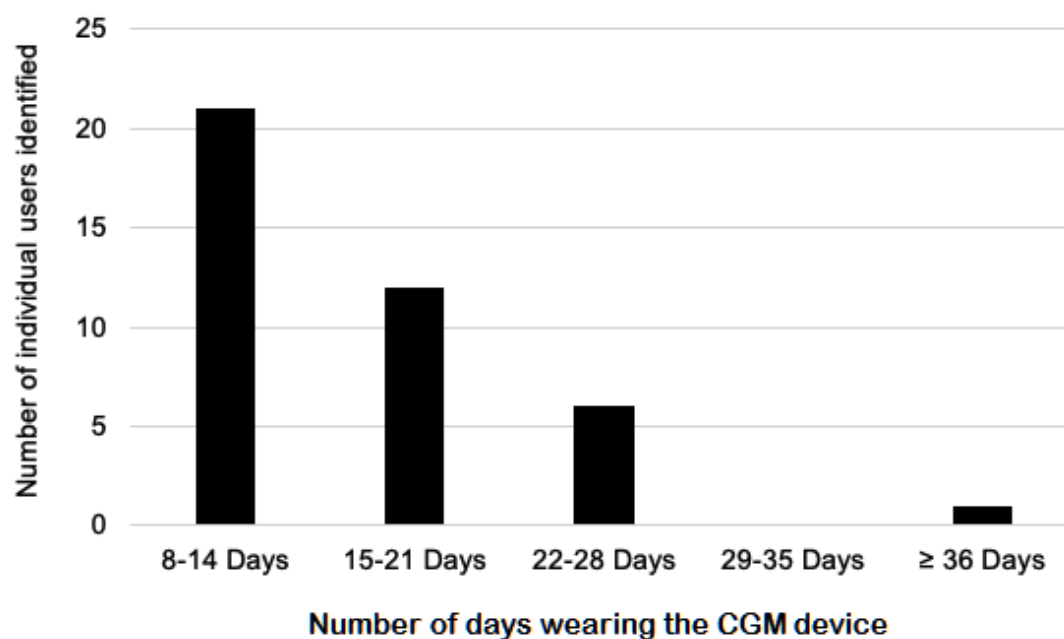


Table 3. Engagement with off-label and FDA^a-approved CGM^b site posts.

Engagement	Off-label, n	FDA approved, n	<i>P</i> value ^c
Likes	24,534	9152	.85
Comments	1500	707	.16

^aFDA: US Food and Drug Administration.

^bCGM: continuous glucose monitoring.

^c*P* values are based on *t* tests.

Qualitative Thematic Analysis

Five themes emerged from the analysis of original Instagram posts and the corresponding comments: (1) endorsement of CGM as providing a sense of well-being; (2) reciprocating information, encouragement, and support; (3) reciprocating CGM-related frustrations; (4) life hacks to optimize CGM use; and (5) sharing and learning about off-label CGM activity.

Endorsement of Continuous Glucose Monitoring Providing a Sense of Well-Being

CGM was described as a “life saver.” Hypoglycemia alarms were valued more than hyperglycemia alarms with regard to immediate safety concerns. Several posts mentioned the CGM *share* feature providing peace of mind for loved ones. Care partners who visualized glucose levels provided support to those with diabetes who reported that they could confidently engage in day-to-day activities, such as exercise and sleep. In fact, individuals felt so strongly about CGM, they used social media as a platform to encourage others with diabetes to use the technology. Social media was also used to teach Instagram followers, who may not have had diabetes, about CGM technology.

Many users shared that using CGM allowed them to think less about their diabetes and confidently manage their health in a less intrusive way. Women who were already pregnant or trying to become pregnant described the value of CGM in having a healthy baby. One person posted, “We've been trying to get pregnant so it feels like such a victory to be closer to my goal! I almost cried [when my hemoglobin A1c was] at 6%!” In addition to CGM, women who were pregnant also described the value of diabetes educators in helping them make sense of the CGM data. Parents described the value of being able to observe glucose levels without having to be intrusive. One parent posted, “[My child’s CGM] means that I don’t have to wake her up at 1am on nights that she eats pasta or lo mein to check her levels. What a gift.” Both adults with diabetes and parents of children with diabetes valued CGM as improving their quality of life.

Individuals also mentioned that CGM use reduced their worry by alerting them early to trends of high or low glucose levels. Although, there were also mentions about alarms being too frequent and some described alarm fatigue. CGM, in some cases, also provided insight into daily trends like overnight hypo- or hyperglycemia and caused them to change their insulin doses or behaviors in an attempt to remedy out-of-range glucose levels. Several posts attributed the use of CGM to a reduction in users’ glycosylated hemoglobin A1c levels and even in prolonging their life.

Reciprocating Information, Encouragement, and Support

Posters and commenters used Instagram as a space to share information based on experiential knowledge. Successes, frustrations, and daily life—#t1dlooklikeme and #thisiswhatdiabeteslookslike—with diabetes were discussed, all of which resulted in encouragement, support, and sometimes tips for problem solving. Self-expressions of strength (eg, #type1warrior, #type1strong, #t1dstrong, and #diabadass) were

used to describe individuals who were not going to let diabetes get in the way of living their best life. One person described their willingness to change their perception about their diabetes: “I’m ready to stop resenting my body, and start working with what I’ve got!”

Several posts included individuals engaging in healthy activities, such as exercise or eating a healthy meal. Peers encouraged others to stay on track or provided general health advice, such as drinking more water, getting enough rest, and the importance of administering a bolus before meals. Individuals who shared success stories, such as overcoming an emotional or physical barrier, were seen as inspirational. One person commented, “Happy Dexing to you, my super active diabuddy! Love to see PWD [people with diabetes] doing great things!” Recipients of these posts of encouragement wrote messages thanking individuals for their support. When individuals were experiencing inaccurate CGM, peers would inquire about recent Tylenol use or frequency of calibrations, reminding them about standard considerations for use of the CGM system.

Emojis and hashtags were used to provide encouragement and support. For example, hearts of all colors were used in a way to exhibit caring. Hand-gesture emojis, such as clapping hands, oncoming fist (ie, fist bump), OK hand (ie, index finger touching thumb to make an open circle), and victory hand (ie, peace sign) were used to exhibit concordance and approval. Hashtags were used to describe their diabetes (ie, #T1D, #type1diabetes, and #type1looklikeme) or tools used to manage diabetes (ie, #insulinpump). Additionally, hashtags were used in fun or humorous tones (ie, #bionicwoman, #insulinjunkie, and #sexybetic).

Continuous Glucose Monitoring-Related Frustrations

Concerns and frustrations were expressed, most often related to costs of managing diabetes or purchase of CGM devices (eg, #WhyDoTheyHaveToBeSoExpensive), failed CGM sites, hypo- or hyperglycemia that may or may not have been accompanied by physical symptoms, or feelings consistent with diabetes burnout. Insurance coverage and out-of-pocket costs were commonly asked about and shared so commenters could draw comparisons. Some individuals expressed their desire to start or restart CGM but were unable to do so because of cost. In other instances, individuals described a willingness to continue use of CGM despite cost and other frustrations they may have experienced because of the assurance CGM provided to them and their loved ones. One commenter noted, “It breaks my heart when my Dex rips out, its like money down the drain.” Failed sensors were not only financially challenging, but also emotionally challenging as one has to troubleshoot both technology and glucose readings simultaneously.

CGM accuracy, at times, was concerning. One person described, “I got the dreaded ‘???’ yesterday and Dex never recovered...I think I was overcalibrating. Sometimes I have a love/hate relationship with my Dex.” Individuals also described “compression lows” as possible causes for inaccurate CGM readings. Some described frustrations from not learning about the impact of sensor compression on glucose levels from their health care provider. Peers suggested avoiding laying on the CGM device (ie, sleeping on the alternate side of the CGM

device) as one solution. Use of alternative CGM site placements was also described as a possible option.

Although mentions of diabetes distress or burnout were uncommon, when they did occur, they were accompanied by reasons for not starting, stopping, or continuing CGM; comments of mutual understanding from those who had experienced similar feelings in the past; and motivational messages from peers. A play-on-words hashtag, #duckfiabetes, was also used to express frustrations related to diabetes.

Life Hacks to Optimize Continuous Glucose Monitoring Use

Instagram users shared diabetes *life hacks* to enhance the use of CGM. For example, individuals posted that if CGM failed in less than 7 days, they should directly contact the CGM company to request a sensor replacement. Discussions took place related to “precooking” or “soaking” the sensor in order to improve accuracy. This means placing a new sensor on the body for several hours prior to starting the official warm-up period on the device. Users reported that the extended warm-up time improves CGM accuracy, addressing that the lower accuracy may be seen within the first 24 hours of CGM use. Some individuals noted concern about the adhesive not working for the full 7-day period or experiences where the CGM device had become inadvertently unattached from the body. The use of various adhesives (ie, GrifGrips, Opsite Flexifit, and medical tape) and adhesive barrier wipes (ie, Skin Tac) were described as a way to improve CGM device adherence to the skin. One person stated, “My [adhesive] stood up well despite my sweating like a pig while running 6 miles.” The addition of adhesives were also used to extend the use of the CGM device. Concerns about CGM device adhesive causing rashes were described. One person stated, “My stomach is at the point that I need to take a break from the Dexcom for the first time in almost 2 years.” In response to concerns about rashes related to CGM device adhesive, several barrier solutions were suggested to prevent skin reactions (ie, Johnson & Johnson Tough Pads and 3M Cavilon No Sting Barrier Film). Off-label use of FLONASE on the skin prior to CGM site insertion was also described as a possible solution to preventing CGM device adhesive-related rashes.

For some, making CGM fun by accessorizing the sensor and receiver helped with the psychosocial aspect of having diabetes. Some individuals shared pictures of personalizing their CGM device with designs as a way to demedicalize the medical device. Extending the wear of the CGM device was also a described benefit. One person commented, “Oh girl! You gotta get some [personalized adhesive]! They help keep your Dexcom or pump sites on longer, AND they come in all kinds of fun shapes and colors!” Personalization occurred with the transmitter or receiver (ie, #PumpPeelz), adhesive overlay (ie, #GrifGrips), and the CGM receiver (ie, #Tallygear).

Sharing and Learning About Off-Label Continuous Glucose Monitoring Activity

Off-label use of CGM was described as solutions to enhance CGM accuracy, decrease cost, and improve comfort. Two types of off-label activity were discussed: CGM device wear on

locations other than the abdomen and extension of CGM device wear beyond 7 days. Off-label CGM activity discussions included those who had engaged in off-label activity often and those who had not. Those who had not engaged in off-label activity sometimes reported that they had only been taught the FDA-approved way of wearing the CGM device and had not thought to self-experiment but sought more information. Some parents expressed that photos viewed on Instagram would be used to explore new CGM site options with their children; one parent stated, “I’m glad you posted a photo [of a CGM site on the thigh]! I’ve been trying to talk my kiddo into putting it on her leg!” However, use of CGM devices on body parts other than the abdomen were oftentimes viewed as exciting and useful and, at times, more practical, less painful, and more accurate. Those who were already using CGM in off-label locations had typically experimented in more than one off-label location until they found one or more sites that were both comfortable and provided accurate readings. One person noted, “I wish my stomach was accurate. If so, it would be my favorite. Alas, my arm is my old faithful.” Sometimes, different sites were used based on the activity the individual planned to engage in. One person commented, “It all depends on activity. My arm is best when I go mountain biking or skiing but not great for arm-balancing yoga exercises.” Comfort was important as well; one person noted, “I always wear my CGM [device] on my thigh! Best spot for working out and comfort for sure!” In other instances, off-label CGM sites were used for practicality. One parent commented, “My 19-month-old daughter loves hers on the back of her arm!! Out of sight out of mind.” Few individuals new to off-label CGM site locations expressed unwillingness to show their CGM device in a location that was more publicly visible (ie, arm); many individuals only wearing the CGM device on the abdomen indicated that they planned to engage in off-label CGM wear in the future. In response to seeing a photo of a CGM device on a forearm, one person commented, “Maybe I’ll try this spot next time I’m due to change [my sensor]!! I’m always open for new spots to put my Dex.” Only one post expressed unwillingness to try off-label CGM site locations, due to it not being FDA approved.

Details on how to extend the use of the CGM was described in multiple posts and comments. Those who were extending the use of CGM for more than 7 days were often proud of this accomplishment. One person noted a new record in wearing the CGM device: “Dexy finally bit the dust today after a beautiful 18 days together.” Curious about this activity, learners sought more specific details from users experienced with this activity. One commenter asked, “How do you make it last 2 weeks? My sensor always tells me I have to change after 7 days.” In response to a request for more information such as this, other commenters would describe the process of resetting the sensor on the receiver without actually removing the sensor from the body and sometimes describe how long they are able to wear a sensor. One person commented, “We leave ours in for 3-4 weeks. Just stop and restart the sensor.” Individuals stated that they extended their use of CGM beyond 7 days for several reasons. First, to save money; CGM was noted to be expensive and not having to change the CGM sensor on a weekly basis allowed them to require fewer sensors each month. One representative quote states, “I’m about to enter my 4th week

with this sensor. #WhyDoTheyHaveToBeSoExpensive.” Additionally, improvement in readings with CGM device wear over a longer period of time was also noted. One person posted, “My back and upper butt gives me the best and most accurate readings for 3-4 weeks.” Similar to off-label locations, several 7-day CGM users expressed interest in extending CGM device wear time in the future.

Discussion

Principal Findings

This is the first study, to our knowledge, that uses photosurveillance, a novel methodology, to examine the real-world use of CGM. We found that CGM users in this sample were successfully engaging in off-label activity related to CGM use in order to improve their experience with the technology. CGM use in location sites that were off-label yielded, in some instances, better success than FDA-approved locations in this sample of Instagram posts and comments reviewed. CGM device wear time was also extended beyond the FDA-approved 7-day period by some individuals. Instagram was used as a resource to provide and reciprocate emotional support and learn about life hacks to optimize CGM use, which included both FDA-approved and off-label activity. Our qualitative findings have several research and clinical implications.

Photosurveillance Methodology

We successfully used photosurveillance, a new method to explore socially shared photographs, to examine real-world trends in diabetes care. This study allowed for the capture of important qualitative metrics of off-label CGM activity that otherwise may be difficult or impossible to capture in the clinical setting. As such, photosurveillance may be a beneficial methodology to explore health-related topics that may be considered taboo. In the future, other research methods may be combined with photosurveillance to examine the spread of health behavior, such as off-label CGM activity, within a network as has been done with other health conditions [10,11].

Hand-searching Instagram was a labor-intensive process. Future research should explore machine learning techniques to examine photographs and corresponding text from social media sources on a larger scale. For instance, computer vision algorithms can be leveraged to automatically process images to identify when individuals are wearing their devices in off-label locations. In particular, convolutional neural networks [24] achieve state-of-the-art accuracy for several computer vision tasks, including object recognition, object detection, and scene labeling [25-28]. Labeled Instagram posts that have been manually annotated by human coders can be used to provide training data to calibrate the model and as test data to evaluate the trained model's accuracy.

Other types of social media platforms, including Twitter, can be utilized to examine public opinion and sentiment around CGM; these data can be processed using natural language techniques developed within the field of computer science. Machine learning and deep learning algorithms can potentially decrease the cost of research and enable research to be conducted

on a larger scale. However, they may not be able to provide as detailed an analysis on features of an image or text that human coders can provide. While a human coder can evaluate potentially hundreds of characteristics of an image or scene, separate algorithms may need to be built for each characteristic to be extracted.

Off-Label Use of Continuous Glucose Monitoring

Information about off-label use of CGM is being shared on Instagram. In this study, individuals wore CGM devices in off-label locations. A small study in a pediatric population found no accuracy difference in sensors worn on the abdomen and buttocks, which are both approved by the FDA in children, and the arm, which is off-label [29]. Further, sensor failure was equal between off-label and FDA-approved locations.

Dexcom is typically used by individuals who are dependent on insulin and therefore need to be protective of the available “real estate” on their body to optimize insulin absorption. The risk of scar tissue and the need to rotate insulin injection and pump sites may increase the desire for multiple areas to place the CGM sensor in addition to the abdomen. Recent research [30] indicates that CGM device placement in areas of lipohypertrophy, which may not necessarily be on the abdomen, has equivalent or superior glucose accuracy to that of normal tissue; however, long-term data are lacking. Currently, 68%-71% of CGM users wear their CGM device at least 75% of the time [31,32]; it is possible that use of off-label CGM locations may increase this number.

We found instances in which CGM devices were being worn successfully beyond the FDA-approved 7 days. While the sample size was relatively small (40/353, 11.3%), the qualitative data indicates that this phenomenon not only contributed to cost savings, but also improved CGM accuracy over time. With the recent FDA approval of the Dexcom G6 in the United States, which is approved for 10-day wear, the extension of use remains to be seen.

Limitations

The interpretation of results should be considered in the context of the study limitations. It is unknown how many people with diabetes use Instagram; therefore, findings are not generalizable. In general, there were few (353/2923, 12.08%) CGM sites identified in the overall sample of photographs screened. While this does not reflect inability to capture data of CGM site placement, it does warrant the question of how valuably individuals perceive the sharing of CGM site photos, compared to other types of CGM photos, with others. Photographs only capture a snapshot in time and may not accurately reflect whether or not CGM was actually used successfully for the full length of the sensor time frame. It is possible we did not capture all reports of off-label CGM activity due to limiting the analysis to only one hashtag, #dexcom, and in only analyzing photographs of actual CGM sites. However, this study provides preliminary evidence that off-label use of CGM exists and appears to be working physiologically, while enhancing flexibility and accuracy for patients. Self-report of CGM accuracy may not reflect actual accuracy of CGM glucose levels. However, our study does confirm previous work [29] focused

on glucose accuracy in off-label locations. To optimize representativeness of our sample, data was collected daily throughout the study period. Due to an inability to assess demographics of Instagram users, we did not dichotomize the data into photographs of adults and children due to risk of error. Therefore, we did not differentiate between photos of children and those of adults in our study. Children do have FDA approval to wear the Dexcom on the abdomen and buttocks, while adults only have FDA approval to wear it on the abdomen. In our study, photographs on the buttocks, a sensitive area of the body, were limited. Additionally, 10% of the sample were categorized as unknown. Future work should make an effort to examine off-label use of CGM derived from survey and/or clinical research.

Conclusions

In this study we used photosurveillance to successfully identify real-world trends in CGM device wear, including placement and length of wear. We found individuals successfully used CGM in off-label locations with greater frequency than on the abdomen, with no indication of sensor failure or significant adverse effects. People with diabetes who are using CGM are finding ways to effectively use the tools in a manner that fits with their individual needs and goals, which is augmented by others on social media. Health care providers, industry, researchers, and even the FDA may want to utilize photosurveillance as a method to discover how people are using diabetes technologies successfully in their daily lives and perhaps adjust care protocols or product design accordingly.

Conflicts of Interest

None declared.

References

1. Hilliard ME, Sparling KM, Hitchcock J, Oser TK, Hood KK. The emerging diabetes online community. *Curr Diabetes Rev* 2015;11(4):261-272 [FREE Full text] [Medline: 25901500]
2. Litchman ML, Rothwell E, Edelman LS. The diabetes online community: Older adults supporting self-care through peer health. *Patient Educ Couns* 2018 Dec;101(3):518-523. [doi: 10.1016/j.pec.2017.08.023] [Medline: 28947360]
3. Litchman ML, Edelman LS, Donaldson GW. Effect of diabetes online community engagement on health indicators: Cross-sectional study. *JMIR Diabetes* 2018 Apr 24;3(2):e8 [FREE Full text] [doi: 10.2196/diabetes.8603] [Medline: 30291079]
4. Sparud-Lundin C, Ohrn I, Danielson E. Redefining relationships and identity in young adults with type 1 diabetes. *J Adv Nurs* 2010 Jan;66(1):128-138. [doi: 10.1111/j.1365-2648.2009.05166.x] [Medline: 20423440]
5. Ravert RD, Hancock MD, Ingersoll GM. Online forum messages posted by adolescents with type 1 diabetes. *Diabetes Educ* 2004;30(5):827-834. [doi: 10.1177/014572170403000518] [Medline: 15510534]
6. Newman M, Lauterbach D, Munson S, Resnick P, Morris M. It's not that I don't have problems, I'm just not putting them on Facebook: Challenges and opportunities in using online social networks for health. In: *Proceedings of the ACM 2011 Conference on Computer-Supported Cooperative Work*. 2011 Presented at: ACM 2011 Conference on Computer-Supported Cooperative Work; March 19-23, 2011; Hangzhou, China p. 341-350. [doi: 10.1145/1958824.1958876]
7. O'Kane AA, Park SY, Mentis H, Blandford A, Chen Y. Turning to peers: Integrating understanding of the self, the condition, and others' experiences in making sense of complex chronic conditions. *Comput Support Coop Work* 2016 Aug 17;25(6):477-501. [doi: 10.1007/s10606-016-9260-y]
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 and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med* 2011 May;40(5 Suppl 2):S154-S158. [doi: 10.1016/j.amepre.2011.02.006] [Medline: 21521589]
10. 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]
11. 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]
12. Stefanidis A, Vraga E, Lamprianidis G, Radzikowski J, Delamater PL, Jacobsen KH, et al. Zika in Twitter: Temporal variations of locations, actors, and concepts. *JMIR Public Health Surveill* 2017 Apr 20;3(2):e22 [FREE Full text] [doi: 10.2196/publichealth.6925] [Medline: 28428164]
13. Hesse CM, Bryan C, Rose AE. Indicators of Suicide Found on Social Networks: Phase 1. Seaside, CA: Defense Personnel and Security Research Center, Defense Manpower Data Center; 2015 Oct 07. URL: <https://apps.dtic.mil/dtic/tr/fulltext/u2/a633259.pdf> [accessed 2019-05-13] [WebCite Cache ID 78LF5oODL]
14. Rose AE, Hesse CM. Indicators of Suicide Found on Social Networks: Phase 2. Seaside, CA: Defense Personnel and Security Research Center, Defense Manpower Data Center; 2015 Oct 08. URL: <https://apps.dtic.mil/dtic/tr/fulltext/u2/1007350.pdf> [accessed 2019-05-13] [WebCite Cache ID 78LFM6040]

15. Pongpaichet S, Tang M, Jalali L, Jain R. Using photos as micro-reports of events. In: Proceedings of the 2016 ACM International Conference on Multimedia Retrieval. 2016 Jun Presented at: 2016 ACM International Conference on Multimedia Retrieval; June 6-9, 2016; New York, NY p. 87-94. [doi: [10.1145/2911996.2912020](https://doi.org/10.1145/2911996.2912020)]
16. Harris P, Taylor R, Thielke R, Payne J, Gonzalez N, Conde J. Research electronic data capture (REDCap): A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
17. Streubert HJ, Rinaldi Carpenter D. *Qualitative Research in Nursing: Advancing the Humanistic Imperative*. 5th edition. Philadelphia, PA: Wolters Kluwer Health | Lippincott Williams & Wilkins; 2011.
18. Cohen MZ, Kahn DL, Steeves RH. *Hermeneutic Phenomenological Research: A Practical Guide for Nurse Researchers*. Thousand Oaks, CA: Sage Publications; 2000.
19. Morgan D. Qualitative content analysis: A guide to paths not taken. *Qual Health Res* 1993 Feb;3(1):112-121. [doi: [10.1177/104973239300300107](https://doi.org/10.1177/104973239300300107)] [Medline: [8457790](https://pubmed.ncbi.nlm.nih.gov/8457790/)]
20. Sandelowski M. Whatever happened to qualitative description? *Res Nurs Health* 2000 Aug;23(4):334-340. [Medline: [10940958](https://pubmed.ncbi.nlm.nih.gov/10940958/)]
21. Miles MB, Huberman AM. *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, CA: Sage Publications; 1994.
22. Tesch R. *Qualitative Research: Analysis Types and Software Tools*. Abingdon, UK: RoutledgeFalmer; 1990.
23. Morse JM. The significance of saturation. *Qual Health Res* 1995 May 01;5(2):147-149. [doi: [10.1177/104973239500500201](https://doi.org/10.1177/104973239500500201)]
24. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems. 2012 Presented at: 25th International Conference on Neural Information Processing Systems; December 3-6, 2012; Lake Tahoe, NV p. 1097-1105 URL: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>
25. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: Proceedings from the International Conference on Learning Representations. 2015 Presented at: International Conference on Learning Representations; May 7-9, 2015; San Diego, CA URL: <https://arxiv.org/pdf/1409.1556.pdf>
26. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015 Presented at: IEEE Conference on Computer Vision and Pattern Recognition; June 7-12, 2015; Boston, MA p. 1-9. [doi: [10.1109/CVPR.2015.7298594](https://doi.org/10.1109/CVPR.2015.7298594)]
27. Zhou B, Lapedriza A, Xiao J, Torralba A, Oliva A. Learning deep features for scene recognition using places database. In: Proceedings of the 27th International Conference on Neural Information Processing Systems. 2014 Presented at: 27th International Conference on Neural Information Processing Systems; December 8-13, 2014; Montreal, QC p. 487-495 URL: <http://papers.nips.cc/paper/5349-learning-deep-features-for-scene-recognition-using-places-database.pdf>
28. Kozinets R. The field behind the screen: Using netnography for marketing research in online communities. *J Mark Res* 2002 Feb 01;39(1):61-72. [doi: [10.1509/jmkr.39.1.61.18935](https://doi.org/10.1509/jmkr.39.1.61.18935)]
29. Faccioli S, Del Favero S, Visentin R, Bonfanti R, Iafusco D, Rabbone I, PedArPan Study Group. Accuracy of a CGM sensor in pediatric subjects with type 1 diabetes. Comparison of three insertion sites: Arm, abdomen, and gluteus. *J Diabetes Sci Technol* 2017 Nov;11(6):1147-1154 [FREE Full text] [doi: [10.1177/1932296817706377](https://doi.org/10.1177/1932296817706377)] [Medline: [28486841](https://pubmed.ncbi.nlm.nih.gov/28486841/)]
30. DeSalvo DJ, Maahs DM, Messer L, Wadwa RP, Payne S, Ly TT, et al. Effect of lipohypertrophy on accuracy of continuous glucose monitoring in patients with type 1 diabetes. *Diabetes Care* 2015 Oct;38(10):e166-e167 [FREE Full text] [doi: [10.2337/dc15-1267](https://doi.org/10.2337/dc15-1267)] [Medline: [26307604](https://pubmed.ncbi.nlm.nih.gov/26307604/)]
31. Pickup JC, Ford Holloway M, Samsi K. Real-time continuous glucose monitoring in type 1 diabetes: A qualitative framework analysis of patient narratives. *Diabetes Care* 2015 Apr;38(4):544-550. [doi: [10.2337/dc14-1855](https://doi.org/10.2337/dc14-1855)] [Medline: [25552422](https://pubmed.ncbi.nlm.nih.gov/25552422/)]
32. Walsh J, Roberts R, Weber D, Faber-Heinemann G, Heinemann L. Insulin pump and CGM usage in the United States and Germany: Results of a real-world survey with 985 subjects. *J Diabetes Sci Technol* 2015 Jun 12;9(5):1103-1110 [FREE Full text] [doi: [10.1177/1932296815588945](https://doi.org/10.1177/1932296815588945)] [Medline: [26071425](https://pubmed.ncbi.nlm.nih.gov/26071425/)]

Abbreviations

CGM: continuous glucose monitoring

FDA: US Food and Drug Administration

PWD: people with diabetes

Edited by T Sanchez; submitted 10.05.18; peer-reviewed by E Gabarron, M Hilliard, L Zhou, T Oser; comments to author 23.07.18; revised version received 14.09.18; accepted 15.03.19; published 24.05.19.

Please cite as:

Litchman ML, Wawrzynski SE, Woodruff WS, Arrington JB, Nguyen QC, Gee PM

Continuous Glucose Monitoring in the Real World Using Photosurveillance of #Dexcom on Instagram: Exploratory Mixed Methods Study

JMIR Public Health Surveill 2019;5(2):e11024

URL: <http://publichealth.jmir.org/2019/2/e11024/>

doi: [10.2196/11024](https://doi.org/10.2196/11024)

PMID: [31127724](https://pubmed.ncbi.nlm.nih.gov/31127724/)

©Michelle L Litchman, Sarah E Wawrzynski, Whitney S Woodruff, Joseph B Arrington, Quynh C Nguyen, Perry M Gee. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 24.05.2019. 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

Early Detection of Adverse Drug Reactions in Social Health Networks: A Natural Language Processing Pipeline for Signal Detection

Azadeh Nikfarjam^{1*}, PhD; Julia D Ransohoff^{1*}, MD; Alison Callahan^{1*}, PhD; Erik Jones², PhD; Brian Loew², BA; Bernice Y Kwong³, MD; Kavita Y Sarin³, MD, PhD; Nigam H Shah¹, MBBS, PhD

¹Stanford Center for Biomedical Informatics Research, Stanford Department of Medicine, Stanford, CA, United States

²Inspire, Arlington, VA, United States

³Department of Dermatology, Stanford University School of Medicine, Stanford, CA, United States

*these authors contributed equally

Corresponding Author:

Julia D Ransohoff, MD

Stanford Center for Biomedical Informatics Research

Stanford Department of Medicine

300 Pasteur Drive, Lane L154

Stanford, CA,

United States

Phone: 1 650 725 6236

Email: ransohoff@stanford.edu

Abstract

Background: Adverse drug reactions (ADRs) occur in nearly all patients on chemotherapy, causing morbidity and therapy disruptions. Detection of such ADRs is limited in clinical trials, which are underpowered to detect rare events. Early recognition of ADRs in the postmarketing phase could substantially reduce morbidity and decrease societal costs. Internet community health forums provide a mechanism for individuals to discuss real-time health concerns and can enable computational detection of ADRs.

Objective: The goal of this study is to identify cutaneous ADR signals in social health networks and compare the frequency and timing of these ADRs to clinical reports in the literature.

Methods: We present a natural language processing-based, ADR signal-generation pipeline based on patient posts on Internet social health networks. We identified user posts from the Inspire health forums related to two chemotherapy classes: erlotinib, an epidermal growth factor receptor inhibitor, and nivolumab and pembrolizumab, immune checkpoint inhibitors. We extracted mentions of ADRs from unstructured content of patient posts. We then performed population-level association analyses and time-to-detection analyses.

Results: Our system detected cutaneous ADRs from patient reports with high precision (0.90) and at frequencies comparable to those documented in the literature but an average of 7 months ahead of their literature reporting. Known ADRs were associated with higher proportional reporting ratios compared to negative controls, demonstrating the robustness of our analyses. Our named entity recognition system achieved a 0.738 microaveraged *F*-measure in detecting ADR entities, not limited to cutaneous ADRs, in health forum posts. Additionally, we discovered the novel ADR of hypohidrosis reported by 23 patients in erlotinib-related posts; this ADR was absent from 15 years of literature on this medication and we recently reported the finding in a clinical oncology journal.

Conclusions: Several hundred million patients report health concerns in social health networks, yet this information is markedly underutilized for pharmacosurveillance. We demonstrated the ability of a natural language processing-based signal-generation pipeline to accurately detect patient reports of ADRs months in advance of literature reporting and the robustness of statistical analyses to validate system detections. Our findings suggest the important contributions that social health network data can play in contributing to more comprehensive and timely pharmacovigilance.

(*JMIR Public Health Surveill* 2019;5(2):e11264) doi:[10.2196/11264](https://doi.org/10.2196/11264)

KEYWORDS

natural language processing; signal detection; adverse drug reactions; social media; drug-related side effects; medical oncology; antineoplastic agents; machine learning

Introduction

Adverse drug reactions (ADRs) are an important public health issue, causing considerable patient harm with a high health care cost [1,2]. Serious ADRs occur in over 2 million patients annually in the United States, resulting in 100,000 deaths [2]. It is difficult to comprehensively characterize ADRs during premarket trials, as many serious ADRs, particularly rare ones, are discovered years after a drug has been on the market [3]. Clinical trials are limited in their ability to detect ADRs because of their relatively small cohort sizes, short durations, and a lack of enrolled patient diversity [4,5]. A significant number of ADRs are recognized in the postmarket period, resulting in over 2 million injuries and US \$75 billion in annual health care costs [6,7]. Ongoing surveillance strategies are therefore necessary to monitor drug safety during the postapproval period.

Spontaneous reporting systems (SRSs) are drug safety surveillance mechanisms designed by regulatory agencies to monitor drug safety during the postmarket period. SRSs can effectively detect rare ADRs by nature of their longitudinal profiling and wide reach of included reports but are limited by their reliance on voluntary patient or provider reporting; it is estimated that more than 90% of ADRs are underreported [8,9]. ADR underdetection has motivated efforts to include complementary, alternative data sources for pharmacovigilance, including electronic health records and administrative claims [10-12], biomedical literature [13,14], Internet search logs [15], patient posts in social media [16-18], and multimodal systems that jointly analyze multiple sources of information for ADR detection [7].

Over 300 million patients seek and share health-related information from online Internet platforms such as health forums, search engine logs, or tweets [18]. These social health networks provide a platform for patients or caregivers to connect in discussing treatment options, drug side effects, and illness trajectories. The valuable health-related content of user posts in social media can be used for public health surveillance purposes [19], also termed *infodemiology* [20], including monitoring the spread of contagious diseases such as influenza [21,22]; monitoring the time and geographical locations of diseases [23]; health outcome measurement [24,25]; discovering associations between health-related concepts such as drugs and diseases; and, particularly, monitoring adverse effects of medications [16,17]. Several studies have highlighted the importance of utilizing social media as a resource for pharmacovigilance [9]. User posts often contain informal, unstructured text from which it is more challenging to extract medical information than from other, more-structured sources. Therefore, the exploration of different natural language processing techniques in ADR concept detection from social media postings has received significant attention from the medical informatics community [9,17], though there are a paucity of studies focusing on drug-ADR, signal-generation methods based on social media postings [26].

Here we use Inspire [27], one of the largest online social health networks, which contains over 12 million health-related patient posts, including discussions of therapy responses, adverse drug reactions, and supplemental treatments. We present an ADR signal-generation pipeline based on patient posts in social health networks and compare the timing and the rate of such ADRs with those published in clinical literature. We demonstrate the capacity for early detection as well as discovery of ADRs using Inspire content.

In this work, we focus on two classes of chemotherapeutics, targeted small molecule inhibitors and immunotherapeutics, which are representative of the dramatic change in the chemotherapy landscape since the early 2000s. These classes of agents are now commonly used in place of more traditional antiproliferative agents and are associated with novel side-effect profiles related to their mechanisms of action. Oncologists have experienced a particularly steep learning curve in recognizing these reactions, which occur in essentially all patients and can be life-threatening [4,28] as there is limited-to-no long-term data with novel agents. A new subfield of oncology has emerged, aimed at recognizing which reactions are reflective of treatment response, which warrant treatment cessation, and managing side effects to permit treatment tolerability. To capture the breadth of reactions seen, here we focus on two representative classes of cancer drugs: (1) epidermal growth factor receptor (EGFR) inhibitors, which are widely used by most oncologists for specific malignancies harboring EGFR mutations, have been in practice for over 15 years, and, therefore, have well-established side-effect profiles; and (2) immune checkpoint inhibitors, which are relatively new, having first gained US Food and Drug Administration (FDA) approval in late 2014 and, therefore, have significantly less data available on their emerging side-effect profiles. We report the construction of a pipeline to study the association of cutaneous ADRs with these selected targeted cancer therapy drugs reported in patient postings in Inspire.

Methods

Overview

We defined a set of common and rare ADRs to study for their association with two classes of drugs: an EGFR inhibitor, erlotinib, and the immune checkpoint programmed cell death 1 (PD-1) inhibitors, nivolumab and pembrolizumab. We focused on eight skin-related ADRs in this work: rash, acne, pruritus (ie, itchy skin), paronychia (ie, nail changes), xerosis (ie, dry skin), hypohidrosis, bullous eruption (ie, blister), and psoriasis.

Drug Corpus Generation

The Inspire dataset consists of 7,320,546 discussion posts from 2005 to 2016. For each drug in the study, we generated a corpus defined as a collection of user posts from Inspire, with every post containing at least one mention of a keyword corresponding to one of the drugs of interest. To retrieve relevant posts from

the Inspire dataset, we used regular expressions, a simple text-processing method for string matching. We created a regular expression pattern that implemented the exact string match. We have provided the list of the drugs and the related regular expression in [Multimedia Appendix 1](#).

We identified 55,778 posts for erlotinib and 15,738 for the combined PD-1 inhibitors, nivolumab and pembrolizumab. Because nivolumab and pembrolizumab were more recently introduced and, therefore, have less discussion content but essentially identical mechanisms, we combined posts for the two drugs referencing the same ADR.

Detection of Adverse Drug Reactions

To extract ADR mentions from user posts, we used DeepHealthMiner (DHM), a neural network-based named entity recognition (NER) system that is specifically trained to extract drug safety-related entities from user-generated content in social media [29]. DHM is a supervised feed-forward neural network that is trained to identify two different entity types: ADRs and drug indications. The original, labeled training data is based on a dataset from DailyStrength [30], an online health community that contains patient-generated reports of treatment experiences very similar to those from Inspire. The labeled training data are the sentences from patient posts manually annotated for the entity span (ie, start and end position offsets) and the type of entity (ie, ADR or indication). More detailed information about the data and annotation details can be found in prior publications [31,32]. The unlabeled sentences from user posts are also used for unsupervised training of the word embeddings, which are used as vectors representing the input tokens to the NER system.

We retrained DHM using both labeled and unlabeled posts from Inspire. We retrained 150-dimensional word embeddings by adding the 7 million Inspire posts to the original unlabeled sentences. We manually labeled an additional 200 Inspire posts, following the original annotation guideline [31,32], to retrain DHM using Inspire content. Some of the rare ADRs in this study are more frequently discussed as indication mentions (eg, psoriasis or blistering) rather than ADRs in user postings; therefore, to improve our system's ability to correctly recognize such cases, we created a subset of 200 Inspire posts from our corpus that particularly contained keywords from the rare seed ADRs. We annotated all the ADR and indication mentions following the original annotation guidelines to retrain DHM using Inspire content. The preliminary results showed that the system could successfully extract mentions of rare ADRs after retraining on 200 posts.

To evaluate the performance of the retrained NER model in identifying ADR entities from Inspire posts, we created a test set—the Inspire test set—consisting of 50 randomly selected posts. Two human annotators labeled all mentions of ADRs, not limited to the selected ADRs in this study, corresponding to any treatment reported in the posts (interannotator agreement: 90.9%). We used the *F*-measure to quantify the interannotator agreement by considering one annotator as the gold standard [33].

The extracted ADR mentions along with the information about the related drugs were stored in a database for further analyses of the associations among drugs and ADRs.

Normalization of Extracted Mentions

To map the extracted ADR mentions to Unified Medical Language System (UMLS) Concept Unique Identifiers (CUIs), we generated a lexicon of the eight ADR concepts. We defined a set of seed CUIs for these ADRs as shown in [Multimedia Appendix 2](#). Using the UMLS hierarchy, we expanded every seed concept by adding synonyms and alternative names. For “nail changes,” we further expanded by including the children of the seed concepts (ie, is-a relation in the hierarchy). Every lexicon entry has a name, UMLS CUI, related seed ADR concept, and an identifier that we defined to group all items corresponding to a specific ADR. As an example, the lexicon entry “ingrown toenail (C0027343)” is an expanded concept based on “Disorder of nail (C0027339),” and “nail changes” is the human readable identifier for the ADR. Additionally, for every ADR, we added a set of colloquial phrases to the lexicon (see [Multimedia Appendix 2](#)). The list of colloquial phrases was generated by physicians, representing the list of words and phrases that they had seen patients use to describe their own cutaneous ADRs. We included these to better capture how patients describe ADRs in their own words.

Finally, we used Lucene [34] to index the lexicon. Lucene has been utilized as a successful tool in previous studies for normalizing biomedical entities [35,36]. We built a Lucene repository from the lexicon by creating one Lucene document for each lexicon entry, indexing the corresponding UMLS concept name and the CUI for each lexicon entry. The concept names are tokenized and lemmatized for generalization.

For mapping an extracted ADR mention to a UMLS concept, we searched the index using the text span extracted from the post by our NER system. We tokenized and lemmatized the extracted text span in the same manner as the indexed concept names and generated a Lucene query to search the index. For example, to map the extracted ADR mention “finger nails have been peeling” to a UMLS concept, we search the index using the query keywords “finger nail have be peel” and retrieve “peel nail (Peeling of nails: C0263531).” We retrieved a ranked list of the matched, relevant ADR concepts and chose the top-ranked concept from this list.

Computing Drug and Adverse Drug Reaction Associations

The co-occurrence of a drug and a reported ADR in a post can be considered as a potential association. To quantify the strength of an association between a drug and an ADR, we calculated the proportional reporting ratio (PRR), a statistic widely used for ADR signal generation from spontaneous reporting databases [26,37] and originally introduced by Evans et al [29]. The PRR quantifies ADR signals by comparing the frequency at which an ADR is reported with a drug of interest compared to the frequency at which the ADR is reported with other drugs (ie, comparison drugs) in the database.

We defined the comparison drugs by sorting drugs mentioned in Inspire by discussion frequency and selecting those most

highly discussed in Inspire discussion posts (frequency >5000). To reduce noise, we excluded over-the-counter drugs with common indications (eg, pain and allergy drugs), as well as several anti-inflammatory drugs (eg, prednisone and Humira) that have indications that match some of our rare ADRs (eg, bullous eruption and psoriasis), given that patients discussing these drugs were almost exclusively discussing their indications rather than experienced ADRs. The final comparison drug list included 27 drugs, which are listed in [Multimedia Appendix 3](#).

The PRR for an adverse reaction, R, and a drug, D, was calculated based on the following formula:

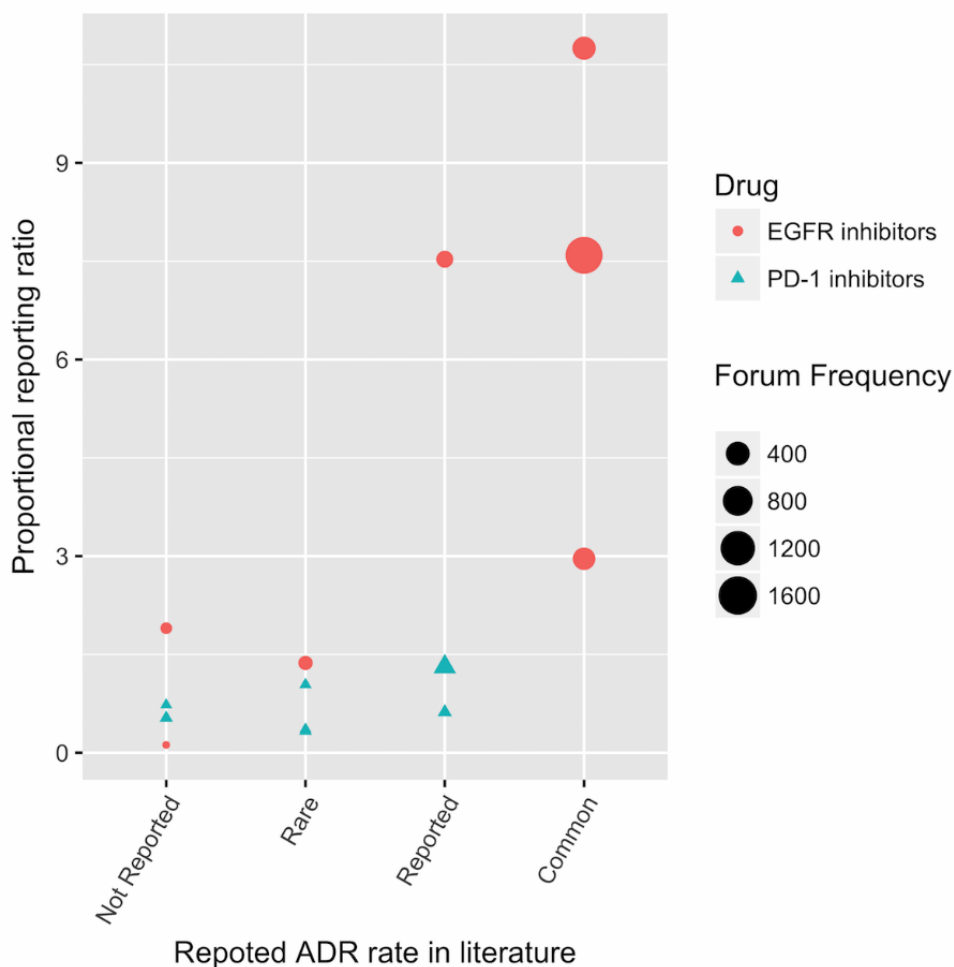
$$PRR(D,R) = (\text{count}[D \cap R] / \text{count}[D]) / (\text{count}[!D \cap R] / \text{count}[!D])$$

We defined count(D∩R) as the number of unique users that have reported both D and R in a post; count(D) as the total number of users reporting any ADR for drug D, including all extracted mentions by DHM (eg, weight gain or fatigue); count(!D∩R) as the total number of unique users that reported R for comparison drugs, excluding D; and count(!D) as the total number of unique users that reported at least one ADR for any drug except D.

Calibrating a Threshold for the Proportional Reporting Ratio

We performed empirical calibration to find the threshold at which the PRR may represent a true ADR signal [38,39]. We considered the distribution of PRR scores for a set of drug-condition pairs with no known associations. To create this negative control set, we first paired all 27 comparison drugs and the eight target ADRs, totaling 216 drug-condition pairs. We then excluded the drug-condition pairs with known associations (ie, ADR or indication) based on the Medi-Span Adverse Drug Effects Database (Wolters Kluwer Health) and side-effect resource (SIDER) [40], leaving 86 pairs without any documented associations. Based on the PRR formula, we expected and validated that drug-ADR pairs with known and common associations have PRR values greater than 1 (see [Figure 1](#)). If a drug-ADR pair has a known association, then the corresponding PRR will be relatively higher than that for pairs with no associations. Our goal in excluding these known pairs from the negative control set was to illustrate the distribution of PRRs for drug-ADR pairs with no known associations and use only these to calibrate the threshold at which a PRR for an unknown drug-ADR pair would represent a significant, novel result.

Figure 1. Comparison of proportional reporting ratio of skin adverse drug reactions (ADRs) reported in social health forums with the ADR rate published in the literature. The ADR rate based on the literature are grouped as follows: Not reported or rare = case report; reported = case series or in clinical trial; common = in significant percentage of patients in large trials. The frequency at which ADRs are talked about is shown by the size of the circle or triangle. The most common ADRs reported in the literature are also the most discussed in the patient posts. EGFR: epidermal growth factor receptor; PD-1: programmed cell death 1.



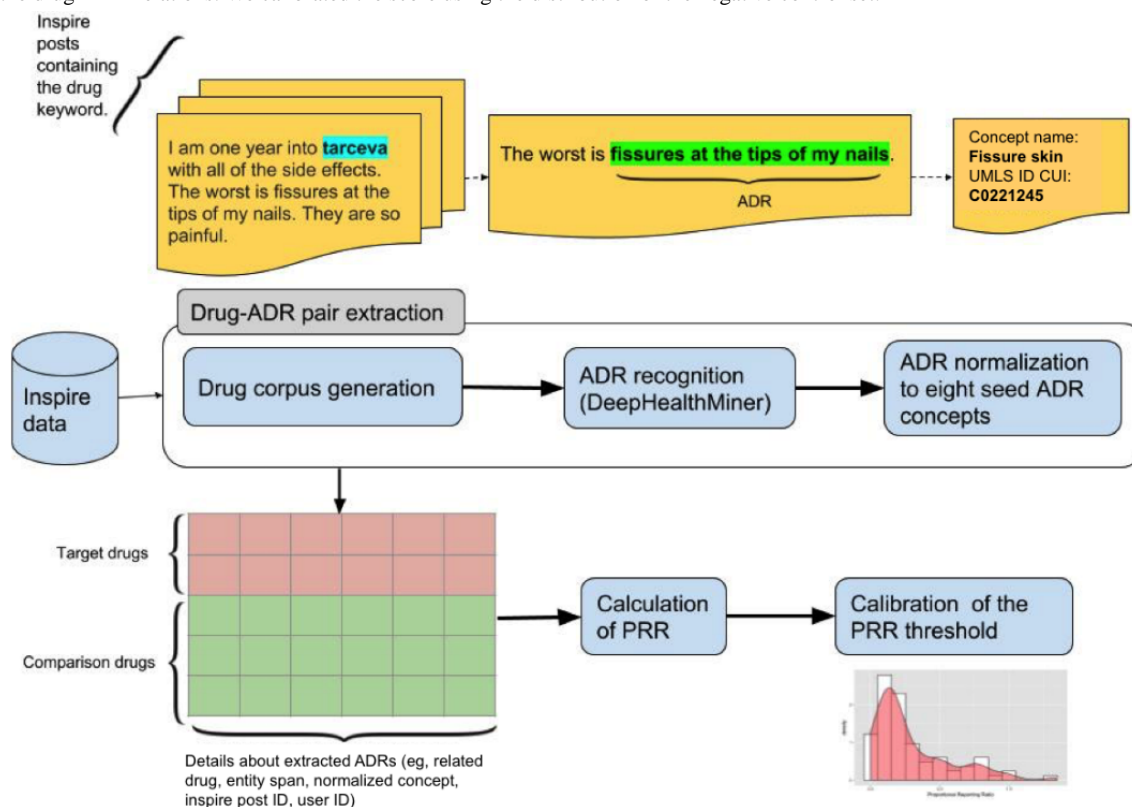
It is important to note that a drug and a condition can be associated for reasons other than an ADR or indication and may, therefore, co-occur frequently within the negative control set. We identified two common reasons that a drug (*drug x*) and a condition (*condition y*) with no documented association co-occur frequently in user posts, therefore representing false positives. The first reason is related to prescribing patterns: if a drug or a set of similar drugs are commonly prescribed with *drug x*, while the reported ADR condition is not attributed to *drug x*, the ADR will appear more frequently with all coprescribed drugs. For example, mentions of nail changes happen frequently with zoledronic acid; however, while a true ADR is associated with capecitabine, which is frequently prescribed with zoledronic acid, zoledronic acid is not responsible for the ADR. The second reason is that if *condition y* is one of the syndromic conditions

related to the indication for *drug x*, the related condition will co-occur with the drug at a similar rate to the true indication. For example, metformin can be prescribed for polycystic ovary syndrome (PCOS), and patients with PCOS often have acne. Acne is not an ADR associated with metformin but will co-occur frequently with mention of metformin by way of acne's association with PCOS.

Because of the potential for observing such drug-condition associations, we manually reviewed the negative control set to exclude them, leaving 81 drug-ADR pairs with no association. The negative control set is available in [Multimedia Appendix 4](#).

Figure 2 depicts our overall pipeline and graphical depiction of the above methods.

Figure 2. Pipeline to identify adverse drug reaction (ADR) signals associated with epidermal growth factor receptor (EGFR) and programmed cell death 1 (PD-1) inhibitors in social health networks. For drug-ADR pair extraction, for each drug, we generate a collection of user posts containing at least one mention of the drug. This drug corpus is then processed via DeepHealthMiner to recognize mentions of ADRs. The extracted mentions are then mapped to the corresponding Unified Medical Language System (UMLS) Concept Unique Identifiers (CUIs). The identified drug-ADR pairs and the related details, for both target and comparison group, are then stored in a relational database. The proportional reporting ratio (PRR) is calculated to quantify the drug-ADR relations. We calibrated the score using the distribution of the negative control set.



Results

Finding Drug-Adverse Drug Reaction Signals From Forum Posts

We identified 50,574 Inspire posts related to EGFR inhibitors and 16,598 related to checkpoint inhibitors; a total number of 13,600 ADR concepts extracted from the former and 812 concepts from the latter mapped to the eight target ADRs. We evaluated the performance of the presented NER method in identifying ADR mentions from the Inspire test set and achieved a microaveraged *F*-measure of 0.738 (0.731 precision and 0.745 recall).

To assess system performance in extraction and normalization of ADR concepts, we performed additional manual validation of the extracted ADR concepts as described in the Methods. Our final validation set for the entire pipeline for identifying drug-ADR pairs consisted of an additional 120 posts which had 15 randomly selected posts for every ADR that the system extracted for one of our target drugs.

We considered an extracted concept as a true positive if the system correctly detected the relevant span of text; classified the entity as an ADR rather than indication; and normalized the ADR to the correct concept. Benchmarking results show that our system achieved a high performance (microaverage precision

of 0.90) in recognizing and normalizing the ADRs in the user posts.

The distribution of the frequencies and the calculated PRR of the skin ADRs extracted from the EGFR inhibitor-related posts correlates closely with published observed rates in patients (see Figure 1), further validating the pipeline for identifying ADRs. The PD-1 inhibitor ADR landscape is largely limited to case reports; forum frequencies and PRR values correlate with published rates, suggesting that those described to date accurately represent the emerging toxicity landscape (see Figure 1).

To determine the threshold at which PRR represents true signal warranting further investigation, we plotted the PRR distribution for the 81 drug-ADR pairs in the negative control set (see Figure 3). More than 95% of the pairs in this set have a PRR of less than 0.82; therefore, drug-ADR pairs with a PRR greater than 1 can be considered as signaling true ADRs.

Adverse Drug Reactions are Described in Internet Forums Prior to Published Reports

To compare the timing of online ADR reports to initial ADR reports in the literature, we looked at both common and rare events associated with our target drugs. Papulopustular (ie, acneiform) rash and fingernail changes are well-reported ADRs associated with erlotinib and were first described in published case reports in September 2005 [41] and September 2006 [42], respectively. Inspire posts for these reactions appeared 5 and 3 months in advance of publication, respectively. Psoriasis in the

setting of PD-1 inhibitor treatment was first documented in case reports in July 2015 [43] and May 2016 [44]. Inspire forum posts describing this ADR preceded the 2015 case report by 9 months (see Figure 4A). Blistering reactions with PD-1 inhibitors were initially published as a case report in June 2015 [45] and as a three-case series in May 2016 [46]. Forum descriptions preceded the first case report by 9 months (see Figure 4B). Taken together, our data suggest a significant and consistent reporting lead-time advantage in online health forum posts. Cutaneous ADRs in both EGFR and PD-1 inhibitors have been reported to be associated with cancer response to therapy [47-50], highlighting the clinical utility of early detection.

Novel Adverse Drug Reaction Discovery in Social Health Networks

In early 2017, we documented three patients on erlotinib reporting hypohidrosis. Our system detected 23 unique Inspire users reporting hypohidrosis in causal association with erlotinib as early as 2006, based on manual review of user posts, with a PRR score of 1.90 (see Figure 3). This predicted ADR is absent from over 15 years of EGFR inhibitor literature, implicating a novel ADR detected through our system extractions from online posts. EGFR is expressed in sweat glands [51] and decreases with pharmacologic EGFR inhibition [52]. Interestingly, the hypohidrotic ectodermal dysplasia phenotype is partially mediated by decreased EGFR signaling [53], which can possibly cause hypohidrosis. These findings demonstrate the potential of ADR extraction from Internet forums to not only detect ADRs earlier, but also to contribute to novel ADR discovery.

Figure 3. Proportional reporting ratio (PRR) distribution for a set of 28 negative example drugs representing 81 drug-adverse drug reaction (ADR) pairs. The mean is 0.12, median is 0.2, and maximum is 1.4, highlighting a PRR threshold of 1, below which <5% of drug-ADR pairs have true associations.

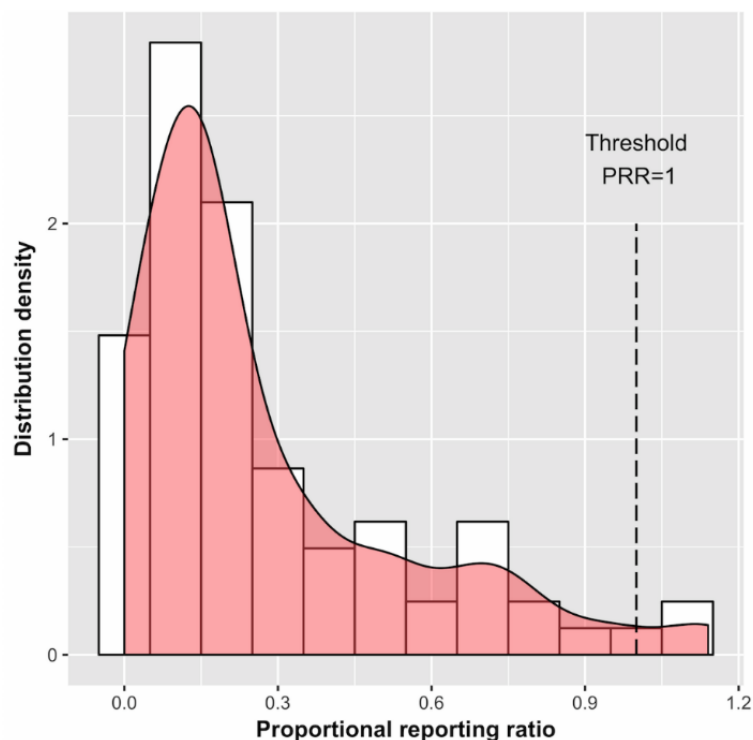
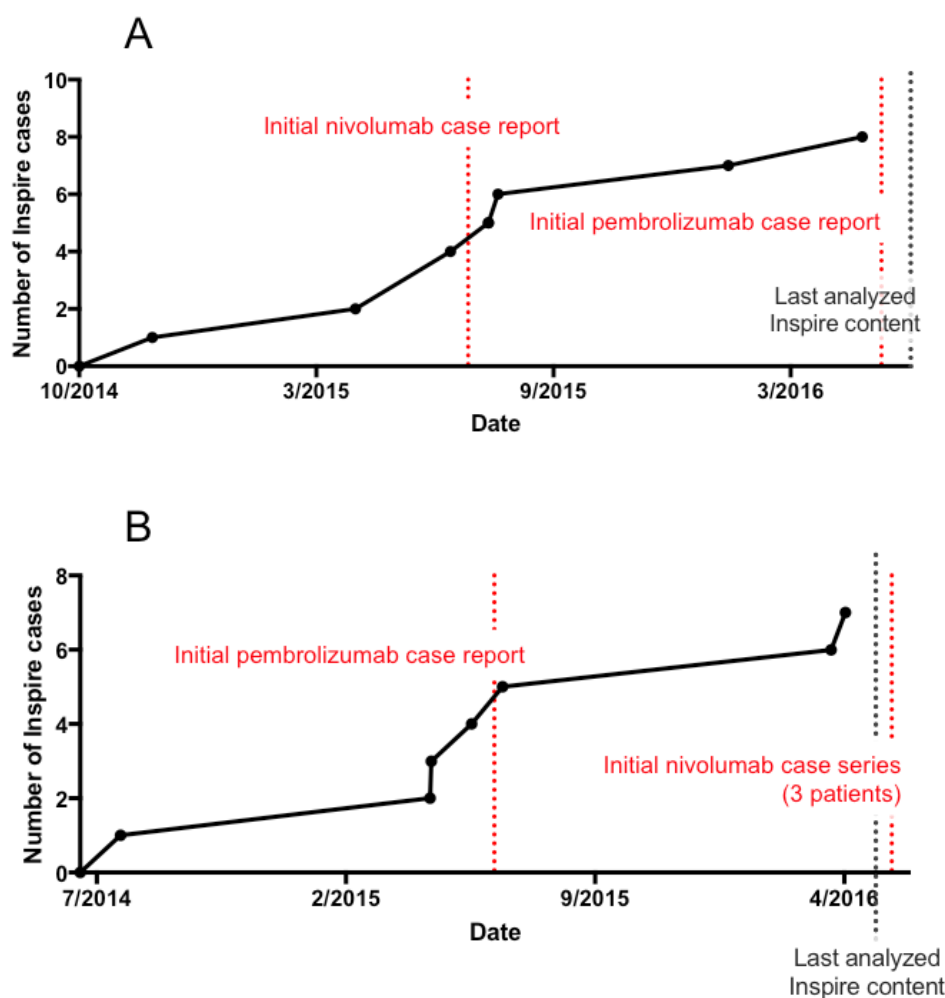


Figure 4. Cutaneous adverse drug reactions (ADRs) identified in Inspire forums precede initial published clinical reports. We plotted cumulative post count (y-axis) at each date (x-axis) for time-to-detection analysis. (A) Psoriasis was first reported in the literature as individual case reports with programmed cell death 1 (PD-1) inhibitors in July 2015 and May 2016. Inspire users began describing psoriasis flares 9 months prior to the first case report. (B) Bullous reactions with checkpoint inhibitors were first reported in the literature as a case report with pembrolizumab in May 2015 and as a three-case series with nivolumab in May 2016. Inspire cases were reported online 9 months before the initial case report.



Discussion

Our retrospective analyses demonstrate the potential of conducting targeted ADR surveillance in real time. Because many chemotherapeutics have nonspecific side-effect profiles, including nausea, vomiting, and fatigue, which are shared with most antineoplastics as well as thousands of other drugs, we see greater relevance in seeking specific signal-to-noise ratios based on clinically generated hypotheses—that is, reactions seen with a drug, with a hypothesized causal association. For example, our impetus for mining for hypohidrosis in association with erlotinib stemmed from a single clinical encounter. As we built and applied our pipeline to the Inspire dataset, we eventually saw two more patients in clinic with the same symptom; together, with the 23 Inspire users we identified, our system reported statistically significant association, supporting the classification of this symptom as a true ADR.

The presented methods in the signal generation pipeline can easily be applied to other drugs and ADRs. Our NER system extracts ADRs independent of a specific drug and, therefore,

the system performance should easily translate to the study of other drugs. In fact, in our previous studies [17,29], we showed that the presented NER methods tested on a dataset of user posts about 81 different drugs achieved high performance in extracting mentions of ADRs, not limited to any specific ADR. Therefore, the NER component can easily be applied to the majority of ADRs. However, if the goal is to identify specific rare ADRs or mentions that often are presented in contexts different than ADRs (eg, psoriasis), the system may be improved by adding additional relevant human-labeled training examples as we did in this study. Future research is needed to further explore transfer learning techniques for improvement in the performance of NER models when applied to the new tasks.

Furthermore, with respect to other ADRs, the only additional effort needed for their inclusion would be normalization and determination of grouping of the UMLS concepts attributed to a single ADR (eg, the concepts that can be considered as “nail changes”). For normalization, we used the seed ADRs to identify and group the concepts that indicated the target ADRs in our study. Overall, the precise normalization of medical mentions in text is a challenging task in biomedical information extraction

[36]; particularly for the medical entities in user posts [54], future research could explore alternative techniques for optimal normalization of medical concepts in social media postings.

Our study demonstrates that it is possible to detect ADR signals from patient-generated social media posts an average of 7 months earlier than literature reporting and at frequencies comparable to their eventual literature descriptions. We describe the proof-of-principle construction and validation of a signal generation pipeline for ADR detection from social health networks. We benchmark our system extractions of known drug events against their literature reports to evaluate our pipeline's accuracy and temporal advantage. Our system is able to extract drug-ADR signals from highly unstructured online patient content with high precision.

In this work, we demonstrate the utility of mining online patient reports to identify signals for both common and rare ADRs with high precision. We envision the potential for social media-based signals to be combined with those derived from alternative modes—electronic health records, insurance claims, and FDA reports—to construct the most comprehensive and dynamic catalogue of ADRs.

Looking ahead, we envision the utility of a live pharmacovigilance system that reports emerging drug-ADR association signals over the course of days to weeks and sends these reports to physicians for review and comparison with clinical experience. This would enable tracking of individual reports and changing PRRs over time, providing the reviewing physicians with access to a significantly broader patient cohort than seen in single treatment centers. We do anticipate potential new challenges in capturing how patients describe ADRs that are not directly visible to them, which will be an engaging and important area of further study.

If a reporting system such as ours were extended to public use, we also anticipate the inherent risks and challenges associated with patient access. While many patients search for their health concerns before presenting to clinicians, they are often misguided, which may impact the decision to seek care or timing of visit to a physician. A system such as ours, which captures the experiences of millions of patients, may be able to offer context for patients to understand their concerns, but it would be a further challenge to offer triaging of ADR severity to patients based solely on the reports of other patients and should instead serve to motivate patients to discuss their concerns in a timely manner with their own doctors.

Conflicts of Interest

BYK serves as a consultant to Genentech and Roche.

Multimedia Appendix 1

List of drug names.

[PDF File (Adobe PDF File), 53KB - [publichealth_v5i2e11264_app1.pdf](#)]

Multimedia Appendix 2

List of seed adverse drug reaction (ADR) concepts used.

[PDF File (Adobe PDF File), 43KB - [publichealth_v5i2e11264_app2.pdf](#)]

Multimedia Appendix 3

List of comparison drugs.

[PDF File (Adobe PDF File), 12KB - [publichealth_v5i2e11264_app3.pdf](#)]

Multimedia Appendix 4

List of negative control drug-adverse drug reaction (ADR) pairs.

[PDF File (Adobe PDF File), 92KB - [publichealth_v5i2e11264_app4.pdf](#)]

References

1. Numerous contributors. In: Lee A, editor. Adverse Drug Reactions. 2nd edition. London, UK: Pharmaceutical Press; 2006.
2. Giacomini K, Krauss R, Roden D, Eichelbaum M, Hayden M, Nakamura Y. When good drugs go bad. *Nature* 2007 Apr 26;446(7139):975-977. [doi: [10.1038/446975a](#)] [Medline: [17460642](#)]
3. Lasser K, Allen PD, Woolhandler SJ, Himmelstein DU, Wolfe SM, Bor DH. Timing of new black box warnings and withdrawals for prescription medications. *JAMA* 2002 May 01;287(17):2215-2220. [Medline: [11980521](#)]
4. Rosen A, Case E, Dusza S, Balagula Y, Gordon J, West D, et al. Impact of dermatologic adverse events on quality of life in 283 cancer patients: A questionnaire study in a dermatology referral clinic. *Am J Clin Dermatol* 2013 Aug;14(4):327-333. [doi: [10.1007/s40257-013-0021-0](#)] [Medline: [23625802](#)]

5. Sultana J, Cutroneo P, Trifirò G. Clinical and economic burden of adverse drug reactions. *J Pharmacol Pharmacother* 2013 Dec;4(Suppl 1):S73-S77 [FREE Full text] [doi: [10.4103/0976-500X.120957](https://doi.org/10.4103/0976-500X.120957)] [Medline: [24347988](https://pubmed.ncbi.nlm.nih.gov/24347988/)]
6. Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. *Arch Intern Med* 2009 Oct 26;169(19):1756-1761. [doi: [10.1001/archinternmed.2009.306](https://doi.org/10.1001/archinternmed.2009.306)] [Medline: [19858432](https://pubmed.ncbi.nlm.nih.gov/19858432/)]
7. Harpaz R, DuMouchel W, Schuemie M, Bodenreider O, Friedman C, Horvitz E, et al. Toward multimodal signal detection of adverse drug reactions. *J Biomed Inform* 2017 Dec;76:41-49 [FREE Full text] [doi: [10.1016/j.jbi.2017.10.013](https://doi.org/10.1016/j.jbi.2017.10.013)] [Medline: [29081385](https://pubmed.ncbi.nlm.nih.gov/29081385/)]
8. Hazell L, Shakir S. Under-reporting of adverse drug reactions: A systematic review. *Drug Saf* 2006;29(5):385-396. [doi: [10.2165/00002018-200629050-00003](https://doi.org/10.2165/00002018-200629050-00003)] [Medline: [16689555](https://pubmed.ncbi.nlm.nih.gov/16689555/)]
9. Sarker A, Ginn R, Nikfarjam A, O'Connor K, Smith K, Jayaraman S, et al. Utilizing social media data for pharmacovigilance: A review. *J Biomed Inform* 2015 Apr;54:202-212 [FREE Full text] [doi: [10.1016/j.jbi.2015.02.004](https://doi.org/10.1016/j.jbi.2015.02.004)] [Medline: [25720841](https://pubmed.ncbi.nlm.nih.gov/25720841/)]
10. Platt R, Wilson M, Chan K, Benner J, Marchibroda J, McClellan M. The new Sentinel Network: Improving the evidence of medical-product safety. *N Engl J Med* 2009 Aug 13;361(7):645-647. [doi: [10.1056/NEJMp0905338](https://doi.org/10.1056/NEJMp0905338)] [Medline: [19635947](https://pubmed.ncbi.nlm.nih.gov/19635947/)]
11. Stang P, Ryan P, Racoosin J, Overhage J, Hartzema A, Reich C, et al. Advancing the science for active surveillance: Rationale and design for the Observational Medical Outcomes Partnership. *Ann Intern Med* 2010 Nov 02;153(9):600-606. [doi: [10.7326/0003-4819-153-9-201011020-00010](https://doi.org/10.7326/0003-4819-153-9-201011020-00010)] [Medline: [21041580](https://pubmed.ncbi.nlm.nih.gov/21041580/)]
12. Norén G, Hopstadius J, Bate A, Star K, Edwards I. Temporal pattern discovery in longitudinal electronic patient records. *Data Min Knowl Discov* 2010 May;20(3):361-387. [doi: [10.1007/s10618-009-0152-3](https://doi.org/10.1007/s10618-009-0152-3)]
13. Wang W, Haerian K, Salmasian H, Harpaz R, Chase H, Friedman C. A drug-adverse event extraction algorithm to support pharmacovigilance knowledge mining from PubMed citations. *AMIA Annu Symp Proc* 2011;2011:1464-1470 [FREE Full text] [Medline: [22195210](https://pubmed.ncbi.nlm.nih.gov/22195210/)]
14. Gurulingappa H, Mateen-Rajput A, Toldo L. Extraction of potential adverse drug events from medical case reports. *J Biomed Semantics* 2012 Dec 20;3(1):15 [FREE Full text] [doi: [10.1186/2041-1480-3-15](https://doi.org/10.1186/2041-1480-3-15)] [Medline: [23256479](https://pubmed.ncbi.nlm.nih.gov/23256479/)]
15. White R, Tatonetti N, Shah N, Altman R, Horvitz E. Web-scale pharmacovigilance: Listening to signals from the crowd. *J Am Med Inform Assoc* 2013 May 01;20(3):404-408 [FREE Full text] [doi: [10.1136/amiajnl-2012-001482](https://doi.org/10.1136/amiajnl-2012-001482)] [Medline: [23467469](https://pubmed.ncbi.nlm.nih.gov/23467469/)]
16. Leaman R, Wojtulewicz L, Sullivan R, Skariah A, Yang J, Gonzalez G. Towards Internet-age pharmacovigilance: Extracting adverse drug reactions from user posts to health-related social networks. In: *Proceedings of the 2010 Workshop on Biomedical Natural Language Processing*. 2010 Presented at: 2010 Workshop on Biomedical Natural Language Processing; July 15, 2010; Uppsala, Sweden p. 117-125 URL: <https://www.aclweb.org/anthology/W10-1915>
17. Nikfarjam A, Sarker A, O'Connor K, Ginn R, Gonzalez G. Pharmacovigilance from social media: Mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. *J Am Med Inform Assoc* 2015 May;22(3):671-681 [FREE Full text] [doi: [10.1093/jamia/ocu041](https://doi.org/10.1093/jamia/ocu041)] [Medline: [25755127](https://pubmed.ncbi.nlm.nih.gov/25755127/)]
18. Berrewaerts J, Delbecq L, Orban P, Deseilles M. Patient participation and the use of eHealth tools for pharmacovigilance. *Front Pharmacol* 2016;7:90 [FREE Full text] [doi: [10.3389/fphar.2016.00090](https://doi.org/10.3389/fphar.2016.00090)] [Medline: [27148052](https://pubmed.ncbi.nlm.nih.gov/27148052/)]
19. Bernardo T, Rajic A, Young I, Robiadek K, Pham M, Funk J. Scoping review on search queries and social media for disease surveillance: A chronology of innovation. *J Med Internet Res* 2013 Jul 18;15(7):e147 [FREE Full text] [doi: [10.2196/jmir.2740](https://doi.org/10.2196/jmir.2740)] [Medline: [23896182](https://pubmed.ncbi.nlm.nih.gov/23896182/)]
20. 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](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](https://pubmed.ncbi.nlm.nih.gov/19329408/)]
21. Lamos V, Cristianini N. Tracking the flu pandemic by monitoring the social Web. In: *Proceedings of the 2nd International Workshop on Cognitive Information Processing*. 2010 Oct 14 Presented at: 2nd International Workshop on Cognitive Information Processing; June 14-16, 2010; Tuscany, Italy p. 411-416. [doi: [10.1109/CIP.2010.5604088](https://doi.org/10.1109/CIP.2010.5604088)]
22. Lamb A, Paul M, Dredze M. Separating fact from fear?: Tracking flu infections on Twitter. In: *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2013 Presented at: 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; June 9-14, 2013; Atlanta, GA p. 789-795.
23. Dredze M, Paul M, Bergsma S, Tran H. Carmen: A Twitter geolocation system with applications to public health. In: *Proceedings of the Workshops at the Twenty-Seventh AAAI Conference on Artificial Intelligence*. 2013 Jul 14 Presented at: Workshops at the Twenty-Seventh AAAI Conference on Artificial Intelligence; July 14-15, 2013; Bellevue, WA p. 20-24.
24. Merolli M, Gray K, Martin-Sanchez F. Health outcomes and related effects of using social media in chronic disease management: A literature review and analysis of affordances. *J Biomed Inform* 2013 Dec;46(6):957-969 [FREE Full text] [doi: [10.1016/j.jbi.2013.04.010](https://doi.org/10.1016/j.jbi.2013.04.010)] [Medline: [23702104](https://pubmed.ncbi.nlm.nih.gov/23702104/)]
25. Naslund J, Aschbrenner K, Marsch L, Bartels S. The future of mental health care: Peer-to-peer support and social media. *Epidemiol Psychiatr Sci* 2016 Apr;25(2):113-122 [FREE Full text] [doi: [10.1017/S2045796015001067](https://doi.org/10.1017/S2045796015001067)] [Medline: [26744309](https://pubmed.ncbi.nlm.nih.gov/26744309/)]

26. Yang C, Yang H, Jiang L, Zhang M. Social media mining for drug safety signal detection. In: Proceedings of the 2012 International Workshop on Smart Health and Well-Being. 2012 Oct 29 Presented at: 2012 International Workshop on Smart Health and Well-Being; October 29, 2012; Maui, Hawaii. [doi: [10.1145/2389707.2389714](https://doi.org/10.1145/2389707.2389714)]
27. Inspire. URL: <https://www.inspire.com/> [accessed 2019-04-21] [WebCite Cache ID 77nwNTHUe]
28. Balagula Y, Rosen S, Lacouture M. The emergence of supportive oncodermatology: The study of dermatologic adverse events to cancer therapies. *J Am Acad Dermatol* 2011 Sep;65(3):624-635. [doi: [10.1016/j.jaad.2010.06.051](https://doi.org/10.1016/j.jaad.2010.06.051)] [Medline: [21777992](https://pubmed.ncbi.nlm.nih.gov/21777992/)]
29. Evans SJW, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol Drug Saf* 2001;10:483-486 [FREE Full text] [doi: [10.1002/pds.677](https://doi.org/10.1002/pds.677)] [Medline: [11828828](https://pubmed.ncbi.nlm.nih.gov/11828828/)]
30. DailyStrength. URL: <https://www.dailystrength.org/> [accessed 2019-04-21] [WebCite Cache ID 77nwVn5WP]
31. Ginn R, Pimpalkhute P, Nikfarjam A, Patki A, Oconnor K, Sarker A, et al. Mining Twitter for adverse drug reaction mentions: A corpus and classification benchmark. In: Proceedings of the Fourth Workshop on Building and Evaluating Resources for Health and Biomedical Text Processing (BioTxtM). 2014 Presented at: Fourth Workshop on Building and Evaluating Resources for Health and Biomedical Text Processing (BioTxtM); May 31, 2014; Reykjavik, Iceland p. 1-8.
32. O'Connor K, Pimpalkhute P, Nikfarjam A, Ginn R, Smith K, Gonzalez G. Pharmacovigilance on Twitter? Mining tweets for adverse drug reactions. In: Proceedings of the American Medical Informatics Association (AMIA) Annual Symposium. 2014 Presented at: American Medical Informatics Association (AMIA) Annual Symposium; November 15-19, 2014; Washington, DC p. 924-933.
33. Hripcsak G, Rothschild A. Agreement, the f-measure, and reliability in information retrieval. *J Am Med Inform Assoc* 2005;12(3):296-298 [FREE Full text] [doi: [10.1197/jamia.M1733](https://doi.org/10.1197/jamia.M1733)] [Medline: [15684123](https://pubmed.ncbi.nlm.nih.gov/15684123/)]
34. Lucene. URL: <http://lucene.apache.org/> [accessed 2019-04-21] [WebCite Cache ID 77nwrX4hR]
35. Huang M, Liu J, Zhu X. GeneTUKit: A software for document-level gene normalization. *Bioinformatics* 2011 Apr 01;27(7):1032-1033 [FREE Full text] [doi: [10.1093/bioinformatics/btr042](https://doi.org/10.1093/bioinformatics/btr042)] [Medline: [21303863](https://pubmed.ncbi.nlm.nih.gov/21303863/)]
36. Leaman R, Islamaj Dogan R, Lu Z. DNorm: Disease name normalization with pairwise learning to rank. *Bioinformatics* 2013 Nov 15;29(22):2909-2917 [FREE Full text] [doi: [10.1093/bioinformatics/btt474](https://doi.org/10.1093/bioinformatics/btt474)] [Medline: [23969135](https://pubmed.ncbi.nlm.nih.gov/23969135/)]
37. van Puijenbroek EP, Bate A, Leufkens H, Lindquist M, Orre R, Egberts A. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. *Pharmacoepidemiol Drug Saf* 2002;11(1):3-10. [doi: [10.1002/pds.668](https://doi.org/10.1002/pds.668)] [Medline: [11998548](https://pubmed.ncbi.nlm.nih.gov/11998548/)]
38. Schuemie M, Ryan P, DuMouchel W, Suchard M, Madigan D. Interpreting observational studies: Why empirical calibration is needed to correct p-values. *Stat Med* 2014 Jan 30;33(2):209-218 [FREE Full text] [doi: [10.1002/sim.5925](https://doi.org/10.1002/sim.5925)] [Medline: [23900808](https://pubmed.ncbi.nlm.nih.gov/23900808/)]
39. Schuemie M, Coloma P, Straatman H, Herings R, Trifirò G, Matthews J, et al. Using electronic health care records for drug safety signal detection: A comparative evaluation of statistical methods. *Med Care* 2012 Oct;50(10):890-897. [doi: [10.1097/MLR.0b013e31825f63bf](https://doi.org/10.1097/MLR.0b013e31825f63bf)] [Medline: [22929992](https://pubmed.ncbi.nlm.nih.gov/22929992/)]
40. Kuhn M, Campillos M, Letunic I, Jensen L, Bork P. A side effect resource to capture phenotypic effects of drugs. *Mol Syst Biol* 2010;6:343 [FREE Full text] [doi: [10.1038/msb.2009.98](https://doi.org/10.1038/msb.2009.98)] [Medline: [20087340](https://pubmed.ncbi.nlm.nih.gov/20087340/)]
41. Martínez de Lagrán Z, Ratón JA, Lasa O, Acebo E, Díaz-Pérez JL. Acneiform eruption from epidermal growth factor receptor inhibitors [Article in Spanish]. *Actas Dermosifiliogr* 2005 Sep;96(7):450-454. [Medline: [16476273](https://pubmed.ncbi.nlm.nih.gov/16476273/)]
42. Roé E, García Muret MP, Marcuello E, Capdevila J, Pallarés C, Alomar A. Description and management of cutaneous side effects during cetuximab or erlotinib treatments: A prospective study of 30 patients. *J Am Acad Dermatol* 2006 Sep;55(3):429-437. [doi: [10.1016/j.jaad.2006.04.062](https://doi.org/10.1016/j.jaad.2006.04.062)] [Medline: [16908348](https://pubmed.ncbi.nlm.nih.gov/16908348/)]
43. Ohtsuka M, Miura T, Mori T, Ishikawa M, Yamamoto T. Occurrence of psoriasiform eruption during nivolumab therapy for primary oral mucosal melanoma. *JAMA Dermatol* 2015 Jul;151(7):797-799. [doi: [10.1001/jamadermatol.2015.0249](https://doi.org/10.1001/jamadermatol.2015.0249)] [Medline: [25875052](https://pubmed.ncbi.nlm.nih.gov/25875052/)]
44. Totonchy M, Ezaldein H, Ko C, Choi J. Inverse psoriasiform eruption during pembrolizumab therapy for metastatic melanoma. *JAMA Dermatol* 2016 Dec 01;152(5):590-592. [doi: [10.1001/jamadermatol.2015.5210](https://doi.org/10.1001/jamadermatol.2015.5210)] [Medline: [26675815](https://pubmed.ncbi.nlm.nih.gov/26675815/)]
45. Carlos G, Anforth R, Chou S, Clements A, Fernandez-Peñas P. A case of bullous pemphigoid in a patient with metastatic melanoma treated with pembrolizumab. *Melanoma Res* 2015 Jun;25(3):265-268. [doi: [10.1097/CMR.000000000000155](https://doi.org/10.1097/CMR.000000000000155)] [Medline: [25831416](https://pubmed.ncbi.nlm.nih.gov/25831416/)]
46. Naidoo J, Schindler K, Querfeld C, Busam K, Cunningham J, Page D, et al. Autoimmune bullous skin disorders with immune checkpoint inhibitors targeting PD-1 and PD-L1. *Cancer Immunol Res* 2016 Dec;4(5):383-389 [FREE Full text] [doi: [10.1158/2326-6066.CIR-15-0123](https://doi.org/10.1158/2326-6066.CIR-15-0123)] [Medline: [26928461](https://pubmed.ncbi.nlm.nih.gov/26928461/)]
47. Pérez-Soler R. Can rash associated with HER1/EGFR inhibition be used as a marker of treatment outcome? *Oncology (Williston Park)* 2003 Nov;17(11 Suppl 12):23-28 [FREE Full text] [Medline: [14682120](https://pubmed.ncbi.nlm.nih.gov/14682120/)]
48. Lee S, Khan I, Upadhyay S, Lewanski C, Falk S, Skales G, et al. First-line erlotinib in patients with advanced non-small-cell lung cancer unsuitable for chemotherapy (TOPICAL): A double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol* 2012 Nov;13(11):1161-1170 [FREE Full text] [doi: [10.1016/S1470-2045\(12\)70412-6](https://doi.org/10.1016/S1470-2045(12)70412-6)] [Medline: [23078958](https://pubmed.ncbi.nlm.nih.gov/23078958/)]

49. Hua C, Boussemart L, Mateus C, Routier E, Boutros C, Cazenave H, et al. Association of vitiligo with tumor response in patients with metastatic melanoma treated with pembrolizumab. *JAMA Dermatol* 2016 Jan;152(1):45-51. [doi: [10.1001/jamadermatol.2015.2707](https://doi.org/10.1001/jamadermatol.2015.2707)] [Medline: [26501224](https://pubmed.ncbi.nlm.nih.gov/26501224/)]
50. Sanlorenzo M, Vujic I, Daud A, Algazi A, Gubens M, Luna S, et al. Pembrolizumab cutaneous adverse events and their association with disease progression. *JAMA Dermatol* 2015 Nov;151(11):1206-1212 [[FREE Full text](#)] [doi: [10.1001/jamadermatol.2015.1916](https://doi.org/10.1001/jamadermatol.2015.1916)] [Medline: [26222619](https://pubmed.ncbi.nlm.nih.gov/26222619/)]
51. Saga K, Jimbow K. Immunohistochemical localization of activated EGF receptor in human eccrine and apocrine sweat glands. *J Histochem Cytochem* 2001 May;49(5):597-602. [doi: [10.1177/002215540104900506](https://doi.org/10.1177/002215540104900506)] [Medline: [11304797](https://pubmed.ncbi.nlm.nih.gov/11304797/)]
52. Albanell J, Rojo F, Averbuch S, Feyerislova A, Mascaro J, Herbst R, et al. Pharmacodynamic studies of the epidermal growth factor receptor inhibitor ZD1839 in skin from cancer patients: Histopathologic and molecular consequences of receptor inhibition. *J Clin Oncol* 2002 Jan 01;20(1):110-124. [doi: [10.1200/JCO.2002.20.1.110](https://doi.org/10.1200/JCO.2002.20.1.110)] [Medline: [11773160](https://pubmed.ncbi.nlm.nih.gov/11773160/)]
53. Vargas G, Fantino E, George-Nascimento C, Gargus J, Haigler H. Reduced epidermal growth factor receptor expression in hypohidrotic ectodermal dysplasia and Tabby mice. *J Clin Invest* 1996 Jun 01;97(11):2426-2432 [[FREE Full text](#)] [doi: [10.1172/JCI118689](https://doi.org/10.1172/JCI118689)] [Medline: [8647934](https://pubmed.ncbi.nlm.nih.gov/8647934/)]
54. Emadzadeh E, Sarker A, Nikfarjam A, Gonzalez G. Hybrid semantic analysis for mapping adverse drug reaction mentions in tweets to medical terminology. *AMIA Annu Symp Proc* 2017;2017:679-688 [[FREE Full text](#)] [Medline: [29854133](https://pubmed.ncbi.nlm.nih.gov/29854133/)]

Abbreviations

ADR: adverse drug reaction
CUI: Concept Unique Identifier
DHM: DeepHealthMiner
EGFR: epidermal growth factor receptor
FDA: US Food and Drug Administration
NER: named entity recognition
PCOS: polycystic ovary syndrome
PD-1: programmed cell death 1
PRR: proportional reporting ratio
SIDER: side-effect resource
SRS: spontaneous reporting system
UMLS: Unified Medical Language System

Edited by G Eysenbach; submitted 08.06.18; peer-reviewed by K Verspoor, A Lambert; comments to author 04.10.18; revised version received 27.02.19; accepted 04.04.19; published 03.06.19.

Please cite as:

Nikfarjam A, Ransohoff JD, Callahan A, Jones E, Loew B, Kwong BY, Sarin KY, Shah NH
Early Detection of Adverse Drug Reactions in Social Health Networks: A Natural Language Processing Pipeline for Signal Detection
JMIR Public Health Surveill 2019;5(2):e11264
URL: <http://publichealth.jmir.org/2019/2/e11264/>
doi: [10.2196/11264](https://doi.org/10.2196/11264)
PMID: [31162134](https://pubmed.ncbi.nlm.nih.gov/31162134/)

©Azadeh Nikfarjam, Julia D Ransohoff, Alison Callahan, Erik Jones, Brian Loew, Bernice Y Kwong, Kavita Y Sarin, Nigam H Shah. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 03.06.2019. 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

Detection of Spatiotemporal Prescription Opioid Hot Spots With Network Scan Statistics: Multistate Analysis

Arinjoy Basak¹, BE; Jose Cadena², PhD; Achla Marathe^{3,4}, PhD; Anil Vullikanti^{3,5}, PhD

¹Biocomplexity Institute of Virginia Tech, Blacksburg, VA, United States

²Lawrence Livermore National Laboratory, Livermore, CA, United States

³Network Systems Science and Advanced Computing Division, Biocomplexity Institute and Initiative, University of Virginia, Charlottesville, VA, United States

⁴Department of Public Health Sciences, University of Virginia, Charlottesville, VA, United States

⁵Department of Computer Science, University of Virginia, Charlottesville, VA, United States

Corresponding Author:

Anil Vullikanti, PhD

Network Systems Science and Advanced Computing Division

Biocomplexity Institute and Initiative

University of Virginia

PO Box 400298

Charlottesville, VA, 22904

United States

Phone: 1 434 982 2200

Fax: 1 434 982 2214

Email: vsakumar@virginia.edu

Abstract

Background: Overuse and misuse of prescription opioids have become significant public health burdens in the United States. About 11.5 million people are estimated to have misused prescription opioids for nonmedical purposes in 2016. This has led to a significant number of drug overdose deaths in the United States. Previous studies have examined spatiotemporal clusters of opioid misuse, but they have been restricted to circular shaped regions.

Objective: The goal of this study was to identify spatiotemporal hot spots of opioid users and opioid prescription claims using Medicare data.

Methods: We examined spatiotemporal clusters with significantly higher number of beneficiaries and rate of prescriptions for opioids using Medicare payment data from the Centers for Medicare & Medicaid Services. We used network scan statistics to detect significant clusters with arbitrary shapes, the Kulldorff scan statistic to examine the significant clusters for each year (2013, 2014, and 2015) and an expectation-based version to examine the significant clusters relative to past years. Regression analysis was used to characterize the demographics of the counties that are a part of any significant cluster, and data mining techniques were used to discover the specialties of the anomalous providers.

Results: We examined anomalous spatial clusters with respect to opioid prescription claims and beneficiary counts and found some common patterns across states: the counties in the most anomalous clusters were fairly stable in 2014 and 2015, but they have shrunk from 2013. In Virginia, a higher percentage of African Americans in a county lower the odds of the county being anomalous in terms of opioid beneficiary counts to about 0.96 in 2015. For opioid prescription claim counts, the odds were 0.92. This pattern was consistent across the 3 states and across the 3 years. A higher number of people in the county with access to Medicaid increased the odds of the county being in the anomalous cluster to 1.16 in both types of counts in Virginia. A higher number of people with access to direct purchase of insurance plans decreased the odds of a county being in an anomalous cluster to 0.85. The expectation-based scan statistic, which captures change over time, revealed different clusters than the Kulldorff statistic. Providers with an unusually high number of opioid beneficiaries and opioid claims include specialties such as physician's assistant, nurse practitioner, and family practice.

Conclusions: Our analysis of the Medicare claims data provides characteristics of the counties and provider specialties that have higher odds of being anomalous. The empirical analysis identifies highly refined spatial hot spots that are likely to encounter prescription opioid misuse and overdose. The methodology is generic and can be applied to monitor providers and their prescription behaviors in regions that are at a high risk of abuse.

KEYWORDS

opiate dependence; spatial temporal analysis; network scan statistics; medical specialty

Introduction

Background

Overuse and misuse of prescription opioids have become significant public health burdens in the United States [1]. More than one thirds of the US civilian population is reported to have used prescription opioids in 2016, as per estimates from the National Survey on Drug Use and Health (NSDUH) [2]. One third of Medicare Part D beneficiaries received prescription opioids in 2017 [3]. Furthermore, about 11.5 million people are estimated to have used prescription opioids for nonmedical purposes in 2016 [1,2]. Pharmaceutical opioids are now among the most popular drugs for nonmedical use in the United States, second only to marijuana [4]. This has led to problems of drug overdose and deaths. According to the Centers for Disease Control and Prevention (CDC), there were more than 63,600 drug overdose deaths in the United States in 2016 [5], and there were over 70,000 deaths in 2017.

One of the major factors driving the opioid crisis is the misuse of prescription opioids; about 40.6% of people reported obtaining them from family or friends [1]. According to the International Narcotics Control Board, in 2009, the United States consumed 99% of the world's hydrocodone, 60% of the world's hydromorphone, and 81% of the world's oxycodone. The number of prescriptions written for opioids increased by 300% between 1991 and 2009 [6], and research by Wisniewski et al (2008) [7] shows high correlations among admissions to addiction treatment facilities, overdose deaths, and the volume of opioids prescribed in the United States.

Public and private health insurance plans have been criticized for lacking consistency in treating chronic pain conditions and in providing viable alternative medical treatments. For example, an extensive study by Lin et al (2018) [8], which considers 15 Medicaid plans, 15 Medicare Advantage plans, and 20 commercial insurers in 2017, which covers half the US population, finds that their health insurance policies offer inconsistent policy terms for nondrug treatments for chronic lower back pain, and they provide little or no coverage for alternative treatments that have scientific backing.

It has also been shown from NSDUH reports that there are significant racial and demographic patterns associated with opioid misuse, for example, white males and lower socioeconomic characteristics [9]. These findings motivate a careful spatiotemporal analysis of prescription rates for opioids to (1) discover regions with unusually high opioid prescription rates and (2) understand what distinct factors in these regions could be associated with higher usage.

To this end, spatial scan statistics [10-12] are among the most common techniques used for analyzing metrics related with opioid use. Linton et al [13] analyze spatial patterns of emergency and nonemergency calls (911 and 311) related to narcotics in Baltimore, Maryland. They identify clusters of calls

in some neighborhoods, coinciding with urban redevelopment. Brownstein et al [14] use patient data from New Mexico at the level of zip codes to identify locations of clusters. They characterize the structure and locations of significant clusters within the state. Cordes [4] analyzes mortality because of prescription opioids, using the North Carolina State Center for Health Statistics' Injury Free North Carolina database.

There has also been some analysis of prescription data at different spatial resolutions. CDC reports prescription rates for different states and demographics [15]. Prescription rates at a county level are reported using the QuintilesIMS Transactional Data Warehouse data [16]. However, these reports only provide visualizations of the spatial data, but they do not identify clusters that are different from the rest of the state or that might have changed over time.

Objectives

This study aimed to fill this gap by identifying spatiotemporal hot spots of opioid users and opioid prescription claims using Medicare Part D data. It applies network scan statistics to detect significant clusters with arbitrary shapes and an expectation-based version to examine the significant clusters relative to past years. Even though this analysis is focused on Virginia, West Virginia, and North Carolina, the methodology is generic and can be applied to other regions and in other contexts for anomaly detection. Its application to spatial epidemiology can help detect unimmunized or underimmunized clusters of individuals in an otherwise well-vaccinated population [17]. Other possible application areas are detection of suicide clusters, teenage pregnancies, and criminal activity [18,19].

Once the clusters are identified, the census data from those spatial regions are analyzed to characterize the demographic attributes of the counties that are a part of the anomalous clusters.

Methods

Ethical Considerations

This study did not require Institutional Review Board approval, as only publicly available, open source, data were used.

Datasets

Centers for Medicare & Medicaid Services Data

We used the Medicare Provider Utilization and Payment Data: Part D Prescriber Public Use File (PUF), made available by the Centers for Medicare & Medicaid Services (CMS) [20]. The Part D Prescriber PUF contains data on all prescription drugs prescribed by individual physicians and other health care providers and paid for under the Medicare Part D Prescription Drug Program. Each provider is identified by a National Provider Identifier, along with the provider's name and address, which are used to map them to counties in the United States.

For each provider, the dataset gives information on the number of beneficiaries seen and the number of prescriptions dispensed at the provider's direction. The dataset also separately lists the number of prescriptions for opioids and the corresponding number of beneficiaries for such prescriptions. These data are available for 2013, 2014, and 2015.

Demographic Data

We used American Community Survey (ACS) census datasets from the US Census Bureau website [21] to analyze the demographic properties of clusters discovered by network scan statistics. Statistics on race, sex, origin, and number of housing units in the counties were covered by the ACS datasets PEPANNHU and PEPSR6H, respectively. Economic characteristics of the population were covered by the DP03 dataset, which gives information on employment status by type of employment, health insurance coverage, subtypes of income characteristics of the population, and some other economic variables. Income distribution for each county were obtained from the S1901 dataset. Table B27019 in ACS provides health insurance coverage characteristics by 2 age groups, that is, ages 25 to 64 and 65 and above; it further breaks down each group by educational attainment. B27002-3 and C27004-7 datasets provide health insurance coverage by age groups, sex, and the kind of health insurance, that is, (1) public or private, (2) Medicare or Medicaid, and (3) in case of private health insurance, whether the plan is employer supported or directly purchased from a private company or exchange. Each one of these datasets is available at different geographic resolutions, starting with the county-level data. Data are available for the years 2013, 2014, and 2015. Different combinations of variables were extracted from these datasets, and the details of the variables used are also mentioned in the *Characterization of Demographics in Clusters* section.

Detecting Clusters Using Network Scan Statistics

Spatial scan statistics are some of the most commonly used approaches for detecting anomalous clusters from spatial data [10-12], and they have been applied in a number of public health problems, including detection of opioid activity and related metrics [4,13,14,22].

We briefly described the Poisson version of the Kulldorff scan statistic [10], which we use in our analysis. Let ζ denote the set of counties in the state. For each county $i \in \zeta$ and time t (which, in our case, is the year), we consider 2 kinds of counts: (1) $\text{tot}_b(i,t)$ and $\text{tot}_p(i,t)$, which are the total number of beneficiaries and prescriptions, respectively, in county i —these are the *baseline* counts for time t —(2) $\text{opioid}_b(i,t)$ and $\text{opioid}_p(i,t)$, which are the total number of beneficiaries and prescriptions for any kind of opioid drug, respectively, for county i —these are the *event* counts for the county at time t . For a cluster $C \subset \zeta$, we define $\text{tot}_p(C,t) = \sum_{i \in C} \text{tot}_p(i,t)$ and $\text{opioid}_p(C,t) = \sum_{i \in C} \text{opioid}_p(i,t)$. Similarly, $\text{tot}_b(C,t)$ and $\text{opioid}_b(C,t)$ are defined. We will drop the dependence on time t from the notation, when it is clear from the context.

The Kulldorff scan statistic is based on hypothesis testing. The null hypothesis H_0 is that the event counts for all counties i are proportional to their baseline counts:

$$H_0: \text{opioid}_p(i) \sim \text{Poisson}(\mu \cdot \text{tot}_p(i))$$

Where μ is the statewide rate for opioid prescriptions: $\text{opioid}_p(\zeta)/\text{tot}_p(\zeta)$. Under the alternative hypothesis $H_1(C)$ for a cluster $C \subset \zeta$, the prescription event counts for counties within C are generated proportionally to some rate η , whereas event counts for counties outside C are generated with some rate $\mu < \eta$:

$$H_1(C) = \text{opioid}_p(i) \sim \text{Poisson}(\eta \cdot \text{tot}_p(i)), i \in C;$$

$$\text{opioid}_p(i) \sim \text{Poisson}(\mu \cdot \text{tot}_p(i)), i \notin C$$

The Kulldorff scan statistic $F(C)$ is defined as a generalized likelihood ratio,

$$F(C) = \Pr[\text{data}|H_1(C)]/\Pr[\text{data}|H_0(C)]$$

and the objective is to find clusters C with maximum $F(C)$ value, as these clusters are likely to have a significantly higher rate than the statewide average.

One of the popular tools for finding such clusters is the SaTScan software (Martin Kulldorff, SaTScan) [23], which considers clusters with circular or elliptical shapes and returns those with the maximum statistic score.

We also searched for clusters where the number of beneficiaries or prescriptions have changed significantly from 1 year to the next. Due to the temporal aspect of this task, we considered the expectation-based Poisson Scan Statistic [24], which compares event counts observed at the present and the past. Under the null hypothesis, the event counts in year t come from the same model as the counts in the year $t-1$ for each county: $\text{opioid}_p(i,t) \sim \text{Poisson}(\text{opioid}_p(i,t-1))$. Under the alternative, there is a cluster $C \subset \zeta$, where the counts are generated at a higher rate than expected: $\text{opioid}_p(i,t) \sim \text{Poisson}(\mu \cdot \text{opioid}_p(i,t-1))$, for some $\mu > 1$.

For both functions, statistical significance of the clusters was assessed via Monte Carlo sampling. We followed the approach from the study by Kulldorff [10] and generated replicates of the data under the null hypothesis, conditioned on the total event counts. The score of any cluster in the original data is compared with the distribution of the optimal solutions on the replicates to get its P value. For example, a P value of .05 indicates that the $F(C)$ score of the cluster we report is in the top 5% of scores obtained in the null replicates; we used 1000 replicates in our experiments.

Network Scan Statistics

We used an extension of scan statistics that is applied to networks. We defined a graph $G=(\zeta,E)$ on the set of counties, where 2 counties $i,j \in \zeta$ are connected by an edge $(i,j) \in E$ if they share a boundary. We focused on clusters $C \subset \zeta$ that induce a connected subgraph in G —this allows us to consider clusters that might be arbitrarily shaped, instead of being restricted to disks or ellipses. We searched for a connected cluster that maximizes the Kulldorff scan statistic. Computationally, this turns out to be much harder, and we used the approach developed in the study by Cadena et al [25] for optimizing different kinds of scan statistics on networks. This method provably gives clusters that are connected subgraphs, with a bounded size, with the maximum likelihood score. We say that a cluster is *significant* if it has a P value of .05, as computed

using the Monte Carlo method described above. We first find clusters with the highest score $F(C)$, computed using the method in the study by Cadena et al [25], and then consider them only if they are also significant.

Regression Method for Characterizing Counties Contained in Opioid Clusters

We performed a logistic regression analysis to identify demographic factors that could help characterize the composition of counties that are identified as anomalous. Separate regressions were run for each of the years, for each of the 3 states, and for both types of clusters, that is, clusters with higher than normal opioid-beneficiary counts and opioid-prescription claim counts.

In each case, the response variable is constructed from the network scan statistics-based clusters that are found to be anomalous in terms of counts. All counties in the state were labeled 1 if they belonged to an anomalous cluster; otherwise, they were labeled 0. The independent variables were extracted from the US Census data for each state. We started with 30 independent variables that included information on race, income, access to private, public, and employer-based insurance, direct purchase insurance, access to Medicare and Medicaid, education level, number of housing units, number of disabled, employment rate, different levels of income-to-poverty ratios, and distribution of male and females in the county.

We measured cross correlations among these variables and removed variables that had 60% or higher correlation among them. If multiple variables are correlated, only one of them is selected. Removing correlated variables addresses the problem of multicollinearity, which results in standard errors of the coefficients to be large [26-28]. After removing all the correlated variables, only 8 independent variables remained.

These 8 variables are the following: (1) percentage of African Americans in the county, (2) percentage of American Indians in the county, (3) percentage of males in the county, (4) number of housing units in the county, (5) percentage of people with access to Medicaid, (6) percentage of people with access to Medicare, (7) percentage of access to direct purchase insurance, and (8) percentage of households with income-to-poverty ratio less than 0.5. Income-to-poverty ratio measures a family's income with respect to the poverty threshold. If it is less than 1, it implies that the household income is below poverty line. Variable (8) represents the percentage of households in the county that have income-to-poverty ratio less than 0.5; it represents the households who live in extreme poverty.

The regression analysis started with these 8 variables. We use the *stepAIC* function given in R software (The R Foundation) to apply the forward selection and backward elimination method to find the most relevant subset of predictors for the regression [29,30]. This method finds the model that minimizes the information loss, as measured by the Akaike Information Criterion (AIC). The *stepAIC* function begins with a full or null model, and method for stepwise regression can be specified in the direction argument with values *forward*, *backward*, or *both*.

Our method used *both*. For each state, each year, and for each type of counts, we separately trained the model. In each case, a subset of the original 8 independent variables was selected, on the basis of the AIC. The selected variables in each case are shown in the tables with regression results. If a variable is missing altogether, it implies it was not selected in the final model in any of the years for the state. If it is used in only some of the 3 years, the coefficient is shown as '-' in the years the variable is not selected. For a variable to be considered significant, the P value should be less than .1.

Results

Overview

The empirical analysis has been performed on 3 states, namely Virginia, North Carolina, and West Virginia. The results for the state of Virginia are described in detail, but for West Virginia and North Carolina, only a brief summary is provided to avoid repetitive discussion. Interested readers can find detailed results on West Virginia and North Carolina in the [Multimedia Appendix 1](#).

Significant Clusters in Virginia

Location and Characteristics Based on Opioid Beneficiary Counts

[Figure 1](#) shows the clusters for different years, discovered using the network-based approach (left) and SaTScan (Martin Kulldorff, SaTScan; right). The clusters are ranked in nonincreasing order of scan statistic score, $F(C)$, mentioned earlier. We observed changes in the number and composition of clusters over time: there is a large cluster of 30 counties in southwest Virginia in 2013, which shrank in 2014 and 2015. On the other hand, significant clusters in other parts of the state emerged in 2015.

[Figure 2](#) shows the clusters with respect to opioid prescription counts for each year, using the network-based approach. The sizes and composition of these clusters are reported in the [Multimedia Appendix 1](#). For 2013 and 2014, the clusters computed using SaTScan (Martin Kulldorff, SaTScan) cover a large part of the state. For 2014 and 2015, SaTScan (Martin Kulldorff, SaTScan) returned more clusters than the network-based approach. However, the likelihood scores of the clusters from SaTScan (Martin Kulldorff, SaTScan) were lower than those using network scan statistics, as the former only considers circular shapes.

We observed a similar layout in 2013 and 2014, as for beneficiary counts ([Figure 1](#)), the most significant cluster is in southwestern Virginia, with a few clusters in other parts. However, in 2015, the most significant cluster in the southwest part has shrunk and become concentrated in the western part. Furthermore, we find more significant clusters with respect to prescription counts. We also find the top clusters exhibit very similar characteristics with respect to different demographic properties, as reported above for the beneficiary counts.

Figure 1. Clusters discovered in Virginia for 2013 (top), 2014 (middle), and 2015 (bottom), using network-based approach (left) and SaTScan (Martin Kulldorff, SaTScan; right) on opioid beneficiary counts.

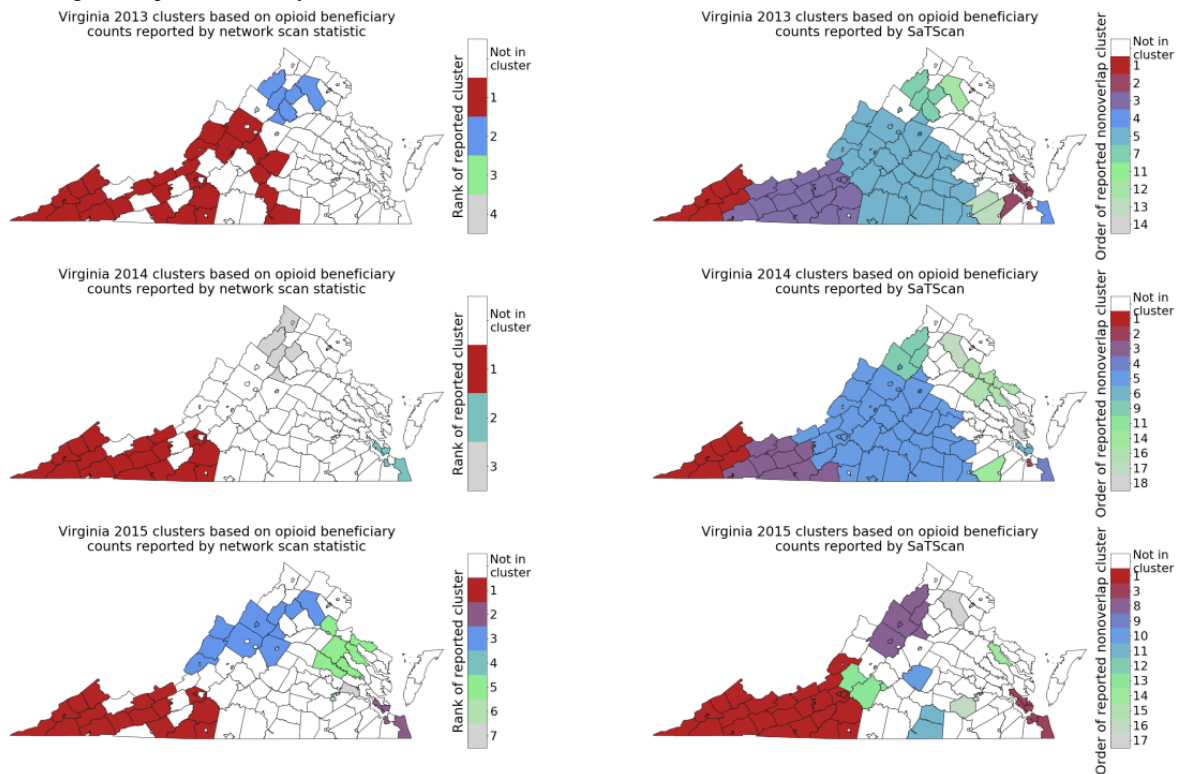
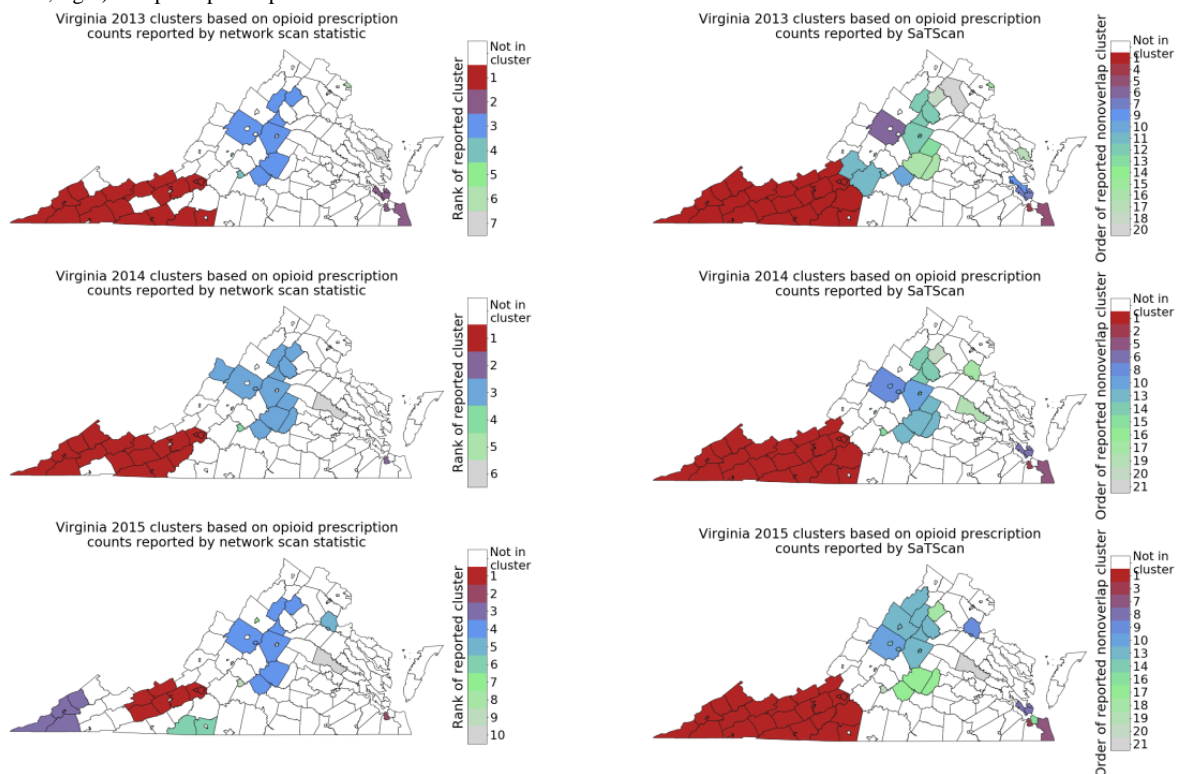


Figure 2. Clusters discovered in Virginia for 2013 (top), 2014 (middle), and 2015 (bottom), using network-based approach (left) and SaTScan (Martin Kulldorff, SaTScan; right) on opioid prescription counts.



Significant Temporal Changes Using the Expectation-Based Scan Statistic

Figure 3 shows the significant clusters using the expectation-based scan statistics for 2015, using an average of

counts from 2013 and 2014 for H_0 . We noted that these are quite different from the clusters computed using the Kulldorff statistics in Figure 1, especially for opioid beneficiary counts. In particular, there is no significant cluster in southwest Virginia in Figure 3, which is consistent with the limited change over

time in Figure 1. Table 1 lists the size and composition of the top 2 clusters. As mentioned above, these are quite different from those reported by the network scan statistic in Figure 1 (left column); the most significant cluster is in central Virginia.

We then focused on the top 10 percentile of the top cluster for years 2013, 2014, and 2015. We selected providers who are in the top 10 percentile of the opioid beneficiary cluster and in the top 10 percentile of the top opioid prescription claims cluster. The scatter plot of Virginia in Figure 4 highlights the specialties of these top providers and plots their 4 attributes, that is, total beneficiaries, percentage of opioid claims, percentage of opioid beneficiaries, and medical specialties. The points that occur close together share common profiles.

Characterizations of Demographics in Clusters

We now provide results from a regression-based analysis, which aims at characterizing all the counties that are a part of the anomalous clusters. Table 2 shows logistic regression-based results for Virginia for the opioid beneficiary data, and Table 3 shows a similar table for opioid prescription claims data. The coefficients represent the odds ratio. The detailed regression results for North Carolina and West Virginia are provided in the Multimedia Appendix 1.

The results show that a high percentage of African Americans in a county lowers the odds of the county being anomalous in terms of high opioid beneficiary count and high prescription count. The odds are fairly stable for opioid prescription claim counts across the 3 years at 0.94, but the odds change from 0.90 in 2013 to 0.96 in 2015 for opioid beneficiary count.

Figure 3. Clusters computed using the expectation-based scan statistic for 2015 using the county-level opioid beneficiary counts (left) and opioid prescription counts (right) in Virginia.

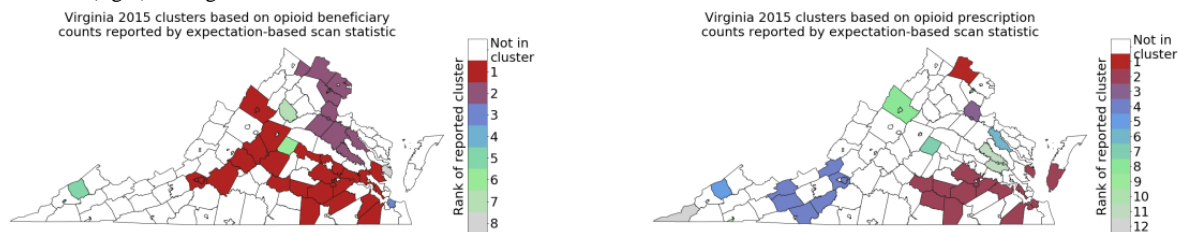


Table 1. Properties of the top 2 significant clusters ($P<.05$) in Virginia for opioid prescription counts and opioid beneficiary counts for the expectation-based scan statistic, for 2015, based on 2013 and 2014.

Beneficiary counts clusters	Prescription claim counts clusters
27 (Harrisonburg city, Rockingham, Albemarle, Nelson, Gloucester, Buckingham, Charles City, Goochland, Henrico, James City, Chesterfield, Isle of Wight, Suffolk city, Newport News city, Prince George, Petersburg city, Brunswick, Nottoway, Dinwiddie, Colonial Heights city, Sussex, Prince Edward, Roanoke city, Roanoke, Bedford, Salem city, and Amherst)	1 (Loudon)
13 (Loudoun, Prince William, Manassas city, Fairfax, Clarke, Falls Church city, Arlington, Fredericksburg city, Stafford, Spotsylvania, Caroline, Essex, and King William)	15 (Mathews, Chesterfield, Northampton, Suffolk city, Norfolk city, Newport News city, Portsmouth city, Prince George, Petersburg city, Brunswick, Nottoway, Dinwiddie, Surry, Prince Edward, and Lunenburg)

Figure 4. Scatter Plots for Virginia (VA) for years 2013, 2014, and 2015, showing the distribution of the provider specialties with respect to the mean of the percentage of opioid beneficiaries and prescriptions served by providers who are in the top 10 percentile of the providers in the anomalous clusters.

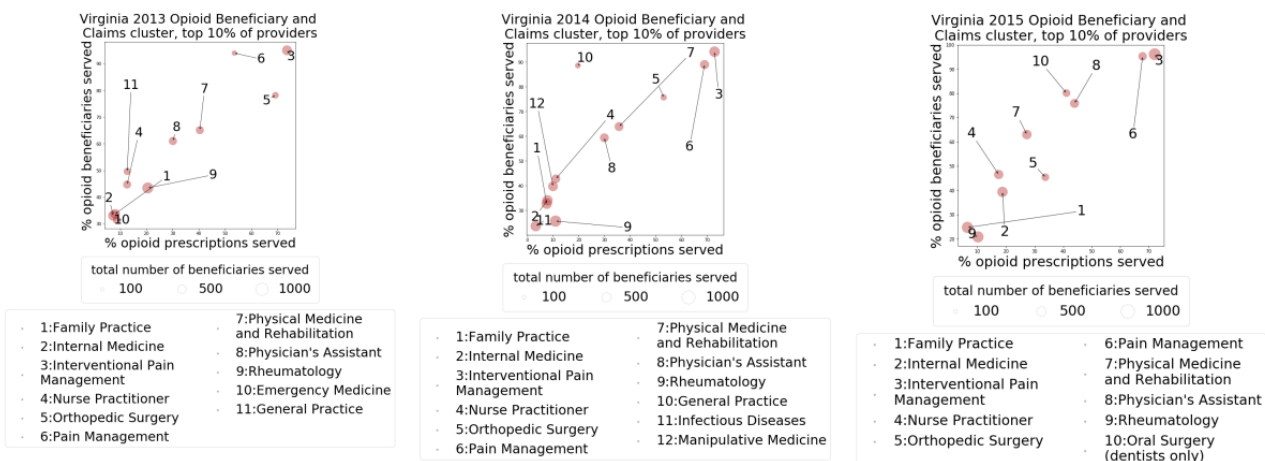


Table 2. Logistic regression results for Virginia opioid beneficiary data. The response variable takes value 1 if the county belongs to an anomalous cluster-based on the opioid beneficiary counts and it takes value 0 if it does not. Empty cells refer to cases when a variable was not selected in the year by AIC.

Variables	2013		2014		2015	
	Coefficient	P value	Coefficient	P value	Coefficient	P value
Intercept	0.32	.22	19.10	.16	0.47	.47
%African American	0.90	<.001 ^a	0.93	.001 ^a	0.96	.005 ^b
%American Indian	0.13	.15	0.004	.01 ^b	1.55	.20
%AccessToMedicaid	1.16	.003 ^b	0.14	.009 ^b	—	—
%AccessToMedicare	—	—	—	—	1.08	.03 ^c
IncomePoverty <0.5	1.09	.13	—	—	1.09	.06 ^c
%AccessToDirCare	—	—	0.83	.05 ^c	0.89	.07 ^c

^aSignificance code ≤ 0.001 .

^bSignificance code ≤ 0.01 .

^cSignificance code ≤ 0.10 .

Table 3. Logistic regression results for Virginia opioid prescription claims data. The response variable takes value 1 if the county belongs to an anomalous cluster based on the opioid prescription claim counts, and it takes value 0 if it does not. Empty cells refer to cases when a variable was not selected in the year by AIC.

Variables	2013		2014		2015	
	Coefficient	P value	Coefficient	P value	Coefficient	P value
Intercept	0.27	.12	0.34	.22	3.19	.38
%AfricanAmerican	0.94	.001 ^a	0.92	<.001 ^a	0.94	.002 ^b
%AmericanIndian	0.13	.14	0.06	.06 ^c	0.18	.17
%AccessToMedicaid	1.15	.001 ^a	1.19	<.001 ^a	1.15	.009 ^b
%AccessToMedicare	—	—	—	—	0.88	.04 ^c
NumHousingUnits	1.0	.12	—	—	—	—

^aSignificance code ≤ 0.001 .

^bSignificance code ≤ 0.01 .

^cSignificance code ≤ 0.10 .

Results also show that a higher number of people in the county with access to Medicaid increases the odds of the county being anomalous to about 1.16 in terms of high opioid beneficiary count and high prescription count in Virginia. This relationship is persistent across years.

Similar to Medicaid, access to Medicare also positively correlates with the odds of a county having a higher number of opioid beneficiaries in Virginia, but it lowers the odds of a county being anomalous in opioid prescription claim count. This relationship only holds in 2015; in 2013 and 2014, there is no relationship between access to Medicare and the 2 types of counts.

Income-to-poverty ratio <0.5 is positively correlated with the odds of higher opioid beneficiary counts in Virginia, in 2015. In rest of the years, it does not show any relationship. Individuals with income-to-poverty ratio <0.5 and those who receive Medicaid are extremely poor, and both these variables have a positive relationship with the opioid beneficiary count.

The number of housing units in a county is not a significant factor in impacting the odds of opioid beneficiary or opioid prescription claims count in that county in Virginia. A higher percentage of people with access to direct purchase of insurance plans decreases the odds of a county having high opioid beneficiary count. Between 2014 and 2015, these odds change from 0.83 to 0.89.

A similar regression analysis of the expectation-based clusters in 2015 did not find any significant demographic variables that were correlated with the odds of a county being anomalous for both types of counts in all 3 states.

Discussion

Principal Findings

Our analysis highlights several similarities and differences among the spatial clusters in Virginia, North Carolina, and West Virginia, with respect to county-level opioid prescription and

beneficiary counts. Some of our key observations were the following: (1) in all 3 states, *the top cluster* with respect to beneficiary counts was fairly stable in 2014 and 2015, and it has shrunk from 2013; (2) the clusters discovered using the Kulldorff scan statistic are different from the expectation-based statistic, partly because of this reason; (3) the *top clusters* in these states have similarities in terms of demographic characteristics, such as racial features (predominantly white, except in North Carolina), income and employment levels (lower than statewide levels), number of housing units (significantly lower than statewide levels), and health care coverage (generally more people have public insurance, including Medicare and Medicaid, fewer have private insurance, and fewer people have employee paid plans).

The regression analysis, which considers all significantly anomalous clusters in the state, identifies factors that are common among counties that belong to these clusters. The regression results for Virginia show that the percentage of African Americans in each county are negatively correlated with the odds of that county being anomalous in both types of counts. As there is over 95% negative correlation between the percentage of whites in the county and percentage of African Americans in each county, these results also suggest that a higher percentage of whites is positively correlated with the odds of the county being anomalous. These results are consistent with the findings of the President's commission report on opioid crisis [1], the CDC, and NSDUH [2,9], which show that the prevalence of opioid-use disorder (OUD) is the highest among whites (72.29%) and only 9.23% among blacks. Other factors overrepresented among those reporting OUDs are males (57.39%) and low-income individuals (<50,000, 67.12%) [9]. Finally, we observe that some of the predictors exhibit significant changes from 2013 to 2015 in Table 2. For instance, the variable *%AmericanIndian* is significant only in year 2014. The coefficient of *%AccessToMedicaid* is stable and significant across all years in Table 3, but in Table 2, the odds drop from 1.16 in 2013 to 0.14 in 2014. The reasons for this are not clear, and there might be other factors at play, which are not evident in the datasets used in this study.

In Virginia, there is a persistent correlation between a higher number of people in the county with access to Medicaid and the increase in the odds of the county being anomalous in both types of counts. As more people in a county depend on Medicaid services, the odds of having higher counts of opioid beneficiaries and prescriptions claim count go up. This may not be surprising as many of the government-run Medicaid programs have often come under fire for various types of waste, fraud, and abuse [31]. These are often because of improper payments to an ineligible beneficiary or an ineligible service or an ineligible provider. States have little incentive to dedicate their limited resources to track and recover these payments as states only get to keep their share of funding which is less than half. In 2017, US Department of Health and Human Services made \$64 billion in improper payments [32].

The public health insurance programs may be more at risk of prescribing opioids because of lack of consistent policies toward *utilization management* such as *step-therapy* and *quantity limits* [8]. These utilization management rules require that the

treatment start with a less risky drug that may be available over the counter and restrict the number of pills that can be prescribed at any given time. This is also consistent with the findings from previous research that low-income individuals are at a much higher risk of OUDs, as many of the Medicaid recipients are extremely poor [1,9].

Our results also show that network-based scan statistics [25] can give new insights into spatiotemporal hot spots of prescription opioids. The clusters we found were at a higher resolution than the circular clusters computed using SaTScan (Martin Kulldorff, SaTScan), and they have a higher likelihood score; thus, combining our approach with SaTScan (Martin Kulldorff, SaTScan) can be useful in practice. Finally, the Kulldorff and expectation-based scan statistics give different insights. In particular, the expectation-based approach can help identify clusters with significant change over time, which is not easy to determine from the Kulldorff clusters, which seem to be fairly stable in 2014 and 2015.

An analysis of the top 10 percentile of the top cluster in Virginia for years 2013, 2014, and 2015 shows some unusual patterns. We focus on the specialties of the providers who are in the top 10 percentile of the opioid beneficiary cluster and in the top 10 percentile of the top opioid claims cluster. The scatter plot of Virginia in Figure 4 shows that the points that occur close together share common profiles. For example, in years 2013 and 2014, physician assistants' and physical medicine and rehabilitation specialists' profiles were very similar. This seems unusual as these professions attend to very different sets of patients [33,34]. Physical medicine and rehabilitation specialists often treat patients with disabilities affecting the brain, spinal cord, nerves, bones, joints, ligaments, muscles, and tendons, whereas physician assistants have more comprehensive responsibilities, which include preventive health care, primary health care, and counseling of patients, etc. Similarly, in 2015, physician assistants and dental oral surgeons had a very similar profile, which is also unexpected as oral surgeons often deal with tooth extraction, implants, oral cancer, jaw surgery, etc—conditions that require significant pain management on the surgeon's part [35].

Finally, a word cloud-based analysis (not shown here) found that a majority of the opioid beneficiaries in the most anomalous cluster are being served by physician assistants, nurse practitioners, and family practitioners. This holds true across the 3 years. We found this to be curious, as provider services in these specialties are often not focused on pain management, although they may be treating chronic pain [36,37]. High occurrences of these medical specialties in the top cluster may require further investigation by domain experts.

This research provides a generic methodology for identifying highly refined spatial hot spots that are likely to encounter prescription opioid misuse and overdose. This information can be used to monitor providers and their prescription behaviors in regions that are at a high risk of abuse. The anomalous regions can be targeted for better public health surveillance, access to treatment and recovery services, and availability and distribution of overdose-reversing drugs.

Limitations

The Medicare Part D Prescription Drug data (*Methods/Datasets/CMS Data* section) has several limitations in the context of revealing hot spots of prescription opioid overdose. These data only comprise prescriptions paid by Medicare; only about 39.5 million people (a little over 10% of the total population) were covered in 2015 [38]. Furthermore, Medicare only covers a specific part of the population, namely people over 65 and those with certain disabilities. Finally, our analysis only considers counts of prescriptions and beneficiaries associated with opioids—the clusters we find are not necessarily hot spots of opioid overdose.

However, the demographics of people covered by Medicare are often associated with use of prescription opioids (though not necessarily their misuse); furthermore, about 40.6% of people reported obtaining prescription opioids from family or friends [1]. Therefore, we believe such an analysis using the CMS data is still valuable.

Another limitation of this study is that it does not account for spatial correlations of observations, which may exist at spatial boundaries and may impact the estimates of the standard errors.

Conclusions

Though misuse of prescription opioids has been recognized as one of the significant factors driving the opioid overdose crisis, there has been limited understanding of the spatiotemporal characteristics of opioid prescription rates. All previous spatiotemporal analysis of metrics directly or indirectly related to opioid prescription rates and overdose has been restricted to a single state [4,13,14,22]. Our analysis using Medicare prescription rates is the first multistate spatiotemporal analysis of prescription rates for opioids. Even though this does not directly contain opioid overdose data, the clusters are likely to be relevant, as people represented in Medicare are associated with greater use of prescription opioids.

The regression analysis characterizes the demographics of the significant clusters associated with high opioid beneficiaries and prescription claims. Scan statistics is a useful technique for spatiotemporal analysis, but the network-based approach gives higher resolution clusters than SaTScan (Martin Kulldorff, SaTScan); the clusters we found have higher likelihood scores. Furthermore, the expectation-based statistic reveals different clusters than the Kulldorff scan statistic.

Acknowledgments

This study has been partially supported by the National Institutes of Health (grant 1R01GM109718), National Science Foundation (NSF) Research Traineeship (grant NRT-DESE-154362), NSF grant IIS-1633028, NSF grant ACI-1443054, and Defense Threat Reduction Agency (DTRA) grants HDTRA1-11-1-0016 and HDTRA1-17-D-0023. This work was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344 and LLNL-JRNL-772856.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary appendix section to paper with studies on counties in North Carolina and West Virginia.

[PDF File (Adobe PDF File), 967KB - [publichealth_v5i2e12110_app1.pdf](#)]

References

1. Christie GC, Baker CG, Cooper GR, Kennedy CP, Madras B, Bondi FA. The White House. 2017. The President's Commission on Combating Drug Addiction And The Opioid Crisis URL: https://www.whitehouse.gov/sites/whitehouse.gov/files/images/Final_Report_Draft_11-1-2017.pdf [WebCite Cache ID 70XexIOIH]
2. Substance Abuse and Mental Health Services Administration. 2017. Results from the 2016 National Survey on Drug Use and Health: Detailed Tables. Table 1.28A and 1.28B URL: <https://www.samhsa.gov/data/sites/default/files/NSDUH-DefTabs-2016/NSDUH-DefTabs-2016.htm> [WebCite Cache ID 77jmgW9NB]
3. US Department of Health and Human Services. Office of Inspector General. Opioid Use in Medicare Part D Remains Concerning URL: <https://oig.hhs.gov/oei/reports/oei-02-18-00220.pdf> [WebCite Cache ID 74SPgbyAr]
4. Cordes J. Carolina Digital Repository. 2017. Spatial Trends in Opiate Overdose Death in North Carolina: 1999-2015 URL: <https://cdr.lib.unc.edu/indexablecontent/uuid:323a68df-cc42-4191-983c-225cb7134249> [WebCite Cache ID 70XIfUIt]
5. Hedegaard H, Warner M, Miniño AM. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. 2017 Dec. Drug overdose deaths in the United States, 1999-2016 URL: <https://www.cdc.gov/nchs/data/databriefs/db294.pdf> [WebCite Cache ID 70XluCvNe]
6. Lyapustina T, Alexander GC. The Pharmaceutical Journal. 2015 Jun. The prescription opioid addiction and abuse epidemic: how it happened and what we can do about it URL: <https://jhu.pure.elsevier.com/en/publications/the-prescription-opioid-abuse-epidemic-4> [WebCite Cache ID 75vMHIEBM]

7. Wisniewski AM, Purdy CH, Blondell RD. The epidemiologic association between opioid prescribing, non-medical use, and emergency department visits. *J Addict Dis* 2008;27(1):1-11. [doi: [10.1300/J069v27n01_01](https://doi.org/10.1300/J069v27n01_01)] [Medline: [18551883](https://pubmed.ncbi.nlm.nih.gov/18551883/)]
8. Lin DH, Jones CM, Compton WM, Heyward J, Losby JL, Murimi IB, et al. Prescription drug coverage for treatment of low back pain among US Medicaid, Medicare Advantage, and commercial insurers. *JAMA Network Open* 2018 Jun 22;1(2). [doi: [10.1001/jamanetworkopen.2018.0235](https://doi.org/10.1001/jamanetworkopen.2018.0235)] [Medline: [25996397](https://pubmed.ncbi.nlm.nih.gov/25996397/)]
9. Wu LT, Zhu H, Swartz MS. Treatment utilization among persons with opioid use disorder in the United States. *Drug Alcohol Depend* 2016 Dec 1;169:117-127 [FREE Full text] [doi: [10.1016/j.drugalcdep.2016.10.015](https://doi.org/10.1016/j.drugalcdep.2016.10.015)] [Medline: [27810654](https://pubmed.ncbi.nlm.nih.gov/27810654/)]
10. Kulldorff M. A spatial scan statistic. *Commun Stat Theory Methods* 2007 Jun 27;26(6):1481-1496. [doi: [10.1080/03610929708831995](https://doi.org/10.1080/03610929708831995)]
11. Neill D. Fast subset scan for spatial pattern detection. *J R Stat Soc Series B* 2012;74(2):337-360. [doi: [10.1111/j.1467-9868.2011.01014.x](https://doi.org/10.1111/j.1467-9868.2011.01014.x)]
12. Jung I, Kulldorff M, Richard OJ. A spatial scan statistic for multinomial data. *Stat Med* 2010 Aug 15;29(18):1910-1918 [FREE Full text] [doi: [10.1002/sim.3951](https://doi.org/10.1002/sim.3951)] [Medline: [20680984](https://pubmed.ncbi.nlm.nih.gov/20680984/)]
13. Linton SL, Jennings JM, Latkin CA, Gomez MB, Mehta SH. Application of space-time scan statistics to describe geographic and temporal clustering of visible drug activity. *J Urban Health* 2014 Oct;91(5):940-956 [FREE Full text] [doi: [10.1007/s11524-014-9890-7](https://doi.org/10.1007/s11524-014-9890-7)] [Medline: [25078036](https://pubmed.ncbi.nlm.nih.gov/25078036/)]
14. Brownstein JS, Green TC, Cassidy TA, Butler SF. Geographic information systems and pharmacoepidemiology: using spatial cluster detection to monitor local patterns of prescription opioid abuse. *Pharmacoepidemiol Drug Saf* 2010 Jun;19(6):627-637 [FREE Full text] [doi: [10.1002/pds.1939](https://doi.org/10.1002/pds.1939)] [Medline: [20535759](https://pubmed.ncbi.nlm.nih.gov/20535759/)]
15. Centers for Disease Control and Prevention. Opioid Overdose Prescribing Data URL: <https://www.cdc.gov/drugoverdose/data/prescribing.html> [WebCite Cache ID 70Xq7R89i]
16. Centers for Disease Control and Prevention. 2016. U.S. County Prescribing Rates URL: <https://www.cdc.gov/drugoverdose/maps/rxcounty2016.html> [WebCite Cache ID 70XqQ3QOc]
17. Cadena J, Falcone D, Marathe A, Vullikanti A. Discovery of under immunized spatial clusters using network scan statistics. *BMC Med Inform Decis Mak* 2019 Dec 4;19(1):28 [FREE Full text] [doi: [10.1186/s12911-018-0706-7](https://doi.org/10.1186/s12911-018-0706-7)] [Medline: [30717725](https://pubmed.ncbi.nlm.nih.gov/30717725/)]
18. Cheung YT, Spittal MJ, Williamson MK, Tung SJ, Pirkis J. Application of scan statistics to detect suicide clusters in Australia. *PLoS One* 2013 Jan 14;8(1). [doi: [10.1371/journal.pone.0054168](https://doi.org/10.1371/journal.pone.0054168)] [Medline: [23342098](https://pubmed.ncbi.nlm.nih.gov/23342098/)]
19. Runadi T. Application of hotspot detection using spatial scan statistic: study of criminality in Indonesia. *AIP Conf Proc* 2017 Mar;1827(1):020011 [FREE Full text] [doi: [10.1063/1.4979427](https://doi.org/10.1063/1.4979427)]
20. Centers for Medicare & Medicaid Services. 2017. Part D Prescriber Data CY 2015 URL: <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/PartD2015.html> [accessed 2019-04-18] [WebCite Cache ID 77jr1qZbH]
21. FactFinder USCBA. 2013 Jan 1. 2009-2016 American Community Survey URL: <https://factfinder.census.gov/> [WebCite Cache ID 70XqnDIFW]
22. Modarai F, Mack K, Hicks P, Benoit S, Park S, Jones C, et al. Relationship of opioid prescription sales and overdoses, North Carolina. *Drug Alcohol Depend* 2013 Sep 1;132(1-2):81-86. [doi: [10.1016/j.drugalcdep.2013.01.006](https://doi.org/10.1016/j.drugalcdep.2013.01.006)] [Medline: [23399467](https://pubmed.ncbi.nlm.nih.gov/23399467/)]
23. Kulldorff M. SaTScan-Software. 2010. URL: <https://www.satscan.org/> [WebCite Cache ID 70Xr4Bhxx]
24. Neill DB. Expectation-based scan statistics for monitoring spatial time series data. *Int J Forecast* 2009 Jul;25(3):498-517. [doi: [10.1016/j.ijforecast.2008.12.002](https://doi.org/10.1016/j.ijforecast.2008.12.002)]
25. Cadena J, Chen F, Vullikanti A. Near-Optimal and Practical Algorithms for Graph Scan Statistics. In: *SIAM Data Mining*. 2017 Presented at: *SIAM Data Mining (SDM)*; April 27-29, 2017; Houston, Texas, USA. [doi: [10.1137/1.9781611974973.70](https://doi.org/10.1137/1.9781611974973.70)]
26. Hill RC, Adkins LC. Collinearity. In: Baltagi BH, editor. *A Companion to Theoretical Econometrics*. Hoboken, NJ: Blackwell Publishing Ltd; 2001.
27. Kmenta J. *Elements of Econometrics*. New York: Macmillan; 1986.
28. Maddala GS, Lahiri K. *Introduction to Econometrics*. Hoboken, NJ: Wiley, John & Sons, Incorporated; 2010.
29. Bursac Z, Gauss CH, Williams DK, Hosmer DW. Purposeful selection of variables in logistic regression. *Source Code Biol Med* 2008 Dec 16;3:17 [FREE Full text] [doi: [10.1186/1751-0473-3-17](https://doi.org/10.1186/1751-0473-3-17)] [Medline: [19087314](https://pubmed.ncbi.nlm.nih.gov/19087314/)]
30. Zhang Z. Variable selection with stepwise and best subset approaches. *Ann Transl Med* 2016 Apr;4(7):136 [FREE Full text] [doi: [10.21037/atm.2016.03.35](https://doi.org/10.21037/atm.2016.03.35)] [Medline: [27162786](https://pubmed.ncbi.nlm.nih.gov/27162786/)]
31. CBS. 2016 Oct 4. Study: Medicaid abuse persistent as boomers retire URL: <https://www.cbsnews.com/news/medicaid-inspector-general-office-personal-care-aides-fraud-patient-neglect-abuse/> [WebCite Cache ID 70XsCW5AX]
32. Beaton T. HealthPayerIntelligence. 2018 Jun 4. HHS Made Nearly \$90B in Improper Payments to Medicaid, Medicare URL: <https://healthpayerintelligence.com/news/hhs-made-nearly-90b-in-improper-payments-to-medicaid-medicare> [WebCite Cache ID 75vMvSqCp]
33. Wei EN. Harvard Health Publishing. Navigating back pain treatments: can a physiatrist help? URL: <https://www.health.harvard.edu/blog/navigating-back-pain-treatments-can-a-physiatrist-help-2018101014969> [accessed 2019-04-18] [WebCite Cache ID 77jsU4akD]

34. Salcido R, Fisher SB, Reinstein L, Willis JB. Underutilization of physician assistants in physical medicine and rehabilitation. *Arch Phys Med Rehabil* 1993 Aug;74(8):826-829. [Medline: [8102228](#)]
35. Becker DE. Pain management: part 1: managing acute and postoperative dental pain. *Anesth Prog* 2010;57(2):67-78; quiz 79 [FREE Full text] [doi: [10.2344/0003-3006-57.2.67](#)] [Medline: [20553137](#)]
36. Shepperd JR, Hale SE. Nurse practitioners in the world of pain management: a cautionary tale. *J Nurse Pract* 2016 Feb;12(2):102-108. [doi: [10.1016/j.nurpra.2015.10.006](#)]
37. Kaasalainen S, DiCenso A, Donald FC, Staples E. Optimizing the role of the nurse practitioner to improve pain management in long-term care. *Can J Nurs Res* 2007 Jun;39(2):14-31. [Medline: [17679583](#)]
38. Centers for Medicare & Medicaid Services. CMS Fast Facts URL: <https://www.cms.gov/fastfacts/> [accessed 2019-04-18] [WebCite Cache ID 77jsOaX51]

Abbreviations

ACS: American Community Survey
AIC: Akaike Information Criterion
CDC: Centers for Disease Control and Prevention
CMS: Centers for Medicare & Medicaid Services
NSDUH: National Survey on Drug Use and Health
NSF: National Science Foundation
OD: opioid-use disorder
PUF: Public Use File

Edited by T Sanchez; submitted 10.09.18; peer-reviewed by P Wang, A Algarin, K Card; comments to author 12.10.18; revised version received 10.12.18; accepted 05.04.19; published 17.06.19.

Please cite as:

Basak A, Cadena J, Marathe A, Vullikanti A

Detection of Spatiotemporal Prescription Opioid Hot Spots With Network Scan Statistics: Multistate Analysis

JMIR Public Health Surveill 2019;5(2):e12110

URL: <http://publichealth.jmir.org/2019/2/e12110/>

doi: [10.2196/12110](https://doi.org/10.2196/12110)

PMID: [31210142](https://pubmed.ncbi.nlm.nih.gov/31210142/)

©Arinjoy Basak, Jose Cadena, Achla Marathe, Anil Vullikanti. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 17.06.2019. 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

Text-Based Illness Monitoring for Detection of Novel Influenza A Virus Infections During an Influenza A (H3N2)v Virus Outbreak in Michigan, 2016: Surveillance and Survey

Rebekah J Stewart^{1,2}, MSN, MPH, NP; John Rossow^{3,4}, MPH, DVM; Seth Eckel⁵, MPH; Sally Bido⁵, MPH; Grant Ballew⁶; Kimberly Signs⁵, MPH, DVM; Julie Thelen Conover⁷, MS; Erin Burns¹, MA; Joseph S Bresee¹, MD, MPH; Alicia M Fry¹, MD, MPH; Sonja J Olsen¹, PhD; Matthew Biggerstaff¹, ScD, MPH

¹Influenza Division, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, GA, United States

²Epidemic Intelligence Service, Centers for Disease Control and Prevention, Atlanta, GA, United States

³Epidemiology Elective Program, Division of Scientific Education and Professional Development, Center for Surveillance, Epidemiology, and Laboratory Services, Centers for Disease Control and Prevention, Atlanta, GA, United States

⁴College of Veterinary Medicine, University of Georgia, Athens, GA, United States

⁵Michigan Department of Health and Human Services, Lansing, MI, United States

⁶Compliant Campaign, Scottsdale, AZ, United States

⁷Michigan State University Extension, East Lansing, MI, United States

Corresponding Author:

Rebekah J Stewart, MSN, MPH, NP

Influenza Division

National Center for Immunization and Respiratory Diseases

Centers for Disease Control and Prevention

1600 Clifton Road NE

MS E-10

Atlanta, GA, 30333

United States

Phone: 1 404 718 4580

Email: yxp5@cdc.gov

Abstract

Background: Rapid reporting of human infections with novel influenza A viruses accelerates detection of viruses with pandemic potential and implementation of an effective public health response. After detection of human infections with influenza A (H3N2) variant (H3N2v) viruses associated with agricultural fairs during August 2016, the Michigan Department of Health and Human Services worked with the US Centers for Disease Control and Prevention (CDC) to identify infections with variant influenza viruses using a text-based illness monitoring system.

Objective: To enhance detection of influenza infections using text-based monitoring and evaluate the feasibility and acceptability of the system for use in future outbreaks of novel influenza viruses.

Methods: During an outbreak of H3N2v virus infections among agricultural fair attendees, we deployed a text-illness monitoring (TIM) system to conduct active illness surveillance among households of youth who exhibited swine at fairs. We selected all fairs with suspected H3N2v virus infections. For fairs without suspected infections, we selected only those fairs that met predefined criteria. Eligible respondents were identified and recruited through email outreach and/or on-site meetings at fairs. During the fairs and for 10 days after selected fairs, enrolled households received daily, automated text-messages inquiring about illness; reports of illness were investigated by local health departments. To understand the feasibility and acceptability of the system, we monitored enrollment and trends in participation and distributed a Web-based survey to households of exhibitors from five fairs.

Results: Among an estimated 500 households with a member who exhibited swine at one of nine selected fairs, representatives of 87 (17.4%) households were enrolled, representing 392 household members. Among fairs that were ongoing when the TIM system was deployed, the number of respondents peaked at 54 on the third day of the fair and then steadily declined throughout the rest of the monitoring period; 19 out of 87 household representatives (22%) responded through the end of the 10-day monitoring period. We detected 2 H3N2v virus infections using the TIM system, which represents 17% (2/12) of all H3N2v virus infections

detected during this outbreak in Michigan. Of the 70 survey respondents, 16 (23%) had participated in the TIM system. A total of 73% (11/15) participated because it was recommended by fair coordinators and 80% (12/15) said they would participate again.

Conclusions: Using a text-message system, we monitored for illness among a large number of individuals and households and detected H3N2v virus infections through active surveillance. Text-based illness monitoring systems are useful for detecting novel influenza virus infections when active monitoring is necessary. Participant retention and testing of persons reporting illness are critical elements for system improvement.

(*JMIR Public Health Surveill* 2019;5(2):e10842) doi:[10.2196/10842](https://doi.org/10.2196/10842)

KEYWORDS

influenza; surveillance; novel; agricultural; fairs; texting

Introduction

Novel influenza A viruses are different from currently circulating human influenza A (H1 and H3) viruses and have the potential to cause a pandemic if viruses gain the capacity to infect and transmit efficiently from person to person and cause clinical illness in humans [1]. In the United States, human infection with a novel influenza A virus is nationally notifiable [1]; swine are the primary source of reported novel influenza A virus infections in humans, and the vast majority of persons are infected after swine exposure at an agricultural fair [2-9]. Some influenza A viruses are endemic pathogens in swine populations [10-13] and swine can be infected without displaying clinical signs of illness [14]. Monitoring for novel influenza virus infections in humans is important to quickly identify viruses with pandemic potential and to speed implementation of an effective public health response. However, traditional forms of active monitoring for illness (eg, through daily phone calls to persons with possible exposure) can be very labor intensive for health department staff; during an outbreak, the need for monitoring may overwhelm a health department's ability to respond in a timely manner. The US Centers for Disease Control and Prevention (CDC) recommends that people participating in avian influenza outbreak response efforts be monitored during the response and for 10 days after their last possible exposure [15], which can result in a substantial number of responders under monitoring per state at any given time. For example, during an avian influenza outbreak in 2014-2015 in Minnesota, USA, public health officials conducted active postexposure symptom monitoring for 459 responders by making daily phone calls for 10 days (Karen Martin, epidemiologist in Minnesota, personal communication). Short message service (SMS) text messaging has been used previously in research [16] and outbreak settings [17] to facilitate the monitoring of multiple individuals for respiratory illness.

To assist US states with monitoring individuals at potential risk of a novel influenza virus infection, CDC, in coordination with the National Association of County and City Health Officials (NACCHO) and Compliant Campaign, developed a text-message illness monitoring (TIM) platform. Although several state health departments had pilot-tested the TIM system, none had previously used it among members of the public during an outbreak. On August 4, 2016, public health officials in Michigan, USA, notified CDC of a laboratory-confirmed influenza A (H3N2) variant (H3N2v) virus infection in a 9-year-old child who had been exhibiting swine at an agricultural

fair in Michigan [18,19]. Public health officials in Michigan reported a second infection the next day along with reports of ill swine at a second agricultural fair.

On August 9, 2016, the Michigan Department of Health and Human Services (MDHHS) asked CDC to pilot-test the TIM system to monitor families of individuals exhibiting swine at county fairs and evaluate the functionality during an outbreak. The objectives of the pilot were to determine the ability of the TIM system to enhance detections of H3N2v virus infections and to evaluate the feasibility and acceptability of the TIM system for use in future outbreak investigations of novel influenza viruses.

Methods

Enrollment, Retention, and Illness Detection

The details of the H3N2v outbreak have been described previously (see [Multimedia Appendix 1](#)) [18]. For the TIM system pilot, we targeted the following:

1. County agricultural fairs that had reported suspected H3N2v infections in the last 10 days during the 2016 H3N2v outbreak.
2. County fairs, to ensure illness follow-up and testing could be coordinated by a single county health department, that started during our enrollment period—August 15 to September 9, 2016—and had large (>50) numbers of swine in exhibit. Large numbers of swine increase the likelihood of swine influenza outbreaks; we chose a number of more than 50 to reach the upper 50th percentile of swine per fair as determined by Bowman and colleagues [20].

We identified three fairs with suspected H3N2v virus infections and approached fair organizers about participating in the TIM pilot. Of those, organizers of two fairs agreed to participate (see [Figure 1](#)). We identified 21 fairs occurring between August 15 and September 9, 2016, from which no human H3N2v infections were reported. Of those, 13 were county fairs and seven of the county fairs had more than 50 exhibition swine. Organizers of these seven fairs agreed to participate in the pilot, bringing the total number of fairs where the TIM system was deployed to nine fairs (see [Figure 1](#)).

As studies have shown that young age and direct contact with swine are risk factors for variant influenza virus infection [4,6,9], we contacted coordinators of youth agricultural clubs to invite households of swine exhibitors to participate. For fairs with reported infections, club coordinators sent recruitment

information solely through email. For the fairs without reported infections, we attended exhibitor meetings or other fair events to describe variant influenza viruses and the TIM system. We also distributed informational flyers about the TIM system and sent recruitment information by email.

We assigned each fair a unique texting code word. Parents of swine exhibitors were asked to text in the unique code word using their mobile phone, which would automatically enroll them in the system (see [Figures 2](#) and [3](#)). After texting the code word, enrollees were asked how many household members were attending or planning to attend the county fair. Next, they were asked if any fair attendees had a fever, cough, sore throat, or other symptoms. Respondents who replied “No” were advised to contact the health department if anyone started to have symptoms. Respondents who replied “Yes” were advised that a representative from the health department would contact them to learn more about the illness. If a response was not received by 6 PM, a reminder text was automatically sent. Texts asking about illness were sent daily at noon during the fair and for 10 days following the fair. Because CDC does not recommend a specific monitoring period for persons exposed to variant influenza viruses, we used the monitoring period recommended by CDC for avian influenza responders [21].

“Yes” responses and any response other than “No” (eg, “OK”) generated an immediate email alert to designated local health

department staff to contact the respondent and determine if additional follow-up was necessary. Health department staff were asked to refer anyone for respiratory testing who reported any symptoms associated with variant influenza virus infection, including mild symptoms such as runny nose and less common symptoms such as vomiting, and who had swine exposure or contact with a person with a confirmed H3N2v virus infection during the week prior to illness onset [6]. Collected specimens underwent real-time reverse-transcription polymerase chain reaction (rRT-PCR) or genetic sequencing at CDC or the state public health laboratory to confirm H3N2v virus isolation.

We monitored retention of respondents throughout the monitoring period and assessed the impact of reminder texts on daily participation. Respondents could opt out at any time by texting “STOP.” Estimates of user retention were based on formal withdrawals and respondents who ceased responding. To understand the impact of the type of texting language on respondent retention, we randomly assigned fairs to one of two groups where attendees would receive texts with either more formal or less formal language (see [Figures 2](#) and [3](#)). We applied a Poisson regression using generalized estimating equations to examine statistically significant differences in daily respondent counts between the two types of texting language. The threshold for statistical significance was set at $P < .05$.

Figure 1. Fair selection. TIM: text-message illness monitoring.

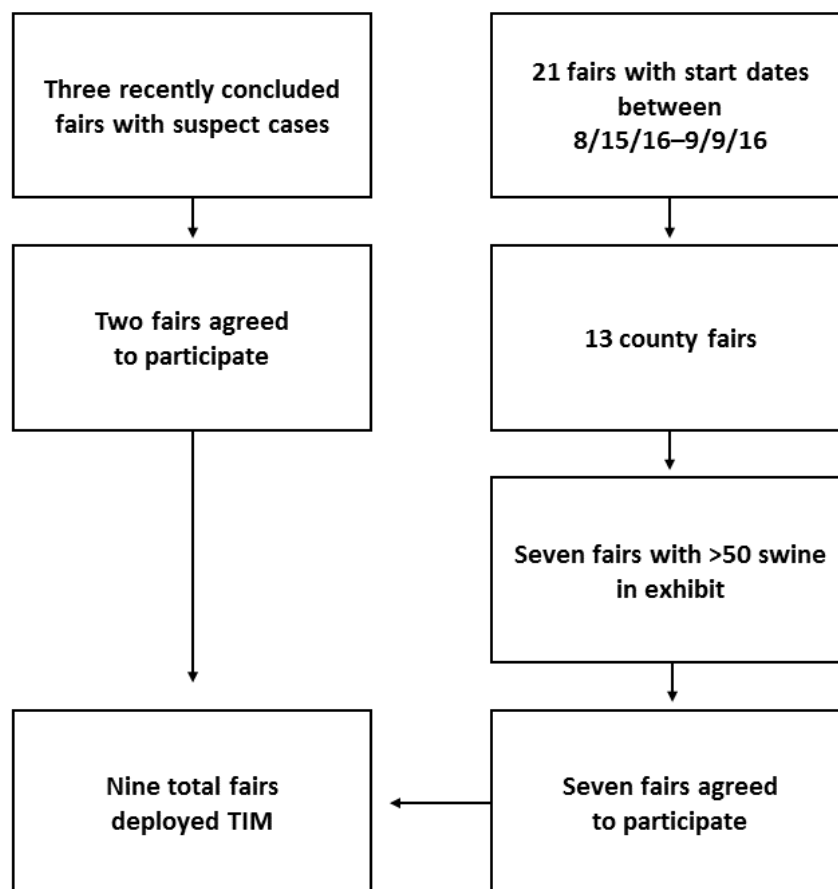


Figure 2. Texting language flow (formal).

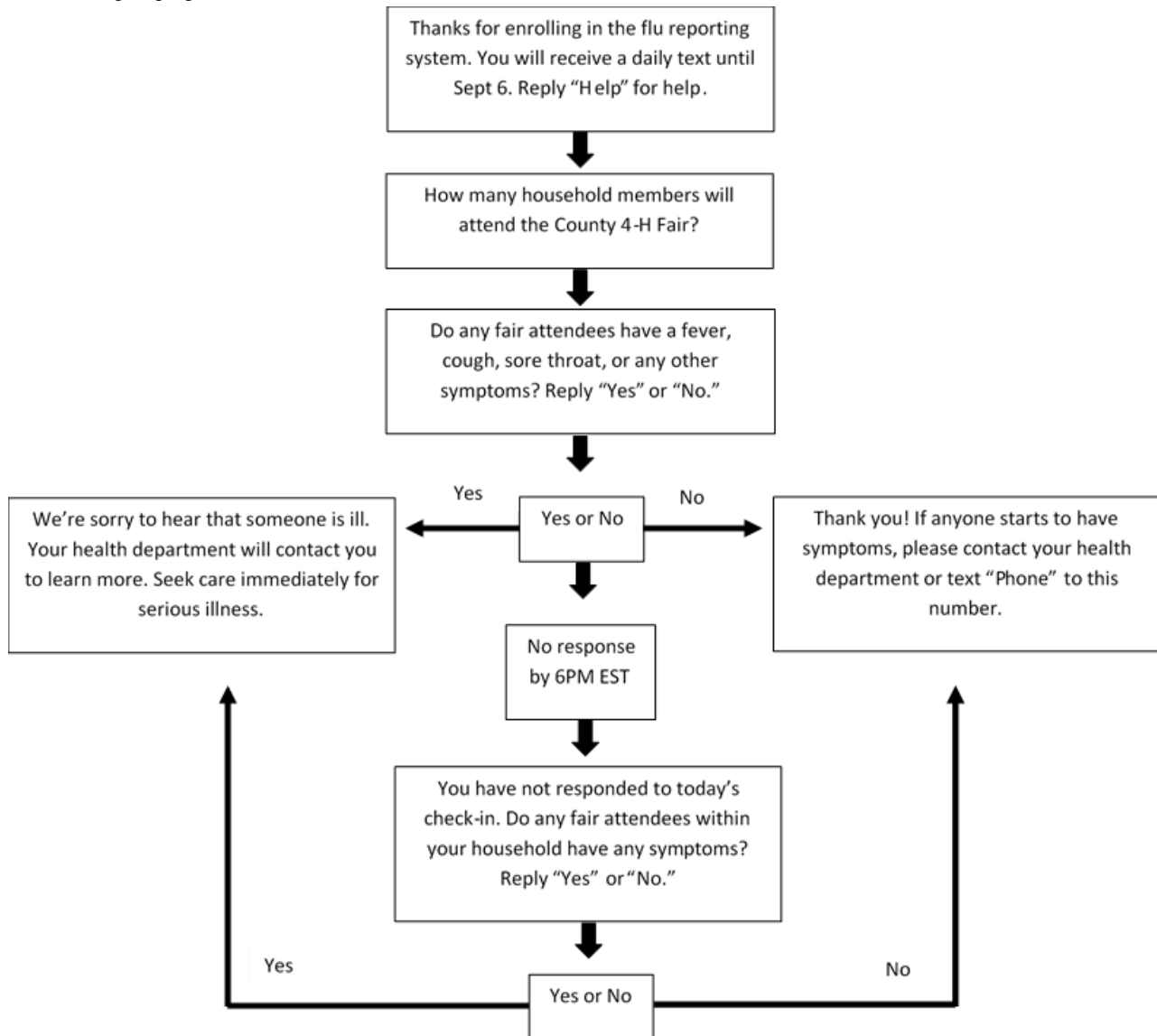
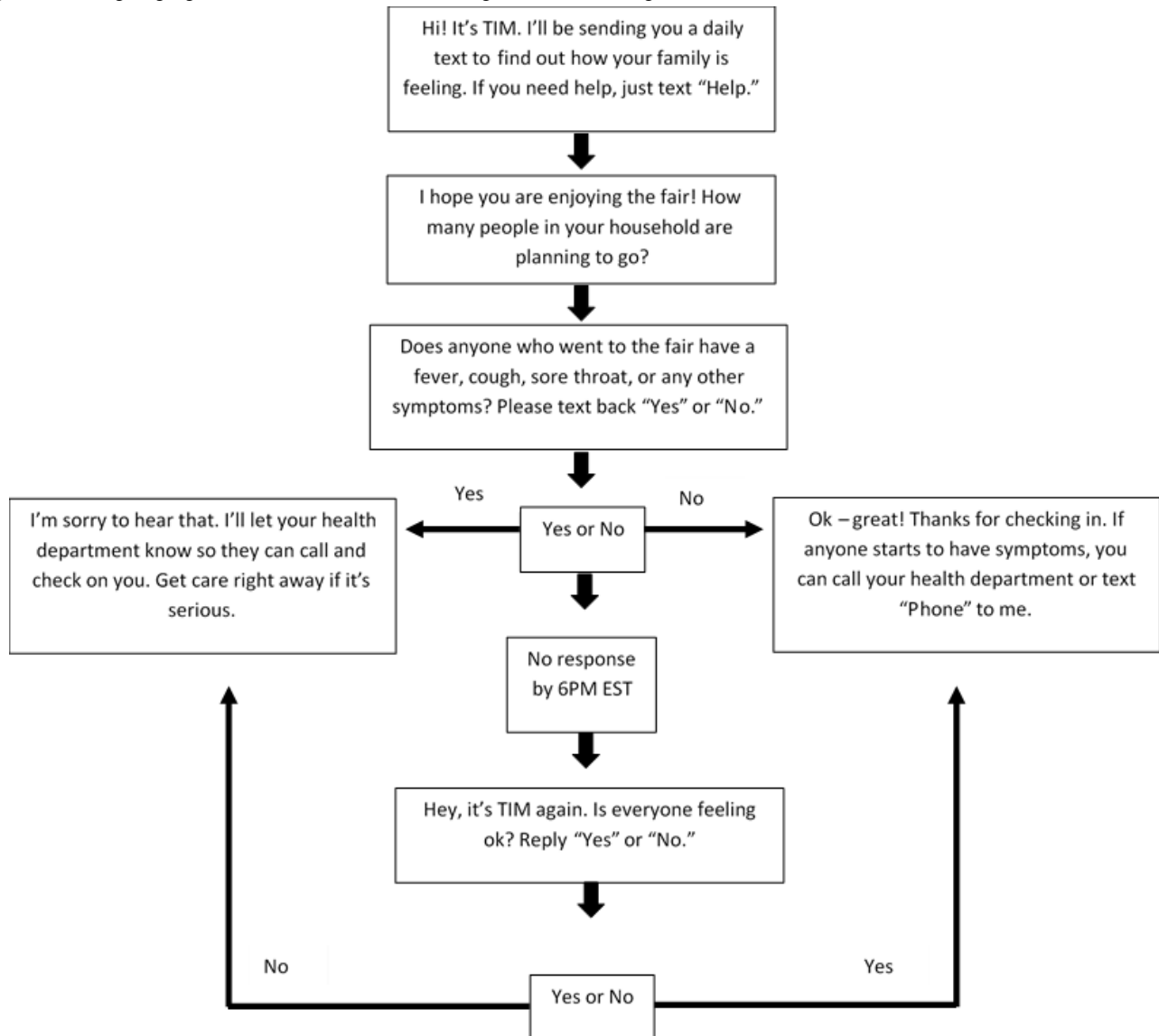


Figure 3. Texting language flow (informal). TIM: text-message illness monitoring.

Web-Based Survey

We developed a Web survey to evaluate the use and acceptability of the TIM system using Epi Info software, version 7.2.0.1 (Centers for Disease Control and Prevention). The youth program club coordinators distributed a link to the survey by email to all swine exhibitor families after the conclusion of the fair season. Questions included reasons for participating or not participating in the TIM system and likelihood of participating again if given the opportunity.

Health Department Interviews

We conducted five unstructured interviews with health officials at participating health departments. These were conducted to determine how much time was spent following up on alerts generated by the system and to determine how many respondents with reports of illness sought care and were tested for influenza.

Human Subjects Determination

This investigation was determined to be part of a public health response; in accordance with federal human subjects' protection regulations, it was not considered human subjects research. Participation in the survey was voluntary and anonymous.

Results

Enrollment, Retention, and Illness Detection

Out of an estimated 500 households contacted at the nine selected fairs, 87 (17.4%) households enrolled, reporting for 392 household members (see [Table 1](#)). Household participation rates varied by fair and by enrollment method (range 3%-86%, median 13%).

Table 1. Recruitment methods and enrollment by fair.

Fair	Primary recruitment method	Recruitment occurred after fair conclusion	H3N2v ^a infections associated with fair	Swine exhibitors (N=1052), n (%)	Estimated households ^b (N=528), n (%)	Peak household enrollment, n/N (%)	Total under TIM ^c surveillance (N=392), n (%)
Fair A	Meeting	No	No	230 (21.9)	115 (21.8)	21/115 (18.3)	92 (23.5)
Fair B	Email	Yes	Yes	200 (19.0)	100 (18.9)	11/100 (11.0)	54 (13.8)
Fair C	Local health department visit	No	No	125 (11.9)	63 (11.9)	3/63 (5.0)	13 (3.3)
Fair D	Meeting	No	No	122 (11.6)	61 (11.6)	7/61 (11.4)	31 (7.9)
Fair E	Email	Yes	No	120 (11.4)	60 (11.4)	8/60 (13.3)	40 (1.0)
Fair F	Local health department visit	No	No	75 (7.1)	38 (7.2)	1/38 (2.6)	4 (1.0)
Fair G	Meeting	No	No	70 (6.7)	35 (6.6)	30/35 (85.7)	132 (33.7)
Fair H	Flyers	No	No	55 (5.2)	28 (5.3)	1/28 (3.6)	4 (1.0)
Fair I	Flyers	No	No	55 (5.2)	28 (5.3)	5/28 (17.9)	22 (5.6)
Total	N/A ^d	N/A	N/A	1052 (100)	528 (100)	87/528 (16.5)	392 (100)

^aInfluenza A (H3N2) variant.

^bDerived by dividing the number of reported swine exhibitors by 2, based on experts' estimates of the maximum likely number of swine exhibitors per household.

^cTIM: text-message illness monitoring.

^dN/A: not applicable.

The number of household respondents for ongoing fairs peaked at 54 respondents on the third day of the fair; after that, the number steadily declined throughout the rest of the monitoring period (see [Figure 4](#)). Approximately 22% (19/87) of households remained enrolled and responded at the end of the 10-day monitoring period; 54% (47/87) made a report every day they were enrolled. Retention of respondents receiving informal texts declined faster than those receiving the formal texting language ($P<.01$); 7% (2/29) of respondents receiving the informal texts remained enrolled through the end of the monitoring period (see [Figure 4](#)). The initial daily text generated 83% of all responses, while 17% of responses followed the second (reminder) text.

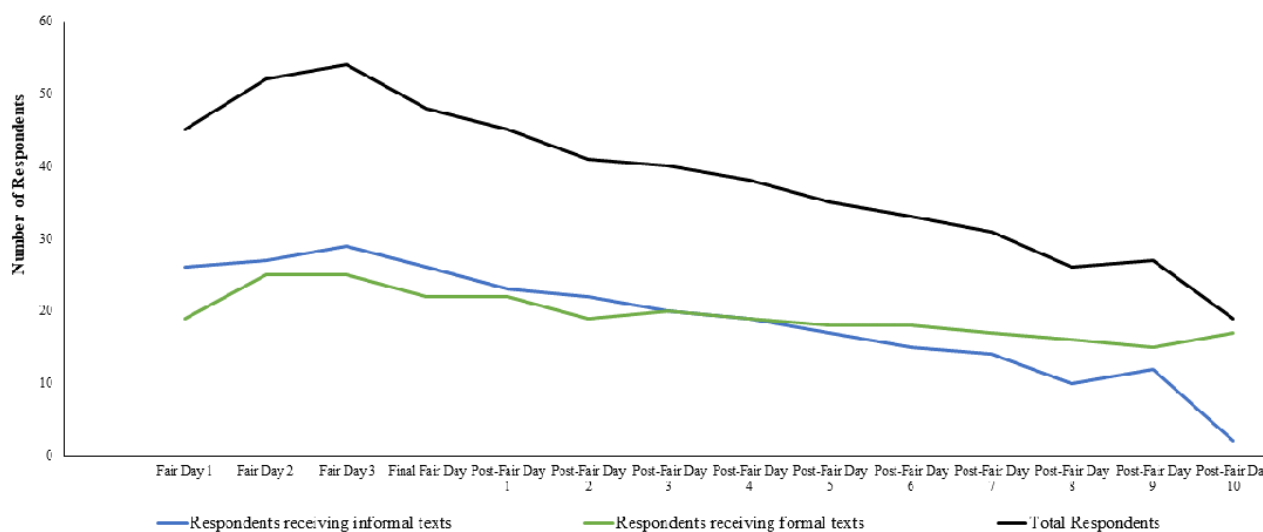
Of the 392 persons who were actively monitored, illness was reported for 22 (5.6%) through the TIM system. Of those 22, 9 (41%) sought care; 5 out of the 9 that sought care (56%) had a specimen tested and 2 out of those 5 (40%) tested positive for H3N2v virus, representing 17% (2/12) of all H3N2v virus infections detected in Michigan during this outbreak [18]. Of the 13 symptomatic persons who did not seek care, 2 (15%) could not be reached by the health department, 1 (8%) was instructed not to seek care by the county medical director, 1 (8%) declined seeking care due to the cost, and 1 (8%) declined because another diagnosis was believed to be more likely; the other 8 (62%) did not provide a reason.

Web-Based Survey

Among five fairs that distributed the survey to swine exhibitor families through an estimated 500 email addresses, 70 households (14.0%) who exhibited swine responded. Respondents from 16 out of 70 (23%) households reported participating in TIM. Among the 15 respondents who answered the question regarding motivation for participation, 11 (73%) participated because it was recommended by the fair and 12 (80%) indicated that they would participate again. Among 50 respondents who reported not participating in TIM and responded to the question, 34 (68%) said it was because they had not heard about it, 7 (14%) said they did not understand the necessity, 2 (4%) said they did not have a phone to use, and 7 (14%) said they did not know why they chose not to participate. None of the respondents reported the cost of text messages as a barrier to enrollment.

Health Department Interviews

Health department staff reported spending between 3 and 60 minutes per day following up on alerts generated through the system. Health systems interacting with respondents receiving the informal texting language indicated that they spent substantially more time (ie, up to 60 minutes) following up on alerts than those interacting with respondents receiving the formal texting language (ie, 3-5 minutes), likely due to a substantial number of false alerts from respondents using colloquial language such as "Aok. All fine here," rather than using a system-recognized response of "Yes" or "No."

Figure 4. Respondent retention throughout monitoring period: total and by texting language type.

Discussion

Principal Findings

During an outbreak of variant influenza virus associated with swine at agricultural fairs, we successfully deployed a text-based system for illness monitoring among potentially exposed people over a 4-week period. Two H3N2v virus infections were detected among 392 individuals monitored for illness during this outbreak [18], suggesting that text messaging for active surveillance was a valuable tool to complement traditional surveillance methods and enhance detection during this outbreak. Both cases identified through texting were among persons attending fairs with confirmed cases of H3N2v virus infections, suggesting that large numbers of variant influenza cases were not occurring at the other fairs included in this pilot study.

Investigation of novel influenza A virus infection is important for early detection, treatment, and prevention of the spread of influenza viruses with pandemic potential, but passive surveillance methods may miss many infections and underestimate the burden of infection associated with exposure. Multiplier models based on an H3N2v virus outbreak in 2011 estimated that for every pediatric H3N2v virus infection detected and reported to CDC, there were approximately 200 infections in the community; for every one adult H3N2v virus infection, there were approximately 255 infections. Reasons for the gap between reported and actual infections were related to the low likelihood that someone who was ill would seek health care and be tested for a variant influenza virus [22]. While syndromic monitoring lacks the ability to provide direct viral confirmation, active surveillance ensures continued reporting and increases the opportunity for an ill person to interact with health officials and to undergo appropriate testing for detecting variant influenza viruses. In our study, the TIM system was able to detect two additional H3N2v virus infections during an ongoing outbreak and provided a direct way for exposed persons to report symptoms to a health official.

Innovative methods, such as participatory, syndromic surveillance systems for influenza, are increasingly being explored as a means to better capture the burden of influenza at the community level [16,23-26] or in a specific population of interest [17,27]. National and international systems using weekly emails for automated two-way communication between system coordinators and volunteer participants have been implemented in the United States [25], Australia [23], and Europe [26] to capture and track trends in influenza-like illness at a population level. In Australia, another system used SMS text messaging to collect adverse-event data among vaccine recipients [27]. In that initiative, the response rate and timeliness were significantly improved with SMS text messaging in comparison to telephone interviews. Australian health officials also used SMS text messaging to actively monitor responders to an outbreak of avian influenza virus on a poultry farm in 2013. Public health officials found the use of SMS text messaging to be less time-consuming and 2.5 times more cost-effective than conducting telephone follow-up interviews [17].

Systems like TIM could be an important tool in future outbreak investigations. They provide a quick, scalable, and cost-effective method to actively monitor a select group of at-risk people, thereby filling a gap in current surveillance capabilities. Once adopted by a health department, the TIM system can monitor at-risk individuals in an outbreak investigation and be integrated into the outbreak response within hours. As they are intended for detection of illness among a select group of individuals, they complement current influenza surveillance systems that are better adapted for capturing population-level trends and impact [28]. Although the cost of texting was not identified as a barrier in this study, application of this method in less-wealthy populations may benefit from participants being provided phone credit [24].

While the TIM system was well-accepted by both respondents and health departments, we identified three areas where improvements were needed. The first was with enrollment, which varied by fair and ranged from 3% to 86%. The vast

majority of households that did not enroll had not heard about the system and those that did enroll did so primarily because it was recommended by the agricultural club coordinator. This highlights the importance of effective communication by people who can reach, and are trusted by, the intended target community. Future deployments of systems like TIM should focus on identifying trusted groups and individuals early and working with them to encourage enrollment. Holding meetings to describe the system and encouraging attendants to enroll during the meeting resulted in higher enrollment, perhaps due to the convenience and personal nature of the recommendation. While strategies to increase enrollment will differ by setting, using personal communication to provide simultaneous rationale and enrollment information appeared to be an effective strategy that could be adapted and applied in most settings.

The second identified area for improvement was testing of symptomatic exposed persons; in our study, only about 20% of ill individuals were tested. Appropriate testing is the only way to confirm infection with a novel influenza virus, as the symptoms are similar to those caused by other respiratory viruses, including seasonal influenza. There are many potential barriers to care-seeking, including lack of motivation to seek influenza testing, the availability of testing at local health care centers or health departments, and the cost associated with visiting a health care provider, especially among people without insurance or with high out-of-pocket health costs. There are several ways to address these barriers and the mechanism depends on the resources and capabilities of the state and local health care systems. In future situations, health officials could consider offering testing for novel influenza virus infections free of charge at health departments or sending clinical staff to collect specimens from ill persons in their homes as a way to address some of these barriers. The latter is a strategy that has been shown to be effective in New York City for longitudinal surveillance of influenza-like illness [16,29]. In that study, home visits were conducted to obtain specimens from persons with reports of illness. Additionally, increased education among exposed persons about the risk of novel influenza infections and the importance of detection also may increase the likelihood of testing.

The third issue identified was with respondent retention. Monitoring for a novel influenza virus requires up to 10 days of monitoring after the last day of exposure, which in the case of a 10-day county fair may necessitate as many as 20 days of monitoring. Maintaining respondents throughout the monitoring period was a challenge in this pilot and only 30% remained enrolled through the end of the recommended monitoring period.

We deployed two types of texting language to see if the formality of the language could increase retention. We found that respondents receiving the informal language texts stopped responding sooner than those receiving the formal texting language and fewer remained enrolled through the end of the monitoring period. Respondents of the informal texting language also responded to texts in a conversational tone, often with abbreviations or slang terms, which generated alerts through the system that required unnecessary follow-up by health department staff. This resulted in some health department staff reporting that they spent as much as 60 minutes per day following up on alerts versus 3-5 minutes for those interacting with respondents receiving the formal texting language. We recommend that future deployments of TIM-like platforms should use the formal texting language and consider the optimal length of monitoring that weighs the incubation of variant and avian influenza against the likelihood of attrition among the respondents.

Limitations

This investigation has certain limitations. First, we conducted it in the middle of an outbreak response and relied on health departments and club coordinators to recruit participants in the middle of a busy fair season. In some situations, recruitment occurred solely through email or delivering flyers, which may have limited the number of people who were aware of the system and its importance. Second, this investigation involved adult participants who responded for the entire household. Other studies also have used this method with success [16]; however, we cannot be certain that the respondent was accurately reporting for all household members. Third, we had a low participation rate to the Web-based survey (14%) and respondents may not have shared the same opinions as those who did not respond. Fourth, our selection of fairs was a convenience sample based on timing and fair attributes. Findings from these fairs may not be representative of all agricultural fairs.

Conclusions

In summary, we successfully piloted a new text-based monitoring tool for detecting variant influenza virus infections and identified two H3N2v virus infections during an outbreak in Michigan. This pilot demonstrates how text messaging can complement traditional surveillance methods during an outbreak. Future activations of the system should work to improve systematic testing of exposed persons who develop symptoms and should continue exploring methods to improve participation and retention of respondents.

Acknowledgments

We gratefully acknowledge the many people who assisted with the outbreak response and the evaluation of this monitoring tool, including James Averill, LaShondra Berman, Sally Bidol, Lenee Blanton, Andrew S Bowman, Lori Cargil, Beth Ferry, Pat Fralick, Colleen Harns, Britney James, Yungo Jang, Kathy Kacynski, Lisa Koonin, Leah Lavanway, Debra Laws, Leslie Lee, Dian Liepe, Stephen Lindstrom, Amanda Lubit, Mark McCorkle, Michelle McDonald, Guy Miller, Anita Patel, Tracy Payne, Trevor Quinlan, Erin Radke, Carrie Reed, Ruby Rodgers, Mary-Grace Stobierski, Arlee Sutton, Susan Trock, Renee Tschrihart, Kristy VanSickle, Colleen Wallace, and David Wentworth. The findings and conclusions in this report are those of the authors and do not necessarily

represent the official views of CDC. Use of trade names and commercial sources is for identification only and does not imply endorsement by CDC, the Public Health Service, or the US Department of Health and Human Services.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Epidemic curve of influenza A (H3N2) variant virus infections by date of symptom onset and detection method—Michigan, 2016.

[[DOCX File, 110KB - publichealth_v5i2e10842_app1.docx](#)]

References

1. Cartter ML. National Reporting for Initial Detections of Novel Influenza A Viruses. Atlanta, GA: Council of State and Territorial Epidemiologists; 2007 Jan 05. URL: <http://c.ymcdn.com/sites/www.cste.org/resource/resmgr/PS/07-ID-01.pdf> [accessed 2019-02-20] [WebCite Cache ID 76KeYopop]
2. Centers for Disease Control and Prevention (CDC). Notes from the field: Outbreak of influenza A (H3N2) virus among persons and swine at a county fair--Indiana, July 2012. *MMWR Morb Mortal Wkly Rep* 2012 Jul 27;61(29):561 [FREE Full text] [Medline: 22832938]
3. Bowman AS, Nelson SW, Page SL, Nolting JM, Killian ML, Sreevatsan S, et al. Swine-to-human transmission of influenza A (H3N2) virus at agricultural fairs, Ohio, USA, 2012. *Emerg Infect Dis* 2014 Sep;20(9):1472-1480 [FREE Full text] [doi: 10.3201/eid2009.131082] [Medline: 25148572]
4. Epperson S, Jung M, Richards S, Quinlisk P, Ball L, Moll M, Influenza A (H3N2)v Virus Investigation Team. Human infections with influenza A (H3N2) variant virus in the United States, 2011-2012. *Clin Infect Dis* 2013 Jul;57 Suppl 1:S4-S11. [doi: 10.1093/cid/cit272] [Medline: 23794729]
5. Greenbaum A, Quinn C, Bailer J, Su S, Havers F, Durand LO, et al. Investigation of an outbreak of variant influenza A (H3N2) virus infection associated with an agricultural fair-Ohio, August 2012. *J Infect Dis* 2015 Nov 15;212(10):1592-1599 [FREE Full text] [doi: 10.1093/infdis/jiv269] [Medline: 25948864]
6. Jung MA, Epperson S, Biggerstaff M, Allen D, Balish A, Barnes N, et al. Outbreak of variant influenza A (H3N2) virus in the United States. *Clin Infect Dis* 2013 Dec;57(12):1703-1712 [FREE Full text] [doi: 10.1093/cid/cit649] [Medline: 24065322]
7. Lindstrom S, Garten R, Balish A, Shu B, Emery S, Berman L, et al. Human infections with novel reassortant influenza A (H3N2)v viruses, United States, 2011. *Emerg Infect Dis* 2012 May;18(5):834-837 [FREE Full text] [doi: 10.3201/eid1805.111922] [Medline: 22516540]
8. Wong KK, Gambhir M, Finelli L, Swerdlow DL, Ostroff S, Reed C. Transmissibility of variant influenza from Swine to humans: A modeling approach. *Clin Infect Dis* 2013 Jul;57 Suppl 1:S16-S22 [FREE Full text] [doi: 10.1093/cid/cit303] [Medline: 23794727]
9. Wong KK, Greenbaum A, Moll ME, Lando J, Moore EL, Ganatra R, et al. Outbreak of influenza A (H3N2) variant virus infection among attendees of an agricultural fair, Pennsylvania, USA, 2011. *Emerg Infect Dis* 2012 Dec;18(12):1937-1944 [FREE Full text] [doi: 10.3201/eid1812.121097] [Medline: 23171635]
10. Vincent AL, Ma W, Lager KM, Janke BH, Richt JA. Swine influenza viruses a North American perspective. *Adv Virus Res* 2008;72:127-154. [doi: 10.1016/S0065-3527(08)00403-X] [Medline: 19081490]
11. Bliss N, Nelson SW, Nolting JM, Bowman AS. Prevalence of influenza A virus in exhibition swine during arrival at agricultural fairs. *Zoonoses Public Health* 2016 Dec;63(6):477-485. [doi: 10.1111/zph.12252] [Medline: 26750204]
12. Hinshaw VS, Bean WJ, Webster RG, Easterday BC. The prevalence of influenza viruses in swine and the antigenic and genetic relatedness of influenza viruses from man and swine. *Virology* 1978 Jan;84(1):51-62. [Medline: 74123]
13. Olsen CW. The emergence of novel swine influenza viruses in North America. *Virus Res* 2002 May 10;85(2):199-210. [Medline: 12034486]
14. Bowman AS, Nolting JM, Nelson SW, Slemons RD. Subclinical influenza virus A infections in pigs exhibited at agricultural fairs, Ohio, USA, 2009-2011. *Emerg Infect Dis* 2012 Dec;18(12):1945-1950 [FREE Full text] [doi: 10.3201/eid1812.121116] [Medline: 23171654]
15. Centers for Disease Control and Prevention. 2018. Information for people exposed to birds infected with avian influenza viruses of public health concern URL: <https://www.cdc.gov/flu/avianflu/h5/infected-birds-exposure.htm> [accessed 2018-11-09] [WebCite Cache ID 73oB4dZDs]
16. Stockwell MS, Reed C, Vargas CY, Camargo S, Garretson AF, Alba LR, et al. MoSAIC: Mobile Surveillance for Acute Respiratory Infections and Influenza-Like Illness in the Community. *Am J Epidemiol* 2014 Dec 15;180(12):1196-1201. [doi: 10.1093/aje/kwu303] [Medline: 25416593]

17. Stephenson LM, Biggs JS, Sheppard V, Oakman TL. An evaluation of the use of short message service during an avian influenza outbreak on a poultry farm in Young. *Commun Dis Intell Q Rep* 2016 Jun 30;40(2):E195-E201 [[FREE Full text](#)] [Medline: [27522128](#)]
18. Schicker RS, Rossow J, Eckel S, Fisher N, Bido S, Tatham L, et al. Outbreak of influenza A (H3N2) variant virus infections among persons attending agricultural fairs housing infected swine - Michigan and Ohio, July-August 2016. *MMWR Morb Mortal Wkly Rep* 2016 Oct 28;65(42):1157-1160 [[FREE Full text](#)] [doi: [10.15585/mmwr.mm6542a1](#)] [Medline: [27787493](#)]
19. Bowman AS, Walia RR, Nolting JM, Vincent AL, Killian ML, Zentkovich MM, et al. Influenza A (H3N2) virus in swine at agricultural fairs and transmission to humans, Michigan and Ohio, USA, 2016. *Emerg Infect Dis* 2017 Dec;23(9):1551-1555 [[FREE Full text](#)] [doi: [10.3201/eid2309.170847](#)] [Medline: [28820376](#)]
20. Bowman AS, Workman JD, Nolting JM, Nelson SW, Slemons RD. Exploration of risk factors contributing to the presence of influenza A virus in swine at agricultural fairs. *Emerg Microbes Infect* 2014 Jan;3(1):e5 [[FREE Full text](#)] [doi: [10.1038/emi.2014.5](#)] [Medline: [26038494](#)]
21. Centers for Disease Control and Prevention. 2018. Self-observation instructions for demobilizing bird flu responders URL: <https://www.cdc.gov/flu/avianflu/h5/demobilizing-responders.htm> [accessed 2018-11-09] [[WebCite Cache ID 73oAX98Ep](#)]
22. Biggerstaff M, Reed C, Epperson S, Jung MA, Gambhir M, Bresee JS, et al. Estimates of the number of human infections with influenza A (H3N2) variant virus, United States, August 2011-April 2012. *Clin Infect Dis* 2013 Jul;57 Suppl 1:S12-S15 [[FREE Full text](#)] [doi: [10.1093/cid/cit273](#)] [Medline: [23794726](#)]
23. Dalton C, Carlson S, Butler M, Cassano D, Clarke S, Fejsa J, et al. Insights from Flutracking: Thirteen tips to growing a Web-based participatory surveillance system. *JMIR Public Health Surveill* 2017 Aug 17;3(3):e48 [[FREE Full text](#)] [doi: [10.2196/publichealth.7333](#)] [Medline: [28818817](#)]
24. Prieto JT, Jara JH, Alvis JP, Furlan LR, Murray CT, Garcia J, et al. Will participatory syndromic surveillance work in Latin America? Piloting a mobile approach to crowdsourcing influenza-like illness data in Guatemala. *JMIR Public Health Surveill* 2017 Nov 14;3(4):e87 [[FREE Full text](#)] [doi: [10.2196/publichealth.8610](#)] [Medline: [29138128](#)]
25. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wójcik O, et al. Flu Near You: Crowdsourced symptom reporting spanning 2 influenza seasons. *Am J Public Health* 2015 Oct;105(10):2124-2130. [doi: [10.2105/AJPH.2015.302696](#)] [Medline: [26270299](#)]
26. Koppeschaar CE, Colizza V, Guerrisi C, Turbelin C, Duggan J, Edmunds WJ, et al. Influenzanet: Citizens among 10 countries collaborating to monitor influenza in Europe. *JMIR Public Health Surveill* 2017 Dec 19;3(3):e66 [[FREE Full text](#)] [doi: [10.2196/publichealth.7429](#)] [Medline: [28928112](#)]
27. Regan AK, Blyth CC, Tracey L, Mak DB, Richmond PC, Effler PV. Comparison of text-messaging to voice telephone interviews for active surveillance of adverse events following immunisation. *Vaccine* 2015 Jul 17;33(31):3689-3694. [doi: [10.1016/j.vaccine.2015.06.022](#)] [Medline: [26079616](#)]
28. Centers for Disease Control and Prevention. 2018. Overview of influenza surveillance in the United States URL: <https://www.cdc.gov/flu/weekly/overview.htm> [accessed 2018-11-09] [[WebCite Cache ID 73oBK8rJ1](#)]
29. Russell KE, Fowlkes A, Stockwell MS, Vargas CY, Saiman L, Larson EL, et al. Comparison of outpatient medically attended and community-level influenza-like illness-New York City, 2013-2015. *Influenza Other Respir Viruses* 2018 Dec;12(3):336-343 [[FREE Full text](#)] [doi: [10.1111/irv.12540](#)] [Medline: [29350791](#)]

Abbreviations

CDC: Centers for Disease Control and Prevention

H3N2v: influenza A (H3N2) variant

MDHHS: Michigan Department of Health and Human Services

NACCHO: National Association of County and City Health Officials

rRT-PCR: real-time reverse-transcription polymerase chain reaction

SMS: short message service

TIM: text-message illness monitoring

Edited by T Sanchez; submitted 11.05.18; peer-reviewed by K Baltrusaitis, L Geneviève; comments to author 22.06.18; revised version received 13.11.18; accepted 20.12.18; published 26.04.19.

Please cite as:

Stewart RJ, Rossow J, Eckel S, Bido S, Ballew G, Signs K, Conover JT, Burns E, Bresee JS, Fry AM, Olsen SJ, Biggerstaff M
Text-Based Illness Monitoring for Detection of Novel Influenza A Virus Infections During an Influenza A (H3N2)v Virus Outbreak in Michigan, 2016: Surveillance and Survey

JMIR Public Health Surveill 2019;5(2):e10842

URL: <http://publichealth.jmir.org/2019/2/e10842/>

doi: [10.2196/10842](#)

PMID: [31025948](#)

©Rebekah J Stewart, John Rossow, Seth Eckel, Sally Bidol, Grant Ballew, Kimberly Signs, Julie Thelen Conover, Erin Burns, Joseph S Bresee, Alicia M Fry, Sonja J Olsen, Matthew Biggerstaff. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 26.04.2019. 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

Informing HIV Prevention Programs for Adolescent Girls and Young Women: A Modified Approach to Programmatic Mapping and Key Population Size Estimation

Eve Cheuk¹, PhD; Shajy Isac², PhD; Helgar Musyoki³, MPH; Michael Pickles¹, PhD; Parinita Bhattacharjee^{1,4}, MA; Peter Gichangi⁵, MD, PhD; Robert Lorway¹, PhD; Sharmistha Mishra^{6,7}, MD, PhD; James Blanchard¹, MD, PhD; Marissa Becker¹, MSc, MD

¹Centre for Global Public Health, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada

²India Health Action Trust, Bangalore, India

³National AIDS and Sexually Transmitted Infection Control Programme, Ministry of Health, Nairobi, Kenya

⁴Partners for Health and Development in Africa, Nairobi, Kenya

⁵International Centre for Reproductive Health Kenya, Mombasa, Kenya

⁶Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, ON, Canada

⁷Department of Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:

Eve Cheuk, PhD

Centre for Global Public Health

Rady Faculty of Health Sciences

University of Manitoba

R070 Medical Rehabilitation Building

771 McDermot Avenue

Winnipeg, MB,

Canada

Phone: 1 204 789 3396

Fax: 1 204 789 3718

Email: eve.cheuk@umanitoba.ca

Abstract

Background: Standard programmatic mapping involves identifying locations where key populations meet, profiling of these locations (hotspots), and estimating the key population size. Information gained from this method has been used for HIV programming—resource allocation, program planning, service delivery, and monitoring and evaluation—for people who inject drugs, men who have sex with men, and female sex workers (FSWs). With an increasing focus on adolescent girls and young women (AGYW) as a priority population for HIV prevention, programs need to know the location of and how to effectively reach individuals who are at increased risk for HIV but were conventionally considered part of the general population. We hypothesize that AGYW who engage in transactional and casual sex also congregate at sex work hotspots to meet sex partners. Therefore, we adapted the standard programmatic mapping approach to understand the geographic distribution and population size of AGYW at increased HIV risk in Mombasa County, Kenya.

Objectives: The objectives are several-fold: (1) detail and compare the modified programmatic mapping approach used in this study to the standard approach, (2) estimate the number of young FSWs, (3) estimate the number of AGYW who congregate in sex work hotspots to meet sex partners other than clients, (4) estimate the overlap in sexual network in hotspots, (5) describe the distribution of sex work hotspots across Mombasa and its four subcounties, and (6) compare the distribution of hotspots that were known to the local HIV prevention program prior to this study and those newly identified.

Methods: The standard programmatic mapping approach was modified to estimate the population of young women aged 14 to 24 years who visit sex work hotspots in Mombasa to meet partners for commercial, transactional, and casual sex.

Results: We estimated that there were 11,777 FSWs (range 9265 to 14,290) in Mombasa in 2014 among whom 6127 (52.02%) were 14 to 24 years old. The population estimates for women aged 14 to 24 years who engaged in transactional and casual sex and congregated at the hotspots were 5348 (range 4185 to 6510) and 4160 (range 3194 to 5125), respectively. Of the 1025 validated

sex work hotspots, 870 (84.88%) were locations also visited by women engaged in transactional and casual sex. Only 47 (4.58%) hotspots were exclusive sex work locations. The geographic and typological distribution of hotspots were significantly different between the four subcounties ($P < .001$). Of the 1025 hotspots, 419 (40.88%) were already known to the local HIV prevention program and 606 (59.12%) were newly identified.

Conclusions: Using the adapted programmatic mapping approach detailed in this study, our results show that HIV prevention programs tailored to AGYW can focus delivery of their interventions to sex work hotspots to reach subgroups that may be at increased risk for HIV.

(*JMIR Public Health Surveill* 2019;5(2):e11196) doi:[10.2196/11196](https://doi.org/10.2196/11196)

KEYWORDS

programmatic mapping; key population size estimation; female sex workers; adolescent girls and young women; sex work; transactional sex; casual sex; HIV prevention

Introduction

Key populations are groups of individuals who are at increased risk for, or who are disproportionately burdened by, a particular disease. Therefore, it is important for individuals of a key population to have access to necessary prevention resources, care, and support. In the context of HIV, key populations include people living with HIV, people who inject drugs, men who have sex with men, transgender people, and sex workers [1]. One of the earliest steps in designing a focused HIV program for a key population is knowing who and how many within a key population are at risk, where they are at risk, and what puts them at risk (eg, condomless sex, experience of violence, needle sharing). Various methods exist for estimating key population size such as capture-recapture, multiplier, population survey, and network scale-up [2,3]. However, these approaches only provide an estimate for the size of a key population, which in itself may not be sufficient for comprehensive program design, planning, and delivery. In recognition of the importance of how place, environment, and sexual network formation contribute to the epidemiology of HIV [4-6], programmatic mapping was developed to shed light on the types (or typology) and distribution of locations where key populations congregate. Programmatic mapping involves locating and profiling geographically defined locations (or hotspots) where key populations meet and risk behaviors take place and estimating the size of key populations in each hotspot. Individual hotspot-level estimates can be aggregated to generate population size estimates for a district, region, and entire country. Many cities and countries, including Kenya, have used programmatic mapping to determine the geographic distribution and population size for people who use injection drugs; men who have sex with men; and male, female, and transgender sex workers [5,7,8].

Given the disproportionately high burden of HIV shouldered by female sex workers (FSWs) [9], many HIV prevention programs have been implemented to provide focused services to this key population, and there have been successes in reducing new HIV infections among FSWs in some regions of the world [10-12]. Nonetheless, in a clinic-based survey of Kenyan FSWs between 18 and 57 years, HIV prevalence was highest among women who had been in sex work for less than two years [13], highlighting a program gap in addressing HIV risk early. In fact, this high HIV prevalence may indeed signal existing HIV risk factors prior to entry into sex work that remain unmitigated

and that are likely compounded by other factors in the early stages of a sex work career.

Conventional FSW programs are not designed to reach women who engage in other types of sexual partnerships that may also be associated with increased risk of HIV acquisition—namely, condomless sex in the context of transactional sex and casual sex [14]. Adolescent girls and young women (AGYW) aged 15 to 24 years are particularly burdened by HIV. In 2016 there were an estimated 790,000 (range 680,000 to 910,000) new HIV infections among adult women globally, among whom 46% were AGYW. AGYW also represented approximately 60% of newly infected young people and of all young people living with HIV globally [15]. In Kenya, 32.8% of all new adult HIV infection in 2015 occurred among AGYW [16].

The shared vulnerabilities among AGYW associated with sexual debut and their early sexual experiences and among young FSWs who have newly entered sex work have led to a global call to refocus and Fast-Track HIV prevention among AGYW [17]. But how does one expand coverage of targeted HIV prevention programs to a subpopulation that has traditionally been regarded as part of the general population?

AGYW are a heterogeneous population. We hypothesize that AGYW who engage in transactional and casual sex and who congregate at sex work hotspots to specifically meet sex partners have higher vulnerabilities and are at an increased risk for HIV compared to their counterparts in the general population. Therefore, understanding whether and to what extent potential overlap in sexual networks exists between "high-risk" AGYW and FSWs could provide valuable information for prevention programs to reach and provide services for AGYW subgroups in the relevant settings. In this paper, we describe how we modified the current approach to FSW programmatic mapping [5,7,8] to gather information on the geographic distribution and estimate the population size of AGYW congregating at established sex work hotspots by the spectrum of sexual partnerships, including sex work and transactional and casual sex. The objectives of our research were several-fold:

- Detail and compare the modified programmatic mapping approach used in this study to the standard approach
- Estimate the number of young FSWs
- Estimate the number of AGYW who congregate in sex work hotspots to meet sex partners other than clients
- Estimate the overlap in sexual network in hotspots

- Describe the distribution of sex work hotspots across Mombasa and its four subcounties
- Compare the distribution of hotspots that were known to the local HIV prevention program prior to this study with those newly identified

Finally, we also reviewed how this modified approach to programmatic mapping and population size estimation could provide strategic guidance for HIV/sexually transmitted infection (STI) prevention programs looking to adapt and expand their services to address a wider spectrum of risk among AGYW within sex work hotspots.

Methods

Study Setting

Mombasa County is the smallest of 47 counties in Kenya, covering a territory of 294.9 km², situated along the coast of the Indian Ocean [18]. In 2014, Mombasa had a population of 1,106,444 (with 319,032 women of reproductive age [14 to 44 years] and 134,885 AGYW [14 to 24 years]) [18]. Mombasa city is an important regional economic hub with a robust tourism sector and a port that handles millions of tons of cargo annually [19]. Kenya has a mixed HIV epidemic, and in 2015 the HIV prevalence in Mombasa County was 7.9% [20], 1.3 times higher than the national average of 5.9% [16]. The HIV prevalence among FSWs in Kenya was 29.3% in 2015 compared with 6.3% among women in the general population [16]. The HIV prevalence among FSWs in Mombasa is currently unknown. In 2014, HIV prevalence was estimated at 6.0% and 10.0% among AGYW in the age groups 15 to 19 years and 20 to 24 years, respectively [21].

Definitions

Sex work is defined as the exchange of money (or gifts or other resources) between individuals for sex, whereby the negotiation of the price of sex is explicit between the sex worker and the client before any exchange takes place. Transactional sex is a more nuanced exchange between individuals who engage in sex with the expectation of receiving money, gifts, or other resources in return; however, the price of sex is not prenegotiated [22,23]. Casual sex occurs when individuals engage in sex but neither party expects to receive money, gifts, or other resources in return. A hotspot is defined as an indoor or outdoor venue or location where FSWs congregate to solicit clients and/or where sex work-related sexual activities take place.

Standard Approach to Programmatic Mapping and Estimation of Key Population Size

Overview

The standard mapping method [5,7,8] (Figure 1) involves (1) premapping and planning exercise, (2) level one secondary and

tertiary key informant (KI) interviews to generate an exhaustive list of hotspots where members of a key population congregate, (3) level two interviews (or group discussion) with primary KIs on site to validate and profile hotspots, and (4) data analysis. Primary KIs are members of the key population (in this case, FSWs), secondary KIs include persons who have close association with and an intimate knowledge about sex work (eg, pimps, brothel owners), and tertiary KIs include those who have a professional knowledge and/or interest in sex work (eg, taxi drivers, club bouncers, police).

Premapping and Planning Exercise

In the first stage of the mapping exercise, the local team of program implementers and peers—former or current members of the sex worker community—are trained and prepared for data collection. The local team discusses and finalizes a list of potential types of secondary and tertiary KIs for the level one interviews and the most common typologies of hotspots where members of the key population congregate. The team then segments the city into divisions within which smaller data collection zones are delineated, often based on existing administrative units, and develops an action plan with timelines. Community engagement with peers from local FSW collectives and nongovernmental organizations (NGOs) serving FSWs is a pivotal part of the premapping exercise. The study team needs to work with peers and NGOs because they hold invaluable knowledge about the sex work context and surroundings that an outsider does not have. More importantly, their involvement will ensure development of a mapping process that is well informed, socially and ethically responsible, and acceptable to members of the key population.

Level One Mapping Activity

The second stage of the mapping exercise involves systematically walking through the streets and open areas (eg, parks, market) of each data collection zone and interviewing secondary and tertiary KIs (and primary KIs as well if possible) as they are being approached by the local team in these public spaces with the aim to identify and generate an exhaustive list of hotspots by names and address. Through this process, KIs also provide information about the typology of the hotspots and seasons and hours of operation. The number of interviews with different types of KIs is tracked to ensure a broad range of perspectives is captured. All mentioned locations are collated and deduplicated daily and checked against the program listing to determine if they are already known to the local program. A final unique hotspot list is created by consolidating data from program listing and previous mapping findings, if available.

Figure 1. Comparison of the standard programmatic mapping approach with the modified approach used in this study. FSW: female sex worker; AGYW: adolescent girls and young women; KI: key informant.

	a) Standard Mapping Approach	b) <i>Modified</i> Mapping Approach
<p>PRE-MAPPING AND PLANNING</p> <p>↓</p>	<ul style="list-style-type: none"> • Train local implementation team • Segment target mapping area into data collection zones • Finalize key informant types for Level One interviews • Finalize hotspot typology • Develop data collection tools, interview guide and script • Develop mapping action plan with timelines for completion • Community engagement with peers from the local FSW collectives 	<p>Standard approach <i>plus</i>:</p> <ul style="list-style-type: none"> • Discuss with local implementation team and peers about the expanded scope of the modified approach to also estimate the population size of AGYW who engage in sex work, transactional and casual sex • Discuss the study's hypothesis about co-location of AGYW subgroups at sex work hotspots seeking different types of sexual partnerships
<p>LEVEL ONE ACTIVITY</p> <p>↓</p>	<ul style="list-style-type: none"> • Systematically walk through streets of each data collection zone • Conduct interviews with secondary and tertiary KIs (and primary KIs if possible) on streets – ensure KI types are evenly represented • Collect data on hotspot name, address, typology, hours of operation, min. and max. # of FSWs visiting these hotspots • Collate data and consolidate with data from program listing and previous mapping, if available, to develop a final unique list of hotspots • Select hotspots for Level Two activity 	<ul style="list-style-type: none"> • No change
<p>LEVEL TWO ACTIVITY</p> <p>↓</p>	<ul style="list-style-type: none"> • Validate and profile hotspots • Conduct interviews or group discussions with FSWs (i.e., primary KIs) at hotspot • Collect data on hotspot typology, peak day/time when most FSWs visit, min. and max. # of all women and FSWs who visit on a normal and peak day, other hotspots in the area, and specific locations where respondents connect with clients • Collate data for triangulation, compilation and analysis 	<p>Standard approach <i>plus</i>:</p> <ul style="list-style-type: none"> • Collect additional data on the min. and max. # of women and FSWs aged 14-24 years, min. and max. # of women aged 14-24 years who seek transactional and casual sex partners, min. and max. # of men who visit and the proportion of whom come to seek sex partners/FSWs, and other places where women visit to meet partners for transactional and casual sex
<p>DATA ANALYSIS</p>	<ul style="list-style-type: none"> • Calculate regional key population size estimate based on hotspot-level population numbers 	<ul style="list-style-type: none"> • No change

Level Two Mapping Activity

In the third stage of the mapping exercise, the implementation team visits each hotspot on the list generated from the level one activity and validates the hotspot as active if it is currently operational and it is a location where FSWs congregate. The team also conducts group discussions with FSWs (ie, primary KIs) on site if more than one FSW can be mobilized. If only one FSW is available and willing to participate, a one-on-one interview is conducted instead. Detailed information was

gleaned about the typology of the hotspot, the peak day of the week and peak time of the day when the majority of FSWs visit the hotspot, and the minimum and maximum number of women and FSWs of all ages who come to visit the hotspot on a normal and peak day. FSWs are also asked to list other sex work hotspots in the area, so any new locations that were missed in the level one activity will be added to the list and followed up. To account for the extent of sex worker mobility, FSWs are asked about other specific locations where they connect with clients.

Figure 2. Population size estimation formula used in this study, where C_r is the crude population size estimate of female sex workers in the county; E_i is the estimated number of female sex workers in hotspot i ; and $i=1,2,3,\dots,n$ is the number of hotspots in the county.

$$C_r = \sum_{i=1}^n E_i$$

Data Analysis

Since all the identified sex work hotspots were validated during the level two activity, we used only level two data for analysis. To obtain a crude population estimate for FSWs in the entire county, the minimum and maximum estimates of all hotspots were summed and the mean was calculated. See Figure 2 for the formula used for population size estimation, where C_r is the crude population size estimate of FSWs in the county; E_i is the estimated number of FSWs in hotspot i ; and $i=1,2,3,\dots,n$ is the number of hotspots in the county.

To account for the extent of duplication due to mobility of FSWs visiting multiple hotspots looking for clients, the following mathematical function was used: $N_i = C_r(1-p_i) + (C_r * p_i / m_i)$, where the adjusted estimate (N_i) is a function of the crude estimate (C_r), the proportion of FSWs who visited other hotspots (p_i), and the mean number of hotspots FSWs visited (m_i). Mobility was adjusted at the subcounty level; adjusted estimates for all subcounties were then summed to generate a county-wide estimate.

Modifying the Standard Approach to Estimate the Population Size of Young Women Engaged in Transactional and Casual Sex

To answer the question whether and if so to what extent AGYW also congregate at sex work hotspots to meet casual and transactional sex partners, we built on the standard approach to programmatic mapping used for FSWs by first identifying sex work venues in Mombasa and then by asking the relevant questions about women who engaged in casual and transactional sex at these locations. Namely, we modified the premapping and planning exercise and level two interviews of the standard programmatic mapping approach but adhered to the standard practice for all other phases (Figure 1). In premapping and planning, the expanded scope of the modified method to also estimate the population size of AGYW engaging in sex work and transactional and casual sex was raised with the implementation team and peers during community engagement to prepare the team for the later phases of the mapping exercise. The study's hypothesis about colocation of AGYW subgroups at sex work hotspots seeking different types of sexual partnership was specifically discussed. In level two interviews, the team asked primary KIs at each hotspot about the number of young women aged 14 to 24 years engaging in sex work and transactional and casual sex, the number of men who frequent the hotspot seeking FSWs and other sex partners, and other places where women visit to meet partners for transactional or casual sex.

In contrast to FSWs, mobility of women engaging in transactional and casual sex was assumed to be limited. As such, the mean peak day population estimate for women engaging in transactional and casual sex at each hotspot was taken as the best estimate for that hotspot. All hotspot-level estimates were summed to generate an estimate for the entire county.

Implementation of the Modified Programmatic Mapping Approach in Mombasa, Kenya

The modified mapping approach was implemented between May 15, 2014, and June 20, 2014, by a 30-member team comprising International Centre for Reproductive Health Kenya (ICRH) research staff and peer educators with experience linking FSWs within their community to ICRH and other local HIV prevention and support programs. Based on existing administrative divisions, Mombasa County was segmented into 4 subcounties encompassing 9 data collection zones. The types of KIs approached for level one interviews included FSWs; men who have sex with men (including male sex workers); drug peddlers; beach boys and girls; public transport drivers (eg, taxi, taxi motorcycles, auto rickshaws, and minibuses); owners, staff, and patrons of internet shops/video dens; security guards, watchmen, and community policing groups; bouncers; bar owners, staff, and patrons; massage parlor owners and staff; brothel/sex den owners and staff; hotel/lodge owners, staff, and patrons; khat vendors; local brew sellers and patrons; staff of the Government of Kenya and NGOs; community health workers and health facility service providers; police/law enforcement agents; pharmacists; and village chiefs, assistant chiefs, and elders.

Sex work hotspots in Mombasa can be categorized into 8 general typologies: (1) public place (eg, beach, park); (2) street; (3) bar, nightclub, casino, and hotel (ie, venues with rooms); (4) bar, restaurant, and café (ie, venues without rooms); (5) guesthouse and lodge (ie, venues without bars); (6) sex den/brothel; (7) local brew den (ie, street kiosks selling *mnazi*, palm wine made from naturally fermented coconut tree sap); and (8) other (eg, home, massage parlors, saunas, video dens, and truck stops). These typology categories were developed in partnership with ICRH.

Statistical Analysis

Chi-square test was used to compare the distribution of sex work hotspots by subcounties and the distribution of previously known and newly identified hotspots by typology. A difference is considered significant if $P < .05$.

Ethical Approval

This study was approved in Kenya by the Kenyatta National Hospital/University of Nairobi Ethics and Research Committee;

the Research Permit Committee of the National Commission for Science, Technology, and Innovation; and in Canada by the Human Research Ethics Board at the University of Manitoba.

Results

Population Size Estimates

We estimated that there were 11,777 FSWs (range 9265 to 14,290) in Mombasa in 2014 among whom 6127 (range 4793 to 7462) (52.03%) were between 14 and 24 years and represented about 4.54% (6127/134,885) of the general female population of the same age group. The population estimates for women aged 14 to 24 years who engaged in transactional and casual sex and congregated at the hotspots were 5348 (range 4185 to 6510) and 4160 (range 3194 to 5125), respectively.

Combined, young women who engaged in transactional and casual sex represented about 7.05% (9508/134,885) of the female population of the same age group (Table 1).

Sexual Network Overlap

A total of 1183 sex work hotspots were named during level one interviews, of which 1025 (86.64%) were validated to be active. Among these validated hotspots, 102 (9.95%), 6 (0.59%), and 870 (84.88%) were also locations visited by women engaged in transactional sex, casual sex, or both, respectively. Only 47 (4.58%) hotspots were exclusive sex work locations (Figure 3).

Hotspot Typology and Distribution

Of the 1025 sex work hotspots, local brew dens were the most numerous overall, followed by bars/restaurants/cafés (Table 2).

Table 1. Estimated population size of women engaging in sex work, transactional sex, and casual sex in Mombasa County, Kenya.

Women engaging in:	Point estimate	Range	Proportion of general female population of the same age group, %
Sex work			
All ages	11,777	9265-14,290	3.69 ^a
14-24 years	6127	4793-7462	4.54
Transactional sex			
14-24 years	5348	4185-6510	4.54
Casual sex			
14-24 years	4160	3194-5125	3.08

^aProportion of general female population of reproductive age: 319,032 (14 to 44 years).

Figure 3. Overlap in the 1025 sex work hotspots in Mombasa County, Kenya, that were also visited by women engaging in transactional and/or casual sex.

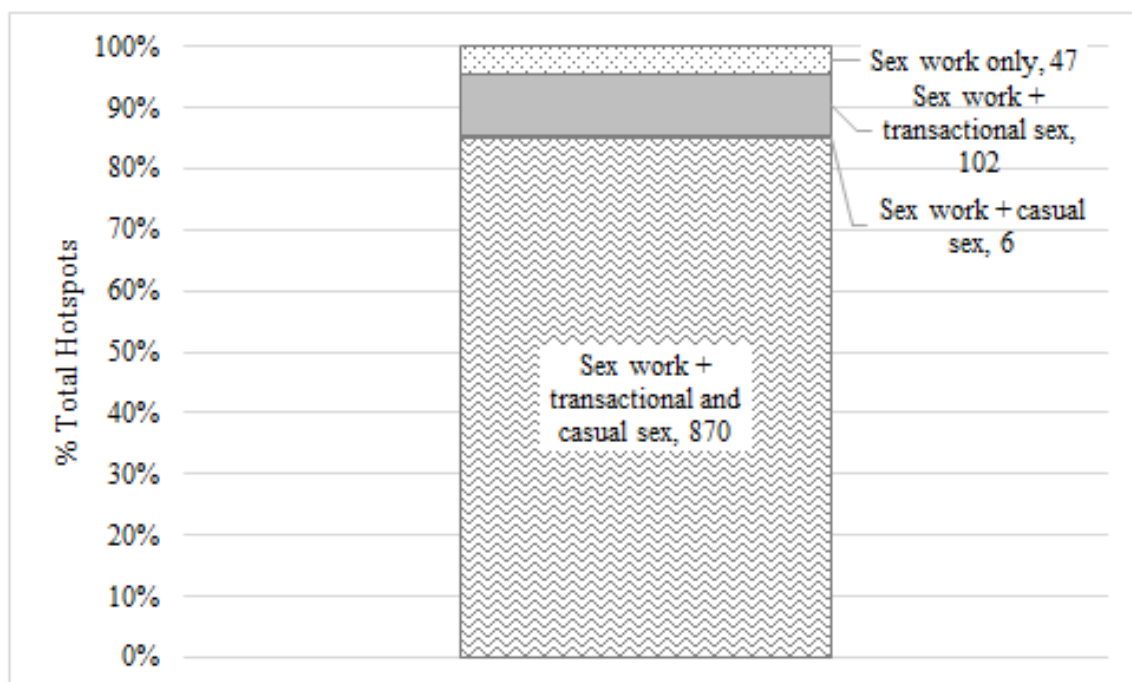


Table 2. Distribution of sex work hotspots by typology and subcounty in Mombasa County, Kenya.

Sex work hotspot typology	Subcounty A ^a (n=280), n (%)	Subcounty B ^a (n=164), n (%)	Subcounty C ^a (n=466), n (%)	Subcounty D ^a (n=115), n (%)	Mombasa County (N=1025), n (%)
Public place	12 (4.3)	6 (3.7)	13 (2.8)	2 (1.7)	33 (3.2)
Street	21 (7.5)	29 (17.7)	31 (6.7)	2 (1.7)	83 (8.1)
Bar/nightclub/casino/hotel (with rooms)	63 (22.5)	40 (24.4)	81 (17.4)	26 (22.6)	210 (20.5)
Bar/restaurant/café (without rooms)	82 (29.3)	49 (29.9)	118 (25.3)	22 (19.1)	271 (26.4)
Guesthouse/lodge (without bars)	19 (6.8)	10 (6.1)	38 (8.2)	6 (5.2)	73 (7.1)
Sex den/brothel	8 (2.9)	6 (3.7)	2 (0.4)	0 (0.0)	16 (1.6)
Local brew den	68 (24.3)	19 (11.6)	173 (37.1)	57 (49.6)	317 (30.9)
Other ^b	7 (2.5)	5 (3.0)	10 (2.1)	0 (0.0)	22 (2.1)

^aThe distribution of hotspots between the four subcounties was significantly different ($P<.001$).

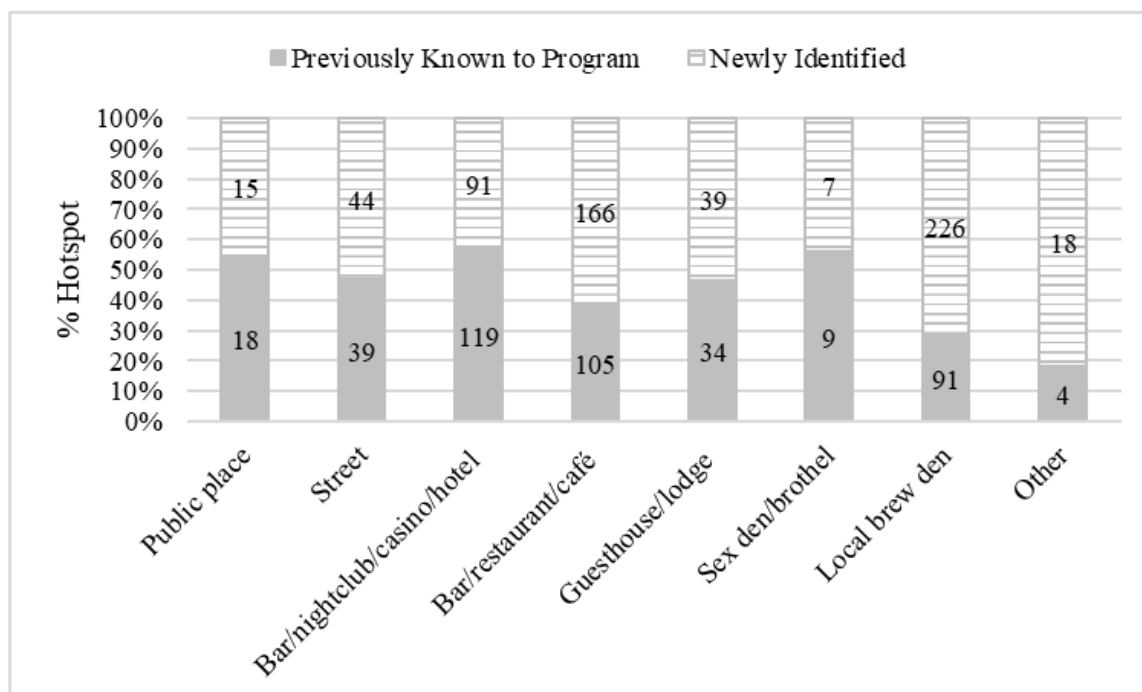
^bThe Other category includes hotspot types such as home, massage parlor/sauna, video den, and truck stop.

Of these 1025 hotspots, 280 (27.32%) were located in subcounty A, 164 (16.00%) in subcounty B, 466 (45.46%) in subcounty C, and 115 (11.22%) in subcounty D. The distribution of the different types of hotspots were significantly different between the four subcounties (Table 2, $P<.001$).

Figure 4 shows the typological distribution of sex work hotspots that were already known to the local HIV prevention program

prior to this study and those that were newly identified. Of the 1025 validated active hotspots, 419 (40.88%) were already known to the program and 606 (59.12%) were newly identified. The typological distribution of previously known hotspots and newly identified hotspots varied significantly ($P<.001$). Bars/nightclubs/casinos/hotels were the most common among previously known hotspots, whereas local brew dens were the most common among newly identified hotspots.

Figure 4. Proportion of sex work hotspots within each typology that were known to the local HIV prevention program prior to this study and those that were newly identified in Mombasa County, Kenya. The Other category includes hotspot types such as home, massage parlors/saunas, video dens, and truck stops.



Discussion

Principal Findings

Using an established mapping method as the foundation, we developed a more nuanced approach to identify the distribution and population size of subgroups at risk of HIV. Our subgroups of interest were AGYW engaged in a spectrum of sexual partnerships who met their sex partners at sex work hotspots. We estimated that the 2014 FSW population size in Mombasa was 11,777 (range 9265 to 14,290), and our estimate largely falls within the range of the estimates from the 2012 mapping round using the standard approach outlined in this paper (point estimate 9288; range 6917 to 11,660) [5,24]. We also estimated that half of the FSW population was between the ages of 14 and 24 years, while a further 9500 young women in the same age group practiced transactional and casual sex. We found that there was a substantial overlap in spaces where young women met sex work clients and partners for transactional and casual sex, thereby providing opportunities for women engaging in transactional and casual sex to interact with higher risk social and sexual networks. Using this modified mapping approach, our results show that local HIV programs can use a sex work hotspot-based mechanism, as opposed to a general population approach, to reach and deliver services to AGYW subgroups at increased risk of HIV. Moreover, this modified mapping approach provides a deeper layer of strategic information for HIV/STI service design and planning and for developing representative sampling frames for surveillance and research studies.

We found that the hotspot typology profile varies significantly geographically between the four subcounties. When comparing the typologies of sex work hotspots known to the local program prior to this study and those newly identified, we additionally found that hotspot turnover was high (ie, high rates of hotspot closure and opening of new ones). Because Mombasa is a vacation destination for both Kenyan and foreign tourists alike, the vibrancy of associated industries, including the sex industry, is directly affected by the boom and slump of tourism at large [25,26]. Aside from broad economic factors, social and political pressure can also affect the viability of sex work-related businesses. Whether they are exclusive to sex work or not, businesses may suffer or may no longer exist as a sex work hotspot if there have been police crackdowns or if their popularity has dwindled among sex work clients. In both scenarios, activities between sex workers and their clients will be displaced to other locations. We have seen substantial changes among local brew dens as well as other sex work hotspots over time. Specifically, we identified a total of 317 local brew dens in 2014, among which 226 (71.3%) were new. Besides an evident increase in number, because roadside local brew dens are highly mobile, the same business can move to more desirable neighborhoods where the demand for their product is high. As a result, sex workers and their clients also tend to move with the local brew dens, in which case the new location would be registered as a new sex work hotspot in the 2014 mapping round. Furthermore, we validated 1025 active sex work hotspot in this study, whereas the 2012 mapping exercise identified only 774. This overall high hotspot turnover

suggests that a significant proportion of active hotspots were not known to the program and therefore no outreach was provided to women at these hotspots. Taken together, the geographic heterogeneity in hotspot typology, high hotspot turnover, and notable increase in the number of hotspots suggest that mapping should be performed iteratively and updated over time in order for local HIV prevention programs to make the necessary service linkage and delivery adaptations responsive to the changing landscape of sex work. Recognizing the resource and time demands on local HIV prevention programs to conduct a full-scale programmatic mapping exercise, regular updates on hotspots and population size estimates can be performed by concentrating effort and resources on the level two activity during which hotspots are validated and primary KIs are interviewed to identify other nearby sex work locations. All locations new to the program should be visited and validated. This validation process needs not be one-off or intensive, but rather it can be built into the program as part of a routine practice during visits by peer educators and outreach workers to hotspots in the community to connect with members of the key population. This method has been used by the local HIV prevention program in Karnataka State, South India, implemented by Karnataka Health Promotion Trust and the Karnataka State AIDS Prevention Society. Since the initial mapping and population size estimation of FSWs and men who have sex with men in Karnataka State in 2004, the program has performed regular annual updates based on the number of key population members who were contacted by outreach and reached for services. If comprehensive mapping information is not available as the foundation, the initial hotspot list can be based upon existing program knowledge (ie, the list of locations to where programs provide outreach services). It is important to note that an estimate given for any hotspot is specific to that hotspot, whereas the size estimate of a key population in a greater geographic area (eg, town, district, county) has accounted for the movement of that key population. Therefore, while a program's capacity to conduct comprehensive mapping may be limited, keeping track of hotspot locations and estimating key population size at a basic level would still be informative for service planning and delivery from a program's perspective.

Programmatic mapping is a population size estimation method that is based on the collective knowledge of individuals associated with a key population or who are members themselves about where the key population of interest is located within a defined geographic area. Unlike many other population size estimation methods that generate estimates that can be wildly variable and difficult to validate [2,3], by tethering population size estimates to spaces where vulnerability and risk cluster, mapping can provide programs with reasonable estimates that can be refined as a program is rolled out and monitored over time. By overlaying data regarding HIV prevalence, sexual network, and risk behavior on top of the key population size data linked to space, this clustering of multiple factors that contribute to HIV transmission risk will help local HIV programs make strategic decisions related to geographic allocation of prevention resources, design the right intervention mix appropriate for the level of risk, and reach key populations for delivery of services to ensure program coverage and efficiency [27].

Limitations

There are limitations to the use of programmatic mapping for identifying hotspots and estimating key population size. Because mapping involves collection of empirical data that are often sensitive, the quality, accuracy, and comprehensiveness of the data are highly dependent on connecting with informants who are knowledgeable about the structure and operation of sex work. Therefore, during the premapping phase, it is important to dedicate time, as part of project planning, to engage with the FSW community and NGOs serving FSWs in order to understand the context of sex work and adapt the mapping approach that fits with the situation. For the purposes of this study, we worked with a well-established NGO and engaged a trusted network of peer educators.

Despite a well-planned and well-executed mapping exercise, the population size estimates generated by this method still likely underestimate the true population size due to the method's limitation to account for populations who congregate in private and virtual spaces. With the growing popularity of social networking websites and mobile apps, virtual spaces have opened up new communication channels for people to connect for sexual activities. To begin to address this question, our team has piloted a new mapping method in conjunction with analysis of internet-based networks to generate crude size estimates for online key populations [28].

Pertinent to the modifications introduced in this study to the standard programmatic mapping approach, another limitation is the accuracy in estimating the number and age of AGYW who might be at hotspots connecting with potential partners for transactional and casual sex. More work, including another round of mapping, will need to be done to validate the current estimates and refine the modified method. Nevertheless, these populations size estimates for AGYW provide a starting place for HIV prevention programs focusing on AGYW to assess the coverage of their current services and inform program adjustments as required.

Conclusions

Programmatic mapping is an effective method that can be adapted to understand the size and geographical organization of diverse priority populations for resource allocation, program planning, service delivery, and monitoring and evaluation. Mapping has been extensively used for planning HIV/STI prevention and control programs with key populations. By adapting the standard programmatic mapping approach used traditionally for key populations, our results show that HIV prevention programs tailored to AGYW can focus delivery of interventions to traditional sex work hotspots to reach subgroups that may be at increased risk for HIV. Effective HIV prevention and control programs need to be responsive to the evolutions of an HIV epidemic, and mapping is an integral tool for the iterative process of program planning, adaptation, and evaluation.

Acknowledgments

We would like to thank all the participants of this study. We also wish to acknowledge the efforts of the Transitions Study team and our partners, the National AIDS and STI Control Programme of the Ministry of Health of Kenya and the International Centre for Reproductive Health Kenya. This work was supported by the Canadian Institutes of Health Research (CIHR) (funding reference numbers MOP-130441 and FDN-13455). MLB is supported by a CIHR New Investigator Award, SM by a CIHR–Ontario HIV Treatment Network New Investigator Award, MP by a Tier 2 Canada Research Chair, and JFB by a Tier 1 Canada Research Chair.

Conflicts of Interest

None declared.

References

1. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach, 2nd Edition.: World Health Organization; 2016. URL: http://apps.who.int/iris/bitstream/handle/10665/208825/9789241549684_eng.pdf [accessed 2019-02-14]
2. 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](https://pubmed.ncbi.nlm.nih.gov/24393694/)]
3. Guidelines on estimating the size of populations most at risk to HIV.: UNAIDS/WHO Working Group on Global HIV/AIDS/STI Surveillance; 2010. URL: http://www.unaids.org/sites/default/files/media_asset/2011_Estimating_Populations_en_0.pdf [accessed 2019-02-14] [WebCite Cache ID 6zorzXt5n]
4. Lorway R, Lazarus L, Chevrier C, Khan S, Musyoki HK, Mathenge J, et al. Ecologies of security: on the everyday security tactics of female sex workers in Nairobi, Kenya. *Glob Public Health* 2018 Mar 05:1-14. [doi: [10.1080/17441692.2018.1442487](https://doi.org/10.1080/17441692.2018.1442487)] [Medline: [29506439](https://pubmed.ncbi.nlm.nih.gov/29506439/)]
5. 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/)]
6. 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](https://pubmed.ncbi.nlm.nih.gov/23627850/)]

7. Chase R, Foran J, Rasheed A, O'Neil J, Moses S, Blanchard J. Mapping and situation assessment of key populations at high risk of HIV in three cities of Afghanistan.: The World Bank, South Asia Human Development Sector; 2008 Apr 20. URL: <http://documents.worldbank.org/curated/en/300611468197940696/pdf/437330NWPOBox31gAfghanistanApril708.pdf> [accessed 2019-02-14] [WebCite Cache ID 6zos9ZLKS]
8. Emmanuel F, Blanchard J, Zaheer HA, Reza T, Holte-McKenzie M, HASP team. The HIV/AIDS Surveillance Project mapping approach: an innovative approach for mapping and size estimation for groups at a higher risk of HIV in Pakistan. *AIDS* 2010 Jul;24 Suppl 2:S77-S84. [doi: [10.1097/01.aids.0000386737.25296.c4](https://doi.org/10.1097/01.aids.0000386737.25296.c4)] [Medline: [20610953](https://pubmed.ncbi.nlm.nih.gov/20610953/)]
9. 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](https://doi.org/10.1016/S1473-3099(12)70066-X)] [Medline: [22424777](https://pubmed.ncbi.nlm.nih.gov/22424777/)]
10. Wariki WMV, Ota E, Mori R, Koyanagi A, Hori N, Shibuya K. Behavioral interventions to reduce the transmission of HIV infection among sex workers and their clients in low- and middle-income countries. *Cochrane Database Syst Rev* 2012 Feb 15(2):CD005272. [doi: [10.1002/14651858.CD005272.pub3](https://doi.org/10.1002/14651858.CD005272.pub3)] [Medline: [22336811](https://pubmed.ncbi.nlm.nih.gov/22336811/)]
11. Kerrigan DL, Fonner VA, Stromdahl S, Kennedy CE. Community empowerment among female sex workers is an effective HIV prevention intervention: a systematic review of the peer-reviewed evidence from low- and middle-income countries. *AIDS Behav* 2013 Jul;17(6):1926-1940. [doi: [10.1007/s10461-013-0458-4](https://doi.org/10.1007/s10461-013-0458-4)] [Medline: [23539185](https://pubmed.ncbi.nlm.nih.gov/23539185/)]
12. Moses S, Plummer FA, Ngugi EN, Nagelkerke NJ, Anzala AO, Ndinya-Achola JO. Controlling HIV in Africa: effectiveness and cost of an intervention in a high-frequency STD transmitter core group. *AIDS* 1991 Apr;5(4):407-411. [Medline: [1905555](https://pubmed.ncbi.nlm.nih.gov/1905555/)]
13. Fonck K, Kaul R, Kimani J, Keli F, MacDonald KS, Ronald AR, et al. A randomized, placebo-controlled trial of monthly azithromycin prophylaxis to prevent sexually transmitted infections and HIV-1 in Kenyan sex workers: study design and baseline findings. *Int J STD AIDS* 2000 Dec;11(12):804-811. [doi: [10.1258/0956462001915327](https://doi.org/10.1258/0956462001915327)] [Medline: [11138916](https://pubmed.ncbi.nlm.nih.gov/11138916/)]
14. Wilson D. HIV programs for sex workers: lessons and challenges for developing and delivering programs. *PLoS Med* 2015 Jun;12(6):e1001808 [FREE Full text] [doi: [10.1371/journal.pmed.1001808](https://doi.org/10.1371/journal.pmed.1001808)] [Medline: [26079267](https://pubmed.ncbi.nlm.nih.gov/26079267/)]
15. UNAIDS Data 2017. URL: http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf [accessed 2019-02-14] [WebCite Cache ID 6zmsEynir]
16. Kenya AIDS Response Progress Report 2016.: Ministry of Health of Kenya and National AIDS Control Council URL: <https://www.medbox.org/ke-studies-reports/kenya-aids-response-progress-report-2016/preview?> [accessed 2019-02-14] [WebCite Cache ID 6zosWw1e3]
17. Get on the fast-track: the life-cycle approach to HIV.: UNAIDS URL: http://www.unaids.org/sites/default/files/media_asset/Get-on-the-Fast-Track_en.pdf [accessed 2019-02-14] [WebCite Cache ID 6zIMUCp66]
18. County statistical abstract—Mombasa County 2015.: Kenya National Bureau of Statistics URL: <https://www.knbs.or.ke/download/mombasa/> [accessed 2018-05-31] [WebCite Cache ID 6zos1yUXf]
19. Economic Survey 2017.: Kenya National Bureau of Statistics URL: <https://www.knbs.or.ke/download/economic-survey-2017/> [accessed 2018-05-31] [WebCite Cache ID 6zoscqWjR]
20. Kenya HIV County Profiles 2016.: Ministry of Health of Kenya and National AIDS Control Council URL: <https://www.fast-trackcities.org/sites/default/files/Kenya%20HIV%20County%20Profiles%20%282016%29.pdf> [accessed 2018-05-31] [WebCite Cache ID 6zosoQxt7]
21. Mombasa County HIV & AIDS Strategic Plan 2016-2020.: Mombasa County Government URL: <https://nacc.or.ke/county-aids-strategic-plans-casps/> [accessed 2018-11-08] [WebCite Cache ID 73mljgGkZ]
22. Stoebenau K, Heise L, Wamoyi J, Bobrova N. Revisiting the understanding of "transactional sex" in sub-Saharan Africa: a review and synthesis of the literature. *Soc Sci Med* 2016 Dec;168:186-197 [FREE Full text] [doi: [10.1016/j.socscimed.2016.09.023](https://doi.org/10.1016/j.socscimed.2016.09.023)] [Medline: [27665064](https://pubmed.ncbi.nlm.nih.gov/27665064/)]
23. Stoebenau K, Nixon SA, Rubincam C, Willan S, Zembe YZ, Tsikoane T, et al. More than just talk: the framing of transactional sex and its implications for vulnerability to HIV in Lesotho, Madagascar and South Africa. *Global Health* 2011 Sep 30;7:34 [FREE Full text] [doi: [10.1186/1744-8603-7-34](https://doi.org/10.1186/1744-8603-7-34)] [Medline: [21961516](https://pubmed.ncbi.nlm.nih.gov/21961516/)]
24. Geographic mapping of most at risk populations for HIV (MARPs) in Kenya.: National AIDS and STI Control Programme URL: <http://www.icop.or.ke/wp-content/uploads/2016/09/Kenya-mapping-report-NASCOP-and-UoM-2013.pdf> [accessed 2018-05-31] [WebCite Cache ID 6zot2teXy]
25. Kibicho W. A critical evaluation of how tourism influences commercial sex workers' operations in the Malindi area, Kenya. *Ann Leisure Res* 2013 Jan 14;7(3-4):188-201. [doi: [10.1080/11745398.2004.10600951](https://doi.org/10.1080/11745398.2004.10600951)]
26. Kibicho W. Tourism and the sex trade in Kenya's coastal region. *J Sustainable Tourism* 2005 May 15;13(3):256-280. [doi: [10.1080/01434630508668556](https://doi.org/10.1080/01434630508668556)]
27. Aral SO, Torrone E, Bernstein K. Geographical targeting to improve progression through the sexually transmitted infection/HIV treatment continua in different populations. *Curr Opin HIV AIDS* 2015 Nov;10(6):477-482. [doi: [10.1097/COH.000000000000195](https://doi.org/10.1097/COH.000000000000195)] [Medline: [26352392](https://pubmed.ncbi.nlm.nih.gov/26352392/)]
28. Emmanuel F, Isac S. Programmatic Mapping of Key Populations in Mali. *Plan Mali* 2015. [doi: [10.2196/11196](https://doi.org/10.2196/11196)]

Abbreviations

AGYW: adolescent girls and young women
CIHR: Canadian Institutes for Health Research
FSW: female sex worker
ICRH: International Centre for Reproductive Health Kenya
KI: key informant
NGO: nongovernmental organization
STI: sexually transmitted infection

Edited by T Sanchez; submitted 31.05.18; peer-reviewed by J Zhao, J Clark; comments to author 21.09.18; revised version received 17.01.19; accepted 25.01.19; published 01.04.19.

Please cite as:

*Cheuk E, Isac S, Musyoki H, Pickles M, Bhattacharjee P, Gichangi P, Lorway R, Mishra S, Blanchard J, Becker M
Informing HIV Prevention Programs for Adolescent Girls and Young Women: A Modified Approach to Programmatic Mapping and
Key Population Size Estimation
JMIR Public Health Surveill 2019;5(2):e11196
URL: <https://publichealth.jmir.org/2019/2/e11196/>
doi: [10.2196/11196](https://doi.org/10.2196/11196)
PMID: [30932868](https://pubmed.ncbi.nlm.nih.gov/30932868/)*

©Eve Cheuk, Shajy Isac, Helgar Musyoki, Michael Pickles, Parinita Bhattacharjee, Peter Gichangi, Robert Lorway, Sharmistha Mishra, James Blanchard, Marissa Becker. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 01.04.2019. 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

Prevalence of *Schistosoma Haematobium* Measured by a Mobile Health System in an Unexplored Endemic Region in the Subprefecture of Torrock, Chad

Didier Lalaye^{1*}, MD; Mirjam E de Bruijn^{2*}, MSc, PhD; Tom PVM de Jong^{3*}, MD, PhD

¹Julius Global Health Center, University Medical Center Utrecht, Utrecht, Netherlands

²African Studies Centre, Leiden University, Leiden, Netherlands

³Department of Pediatric Urology, University Children's Hospitals University Medical Center Utrecht and Amsterdam University Medical Center, Utrecht, Netherlands

* all authors contributed equally

Corresponding Author:

Didier Lalaye, MD

Julius Global Health Center

University Medical Center Utrecht

Heidelberglaan 100

Utrecht, 3584CX

Netherlands

Phone: 31 887555555

Fax: 31 887555384

Email: lalaye84@gmail.com

Abstract

Background: *Schistosoma haematobium* is a parasitic digenetic trematode responsible for schistosomiasis (also known as bilharzia). The disease is caused by penetration of the skin by the parasite, spread by intermediate host molluscs in stagnant waters, and can be treated by administration of praziquantel. Schistosomiasis is considered to be an important but neglected tropical disease.

Objective: The aim of this pilot study was to investigate the prevalence of schistosomiasis in the subprefecture of Torrock, an endemic area in Chad where no earlier investigation had been conducted and no distribution system for pharmacotherapy has ever existed.

Methods: This study examined 1875 children aged 1 to 14 years over a period of 1 year. After centrifugation, urine examination was performed by a direct microscopic investigation for eggs. The investigation was conducted with a mobile health (mHealth) approach, using short message service (SMS) for communication among parents, local health workers, a pharmacist, and a medical doctor. An initial awareness campaign requested parents to have their children examined for schistosomiasis. Urine was then collected at home by the parents following the SMS request. Urine results that proved positive were sent to a medical doctor by SMS, who in turn ordered a pharmacist by SMS to distribute praziquantel to the infected children.

Results: Direct microscopic examination of urine found 467 positive cases (24.9% of the total sample). Of all male and female samples, 341 (34%) and 127 (14.4%) samples were positive, respectively. The infection rate was equally distributed over age groups. The newly developed mHealth system had a limited level of participation (8%) from an estimated total of 25,000 children in the target group.

Conclusions: The prevalence of schistosomiasis in children in the subprefecture of Torrock is moderately high. Efforts will be required to enhance the awareness of parents and to reach a larger percentage of the population. Systematic governmental measures should be put in place as soon as possible to increase awareness in the area and to diagnose and treat cases of schistosomiasis.

(*JMIR Public Health Surveill* 2019;5(2):e13359) doi:[10.2196/13359](https://doi.org/10.2196/13359)

KEYWORDS

Schistosoma haematobium; prevalence; Chad; neglected tropical diseases; mobile health

Introduction

Schistosomiasis of the urogenital tract is a parasitic disease caused by *Schistosoma haematobium* (Bilharz), a flatworm of the class trematode of the genus *Schistosoma*, whose habitat is the venous circulatory system of the urinary tract and internal genitalia. The disease is spread by snails that live in stagnant infected water and act as intermediate hosts for the trematode. The snails expel Cercaria, a free swimming larval stage, in the water, which are able to penetrate the skin of people or cattle that are in the water. Once in the body, the Cercaria migrate to the blood vessels surrounding the urogenital system and develop into adult worms. The female part of a pair of worms produces eggs that penetrate into the bladder producing symptoms of lower urinary tract disease and changes of the bladder wall that may lead to malignancies of the bladder at a later age. Clinical signs may vary from hematuria with few extra complaints to severe forms of irritative bladder complaints and obstruction of the upper urinary tract caused by fibrosis of ureters. The eggs are excreted with the urine into the water and enter the snail to finish the cycle. To diagnose, the disease direct detection of eggs under a microscope in a centrifuged urine sample and detection of hematuria by urine test strips are the most commonly used tests [1].

Schistosomiasis often results in severe physical, social, and economic disabilities [2]. Schistosomiasis is a major public health problem in endemic tropical and subtropical countries. After malaria, it occupies the first place in importance in public health with regard to its prevalence [3]. In contrast to malaria, which affects people without discrimination of social position or income, schistosomiasis primarily affects poor people in remote areas [1]. This is a consequence of inadequate water, Sanitation, and hygiene conditions, based on poor infrastructure and habits in agriculture and livestock farming. This may be one of the important reasons for schistosomiasis being considered a neglected tropical disease.

In Africa, schistosomiasis affects more than 200 million people in rural or periurban areas, of whom 120 million show symptoms of the disease. For 20 million of those infected, this may have severe consequences because of increased mortality risk. In many areas, many adults and a high proportion of children under the age of 14 are infected. It is estimated that 650 million people worldwide live in endemic areas [4].

Urogenital schistosomiasis occurs in tropical and subtropical areas of the world where climatic, ecological, and socioeconomic conditions favor the spread of the disease. The risk of infection is high particularly in flood and irrigation areas [5].

In most endemic areas, the disease is diagnosed only incidentally when the patient consults a health worker for other reasons. As a result, many cases go unnoticed and the disease is rarely recognized as a cause of death in Sub-Saharan Africa, although there are reports of 280,000 deaths per year [6]. Treatment of

schistosomiasis is relatively simple, by the administration of the antihelminthic medication praziquantel [7].

A survey in 2010 by the Koyom District Hospital in Chad on the prevalence of helminthiasis and schistosomiasis among rural students in Chad found a prevalence of 25% of urinary schistosomiasis [8]. In many countries, mass drug administration programs exist where the population at risk is treated routinely with praziquantel. In Chad, no reliable data exist on the prevalence of the disease, given the absence of a national program for the control of schistosomiasis and also because the organization of the health care system does not provide diagnostics to prove this pathology—thus hindering the pathway to treatment. The World Health Organization provides data on the numbers of people at risk for schistosomiasis in Chad [9].

Over 2016, more than 3 million people were considered to be in need for yearly drug administration, none have been reported receiving medication.

Our pilot study aimed to determine the prevalence of schistosomiasis in the population of Torrock and to test the effectiveness of a mobile health (mHealth) system installed locally. This followed a pilot study conducted in 2014 at the Torrock health center, which revealed 157 positive cases (26%) from a total of 611 children aged 1 to 14 years.

Methods

The study was conducted between March 2015 and March 2016, based on a sampling and treatment system and controlled by short message service (SMS) on mobile telephones.

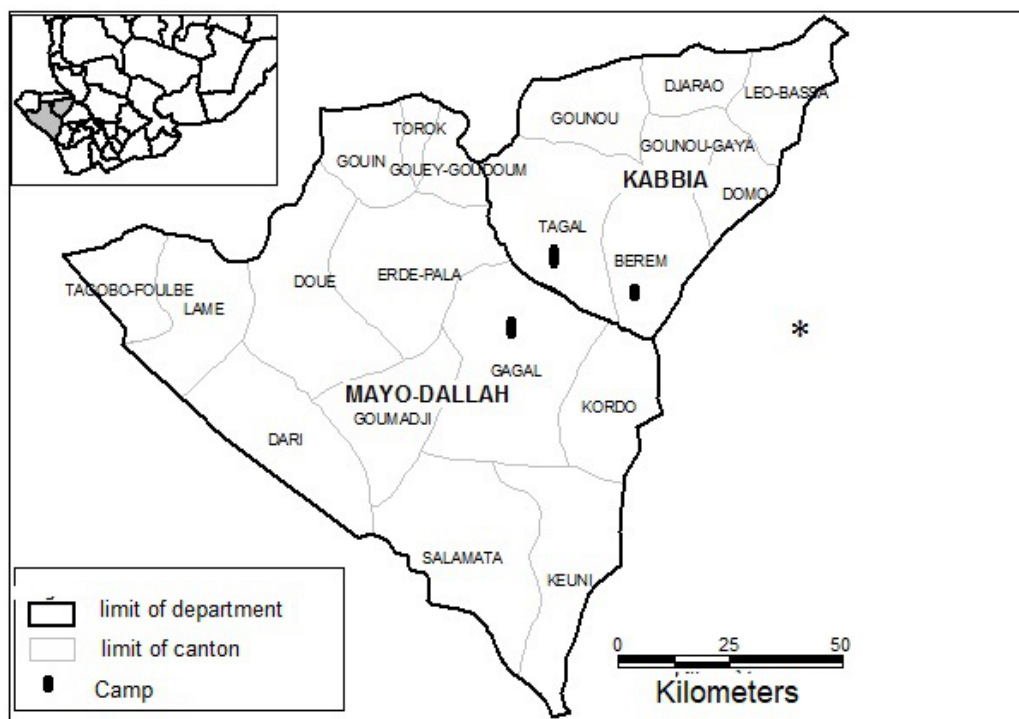
Study Zone: Mayo-Kebbi Ouest

In Chad, the subprefecture of Torrock is in the department of Mayo-Kebbi West, located 70 km north of Pala, the capital of the province (see Figure 1). According to the last census (2009), the population is 50,000 inhabitants, with an estimated percentage of 50.6% for children between 1 and 15 years [10]. Torrock is composed of 3 cantons and has a total of 9 health centers, none of which had laboratory facilities before the start of the study. Running alongside the village of Torrock, where the sample population was chosen, there is a river called El Madorbob, the main source of molluscs that act as intermediate hosts of schistosomiasis.

No earlier studies, except for a small pilot study carried out by the first author, had been conducted in this area, and no program for distribution of medication had ever existed there.

The population had earlier been informed of the possibility to have children tested for schistosomiasis by urine sampling at home, whether they showed symptoms or not. This information was spread via churches, mosques, markets, schools, mouth to mouth, and FM radio stations serving the region both in the local language Moundang and in French. Information was given by the local announcer by megaphone, by imams in mosques, and by priests in churches. Information contained specific items on most aspects of urinary schistosomiasis.

Figure 1. Map of Mayo Kebbi Ouest.



Following this awareness campaign in the first half of March 2015, an SMS relay system was set up, which allowed children to have their urine examined at laboratories in 3 village health centers: Torrock, Gouin, and Goy-Goudoum, all in the subprefecture of Torrock. These laboratories perform limited examinations. Urine analysis for *Schistosoma haematobium* was unavailable before the study. During the study, urine analysis for *Schistosoma haematobium* was provided free of charge for the study group. After positive testing, the correct dosage of praziquantel for a child was sent to his or her home after SMS contact with a medical doctor, who prescribed the medication by SMS to the regional pharmacist.

The chain for diagnostics and treatment (Figure 2 provides a graphic example of the SMS health system. The numbers refer to the different steps in the procedure) starts with a representative in the village with a mobile phone, these were 4 males, appointed by the chiefs of the villages. They are responsible for sending test request messages to the liaison officer from parents who are unable to do so themselves and were paid for each patient by the project. They link the parents to the diagnostic and therapeutic system. The liaison officer receives the message from a parent or a representative, bicycles to the area and takes the urine sample from the child's house to

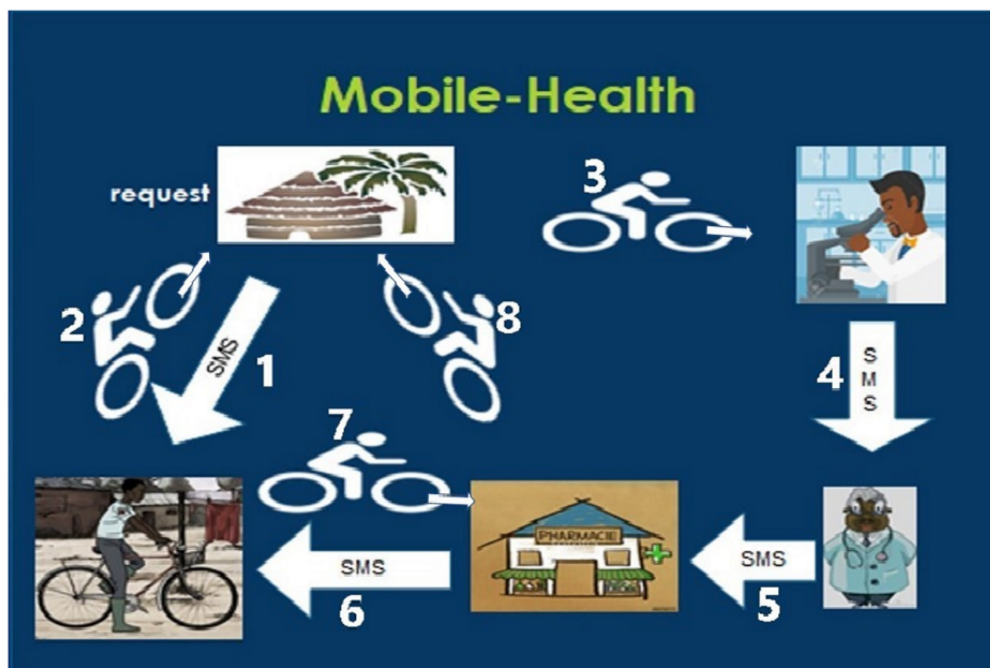
the health center for an analysis. The liaison officer also weighs the child, records the child's name, and creates a personal code. The officer is also responsible for delivery of the praziquantel in cases of a positive diagnosis. The liaison officer was appointed by the health centers and provided with a bicycle and was paid a small amount of money for each patient.

The laboratory worker in the health center performs analyses of the urine collected by the liaison officer. He informs the physician of the positive cases of schistosomiasis by noting name, code, and weight of the child. They did not receive extra payment for the project other than their regular salary.

The physician receives the urine results from the laboratory via internet or SMS, calculates the dosage of the drug according to the weight of the child, and issues orders to the pharmacist.

The pharmacist is responsible for delivering medication with the patient's name and personal code on the package. The pharmacist sends an SMS message to the liaison officer, who delivers the medication to the families. For this study, 1 health worker trained for doing microscopy of urine samples, 1 liaison officer, 1 pharmacist, and 1 doctor were sufficient. Praziquantel was provided for free by the government district hospital.

Figure 2. The mobile health flow diagram.



Laboratory Test

Laboratory equipment consisted of dry tubes for urine collection, a mechanical centrifuge operated by hand, and an optical microscope (Olympus GX 201). The urine samples were collected between 9 am and 1 pm, because this is the period that maximal egg load in the urine can be expected [11]. The urine was collected in dry tubes and transported to the corresponding health center for analysis. The examination technique consisted of centrifugation for 10 min. The supernatant was discarded, and a drop of the centrifugate was examined under the microscope between glass slides. The microscopic reading of the samples was performed at magnification 100, and eggs were confirmed at magnification 400. All samples containing eggs measuring 65 to 145 μm , with a thick, pink, transparent shell with a terminal spur, were considered positive. The number of eggs of *Schistosoma haematobium* was not quantified, so there was no assessment of the intensity of the infection. Reason to do so was that the laboratory workers were not trained in counting eggs per field.

Inclusion Criterion

All children aged 1 to 14 years whose parents had made a request for urine testing by SMS and whose children visited the health center of their own accord or as a result of parental request.

Exclusion Criteria

Children whose parents did not make an SMS request, children under 1 year of age, and individuals 15 years of age and older.

Ethical Considerations

Regional health authority in the province of Mayo Kebbi Ouest granted written permission for this study. Before commencing the study and before the start of the campaign to raise awareness,

an application for authorization was sent to the subprefect of Torrock and the various administrative, religious, and educational authorities of the village. An authorization was granted by the health delegate of Mayo-Dallah for the establishment of 3 laboratories in the region for the analysis of urine. An institutional ethics committee did not exist in Chad at the start of the study.

Oral consent was given by the parents after explanation of the study's aim and procedure. Written consent was impossible because the vast majority of the population is illiterate. Data were collected anonymously.

Results

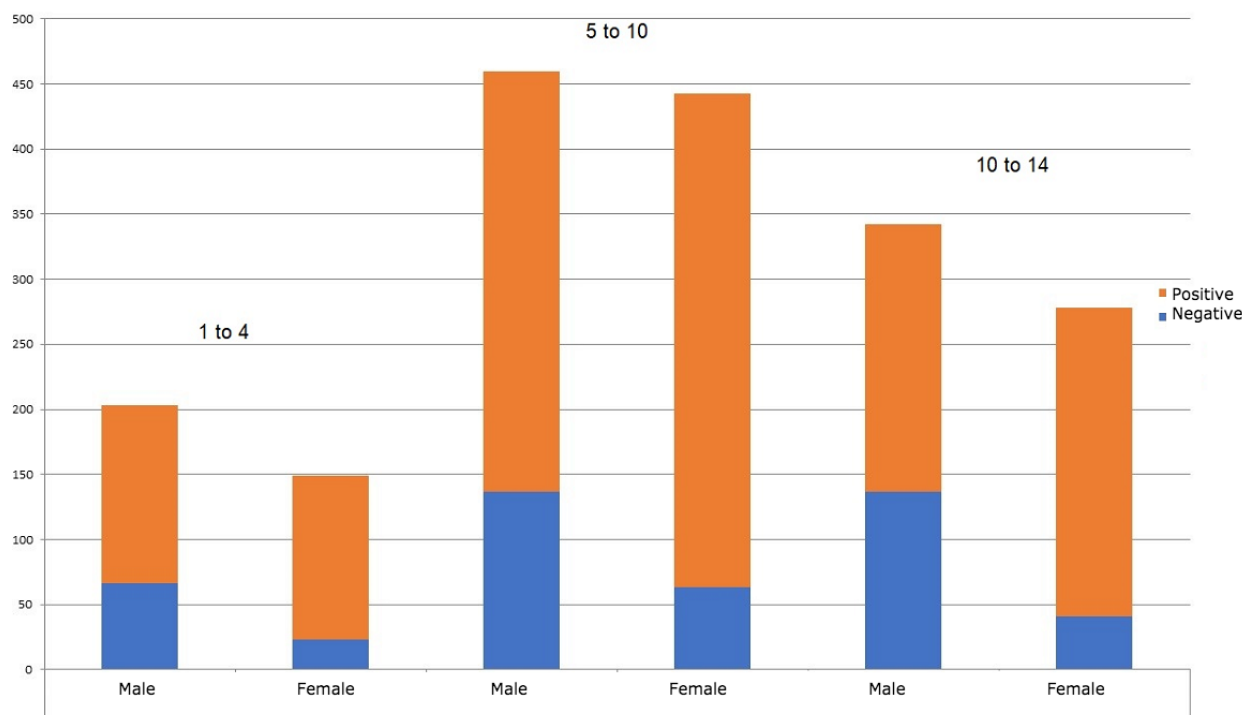
Out of an estimated number of approximately 25,000 children in the target group, 1875 requests for urine investigation were received from parents, which reflects 8% of the target population.

In 1 year, a total of 1875 home-based samples were investigated. Of 1875 urine samples, 467 cases proved positive (24.9%). The results are summarized in Table 1 and Figure 3, showing that prevalence was higher in males than in females.

Total rate of infection was approximately 25% in each age group, with an infection rate of 34% and 14% in males and females, respectively. Specified for age, males, and females 1 to 4 years old combined gave 25% (88/353) positive, 32% (64/202) positive with males, and 16% (24/151) positive with females. Males and females combined 5 to 10 years old gave 25% (200/803) positive, 30% (138/460) positive with males, and 18% (62/343) positive with females. Combined 11 to 14 years old gave 25% (179/718) positive, 33% (139/422) positive with males, and 14% (40/296) positive with females.

Table 1. Distribution of the presence of *Schistosoma haematobium* in the study population by gender.

Gender, n (%)	Positive, n (%)	Negative, n (%)	Total, n (%)
Male	341 (34)	664 (66)	1005 (53.6)
Female	126 (14.6)	743 (85.4)	870 (46.4)
Total	467 (24.9)	1408 (75.1)	1875 (100)

Figure 3. Distribution of the infected population by age group and gender.

Discussion

Principal Findings

This study revealed a prevalence of 24.9% of *Schistosoma haematobium* infection in our cohort aged 1 to 14 years (see Figure 3). This result corresponds to a study in a neighboring province, Koyom, which had an infection rate of 27.4% [8]. The 2 provinces have several points in common (rice fields and stagnant streams). This rate is lower than 2 studies in Nigeria, which had a prevalence of 40% and 47% [12,13]. The differences can be explained by the different landscape and the fact that the research in Nigeria was conducted in flood zones, where the activities of the population are based primarily on fishing and rice cultivation.

Of the total prevalence of 24.9%, the gender distribution is 73% males and 27% females. This distribution between genders is similar to that found by Dankoni and Tchuenté in northern Cameroon in 2014 [14]. They found a prevalence of 2.5% in male subjects versus 1% in females. The higher prevalence in males can be explained by the fact that boys are more active in stagnant water. It is acceptable for boys to be naked in the water up until puberty, whereas girls have to cover themselves at all ages. As a consequence, boys are more exposed to the parasite.

Surprisingly, no differences in prevalence are seen in different age groups. One would expect that a lower infection rate would

be present at an early age, when children are under parental control and their only contact with stagnant water is in the company of their parents. The group aged 5 to 10 years has more freedom and also takes care of livestock, whereas the group aged 11 to 14 years should have less exposure as they are at school. This is in contrast with studies carried out in 2014 by Dankoni and Tchuenté in Cameroon [14] and by Senghor et al in Senegal [15], who found a peak, respectively, in the group aged 7 to 10 years and 7 to 12 years. Although the sample size of the group 1 to 4 years old is relatively small compared with the other age groups, it is possible that parents have been biased by clinical signs in their child. We have no data to support this.

This study in Chad is the first study, after our earlier survey, into *Schistosoma haematobium* in this large endemic area. A limitation of this study was the fact that selection was made by requesting SMS contact from parents. Although all parents in the area have access to a local representative that can send an SMS, only 8% chose to do so. When asked specifically after the information campaign, 60 out of 100 parents claimed they did not receive or did not understand the information well enough. Another limitation of the study is that only microscopy of the urine samples has been used to detect infection. This may underestimate the prevalence because of a lower detection rate compared with dipstick use for hematuria. Finally, the sample size is relatively small by the use of a bicycle. Future studies will investigate methods for optimization of information

dissemination to the local people, the results of treatment, the rate of reinfection, and the difference in infection prevalence in areas with or without mHealth systems. The mHealth system proved to be cheap, quick and effective, useful in such remote contexts, and easy to operate.

Conclusions

This study, based on the use of an mHealth care system, investigated a total of 1875 home samples of urine and found

a prevalence of 24.9% of children aged 1 to 14 years infected with schistosomiasis. This demonstrates the extent of this pathology in the study area, and we hope that this model of *mHealth*, as the first phase of management of schistosomiasis in the region, will prove to be effective. Follow-up studies are currently underway to demonstrate the efficiency of the system.

Acknowledgments

The authors would like to thank Chad Plus for its support. It is a nongovernmental organization for sustainable development of youth, environment, culture, and health care, founded in 2012 and run by Chadians.

Conflicts of Interest

None declared.

References

1. Colley DG, Bustinduy AL, Secor WE, King CH. Human schistosomiasis. *Lancet* 2014 Jun 28;383(9936):2253-2264 [FREE Full text] [doi: [10.1016/S0140-6736\(13\)61949-2](https://doi.org/10.1016/S0140-6736(13)61949-2)] [Medline: [24698483](https://pubmed.ncbi.nlm.nih.gov/24698483/)]
2. Fenwick A. Praziquantel: do we need another antischistosome treatment? *Future Med Chem* 2015;7(6):677-680. [doi: [10.4155/fmc.15.16](https://doi.org/10.4155/fmc.15.16)] [Medline: [25996058](https://pubmed.ncbi.nlm.nih.gov/25996058/)]
3. Savioli L, Renganathan E, Montresor A, Davis A, Behbehani K. Control of schistosomiasis--a global picture. *Parasitol Today* 1997 Nov;13(11):444-448. [doi: [10.1016/S0169-4758\(97\)01141-1](https://doi.org/10.1016/S0169-4758(97)01141-1)] [Medline: [15275147](https://pubmed.ncbi.nlm.nih.gov/15275147/)]
4. Savioli L, Albonico M, Colley DG, Correa-Oliveira R, Fenwick A, Green W, et al. Building a global schistosomiasis alliance: an opportunity to join forces to fight inequality and rural poverty. *Infect Dis Poverty* 2017 Mar 23;6(1):65 [FREE Full text] [doi: [10.1186/s40249-017-0280-8](https://doi.org/10.1186/s40249-017-0280-8)] [Medline: [28330495](https://pubmed.ncbi.nlm.nih.gov/28330495/)]
5. King C, Dickman K, Tisch DJ. Reassessment of the cost of chronic helminthic infection: a meta-analysis of disability-related outcomes in endemic schistosomiasis. *Lancet* 2005;365(9470):1561-1569. [doi: [10.1016/S0140-6736\(05\)66457-4](https://doi.org/10.1016/S0140-6736(05)66457-4)] [Medline: [15866310](https://pubmed.ncbi.nlm.nih.gov/15866310/)]
6. Abou-Zeid A, Abkar T, Mohamed R. Schistosomiasis infection among primary school students in a war zone, Southern Kordofan State, Sudan: a cross-sectional study. *BMC Public Health* 2013 Jul 11;13:643 [FREE Full text] [doi: [10.1186/1471-2458-13-643](https://doi.org/10.1186/1471-2458-13-643)] [Medline: [23845226](https://pubmed.ncbi.nlm.nih.gov/23845226/)]
7. Caffrey CR. Schistosomiasis and its treatment. *Future Med Chem* 2015;7(6):675-676. [doi: [10.4155/fmc.15.27](https://doi.org/10.4155/fmc.15.27)] [Medline: [25996057](https://pubmed.ncbi.nlm.nih.gov/25996057/)]
8. Helena H. Prévalence et enquête sur les connaissances, attitudes et pratiques de la population de Koyom (MAYO-KEBBI EST) concernant la schistosomose urinaire. Thesis, University of N'Djamena, Chad 2012:77.
9. World Health Organization. Schistosomiasis URL: http://www.who.int/neglected_diseases/preventive_chemotherapy/sch/db/?units=minimal®ion=all&country=tcd&countries=tcd&year=2016
10. Plan REDU, Coop CLA, Internationale R. 2012. Resultats Definitifs Par Sous-Prefecture URL: www.1-world-1.com/td/pdf/World_doc.pdf
11. Ephraim R, Duah E, Cybulski J, Prakash M, D'Ambrosio M, Fletcher D, et al. Diagnosis of *Schistosoma haematobium* infection with a mobile phone-mounted Foldscope and a reversed-lens CellScope in Ghana. *Am J Trop Med Hyg* 2015 Jun;92(6):1253-1256 [FREE Full text] [doi: [10.4269/ajtmh.14-0741](https://doi.org/10.4269/ajtmh.14-0741)] [Medline: [25918211](https://pubmed.ncbi.nlm.nih.gov/25918211/)]
12. Abdullahi M, Saidu T. Prevalence of urinary schistosomiasis among school aged children in Wushishi Local Government Area of Niger State, Nigeria. *Bayero J Pure Appl Sci* 2012;4(2):55. [doi: [10.4314/bajopas.v4i2.10](https://doi.org/10.4314/bajopas.v4i2.10)]
13. Houmsou R, Amuta E, Sar T. Profile of an epidemiological study of urinary schistosomiasis in two local government areas of Benue state, Nigeria. *Int J Med Biomed Res* 2012;1(1):48. [doi: [10.14194/ijmbr.117](https://doi.org/10.14194/ijmbr.117)]
14. Dankoni E, Tchuente L. [Epidemiology of schistosomiasis and soil-transmitted helminthiasis in the sub-division of Kékem (West-Cameroon)]. *Int J Innov Appl Stud* 2014;8(4):1790 [FREE Full text]
15. Senghor B, Diallo A, Sylla S, Doucouré S, Ndiath MO, Gaayeb L, et al. Prevalence and intensity of urinary schistosomiasis among school children in the district of Niakhar, region of Fatick, Senegal. *Parasit Vectors* 2014 Jan 03;7:5 [FREE Full text] [doi: [10.1186/1756-3305-7-5](https://doi.org/10.1186/1756-3305-7-5)] [Medline: [24387599](https://pubmed.ncbi.nlm.nih.gov/24387599/)]

Abbreviations

mHealth: mobile health

SMS: short message service

Edited by G Eysenbach; submitted 10.01.19; peer-reviewed by M Navarro, N Gbalégba; comments to author 28.03.19; revised version received 08.04.19; accepted 23.04.19; published 18.06.19.

Please cite as:

Lalaye D, de Bruijn ME, de Jong TPVM

Prevalence of Schistosoma Haematobium Measured by a Mobile Health System in an Unexplored Endemic Region in the Subprefecture of Torrock, Chad

JMIR Public Health Surveill 2019;5(2):e13359

URL: <http://publichealth.jmir.org/2019/2/e13359/>

doi: [10.2196/13359](https://doi.org/10.2196/13359)

PMID: [31215519](https://pubmed.ncbi.nlm.nih.gov/31215519/)

©Didier Lalaye, Mirjam E de Bruijn, Tom PVM de Jong. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 18.06.2019. 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.

Corrigenda and Addenda

Correction: Google for Sexual Relationships: Mixed-Methods Study on Digital Flirting and Online Dating Among Adolescent Youth and Young Adults

James Lykens¹, MA; Molly Pilloton², MPH; Cara Silva², MPH; Emma Schlammm², BA; Kate Wilburn², BS; Emma Pence², BA

¹Center for Research and Education on Gender and Sexuality, San Francisco State University, San Francisco, CA, United States

²Youth Tech Health, Oakland, CA, United States

Corresponding Author:

James Lykens, MA
Center for Research and Education on Gender and Sexuality
San Francisco State University
835 Market Street
5th Floor, Suite 525
San Francisco, CA,
United States
Phone: 1 415 817 4520
Email: james.lykens@etr.org

Related Article:

Correction of: <http://publichealth.jmir.org/2019/2/e10695/>

(*JMIR Public Health Surveill* 2019;5(2):e14815) doi:[10.2196/14815](https://doi.org/10.2196/14815)

The manuscript “Google for Sexual Relationships: Mixed-Methods Study on Digital Flirting and Online Dating Among Adolescent Youth and Young Adults” (*JMIR Public Health Surveill* 2019;5(2):e10695) was initially published with two separate headings for the Methods section. This error has now been corrected and the second heading has been removed.

The correction will appear in the online version of the paper on the JMIR website on June 21, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been resubmitted to those repositories.

Submitted 24.05.19; this is a non-peer-reviewed article; accepted 30.05.19; published 21.06.19.

Please cite as:

Lykens J, Pilloton M, Silva C, Schlammm E, Wilburn K, Pence E

Correction: Google for Sexual Relationships: Mixed-Methods Study on Digital Flirting and Online Dating Among Adolescent Youth and Young Adults

JMIR Public Health Surveill 2019;5(2):e14815

URL: <http://publichealth.jmir.org/2019/2/e14815/>

doi:[10.2196/14815](https://doi.org/10.2196/14815)

PMID:[31228178](https://pubmed.ncbi.nlm.nih.gov/31228178/)

©James Lykens, Molly Pilloton, Cara Silva, Emma Schlammm, Kate Wilburn, Emma Pence. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 21.06.2019. 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

Capture-Recapture Among Men Who Have Sex With Men and Among Female Sex Workers in 11 Towns in Uganda

Kevin Apodaca^{1,2}, MPH; Reena Hemendra Doshi^{2,3}, MPH, PhD; Moses Ogwal⁴, BA; Herbert Kiyingi⁵, MPH, MBChB; George Aluzimbi⁵, MPH; Geoffrey Musinguzi⁴, MPH, PhD; Ibrahim Lutalo⁶, MSc; Evelyn Akello⁶, MSc; Wolfgang Hladik², MD, PhD

¹Public Health Institute, Oakland, CA, United States

²Centers for Disease Control and Prevention, Center for Global Health, Division of Global HIV and TB, Atlanta, GA, United States

³Centers for Disease Control and Prevention, Epidemic Intelligence Service, Atlanta, GA, United States

⁴Makerere University, School of Public Health, Kampala, Uganda

⁵Centers for Disease Control and Prevention, Division of Global HIV and TB, Kampala, Uganda

⁶Makerere University, School of Public Health, Monitoring and Evaluation Technical Support Program, Kampala, Uganda

Corresponding Author:

Reena Hemendra Doshi, MPH, PhD

Centers for Disease Control and Prevention

Center for Global Health

Division of Global HIV and TB

1600 Clifton Road

Atlanta, GA, 30329

United States

Phone: 1 470 298 5583

Email: rdoshi@cdc.gov

Abstract

Background: Key populations at higher risk for HIV infection, including people who inject drugs, men who have sex with men (MSM), and female sex workers (FSWs), are disproportionately affected by the HIV/AIDS epidemic. Empirical estimates of their population sizes are necessary for HIV program planning and monitoring. Such estimates, however, are lacking for most of Uganda's urban centers.

Objective: The aim of this study was to estimate the number of FSWs and MSM in select locations in Uganda.

Methods: We utilized conventional 2-source capture-recapture (CRC) to estimate the population of FSWs in Mbale, Jinja, Wakiso, Mbarara, Gulu, Kabarole, Busia, Tororo, Masaka, and Kabale and the population of MSM in Mbale, Jinja, Wakiso, Mbarara, Gulu, Kabarole, and Mukono from June to August 2017. Hand mirrors and key chains were distributed to FSWs and MSM, respectively, by peers during capture 1. A week later, different FSWs and MSM distributors went to the same towns to collect data for the second capture. Population size estimates and 95% CIs were calculated using the CRC Simple Interactive Statistical Analysis.

Results: We estimated the population of FSWs and MSM using 2 different recapture definitions: those who could present the object or identify the object from a set of photos. The most credible (closer to global estimates of MSM; 3%-5%) estimates came from those who presented the objects only. The FSW population in Mbale was estimated to be 693 (95% CI 474-912). For Jinja, Mukono, Busia, and Tororo, we estimated the number of FSWs to be 802 (95% CI 534-1069), 322 (95% CI 300-343), 961 (95% CI 592-1330), and 2872 (95% CI 0-6005), respectively. For Masaka, Mbarara, Kabale, and Wakiso, we estimated the FSWs population to be 512 (95% CI 384-639), 1904 (95% CI 1058-2749), 377 (95% CI 247-506), and 828 (95% CI 502-1152), respectively. For Kabarole and Gulu, we estimated the FSWs population to be 397 (95% CI 325-469) and 1425 (95% CI 893-1958), respectively. MSM estimates were 381 (95% CI 299-462) for Mbale, 1100 (95% CI 351-1849) for Jinja, 368 (95% CI 281-455) for Wakiso, 322 (95% CI 253-390) for Mbarara, 180 (95% CI 170-189) for Gulu, 335 (95% CI 258-412) for Kabarole, and 264 (95% CI 228-301) for Mukono.

Conclusions: The CRC activity was one of the first to be carried out in Uganda to obtain small town-level population sizes for FSWs and MSM. We found that it is feasible to use FSW and MSM peers for this activity, but proper training and standardized data collection tools are essential to minimize bias.

KEYWORDS

sex worker; population size; men who have sex with men; HIV

Introduction

Background

Key populations such as female sex workers (FSWs) and men who have sex with men (MSM) are disproportionately affected by the HIV epidemic [1,2]. FSWs are estimated to be 13.5 times more likely to become infected with HIV than the general female population, whereas MSM are 19 times more likely to be living with HIV than the general male population [3]. In sub-Saharan Africa, the estimated HIV prevalence among FSWs is 36.9%, whereas it is 17.9% among MSM [4,5]. Sexual risk behavior and number of sexual partners, along with violence, criminalization, and stigma may contribute to their increased HIV vulnerability [6].

Uganda has a generalized epidemic; however, the estimated HIV prevalence among FSWs and MSM (33% [7] and 13.7% [8], respectively) are substantially higher than that in the general population (6.2% [9]). In Uganda, FSWs and MSM face stigma and marginalization, limiting their access to prevention and treatment programs. Thus, targeted HIV services for prevention, care, and treatment need to be specifically planned and evaluated for such populations. Such planning and evaluation requires an accurate estimate of the population of FSWs and MSM.

There is currently no gold standard method on estimating the size of key populations. Several population size estimation (PSE) methods have been applied in different settings, but each method has its own set of potential biases and limitations [10]. Capture-recapture (CRC) has been recommended for use when a census is not feasible and if there are no or poor quality service data, which is the case in Uganda [11]. The size estimate from CRC is based on 2 independent captures: capture 1 and capture 2. In the first capture, participants are tagged and counted. This follows a second independent capture of participants, some of which would have been tagged in the first capture. From the proportion of participants recaptured, an estimate of the entire population is inferred [12]. For the estimate to be unbiased, CRC relies on 4 assumptions: (1) independence of samples, (2) closed population (no migration), (3) matching of individual samples in both captures, and (4) equal likelihood of capture [12].

Objectives

Population estimates for FSWs and MSM in Uganda are particularly scarce. The World Health Organization and other multilateral organizations have made estimating the size of key populations a priority [9,13,14]. There have been challenges on quality and timely reporting, especially at the subnational level, as suggested by the same report, resulting in gaps of size estimation data. A 2-source CRC (2SCRC) has previously been

conducted in Kampala for FSWs and MSM, which yielded estimates of 13,000 FSWs and 8000 MSM [15]. However, population estimates of FSWs and MSM has not been conducted outside of Kampala. To address this gap in knowledge, we utilized the 2SCRC method to estimate the size of FSWs and MSM populations in select towns in Uganda. This is the first time that CRC has been used in these areas and the first population size estimates for MSM outside of Kampala.

Methods

Target Population

We conducted the CRC activity among FSWs in 11 locations (Tororo, Busia, Mbarara, Mbale, Gulu, Kabale, Kabarole, Jinja, Masaka, Wakiso, and Mukono) and among MSM in 7 locations (Mukono, Wakiso, Jinja, Kabarole, Gulu, Mbale, and Mbarara). These locations were chosen based on the availability of funding and the availability of MSM and FSW community-based organizations (CBOs) in these areas. The HIV prevalence in these districts is higher than the national average. FSWs who are 15 years or older and MSM who are 18 years or older were identified by FSW and MSM peers, respectively.

Sample Size and Precision

There were no previous estimates of MSM and FSWs in these locations. To determine the target ranges for the number of unique objects to distribute per population for each town, 0.5% of all adult females and 1.5% of all adult males living in that particular location were assumed to be FSWs and MSM, respectively [16].

Capture-Recapture Implementation

CBOs who operated in each location were consulted before the data collection to discuss the selected objects and recommend peer distributors for each target population. They provided a list of FSW and MSM peers (other members of the target population) who would act as distributors of unique objects. Objects were selected in collaboration with local CBOs and were vetted by target population members. The unique objects had unique phrases printed in various colors. Hand mirrors and keychains were distributed to FSWs and MSM, respectively, as shown in [Figure 1](#).

In total, 2 FSW and 2 MSM peers were chosen to distribute unique objects for the first capture. Another set of 2 FSW and 2 MSM peers were assigned to collect data for the second capture. All distributors participated in a half-day training. Trainings were conducted separately for MSM and FSWs and for captures 1 and 2. The training included the purpose of the activity, a description of the target population, and instructions for data collection and documentation.

Figure 1. Selected unique objects by population (unique phrases not shown).



Data Collection

Estimation was piloted first in Mbale between June 7 and 21, 2017, and then initiated in 9 phases. Phase 1, which included Jinja, Mukono, Busia, and Malaba, was conducted between June 26 and July 13, 2017. Phase 2, which included Masaka, Mbarara, Kabale, and Wakiso, was conducted between July 17 and August 4, 2017. Phase 3, which included Kabarole and Gulu, was conducted between August 7 and 25, 2017.

All data were collected using Open Data Kit Collect, an open source Android app, on a smartphone [17]. The first sample was captured by distributing unique objects by MSM and FSW peers to the target population. Distributors were asked to visit all known hotspots at different times of the day. When the distributors came in contact with a target population member, they were told to estimate the age group of the individual, indicate the Global Positioning System (GPS) coordinates of where the object was given out, and identify if the participant accepted the unique object. Distributors were also asked to randomly give out unique objects at all possible places where target population members might congregate. Per the protocol, any FSW estimated to be under the age of 18 was referred to specialized services.

The second capture took place approximately 5 to 7 days after the first capture, with a different set of peers to collect the data. No objects were distributed during capture 2, and distributors asked target population members if they had received a gift in the previous week. If they indicated that they had received a gift, they were asked to show the distributor the object. If the approached individual claimed to have received an object but did not have the object with them, they were asked to identify the correct object from a piece of paper with pictures of 10 to

15 different objects (some similar to the real objects, some very different). Distributors recorded the picture the individual identified but did not reveal whether they were correct or not. Distributors also electronically collected the same demographic and geographical data as in the first capture.

Data Analysis

All PSE calculations were carried out using a CRC calculator developed by Simple Interactive Statistical Analysis [18]. Nonweighted data were used for analysis. Using the 2 data sources from captures 1 and 2, the CRC method provides an estimate by the following formula:

$$N = MC/R$$

where N is the estimate of the total population size, M is the total number of people captured and marked on the first visit, C is the total number of people captured on the second visit, and R is the number of people captured on the first visit and then recaptured on the second visit.

To calculate the CI to give a range of error for the estimate of total population, the following formula was used:

$$95\%CI = N \pm 1.96 * \text{sqrt}(\text{Var}(N))$$

where variance is calculated as follows:

$$\text{Var}(N) = [MC * (M-R) * (C-R)] / (R^3)$$

Size estimates were calculated using 2 different definitions: (1) those were able to present the object or identify the correct object from a set of pictures and (2) only those who were able to present the object. All values were rounded to the next whole number.

Results

Sampling

We show the sampling results using 2 different recapture definitions: (1) those who are able to present the object or identify the correct object from a set of pictures and (2) those who are only able to present the object. The numbers of objects

offered, how many FSWs/MSM refused objects, the number of MSM/FSWs in captures 1 and 2, and the number of people captured twice (recaptures) are shown (Tables 1 and 2).

We show the number of MSM or FSWs who misidentified unique objects from a set of pictures (Table 3). We see that in some towns, no misidentification of objects occurred. We present the proportion of capture 2 that was recaptures for each distributor for each town (Tables 4 and 5).

Table 1. Target sample size, number of objects distributed, number of refusals, number of female sex workers (FSWs) in capture 1, number of FSWs in capture 2, number of recaptures for recapture definition 1, and number of recaptures for recapture definition 2 for FSWs in 11 Ugandan towns.

Town	Objects offered, n	Refusals, n	Capture 1, n	Capture 2, n	Recapture definition 1, n	Recapture definition 2, n
Mbale	143	32	111	156	31	25
Jinja	103	8	95	194	23	23
Busia	117	9	108	169	22	19
Tororo	66	7	59	146	16	3
Masaka	130	4	126	138	50	34
Wakiso	105	3	102	146	18	18
Mbarara	155	5	150	203	16	16
Kabale	86	21	65	110	22	19
Gulu	141	8	133	225	32	21
Kabarole	112	10	102	183	55	47
Mukono	159	9	150	242	116	113

Table 2. Target sample size, number of objects distributed, number of refusals, number of men who have sex with men (MSM) in capture 1, number of MSM in capture 2, number of recaptures for recapture definition 1, and number of recaptures for recapture definition 2 for MSM in 7 Ugandan towns.

Town	Objects offered, n	Refusals, n	Capture 1, n	Capture 2, n	Recapture definition 1, n	Recapture definition 2, n
Mbale	129	28	101	143	46	38
Jinja	81	0	81	95	17	7
Wakiso	81	2	79	149	62	32
Mbarara	99	0	99	120	51	37
Gulu	109	7	102	153	88	87
Kabarole	92	8	83	133	61	33
Mukono	106	2	104	142	88	56

Table 3. Number of female sex workers and men who have sex with men who misidentified the objects.

Town	Female sex workers who misidentified objects, n	Men who have sex with men who misidentified objects, n
Mbale	5	13
Jinja	6	0
Busia	0	— ^a
Tororo	1	—
Masaka	0	—
Wakiso	0	0
Mbarara	0	10
Kabale	0	—
Gulu	2	5
Kabarole	0	8
Mukono	2	0

^a—: not applicable.

Table 4. Recaptures for distributors 1 and 2 for female sex workers per each town.

Town	Recaptures for distributor 1, n	Recaptures for distributor 1/capture 2, n	Recaptures for distributor 2, n	Recaptures for distributor 2/capture 2, n
Mbale	6	0.04	19	0.12
Jinja	7	0.04	16	0.08
Busia	7	0.04	16	0.09
Tororo	3	0.02	0	0.00
Masaka	15	0.11	19	0.14
Wakiso	1	0.01	17	0.12
Mbarara	16	0.08	0	0.00
Kabale	9	0.08	10	0.09
Gulu	10	0.04	11	0.05
Kabarole	5	0.03	42	0.23
Mukono	5	0.02	109	0.45

Table 5. Recaptures for distributors 1 and 2 for men who have sex with men per each town.

Town	Recaptures for distributor 1, n	Recaptures for distributor 1/capture 2, n	Recaptures for distributor 2, n	Recaptures for distributor 2/capture 2, n
Mbale	6	0.04	32	0.22
Jinja	5	0.05	2	0.02
Wakiso	21	0.14	11	0.07
Mbarara	17	0.14	20	0.17
Gulu	19	0.12	68	0.44
Kabarole	26	0.20	7	0.05
Mukono	36	0.25	20	0.14

Table 6. Population size estimates per town for female sex workers.

Town	Adult female population, N	PSE ^a of recapture definition 1, n (PSE/total Population size as % of recapture definition 1)	95% CI	PSE of recapture definition 2, n (PSE/total population size as % of recapture definition 2)	95% CI
Mbale	161,720	559 (0.35)	410-708	693 (0.43)	474-912
Jinja	144,280	802 (0.56)	534-1069	802 (0.56)	534-1069
Busia	98,210	830 (0.85)	542-1119	961 (0.98)	592-1330
Tororo	155,520	539 (0.35)	326-751	2872 (1.85)	0-6005
Masaka	91,680	348 (0.38)	288-408	512 (0.56)	384-639
Wakiso	786,970	828 (0.11)	503-1152	828 (0.11)	502-1152
Mbarara	156,270	1904 (1.22)	1058-2749	1904 (1.22)	1058-2749
Kabale	75,809	325 (0.43)	227-424	377 (0.50)	247-506
Gulu	88,820	936 (1.05)	674-1197	1425 (1.60)	893-1958
Kabarole	90,307	340 (0.38)	289-391	398 (0.44)	326-469
Mukono	194,920	313 (0.16)	294-333	322 (0.17)	300-343

^aPSE: population size estimation.

Table 7. Population size estimates per town for men who have sex with men.

Town	Adult male population, N	PSE ^a of recapture definition 1, n (PSE/total Population size as % of recapture definition 1)	95% CI	PSE of recapture definition 2, n (PSE/total population size as % of recapture definition 2)	95% CI
Mbale	123,201	314 (0.25)	259-370	381 (0.31)	299-462
Jinja	18,233	453 (2.48)	280-626	1100 (6.03)	351-1849
Wakiso	15,228	190 (1.25)	173-207	368 (2.42)	281-455
Mbarara	48,754	233 (0.48)	200-267	322 (0.66)	253-390
Gulu	38,069	178 (0.47)	169-187	180 (0.47)	170-189
Kabarole	13,569	181 (1.33)	164-199	335 (2.47)	258-412
Mukono	40,686	168 (0.41)	160-177	264 (0.65)	228-301

^aPSE: population size estimation.

Population Size

We present the results of the population size estimates of FSWs and MSM, the 95% CI, the number of adult male or female population of that particular town, and the prevalence of MSM or FSWs (Tables 6 and 7).

Female Sex Workers

We estimated the number of FSWs in Mbale to be 693 (95% CI 474-912). For Jinja, Mukono, Busia, and Tororo, we estimated the number of FSWs to be 802 (95% CI 534-1069), 322 (95% CI 300-343), 961 (95% CI 592-1330), and 2872 (95% CI 0-6005), respectively. For Masaka, Mbarara, Kabale, and Wakiso, we estimated the number of FSWs to be 512 (95% CI 384-639), 1904 (95% CI 1058-2749), 377 (95% CI 247-506), and 828 (95% CI 502-1152), respectively. For Kabarole and Gulu, we estimated the number of FSWs to be 397 (95% CI 325-469) and 1425 (95% CI 893-1958), respectively.

Men Who Have Sex With Men

We estimated the number of MSM in Mbale to be 381 (95% CI 299-462). For Jinja and Mukono, we estimated the number of MSM to be 1100 (95% CI 351-1849) and 264 (95% CI 228-301), respectively. For Mbarara and Wakiso, we estimated the number of MSM to be 322 (95% CI 253-390) and 368 (95% CI 281-455), respectively. For Kabarole and Gulu, we estimated the number of MSM to be 335 (95% CI 258-412) and 180 (95% CI 170-189), respectively.

Discussion

Principal Findings

We conducted multiple size estimation activities using the CRC method and present 2 size estimates for FSWs and MSM in 11 select Ugandan towns.

Comparing the population estimates from the different recapture definitions for FSWs and MSM, we see that recaptures of those who had the object or were able to identify the object were

generally higher, resulting in lower population estimate compared with the results from those who just had the object with them.

Our subnational estimates for FSWs as percentage of the adult female population, or FSWs prevalence, range from 0.3% to 1.2%. In Tororo, we found that women were less likely to keep the object with them. This is evident from a comparison of the number of recaptures in [Table 1](#). Only 3 women had the object with them, but an additional 13 women were able to accurately identify the object from a set of pictures. Perhaps, this is because the women lost the objects more frequently or were selling the objects, which may reflect the socioeconomic status of women in Tororo. When women who could identify the object from a set of pictures were included in the analysis, the size estimate dropped from 2872 (FSWs prevalence of 1.85%) to 539 (FSWs prevalence of 0.35%). Our range of FSW prevalence is similar to subnational FSWs prevalence estimates in other regions of Africa (0.1%-12.0%) [[19](#)]. Generally, in other areas, the results did not differ as much because the majority of FSWs kept the object for a week. Our subnational estimates for MSM as percentage of the adult male population, or MSM prevalence, range from 0.03% to 0.9%. When compared with global estimates of MSM (3%-5%), we find that most of our MSM prevalence estimates are lower [[20,21](#)]. These results reflect the difficulties in executing venue-based population estimates in small settings in countries such as Uganda and may suggest that a large proportion of MSM was not recognizable to distributors or that they do not frequent venues that distributors targeted, or that these smaller towns indeed harbor fewer MSM.

We can compare our FSWs and MSM estimates with results from a 2014 report by the Ugandan Ministry of Health and Uganda AIDS Commission where they estimated 54,549 sex workers (0.57% of the adult female population) and 10,533 MSM (0.12% of the adult male population) in Uganda [[15](#)]. With the exception of our FSWs prevalence estimates from Mbale, our FSW and MSM prevalence estimates are higher compared with the FSWs and MSM prevalence found in the report. However, it is important to note that the report estimates are national, whereas our estimates are subnational. Unfortunately, because of lack of data in the literature, we were not able to compare our subnational estimates to other population estimates.

The Priorities for Local AIDS Control Effort (PLACE) method has been utilized for FSWs in some of these locations [[22](#)]. Though the PLACE report did not provide population size estimates for FSWs, we can compare the numbers of FSWs reported using the PLACE method. Per PLACE method, there were 65 FSWs in Jinja, 62 FSWs in Mbale, 61 FSWs in Kabale, 70 FSWs in Masaka, 119 FSWs in Malaba, and 113 FSWs in Mbarara. The numbers of FSWs found using PLACE are lower than the number of FSWs captures for each corresponding location. This is likely explained by the PLACE methodology, where only 3 areas within the district are visited, whereas our distributors visited multiple hotspots.

Strengths and Limitations

There were a number of limitations in our PSE activity that is inherent to CRC methodology and failing to meet its

assumptions. As we did not want to collect personal identifying information, we used unique objects to identify recaptures. However, not every person carried the unique object with them, which made determining the exact number of recaptures problematic because we could not confirm that the person who received the object was the same person who received it. Estimates would thus be too high if matches were not identified or too low if recaptures were matched incorrectly. We attempted to reduce the bias from using unique objects as a matching mechanism by instructing participants to keep the object with them for as long as the data collection period (5 to 7 days) and limiting the time between captures to 5 to 7 days. However, there were individuals from both populations who did not have the object with them. This is not surprising, given that a population member may have lost the object, given it away, or left it at home. Minimizing the time between captures was also an effort to meet the assumption of a closed population (no in and out migration). If there is an increase in the number of people in the second sample because of migration, it may result in an underestimate, or vice versa. In addition, we tried to meet the assumption of equal likelihood of capture by instructing the peer distributors to visit all known hotspots and randomly sample individuals. In certain towns, this may not have been possible, and our peer distributors may not have been able to visit all hotspots in the allotted time. Our method involved approaching population members at select hotspots and may have excluded specific subgroups that may not attend venues (eg, street- or home-based FSWs or MSM/FSWs who meet individuals or clients using social media or other app). Individuals with higher social visibility are more likely to be captured; thus, our results are likely to be underestimated for all populations. The expansion of captures to various other data sources (not just object distribution) to include service lists and or social media and websites could reach those individuals who are less likely to attend venues. Furthermore, members of the FSWs and MSM populations may have decided to not participate in the activity for fear of being identified. The use of FSW and MSM peers for data collection may have alleviated this issue. Finally, we attempted to meet the assumption of independence of samples by using different staff for captures 1 and 2.

We encountered unique observations and challenges during our size estimation activities. First, the PSE activity was well received by the community with the help of the local CBO. The unique objects were generally well liked by the target population members. In Mukono, the unique objects were too well liked; FSW distributors reportedly were selling objects, and thus the activity was repeated with a new set of peers. Tainted data from Mukono were discarded, and the Mukono data presented relied on the repeated activity. We also found that in some areas, target population members would surround distributors hoping to receive an object, suggesting that objects may not have always been given out at random and that not everyone had an equal chance of being captured. To ensure that these issues do not occur in the future, we recommend consultation with local CBOs and to choose unique objects that are likeable but not overly popular. Furthermore, although we instructed data collectors to go to any place where they might find a member of the target population within the district, the majority of the captures were found in the town centers. In addition, considering that each

distributor had the same training and used the same electronic data collection tool, we expected that the proportion of captures that were recaptures be similar between the 2 FSWs or MSM distributors for capture 2. However, we saw that in many sites, 1 distributor had substantially more (>10%) recaptures than the other. We attempted to mitigate this issue by looking at the GPS data for each of the distributor's captures, to see if they made up certain captures. However, it was difficult to assess which captures were true captures. To the extent that we can check, we did not find any fabricated captures. We also checked to see if the location of the recaptures were in similar locations to those of captures 1 and 2 and found that they were, suggesting that the recaptures were recorded in locations where FSWs and MSM congregate and are thereby legitimate. Regardless, these

issues brought to light that one of the most critical elements of a successful CRC activity using unique objects is choosing the right distributors.

Conclusions

New empirical population size estimates were generated for MSM and FSWs in select towns in Uganda. These estimates are among the first in these locations in Uganda and are important in providing knowledge and insight in planning for HIV programs by stakeholders working in Uganda. More implementation research and more systematic use of CRC or other empirical PSE methods are warranted for Uganda. PSE activities should be included in future HIV surveillance efforts to improve estimates and optimally plan for the provision of services to high-risk populations.

Acknowledgments

This paper was funded by the President's Emergency Plan for AIDS Relief through the Centers of Disease Control and Prevention under the terms of project number U2GGH000466. The authors would like to thank the staff of the Crane Survey, under which umbrella this CRC activity was conducted. They also would like to acknowledge the Monitoring and Evaluation Technical Support for their support. In addition, they would like to thank the CBOs that dedicated their time to this activity: Uganda Harm Reduction Network, Women's Organization Network for Human Rights Advocacy, Lady Mermaid's Bureau, Alliance for Women, Advocating for Change, Women Arise for Change, Organization for Gender Empowerment and Rights, Advocacy (Ogera Uganda), Empower at Dusk, Women's Association, Serving Lives Under Marginalization, Women's Positive Empowerment Initiative, Ice Breakers, Youth on the Rocks Foundation, Spectrum Uganda, Come out Post Test Club, Kuchu Shiners Uganda, Frank and Candy Uganda, and Rainbow Mirrors Uganda. The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the funding agencies

Conflicts of Interest

None declared.

References

1. Ghys PJ, Jenkins C, Pisani E. HIV surveillance among female sex workers. *AIDS* 2001 Apr;15(Suppl 3):S33-S40. [Medline: [11421180](#)]
2. 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](#)]
3. UNAIDS. Geneva; 2014. The Gap Report URL: http://files.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2014/UNAIDS_Gap_report_en.pdf [accessed 2019-03-08] [WebCite Cache ID 76ioCzOls]
4. Baral SB, 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](#)]
5. Beyrer CB, 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](#)] [Medline: [22819660](#)]
6. World Health Organization Regional Office for the Eastern Mediterranean. Cairo; 2016. Estimating size of key populations: Guide for HIV programming in countries of the Middle East and North Africa URL: http://applications.emro.who.int/dsaf/EMROPub_2016_EN_19361.pdf [accessed 2019-03-08] [WebCite Cache ID 76imGJMhm]
7. Hladik WB, Baughman AL, Serwadda D, Tappero JW, Kwezi R, Nakato ND, et al. Burden and characteristics of HIV infection among female sex workers in Kampala, Uganda - a respondent-driven sampling survey. *BMC Public Health* 2017 Jun 10;17(1):565 [FREE Full text] [doi: [10.1186/s12889-017-4428-z](#)] [Medline: [28601086](#)]
8. Baral SG, Grosso A, Holland C, Papworth E. The epidemiology of HIV among men who have sex with men in countries with generalized HIV epidemics. *Curr Opin HIV AIDS* 2014 Mar;9(2):156-167. [doi: [10.1097/COH.000000000000037](#)] [Medline: [24445371](#)]
9. UNAIDS. Northern Uganda Malaria, AIDS & Tuberculosis Programme. Kampala: Uganda AIDS Commission; 2009. Final Report: Uganda HIV Modes of Transmission and Prevention Response Analysis URL: http://numat.jsi.com/Resources/Docs/Unaid UgandaCountryReport_09.pdf [accessed 2019-03-08] [WebCite Cache ID 76ioBjEku]

10. PHIA Project. New York: ICAP; 2017 Mar 3. Uganda Population-Based HIV Impact Assessment Preliminary Findings URL: https://phia.icap.columbia.edu/wp-content/uploads/2018/07/3430%E2%80%A2PHIA-Uganda-SS_NEW.v14.pdf [accessed 2019-03-08] [WebCite Cache ID 76iq7ld14]
11. Abdul-Quader AB, 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](https://pubmed.ncbi.nlm.nih.gov/24393694/)]
12. Thompson SK. Sampling. New York: John Wiley & Sons; 2002.
13. PEPFAR. Accelerating Progress Toward HIV/AIDS Epidemic Control. 2011. Technical Guidance on Combination HIV Prevention URL: <http://www.pepfar.gov/documents/organization/164010.pdf> [accessed 2019-03-03] [WebCite Cache ID 76ikJGnKt]
14. World Health Organization. Geneva; 2008. Prevention and treatment of HIV and other sexually transmitted infections among men who have sex with men and transgender populations : report of a technical consultation, 15-17 September 2008, Geneva, Switzerland URL: <https://apps.who.int/iris/handle/10665/44103> [accessed 2019-03-08] [WebCite Cache ID 76ipJPN4V]
15. Ministry of Health, Uganda AIDS Commission. UNAIDS. 2015 Jun 15. The HIV and AIDS Country Progress Report 2014: Kampala, Uganda URL: http://www.unaids.org/sites/default/files/country/documents/UGA_narrative_report_2015.pdf [accessed 2019-03-08] [WebCite Cache ID 76iqDuYnv]
16. Uganda Bureau of Statistics. Kampala: Uganda Bureau of Statistics; 2016. The National Population and Housing Census 2014-Main Report URL: https://www.ubos.org/wp-content/uploads/publications/03_20182014_National_Census_Main_Report.pdf [accessed 2019-03-08] [WebCite Cache ID 76ipuBJcu]
17. Brunette W, Sundt M, Chaudhri R, Breit N, Borriello G. 2013. Open Data Kit 2.0: Expanding and Refining Information Services for Developing Regions URL: <https://docs.opendatakit.org/odk2/> [accessed 2019-03-08] [WebCite Cache ID 76ipz8Sag]
18. SISA. Quantitative Skills.: Simple Interactive Statistical Analysis; 2017. Capture-recapture URL: <http://www.quantitativeskills.com/sisa/statistics/crc.htm> [accessed 2019-03-08] [WebCite Cache ID 76inU38fA]
19. Vandepitte J, Lyerla 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/)]
20. Marcus UH, Hickson F, Weatherburn P, Schmidt AJ, EMIS Network. Estimating the size of the MSM populations for 38 European countries by calculating the survey-surveillance discrepancies (SSD) between self-reported new HIV diagnoses from the European MSM internet survey (EMIS) and surveillance-reported HIV diagnoses among MSM in 2009. *BMC Public Health* 2013 Oct 3;13:919 [FREE Full text] [doi: [10.1186/1471-2458-13-919](https://doi.org/10.1186/1471-2458-13-919)] [Medline: [24088198](https://pubmed.ncbi.nlm.nih.gov/24088198/)]
21. Cáceres C, Konda K, Pecheny M, Chatterjee A, Lyerla R. Estimating the number of men who have sex with men in low and middle income countries. *Sex Transm Infect* 2006 Jun;82(Suppl 3):iii3-iii9 [FREE Full text] [doi: [10.1136/sti.2005.019489](https://doi.org/10.1136/sti.2005.019489)] [Medline: [16735290](https://pubmed.ncbi.nlm.nih.gov/16735290/)]
22. Measure Evaluation. Kampala; 2013. Priorities for Local AIDS Control Efforts Uganda, 2013-2014 URL: <https://www.measureevaluation.org/resources/publications/tr-14-109-en> [accessed 2019-03-08] [WebCite Cache ID 76inGiKib]

Abbreviations

- CBOs:** community-based organizations
- CRC:** capture-recapture
- FSWs:** female sex workers
- GPS:** Global Positioning System
- MSM:** men who have sex with men
- PLACE:** Priorities for Local AIDS Control Effort
- PSE:** population size estimation

Edited by J Neal; submitted 27.09.18; peer-reviewed by A Abdul-Quader, P Wesson, P Nadol; comments to author 06.11.18; revised version received 18.12.18; accepted 05.02.19; published 03.04.19.

Please cite as:

Apodaca K, Doshi RH, Ogwal M, Kiyangi H, Aluzimbi G, Musinguzi G, Lutalo I, Akello E, Hladik W
Capture-Recapture Among Men Who Have Sex With Men and Among Female Sex Workers in 11 Towns in Uganda
JMIR Public Health Surveill 2019;5(2):e12316
URL: <https://publichealth.jmir.org/2019/2/e12316/>
doi:[10.2196/12316](https://doi.org/10.2196/12316)
PMID:[30942697](https://pubmed.ncbi.nlm.nih.gov/30942697/)

©Kevin Apodaca, Reena Hemendra Doshi, Moses Ogwal, Herbert Kiyangi, George Aluzimbi, Geoffrey Musinguzi, Ibrahim Lutalo, Evelyn Akello, Wolfgang Hladik. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 03.04.2019. 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

Estimation of the Population Size of Men Who Have Sex With Men in Vietnam: Social App Multiplier Method

Vo Hai Son^{1*}, MD, MPH; Ali Safarnejad^{2*}, PhD; Nguyen Thien Nga², MIPH; Vu Manh Linh¹, BPH; Le Thi Cam Tu¹, MSc; Pham Duc Manh¹, MD, PhD; Nguyen Hoang Long¹, MD, PhD; Abu Abdul-Quader³, PhD

¹Vietnam Authority of HIV/AIDS Control, Ministry of Health, Hanoi, Vietnam

²The Joint United Nations Programme on HIV/AIDS, Green One United Nations House, Hanoi, Vietnam

³Centers for Disease Control and Prevention, Hanoi, Vietnam

*these authors contributed equally

Corresponding Author:

Ali Safarnejad, PhD

The Joint United Nations Programme on HIV/AIDS

Green One United Nations House

Hanoi

Vietnam

Phone: 84 +84243 850 1887

Email: ali.safarnejad@gmail.com

Abstract

Background: Although the prevalence of HIV among men who have sex with men (MSM) in Vietnam has been increasing in recent years, there are no estimates of the population size of MSM based on tested empirical methods.

Objective: This study aimed to estimate the size of the MSM population in 12 provinces in Vietnam and extrapolate from those areas to generate a national population estimate of MSM. A secondary aim of this study was to compare the feasibility of obtaining the number of users of a mobile social (chat and dating) app for MSM using 3 different approaches.

Methods: This study used the social app multiplier method to estimate the size of MSM populations in 12 provinces using the count of users on a social app popular with MSM in Vietnam as the first data source and a questionnaire propagated through the MSM community using respondent-driven sampling as the second data source. A national estimation of the MSM population is extrapolated from the results in the study provinces, and the percentage of MSM reachable through online social networks is clarified.

Results: The highest MSM population size among the 12 provinces is estimated in Hanoi and the lowest is estimated in Binh Dinh. On average, 37% of MSM in the provinces surveyed had used the social app Jack'd in the last 30 days (95% CI 27-48). Extrapolation of the results from the study provinces with reliable estimations results in an estimated national population of 178,000 MSM (95% CI 122,000-512,000) aged 15 to 49 years in Vietnam. The percentage of MSM among adult males aged 15 to 49 years in Vietnam is 0.68% (95% CI 0.46-1.95).

Conclusions: This study is the first attempt to empirically estimate the population of MSM in Vietnam and highlights the feasibility of reaching a large proportion of MSM through a social app. The estimation reported in this study is within the bounds suggested by the Joint United Nations Programme on HIV/AIDS. This study provides valuable information on MSM population sizes in provinces where reliable estimates were obtained, which they can begin to work with in program planning and resource allocation.

(*JMIR Public Health Surveill* 2019;5(2):e12451) doi:[10.2196/12451](https://doi.org/10.2196/12451)

KEYWORDS

HIV; AIDS; population size estimation; men who have sex with men; respondent-driven sampling; Vietnam

Introduction

Background

Since the first reported case of HIV in Vietnam in 1990 [1-3], over 400,000 people have been estimated to be infected with HIV in the country and 150,000 are reported to have died of AIDS-related causes [4]. The HIV epidemic in Vietnam is concentrated among 3 key populations (ie, people who inject drugs, men who have sex with men [MSM], and female sex workers) and their sexual partners [5]. The main route of transmission has been through sharing of needles when injecting drugs, followed by sexual transmission; however, the proportion attributable to the latter mode of transmission has been increasing in recent years [4]. By 2015, the estimated number of new infections had decreased by 50% from the peak of the epidemic in 2002, thanks to prevention initiatives for key populations, including provision of clean needles and syringes, condoms, methadone maintenance therapy, and antiretroviral treatment [4,6].

Although the overall number of new infections has declined, the prevalence of HIV among MSM has been increasing in recent years [4]. HIV sentinel surveillance data for MSM in 2014 in 8 provinces found an average HIV prevalence of 6.7%, with higher prevalence reported in major urban areas, and 2017 surveillance in 9 provinces found an average HIV prevalence of 12.2% [7,8]. Despite these alarming estimates, information about the population size of the MSM community is limited and imprecise.

Estimates of the size of populations at risk of HIV are necessary to understand the scale of the epidemic and in planning appropriate interventions and allocation of resources. A number of estimates of the MSM population size have been attempted in Vietnam, targeted to limited provinces [9]. The national estimate of the MSM population size in Vietnam has not been based on any empirical method of estimation. A frequently used method to approximate the range of the MSM population is to take a percentage of the male population as being MSM, based on regional percentages and consensus among experts, and adjust it by the level of urbanization [10-12]. In Vietnam, the accepted MSM population size is concluded on by the national technical working group, which updates the epidemiological model for the country and revises the expected MSM population size, among other input values to the model, until it produces a validated epidemiological prediction. The MSM population size based on this modeling process in Vietnam was around 330,000 in 2016.

Objectives

Besides the inherent imprecision of the methods to arrive at the MSM population size based on profiling of urbanization of regions and modeling, the previous accepted population size of MSM in Vietnam does not reveal much about how this population can be reached, how MSM network together, what regional variations exist, or the age demographics of the reachable MSM. This study attempted to directly estimate the size of the MSM population in 12 provinces in Vietnam using the social app multiplier method and extrapolate from those areas to generate a national population estimate of MSM. A

secondary aim of this study was to compare the feasibility of obtaining the number of users on a social app for MSM using 3 different approaches.

Methods

Social App Multiplier Overview

This study used the social app multiplier method to estimate the population size of MSM in 12 provinces of Vietnam. This method was piloted in Ho Chi Minh City and Nghe An province in 2016 and subsequently updated and improved based on learnings from the pilot study [9]. In the formative phase of this study, 12 out of 63 provinces of Vietnam were selected that were predicted to have the greatest size of MSM population, which represent a diversity of the regions of the country, and where there exist a minimum number of social app users so that the method could be applied. These 12 provinces were An Giang, Bac Giang, Binh Dinh, Can Tho, Da Nang, Dak Lak, Dong Nai, Dong Thap, Hanoi, Hai Phong, Nam Dinh, and Thanh Hoa. The next sections briefly describe the methods used in this study. They follow the general multiplier method, which compares 2 independent data sources to estimate the total number in a population, the first source being a count from program data that include only the population whose size is being estimated and a second source being a representative survey of the population whose size is being estimated [13]. Additional details of the data collection procedures have been previously described in the pilot study [9].

Count of Social Apps' Users

The social app Jack'd was selected to provide the first count. Overall, 3 methods were used to count the total number of users on Jack'd. First, following the method used during the pilot, in each of the 12 provinces, the total number of active users on Jack'd was enumerated over a 1-month period, and the final list of active users was deduplicated using the public profile information of app users such as age, pseudonym, and avatar. When counting active users, only profiles that appeared at least twice and spaced by several days in between were included in the final count to minimize the possibility of counting short-term visitors. A second method was a capture-recapture procedure that matched active users on Jack'd at 2 different time points, using the same public profile information as used in the first method, to estimate the total number of active users on the app in each of the respective provinces. The third method was to procure the total aggregate, unduplicated, and nonidentifiable number of app users in the respective province over a period of 1 month directly from the Jack'd social app administrator.

Respondent-Driven Sampling Online Survey

Immediately after 1 month of counting users on the Jack'd social app, an online survey using respondent-driven sampling (RDS) recruitment strategy was conducted in the MSM community in each of the 12 provinces to find out about their use of Jack'd in the past month. RDS is a form of chain referral sampling that uses a mathematical model to approach a true random sample [14]. The inclusion criteria for participants in the 12 RDS online surveys were men whose gender at birth was male, were at least 18 years old, residing for at least 3 months in the province where

the RDS online survey was being conducted, and who have had anal or oral sex with another man in the last year or do not prefer sex with women only. Regarding the last 2 criteria, the participants were first asked if they had anal or oral sex with another man in the last year and only those who responded *no* were subsequently asked if they prefer sex with women only. To be eligible, participants had to answer *yes* to the former question about having had sex with another man, or if they answered *no* to that question, they would have had to answer *no* to the subsequent question about preferring sex with women only.

The following study procedures were followed for the RDS online survey. Between 6 and 8 seed individuals were identified in each province and given 3 coupons each to recruit their peers to respond to the RDS online survey, who would in turn be given 3 coupons each, and so on. Recruits who did not have access to the internet to complete the RDS online survey or preferred to provide answers offline were provided with a telephone number to contact a member of the investigation team who would collect information from them over the phone or in person and enter it online.

Estimation of Population Size

The nonidentifying information collected from the 12 RDS online surveys was analyzed to calculate the proportion of respondents who answered *yes* to having used Jack'd in the past month with a 95% CI. In each province, 2 different estimators, Gile's sequential sampling (Gile's SS) and Salganik-Heckathorn (RDS-I), were used to generate estimates for the survey respondents' Jack'd social app use [15,16].

Homophily tests were conducted to assess the extent of random referrals from respondents to their personal networks [17]. A homophily test value less than 1 indicates respondents with similar characteristics in successive waves of peer recruitment. A test of the sensitivity of the estimator was conducted on the proportion of survey participants who answered *yes* to having used Jack'd in the past month. When the predicted proportion fell within the range of 0.2 to 0.8, we considered the estimator to be not sensitive [18]. Additional tests were conducted on the provincial data to determine if there are significant social, behavioral, and demographic differences between participants who had used Jack'd and those who had not.

Convergence plots in RDS were 1 indicator of having sufficient data collected to get a reliable estimate. When the key estimator remained stable within 2% of the sample proportion, we predicted that additional responses collected would yield insignificant changes to the estimate [19]. Bottleneck plots were also created to show the differences between the individual seeds in each province and determine if they had converged.

The RDS-I estimator tends to underestimate the result if the sampling does not converge, and the Gile's SS bootstrap tends to underestimate the result if the homophily test value is less

than 1 or there is bias in seed selection [20]. On the basis of the convergence and homophily tests, the results of 1 estimator were selected to estimate the population size in each province. The number of Jack'd users counted in each province over 1 month was divided by the selected estimator results of the proportion of RDS survey respondents who said *yes* to having used Jack'd in the past month in that province, to arrive at the provincial MSM population size estimate.

The population size estimates were converted to a percentage of the general adult male population (15-49 years) and compared with the range of percentages reported in the 2018 Spectrum Quick Start Guide [21,22]. The estimates of 7 provinces in this study were extrapolated to 50 other provinces of Vietnam according to the geographic and socioeconomic regional grouping of provinces [23-25]. The percentage of MSM among the adult male population, weighted based on proportions of Jack'd users within each age group (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, and 45-49 years), was applied to the matching provinces. The sum of 11 directly observed and 50 extrapolated provincial estimates in this study, along with 2 provincial estimates from previous piloting [9], provided the national MSM population size. Steps of the extrapolation process and calculations are detailed in [Multimedia Appendix 1](#).

Ethical approval for this study was obtained from the Institutional Review Board of the Hanoi School of Public Health (Approval no. 298/2016/YTCC-HD3, IORG no. 0003239; FWA number: 0009326). The protocol was also reviewed and approved by the US Centers for Disease Control and Prevention human subjects research office.

Results

Hanoi had the highest number of active Jack'd users among the 12 provinces at 12,848 persons, and Binh Dinh province had the lowest number of active users at 260 persons. The median age of Jack'd users in the provinces was 25 years, with 75% of the users in the 19 to 35 years age group. The median age of the RDS survey respondents was 24 years, with 75% of the users in the 19 to 32 years age group. There was no statistically significant difference between the mean age or the mean network size of respondents who answered *yes* to having used Jack'd in the past month and those who answered *no* ($t_{22}=1.381$; $P=.18$ and $t_{21}=1.055$; $P=.304$, respectively). There was a significant difference between users of Jack'd and nonusers of Jack'd having had sex with another man in the last year ($\chi^2_{1}=41.0$; $P<.001$). The numbers of participants at each stage of the RDS survey, their social and demographic characteristics, and the summary measures of Jack'd use are presented by province in [Table 1](#). Among the 2177 eligible respondents, 5% completed the survey by telephone and the rest completed the survey online.

Table 1. Number of respondent-driven sampling survey participants and respondents, their social and demographic characteristics, and summary measures of their Jack'd usage.

Provinces	Participants		Characteristics of participants			Summary measures	
	Survey participants, n	Eligible survey respondents, n	Mean age (years)	Mean network size	Participants who did not have sex with another man in the last 12 months but who did not prefer sex with women only among eligible survey respondents, n	Participants with a Jack'd account, n	Participants using Jack'd in the last 30 days, n
An Giang	173	167	22	8	39	33	18
Bac Giang	132	125	28	46	28	77	52
Binh Dinh	134	126	27	5	26	101	43
Can Tho	195	167	23	6	28	37	19
Da Nang	193	167	24	5	17	114	80
Dak Lak	123	121	24	8	52	41	21
Dong Nai	238	227	25	7	15	69	17
Dong Thap	217	212	24	9	61	47	33
Hanoi	296	264	23	13	22	209	130
Hai Phong	345	279	25	10	70	155	96
Nam Dinh	132	122	25	10	21	84	65
Thanh Hoa	244	200	25	9	34	107	67
Total	2422	2177	24.3	10.6	413	1074	641

Among the 12 provinces, 4 (Dak Lak, Dong Thap, An Giang, and Can Tho) had sensitivity ratios that were outside the 0.2 to 0.8 range when using Jack'd use in the last 30 days as the estimator. Among the 12 provinces, 2 (Dong Nai and An Giang) had sensitivity ratios that were outside the 0.2 to 0.8 range when using all-time Jack'd usage as the estimator. Among the 12 provinces, 4 (Bac Giang, Can Tho, Dak Lak, and Dong Nai) had homophily values less than 1 when using Jack'd use in the last 30 days as the estimator and 1 (Nam Dinh) had homophily value less than 1 when using all-time Jack'd use as the estimator. The convergence plots in 10 provinces showed that the RDS survey samples converged; however, in 2 provinces, Dak Lak and Thanh Hoa, there was bottlenecking between seeds (see [Multimedia Appendix 2](#)). As there were more provinces with a homophily value less than 1 than provinces whose RDS survey sample did not converge, we used the RDS-I estimator, which may be less prone to underestimate the result. An Giang was the only province with its sensitivity ratio outside of the 0.2 to 0.8 range for both 30-day use of Jack'd and all-time use of Jack'd. It is also the province with a sensitivity ratio furthest outside of the 0.2 to 0.8 range. For these reasons, it was

considered that the RDS estimator failed to produce a reliable estimate for An Giang.

Analysis of the data shows that on average 37.5% of MSM in the 11 provinces with reliable estimates had used Jack'd in the last 30 days (95% CI 27.0-47.9). Among these provinces, Can Tho had the lowest percentage of active Jack'd users in the last 30 days at 11.4% (95% CI 3.3-19.4) and Da Nang had the highest percentage at 42.9% (95% CI 28.3-57.6). The average weighted percentage of MSM ever using Jack'd in the 11 provinces was 56.7% (95% CI 47.8-65.5). In 1 province, Dong Nai, the 30-day RDS-I and Gile's SS estimates did not produce statistically meaningful results, and the all-time Jack'd use estimator was used instead in this province.

The highest population size of MSM aged 18 to 49 years among the 11 provinces with reliable estimates was in Hanoi at 30,417 persons (95% CI 24,656-39,691), and the lowest MSM population size was estimated in Binh Dinh at 743 persons (95% CI 559-1108). The average weighted percentage of MSM among males aged 15 to 49 years in the 11 provinces was 0.96%, with a range of percentages from 0.70% to 2.47%. The complete results are presented in [Table 2](#).

Table 2. Estimated number of men who have sex with men aged 18 to 49 years and weighted percentage of men who have sex with men among males aged 15 to 49 years in 11 provinces of Vietnam.

Province	Participants counted on Jack'd over 30 days, n	Proportion of MSM ^a using Jack'd (last 30 days, RDS-I ^b)	95% CI for proportion of MSM using Jack'd (last 30 days, RDS-I)	Estimated population size of MSM aged 18-49 years	95% CI for estimated population size of MSM aged 18-49 years	Weighted percentage of MSM among males aged 15-49 years
Bac Giang	360	0.42	0.28-0.55	864	653-1274	0.23
Binh Dinh	260	0.35	0.23-0.47	743	559-1108	0.21
Can Tho	713	0.11	0.03-0.19	6276	3677-21,418	1.22
Da Nang	1713	0.43	0.28-0.58	3990	2974-6059	1.70
Dak Lak	279	0.15	0.07-0.23	1895	1226-4180	0.46
Dong Nai	1191 ^c	0.15 ^d	0.02-0.28	7759	4216-68,370	1.10
Dong Thap	303	0.14	0.06-0.21	2181	1420-4711	0.48
Hanoi	12,848 ^e	0.42	0.32-0.52	30,417	24,656-39,691	1.81
Hai Phong	1141	0.34	0.25-0.43	3336	2645-4515	0.73
Nam Dinh	446	0.39	0.26-0.52	1131	850-1687	0.29
Thanh Hoa	670 ^e	0.22	0.11-0.33	3017	2032-5846	0.40
Total	19,924	0.37	0.27-0.48	61,609	44,909-158,860	0.96

^aMSM: men who have sex with men.

^bRDS: respondent-driven sampling.

^cCount is for all-time use of Jack'd using the capture-recapture method.

^dEstimate is for all-time use of Jack'd (RDS-I).

^eCount obtained directly from social app service provider.

Extrapolation of the results from the 11 provinces with reliable estimates to the national MSM population size resulted in an estimate of 178,000 MSM (95% CI 122,000-512,000) in Vietnam. The percentage of MSM among adult males aged 15 to 49 years in Vietnam is 0.68% (95% CI 0.46-1.95).

Discussion

Principal Findings

Our estimates are the first comprehensive national estimation of the MSM population size conducted in Vietnam that use an empirical method. The point estimate of the 15- to 49-year-old MSM population in Vietnam produced in this study is 178,000, with an estimated range from 122,000 to 512,000. The corresponding estimated percentage of MSM among adult males aged 15 to 49 years of 0.68% is within the range of 0.09% to 4.06% suggested for the Asia and Pacific region by the Joint United Nations Programme on HIV/AIDS (UNAIDS) Spectrum guideline; however, the guideline does not provide a functional definition of MSM, which limits the comparability of the results [21]. The estimated percentage of MSM among adult males aged 15 to 49 years in Vietnam also agrees with recent estimations in other low- and middle-income countries (eg, [26-30]).

The specific definition of MSM adopted in this study based on behavior and sexual preference has broader inclusion criteria than internet-based surveys, which only include men who have been sexually active with a man in the past year in their analysis [31,32]. This study is also different from surveys that measure

population sizes based on sexual orientation or gender identity. For example, the 2014 US National Health Statistics Report uses self-identification methods to estimate that 1.6% of adults in the United States are gay or lesbian [33]. Public health researchers prefer the behavioral and temporal definition of MSM over identities because behaviors, not identities, lead to sexual transmission risk [34]. Self-identification of sexual orientation can also be biased by social stigma, either underestimating or overestimating nonheterosexual population sizes [35-37]. The use of an anonymous RDS online survey propagated through social networks of the MSM community reduces the effect of social stigma in our study results.

The provincial estimates in this study were lower in Bac Giang, Binh Dinh, Dak Lak, Dong Thap, Nam Dinh, and Thanh Hoa; within the range in Da Nang, Dong Nai, Hanoi, and Hai Phong; and higher in Can Tho, in comparison with the expected range of MSM population sizes reported in past HIV and AIDS estimates and projection reports [2,38]. The provinces in this study were selected to be representative of other unsampled provinces in the country with similar geographic and socioeconomic characteristics [23-25]. During extrapolation, the age-disaggregated population of unsampled provinces was used to weight the population proportion of MSM among adult males. There may be other factors that influence the proportional population of MSM in the provinces for building the strata for the extrapolation or for weighting of proportions. For instance, there are different levels of social stigma experienced by key populations at risk of HIV in different geographic areas, and this may be a factor in variations in MSM population size in

the provinces [39]. In consideration of variability across provincial estimates, additional studies are required for future extrapolations of MSM population size and further generation of stigma and discrimination data at the provincial level.

According to our estimates, during a given month, nearly 2 out of 3 people with a Jack'd account are actively connecting with other MSM on the social app, and there is a significant relationship between use of the social app and having sex with other men. These results are consistent with other recent studies that indicate increasing use of online social networks by MSM to find partners, while bypassing social stigma [40,41]. These results also imply the potential for rapid, targeted, and cost-effective outreach to the MSM community through the social apps with messaging on health and social services available in their vicinity. Future analysis may look at the relationship between social app usage and sexual risk behavior. Further studies will also elucidate if members of the MSM community who are on social apps are socially well connected to other MSM or not, which will have important implications for the expected yield of different communication campaigns.

Overall, 3 different approaches were used to obtain the count of Jack'd users for the multiplier method in this study. Each of these approaches comes with its own advantages and disadvantages. The direct counting of users on social apps is a resource-intensive process. It requires investigators to manually record characteristics of active users on the social app daily or more frequently to not miss any peak periods when users log into the apps. In peak periods and in large cities with thousands of users, this process may not be feasible. Moreover, this approach also requires the deduplication of users in the records, which is also a resource-intensive step in the process. However, the advantage of this method is that the resulting number of active users during a brief period of 1 month reduces the recall bias when survey respondents in the second part of the multiplier method are asked about their use of the app in the past month.

The second approach to obtain the count of Jack'd users for the social app multiplier method in this study was obtaining the data directly from the app service provider. The advantage of this approach is that it requires little time, there is no need for deduplication, there is higher accuracy in the count than the manual count, and there is greater privacy as 1 integer figure is collected as the sum of all active users. However, these advantages are traded off with the cost of purchasing the aggregate, unduplicated, and nonidentifiable number of active users from the service provider. The third approach to obtain the count of users was a capture-recapture method on Jack'd. This approach required only 2 counts on 2 distinct days on Jack'd and did not require any deduplication. The disadvantage of this method is that it produces a count of all active users at any time on the social app, which in turn requires a less precise question in the RDS survey that is prone to recall bias. Future research should consider the reliability and precision of the data generated by these approaches as additional criteria to decide on the approach for use with the social app multiplier method.

Limitations

The multiplier method requires the independence of the 2 data sources, the population in the 2 data sources to be defined the

same way, and the 2 data sources to have aligned time periods and geographic areas [13]. This study attempts to address these requirements of the multiplier method. For example, we limit the time between the count of users on Jack'd and the RDS online survey. Furthermore, the 2 data sources match in the 1-month timeframe of the data collection and within each of the 12 provinces where data are collected. In terms of population definition, an assumption is made that the sexual behavior or preference of the Jack'd users matches the inclusion criteria of the participants in the RDS online survey. In matching the age of participants in the 2 data sources, the RDS survey participants are asked about their age, and only those reporting to be at least 18 years old are eligible to participate, similar to the Jack'd requirement of being at least 18 years old to download and use the app. However, as verifying ages on social apps and online surveys is enormously complicated and contentious, and considering literature that reports up to one-fifth of dating profiles and survey respondents being inaccurate in reporting their age [42-44], there may be some skewing of the results in this study because of misreporting of age. Future research should validate the assumptions of the multiplier method and other assumptions of the RDS data by inviting a sample of social app users and RDS survey respondents to a follow-up survey based on standard RDS diagnostic questions [45], along with questions on the participants' age and their sexual behavior and preferences.

Although the age profile of Jack'd users and RDS survey respondents in the 12 provinces included in this study was comparable, these age profiles are skewed toward a younger age range than Vietnam's total male population pyramid [46]. This youth bulge and underrepresentation of older MSM on social media are not uncommon, and studies of the MSM population in Europe and South America speculate that lower internet literacy and increasing proportions of older MSM living in settled relationships may be among factors related to the skewed demographics [47,48]. As noted earlier on misreporting of age, it is plausible that in Vietnam there may also be some MSM younger than 18 years reporting an older age to download and use the Jack'd app and older MSM reporting a younger age on their profile. Acknowledging the low proportion of older participants and some bias in our dataset toward the younger population of MSM, we limit the generalization of our results to the maximum age of 49 years; however, we do not have sufficient data on the extent and amplitude of skewing to attempt any corrective adjustments to the age demographic information.

Intraprovince migration of MSM risks some bias in the population size estimates as the RDS online survey excludes individuals who report having moved within the past 3 months even though there is no way to be sure if the Jack'd users had moved to the province in the past 3 months. We attempt to reduce the possibility of counting short-term migrants and visitors by only counting users who appear at least twice on the social app at 2 distinct days spaced over 1 month.

Conclusions

The estimation of MSM in Vietnam reported in this study is within the bounds suggested by UNAIDS for countries in the Asia and Pacific region, and the range produced in this study

comfortably includes the estimated number of MSM in Vietnam arrived at through the national technical working group profiling and modeling process. The current estimation is based on an empirical method that relies on well-known and tested techniques along with innovative use of social apps used by the MSM population in Vietnam. This study highlights the feasibility of reaching a large percentage of MSM through a social app with programmatic and health promotion interventions. It is also the first time that population size estimations have been conducted in the provinces included in this study, and where reliable estimates were obtained, this study provides those provinces with valuable information on MSM population sizes that they can begin to work with in program planning and resource allocation. In other provinces where the population size estimates were extrapolated to but not directly

observed, this study recommends that the extrapolated estimates be validated using locally appropriate, empirical size estimation methods, including the reliable methods and technologies that were introduced in this study. In provinces where there was a degree of homophily, bottlenecks, and sensitivity in the RDS survey results or where the estimators failed to produce reliable results, alternative methods should be attempted to assess and validate the MSM population sizes. Although the national estimation in this study gets closer to defining the potential range of the number of MSM in Vietnam, future studies will be needed to validate the range and further specify the estimated number. As the MSM population size is one of the key inputs to the national AIDS epidemic modeling and projection process, the AIDS epidemic model needs to be reviewed and updated with the new estimation.

Acknowledgments

The authors would like to extend a special thanks to Kirk Dombrowski for guidance and review of the methodology and insights provided during the training and analysis; Jinkou “Button” Zhou for review and advice on the outline of the manuscript; and Marie-Odile Emond for invaluable support, review, and advice throughout the size estimation study. This study was made possible by the generous financial support of The Global Fund to fight AIDS, Tuberculosis and Malaria and the technical support of UNAIDS. The findings and conclusions in this study are those of the authors alone and they do not necessarily represent the views, decisions, policies, or official position of the institutions or funding agencies with which they are affiliated.

Authors' Contributions

VHS was the lead investigator, conceptualizing the research. AS reviewed the literature and drafted the manuscript aided by VHS and AAQ. VHS designed and had the overall responsibility for the quantitative analysis, with technical support from VML, LTCT, and AS. NTN, LTCT, VML, and VHS were responsible for the translation and localization of the survey. All authors contributed to the development of the argument, interpretation of the results, and revising the manuscript critically for important intellectual content. All authors read and approved the final paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Method of extrapolation from the study provinces and calculation of the national population size of men who have sex with men.

[[PDF File \(Adobe PDF File\), 78KB - publichealth_v5i2e12451_app1.pdf](#)]

Multimedia Appendix 2

Convergence, bottlenecks, sensitivity, and homophily in the respondent-driven sampling survey.

[[PDF File \(Adobe PDF File\), 1MB - publichealth_v5i2e12451_app2.pdf](#)]

References

1. Lindan CP, Lieu TX, Giang LT, Lap VD, Thuc NV, Tinh T, et al. Rising HIV infection rates in Ho Chi Minh City herald emerging AIDS epidemic in Vietnam. *AIDS* 1997 Sep;11(Suppl 1):S5-13. [Medline: [9376101](#)]
2. VAAC. Viet Nam HIV/AIDS Estimates and Projections 2007-2012. Ha noi; 2009 Apr. URL: <http://unaids.org.vn/wp-content/uploads/2015/08/EPP-report-EN.pdf> [WebCite Cache ID 72zYnakRD]
3. Hien N, Wolffers I. HIV infection in Vietnam. *Lancet* 1994 Feb;343(8894):410. [doi: [10.1016/S0140-6736\(94\)91239-4](https://doi.org/10.1016/S0140-6736(94)91239-4)]
4. UNAIDS. AIDSinfo. 2015. URL: <http://aidsinfo.unaids.org/> [WebCite Cache ID 72zZ7NDWe]
5. The Global Fund to Fight AIDS, Tuberculosis and Malaria. TB and HIV Concept Note: Investing for impact against tuberculosis and HIV. 2014 Mar 10. URL: http://www.unaids.org/sites/default/files/country/documents/VNM_narrative_report_2015.pdf [WebCite Cache ID 72ywO8X2y]
6. Viet Nam Ministry of Health. Optimizing Viet Nam's HIV Response: An Investment Case. 2014 Oct. URL: https://www.aidsdatahub.org/sites/default/files/publication/Vietnam_investment_case_2014.pdf [WebCite Cache ID 75IDM9R9B]

7. Thanh DC, Ha NT, Ha HT, Tuan NA. Men who have sex with men in Vietnam: an update on HIV epidemic and risk behavior. *Vietnam J Prevent Med* 2017;27(12):5-10 [FREE Full text]
8. UNAIDS. UNAIDS Data 2018. 2018. URL: http://www.unaids.org/sites/default/files/media_asset/unaid-data-2018_en.pdf [WebCite Cache ID 72zbOO2O5]
9. Safarnejad A, Nga NT, Son VH. Population size estimation of men who have sex with men in Ho Chi Minh City and Nghe An using social app multiplier method. *J Urban Health* 2017 Dec;94(3):339-349 [FREE Full text] [doi: [10.1007/s11524-016-0123-0](https://doi.org/10.1007/s11524-016-0123-0)] [Medline: [28138799](https://pubmed.ncbi.nlm.nih.gov/28138799/)]
10. Cáceres C, Konda K, Pecheny M, Chatterjee A, Lyerla R. Estimating the number of men who have sex with men in low and middle income countries. *Sex Transm Infect* 2006 Jun;82(Suppl 3):iii3-iii9 [FREE Full text] [doi: [10.1136/sti.2005.019489](https://doi.org/10.1136/sti.2005.019489)] [Medline: [16735290](https://pubmed.ncbi.nlm.nih.gov/16735290/)]
11. Colby D, Cao NH, Doussantousse S. Men who have sex with men and HIV in Vietnam: a review. *AIDS Educ Prev* 2004 Feb;16(1):45-54. [doi: [10.1521/aeap.16.1.45.27722](https://doi.org/10.1521/aeap.16.1.45.27722)] [Medline: [15058710](https://pubmed.ncbi.nlm.nih.gov/15058710/)]
12. UNAIDS. Ending AIDS: Progress Towards the 90-90-90 Targets. 2017. URL: http://www.unaids.org/sites/default/files/media_asset/Global_AIDS_update_2017_en.pdf [WebCite Cache ID 72zbxo735]
13. UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance. Guidelines on Estimating the Size of Populations Most at Risk to HIV. Geneva: World Health Organization; 2010.
14. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. *Soc Probl* 1997 May;44(2):174-199. [doi: [10.2307/3096941](https://doi.org/10.2307/3096941)]
15. Gile KJ, Handcock MS. 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/)]
16. Salganik MJ, Heckathorn DD. Sampling and estimation in hidden populations using respondent - driven sampling. *Sociol Methodol* 2004 Dec;34(1):193-240. [doi: [10.1111/j.0081-1750.2004.00152.x](https://doi.org/10.1111/j.0081-1750.2004.00152.x)]
17. Heckathorn DD. 6. Extensions of respondent-driven sampling: analyzing continuous variables and controlling for differential recruitment. *Sociol Methodol* 2007 Aug;37(1):151-208. [doi: [10.1111/j.1467-9531.2007.00188.x](https://doi.org/10.1111/j.1467-9531.2007.00188.x)]
18. Fosgate GT. Practical sample size calculations for surveillance and diagnostic investigations. *J Vet Diagn Invest* 2009 Jan;21(1):3-14. [doi: [10.1177/104063870902100102](https://doi.org/10.1177/104063870902100102)] [Medline: [19139495](https://pubmed.ncbi.nlm.nih.gov/19139495/)]
19. Volz E, Heckathorn D. Probability based estimation theory for respondent driven sampling. *J Off Stat* 2008;24(1):79-97 [FREE Full text]
20. Handcock M. UCLA Statistics. 2013. Estimation of Uncertainty: Confidence Intervals and Standard Errors URL: <http://www.stat.ucla.edu/~handcock/hpmrg/files/VIII.pdf> [accessed 2018-10-07] [WebCite Cache ID 72zcEBfQe]
21. UNAIDS. Quick Start Guide for Spectrum. 2018 Jan. URL: http://www.unaids.org/sites/default/files/media_asset/QuickStartGuide_Spectrum_en.pdf [WebCite Cache ID 72zcLdfza]
22. General Statistics Office of Viet Nam. Statistical Data: Population and Employment. 2016. URL: <http://www.gso.gov.vn> [WebCite Cache ID 72zcXPgOd]
23. Tuan NT. [Communist Review]. 2007. [Current management policy of cultural areas in Viet Nam] URL: <http://www.tapchiconsan.org.vn/Home/Tieu-diem/2007/394/Chinh-sach-quan-ly-cac-vung-van-hoa-o-Viet-Nam-hien.aspx> [accessed 2018-12-27] [WebCite Cache ID 74yRp0PjB]
24. H ng L, Ph ng D, Tàì L, Linh N, Duyễn P. [Prevalence of Stroke in Provinces Representing Ecoregions of Viet Nam 2013 - 2014 and Related Factors]. [Medical Research Journal] 2016;104(6) [FREE Full text]
25. Wikimedia Commons. 2016. Labeled map of the regions in Vietnam classified by General Statistics Office of Vietnam since 1999 URL: <https://commons.wikimedia.org/wiki/File:VietnameseRegions.png> [accessed 2018-12-27] [WebCite Cache ID 74yVYxUN0]
26. 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/)]
27. Quaye S, Fisher Raymond H, 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/)]
28. Guo W, Bao S, Lin W, Wu G, Zhang W, Hladik W, et al. Estimating the size of HIV key affected populations in Chongqing, China, using the network scale-up method. *PLoS One* 2013;8(8):e71796 [FREE Full text] [doi: [10.1371/journal.pone.0071796](https://doi.org/10.1371/journal.pone.0071796)] [Medline: [23967246](https://pubmed.ncbi.nlm.nih.gov/23967246/)]
29. Berleva G, Dumchev K, Kasianchuk M, Nikolko M, Saliuk T, Shvab I, et al. AIDS Alliance Ukraine. 2012. Estimation of the Size of Populations Most-at-Risk for HIV Infection in Ukraine URL: http://www.aidsalliance.org.ua/ru/library/our/2013/SE_2012_Eng.pdf [accessed 2018-10-07] [WebCite Cache ID 72ykAAOfa]
30. 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/)]

31. Evans A, Wiggins R, Mercer C, Bolding G, Elford J. Men who have sex with men in Great Britain: comparison of a self-selected internet sample with a national probability sample. *Sex Transm Infect* 2007 Jun;83(3):200-5; discussion 205 [FREE Full text] [doi: [10.1136/sti.2006.023283](https://doi.org/10.1136/sti.2006.023283)] [Medline: [17135330](https://pubmed.ncbi.nlm.nih.gov/17135330/)]
32. Marcus U, Hickson F, Weatherburn P, Schmidt AJ, EMIS Network. Estimating the size of the MSM populations for 38 European countries by calculating the survey-surveillance discrepancies (SSD) between self-reported new HIV diagnoses from the European MSM internet survey (EMIS) and surveillance-reported HIV diagnoses among MSM in 2009. *BMC Public Health* 2013 Oct 3;13:919 [FREE Full text] [doi: [10.1186/1471-2458-13-919](https://doi.org/10.1186/1471-2458-13-919)] [Medline: [24088198](https://pubmed.ncbi.nlm.nih.gov/24088198/)]
33. Ward B, Dahlhamer J, Galinsky A, Joestl S. Centers for Disease Control and Prevention. 2014. Sexual Orientation and Health Among US Adults: National Health Interview Survey, 2013 URL: <https://www.cdc.gov/nchs/data/nhsr/nhsr077.pdf> [accessed 2018-10-07] [WebCite Cache ID 72yjrFuLx]
34. Grey JA, Bernstein KT, Sullivan PS, Purcell DW, Chesson HW, Gift TL, 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/)]
35. Coffman KB, Coffman LC, Ericson KMM. The size of the LGBT population and the magnitude of antigay sentiment are substantially underestimated. *Manage Sci* 2017 Sep;63(10):3168-3186. [doi: [10.1287/mnsc.2016.2503](https://doi.org/10.1287/mnsc.2016.2503)]
36. Newport F. Gallup. 2015. Americans Greatly Overestimate Percent Gay, Lesbian in US URL: <https://news.gallup.com/poll/183383/americans-greatly-overestimate-percent-gay-lesbian.aspx> [accessed 2018-10-07] [WebCite Cache ID 72zcv6bPU]
37. Robison J. Gallup. 2012. What Percentage of the Population Is Gay? URL: <https://news.gallup.com/poll/6961/what-percentage-population-gay.aspx> [accessed 2018-10-07] [WebCite Cache ID 72yixRfbP]
38. VAAC. Vietnam HIV/AIDS Estimates and Projections 2011-2015. 2012. URL: <https://www.fhi360.org/sites/default/files/media/documents/FHI-Layout> [WebCite Cache ID 731uLFu9C]
39. VNP+. Viet Nam Stigma Index. 2015. URL: <http://www.stigmaindex.org/viet-nam> [WebCite Cache ID 72yqkLR10]
40. Baral S, Turner RM, Lyons CE, Howell S, Honermann B, Garner A, et al. Population size estimation of gay and bisexual men and other men who have sex with men using social media-based platforms. *JMIR Public Health Surveill* 2018 Dec 8;4(1):e15 [FREE Full text] [doi: [10.2196/publichealth.9321](https://doi.org/10.2196/publichealth.9321)] [Medline: [29422452](https://pubmed.ncbi.nlm.nih.gov/29422452/)]
41. Krishnan A, Nguyen M, Giang LM, Ha TV, Bhadra M, Nguyen SM, et al. Finding sex partners through social media among men who have sex with men in Hanoi, Vietnam. *J Community Health* 2018 Feb;43(1):146-156. [doi: [10.1007/s10900-017-0397-5](https://doi.org/10.1007/s10900-017-0397-5)] [Medline: [28677027](https://pubmed.ncbi.nlm.nih.gov/28677027/)]
42. Masquelier B, Pison G, Ndiaye C, Lecomte L, Diémé N, Diouf I, et al. Age misreporting in censuses in developing countries: a record linkage study in health and demographic surveillance systems in Senegal. : Office of Population Research at Princeton University; 2016 Presented at: European Population Conference, Session 69: Data sources and methods quality; August 31-September 1, 2016; Mainz, Germany.
43. Toma C, Hancock J, Ellison N. Separating fact from fiction: an examination of deceptive self-presentation in online dating profiles. *Pers Soc Psychol Bull* 2008 Aug;34(8):1023-1036. [doi: [10.1177/0146167208318067](https://doi.org/10.1177/0146167208318067)] [Medline: [18593866](https://pubmed.ncbi.nlm.nih.gov/18593866/)]
44. Oyer PE. Everything I Ever Needed to Know about Economics I Learned from Online Dating. Boston, Massachusetts: Harvard Business Review Press; 2014.
45. 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/)]
46. Central Intelligence Agency. CIA World Factbook. 2016. URL: <https://www.cia.gov/library/publications/the-world-factbook/geos/vm.html> [WebCite Cache ID 759BrW5Us]
47. Marcus U, Hickson F, Weatherburn P, Schmidt AJ, EMIS network. Age biases in a large HIV and sexual behaviour-related internet survey among MSM. *BMC Public Health* 2013 Sep 10;13:826 [FREE Full text] [doi: [10.1186/1471-2458-13-826](https://doi.org/10.1186/1471-2458-13-826)] [Medline: [24020518](https://pubmed.ncbi.nlm.nih.gov/24020518/)]
48. de Lima CE, Garner A, Braga-Orillard G. UNAIDS. 2018. Brazil, PrEP, & gay dating apps: a survey about knowledge, access and future use of PrEP among users of dating app Hornet URL: https://unaids.org.br/wp-content/uploads/2018/08/AIDS2018_PREP.compressed.pdf [accessed 2019-04-02] [WebCite Cache ID 77KDTY9UF]

Abbreviations

Gile's SS: Gile's sequential sampling

MSM: men who have sex with men

RDS: respondent-driven sampling

UNAIDS: Joint United Nations Programme on HIV/AIDS

Edited by J Neal; submitted 09.10.18; peer-reviewed by B Honermann, D Gexha Bunjaku; comments to author 19.11.18; revised version received 10.01.19; accepted 22.02.19; published 17.04.19.

Please cite as:

Son VH, Safarnejad A, Nga NT, Linh VM, Tu LTC, Manh PD, Long NH, Abdul-Quader A

Estimation of the Population Size of Men Who Have Sex With Men in Vietnam: Social App Multiplier Method

JMIR Public Health Surveill 2019;5(2):e12451

URL: <http://publichealth.jmir.org/2019/2/e12451/>

doi: [10.2196/12451](https://doi.org/10.2196/12451)

PMID: [30994469](https://pubmed.ncbi.nlm.nih.gov/30994469/)

©Vo Hai Son, Ali Safarnejad, Nguyen Thien Nga, Vu Manh Linh, Le Thi Cam Tu, Pham Duc Manh, Nguyen Hoang Long, Abu Abdul-Quader. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 17.04.2019. 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

Wet Markets and Food Safety: TripAdvisor for Improved Global Digital Surveillance

Nicole E Kogan^{1,2}, BSc; Isabelle Bolon², DVM, PhD, MPH; Nicolas Ray^{2,3}, PhD; Gabriel Alcoba⁴, MD, MPH; Jose L Fernandez-Marquez⁵, PhD; Martin M Müller⁶, MSc; Sharada P Mohanty⁶, MSc; Rafael Ruiz de Castañeda², PhD

¹Massachusetts Institute of Technology, Cambridge, MA, United States

²Institute of Global Health, Faculty of Medicine, University of Geneva, Geneva, Switzerland

³Institute for Environmental Sciences, University of Geneva, Geneva, Switzerland

⁴Division of Tropical and Humanitarian Medicine, University Hospitals of Geneva, Geneva, Switzerland

⁵Citizen Cyberlab, Centre Universitaire d'Informatique, University of Geneva, Carouge, Switzerland

⁶Digital Epidemiology Lab, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Corresponding Author:

Rafael Ruiz de Castañeda, PhD

Institute of Global Health

Faculty of Medicine

University of Geneva

Campus Biotech

Chemin des Mines 9

Geneva, 1202

Switzerland

Phone: 41 223790951

Email: rafael.ruizdecastaneda@unige.ch

Abstract

Background: Wet markets are markets selling fresh meat and produce. Wet markets are critical for food security and sustainable development in their respective regions. Due to their cultural significance, they attract numerous visitors and consequently generate tourist-gear information on the Web (ie, on social networks such as *TripAdvisor*). These data can be used to create a novel, international wet market inventory to support epidemiological surveillance and control in such settings, which are often associated with negative health outcomes.

Objective: Using social network data, we aimed to assess the level of wet markets' touristic importance on the Web, produce the first distribution map of wet markets of touristic interest, and identify common diseases facing visitors in these settings.

Methods: A *Google* search was performed on 31 food market-related keywords, with the first 150 results for each keyword evaluated based on their relevance to tourism. Of all these queries, *wet market* had the highest number of tourism-related *Google Search* results; among these, *TripAdvisor* was the most frequently-occurring travel information aggregator, prompting its selection as the data source for this study. A Web scraping tool (*ParseHub*) was used to extract wet market names, locations, and reviews from *TripAdvisor*. The latter were searched for disease-related content, which enabled assignment of *GeoSentinel* diagnosis codes to each. This syndromic categorization was overlaid onto a mapping of wet market locations. Regional prevalence of the most commonly occurring symptom group - food poisoning - was then determined (ie, by dividing the number of wet markets per continent with more than or equal to 1 review containing this syndrome by the total number of wet markets on that continent with syndromic information).

Results: Of the 1090 hits on *TripAdvisor* for *wet market*, 36.06% (393/1090) conformed to the query's definition; wet markets were heterogeneously distributed: Asia concentrated 62.6% (246/393) of them, Europe 19.3% (76/393), North America 7.9% (31/393), Oceania 5.1% (20/393), Africa 3.1% (12/393), and South America 2.0% (8/393). Syndromic information was available for 14.5% (57/393) of wet markets. The most frequently occurring syndrome among visitors to these wet markets was food poisoning, accounting for 54% (51/95) of diagnoses. Cases of this syndrome were identified in 56% (22/39) of wet markets with syndromic information in Asia, 71% (5/7) in Europe, and 71% (5/7) in North America. All wet markets in South America and Oceania reported food poisoning cases, but the number of reviews with syndromic information was very limited in these regions (n=2).

Conclusions: The map produced illustrates the potential role of touristically relevant social network data to support global epidemiological surveillance. This includes the possibility to approximate the global distribution of wet markets and to identify diseases (ie, food poisoning) that are most prevalent in such settings.

(*JMIR Public Health Surveill* 2019;5(2):e11477) doi:[10.2196/11477](https://doi.org/10.2196/11477)

KEYWORDS

epidemiology; maps; foodborne diseases; social networking; travel; agriculture

Introduction

Background

Traditional food markets (ie, wet markets) play important roles in food security and local development [1]; however, they also have negative health implications. In 2003, severe acute respiratory syndrome spread globally from a Chinese wet market, causing hundreds of deaths [2] and major economic losses [3]. Avian influenza has also been repeatedly associated with wet markets [1,4]. Foodborne *Campylobacter*, *Salmonella*, *Giardia*, and *Escherichia* are most common in these settings, leading to 18 million disability-adjusted life years annually [5]. These are particularly important in low- and middle-income countries (LMICs), but their true impact is unknown given that many episodes go unreported [6].

Due to wet markets' cultural importance, there exists extensive tourist-gear information on the Web, often on websites such as *Yelp* and *TripAdvisor*, which serve as forums to share experiences. Although the use of these data remains unexploited for wet markets, it has yielded compelling results for restaurants. For example, *iwaspoisoned.com* [7] serves as a platform where individuals can report symptoms of food poisoning alongside the offending eatery. Through citizen participation, the website has identified several foodborne disease outbreaks before traditional epidemiological methods [8].

Objectives

This approach shaped the objectives of this study, which were to show the link between wet markets and tourism as well as to exploit tourist-generated social network data to create the first map of the distribution of wet markets of touristic interest and their associated adverse health events. Rather than function as an epidemiological analysis in which foodborne disease incidence related to wet market visits is calculated, our approach aims to showcase the potential of Web-based social networks to pick up potentially overlooked instances of disease.

Methods

Establishing the Data Source

A *Google* search was performed on different food market types (see [Multimedia Appendix 1](#)). For each keyword, the first 150 results were scraped and characterized based on their relevance to tourism (ie, presence of tourist-gear content and promotion of an area to potential visitors). Of all food market types, *wet market* was linked to the highest proportion of tourism-related

websites—59.3% (89/150) of *Google Search* results. This connection motivated its use in this study. Any social networking websites were flagged, and their touristic importance was assessed. Of these, *TripAdvisor* appeared most frequently and had the most comprehensive wet market-related information, prompting its selection as this study's data source.

Web Scraping

The term *wet market* was inputted into *TripAdvisor* (July 2017). Wet market names and locations were harvested using *ParseHub* [9]. Irrelevant results (eg, waterparks with *wet* in their names) were removed manually. A *Python 2.7* script integrating geocoding library *GeoPy* [10] was developed to convert wet market locations into geographic coordinates for mapping.

Text Mining

For each wet market, the *TripAdvisorReviews* section was parsed for mentions of keywords most often associated with foodborne disease, a list corroborated by the National Institutes of Health National Institute of Diabetes and Digestive and Kidney Diseases Web page on foodborne illness [11]: *diarrhea*, *vomit/vomiting/vomited*, *foodpoisoning*, *stomach ache*, *headache*, *nausea/nauseous*, *upset stomach*, *sick*, *ill*, and *dizzy*. Comments containing at least one of these were manually extracted. When possible, comments were assigned a *GeoSentinel* diagnosis code [12] through medical expert analysis using syndrome keywords and indicators of symptom duration. Wet market and syndromic distribution were analyzed using descriptive statistics.

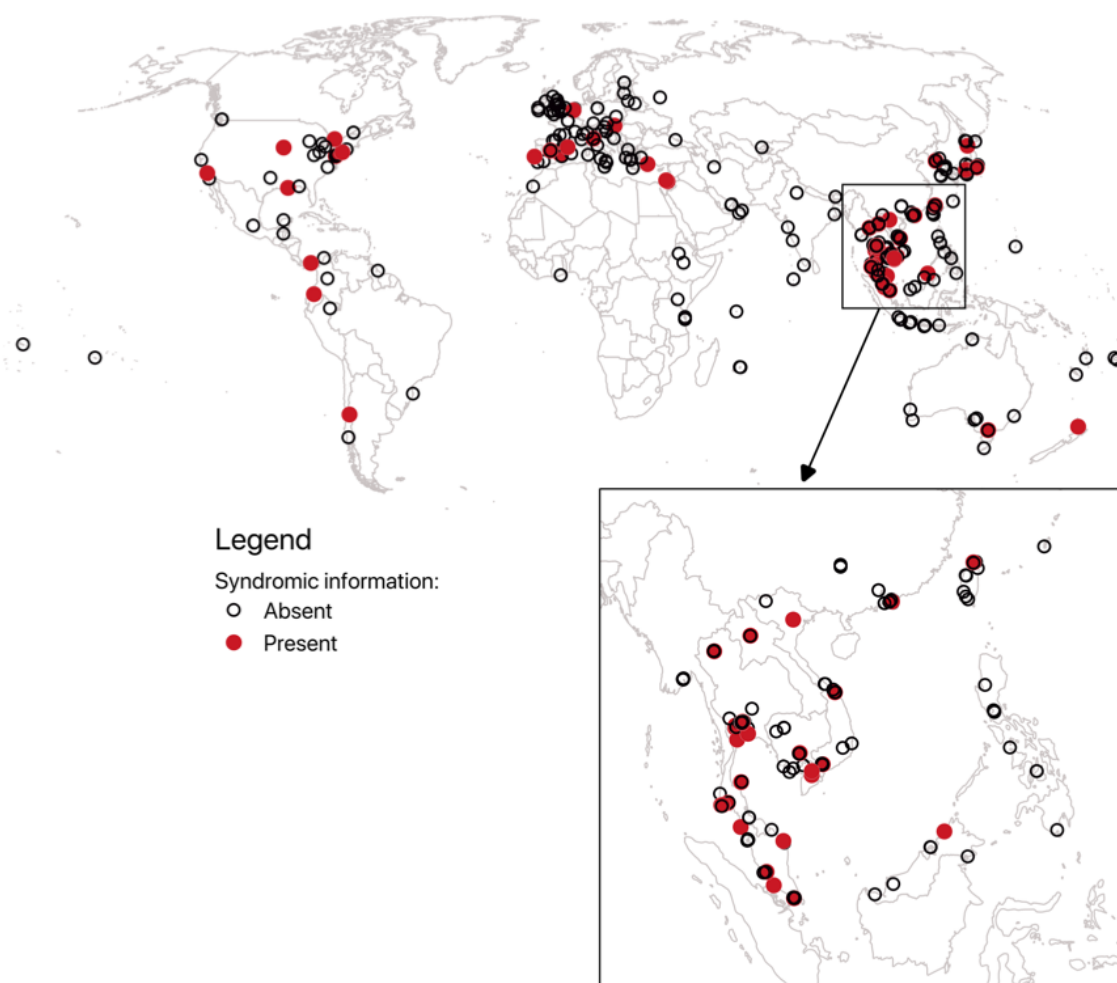
The study represents a passive analysis of information on *TripAdvisor*. The investigators did not participate in nor were involved in *TripAdvisor* communications (ie, there was no posting), so the analysis should not be considered active internet-based research requiring human subject consent.

Results

Global Wet Market Distribution

Wet market yielded 1090 attractions on *TripAdvisor*, 36.06% (393/1090) of which aligned with the term's definition. Mapping revealed Asia as the region with the greatest wet market density, accounting for 62.6% (246/393) of wet markets reviewed ([Figure 1](#)). The second-most wet market-dense region was Europe with 19.3% (76/393) of wet markets reviewed, followed by North America with 7.9% (31/393), Oceania with 5.1% (20/393), Africa with 3.1% (12/393), and South America with 2.0% (8/393).

Figure 1. Locations of TripAdvisor-sourced wet markets, with zoomed inset in Southeast Asia. Red circles represent wet markets where visitors reported adverse health events, whereas empty circles denote locations where such reports were lacking. The background country border is sourced from Natural Earth vector data projected in the World Robinson coordinate reference system using Quantum Geographic Information System (QGIS) 2.18.2.

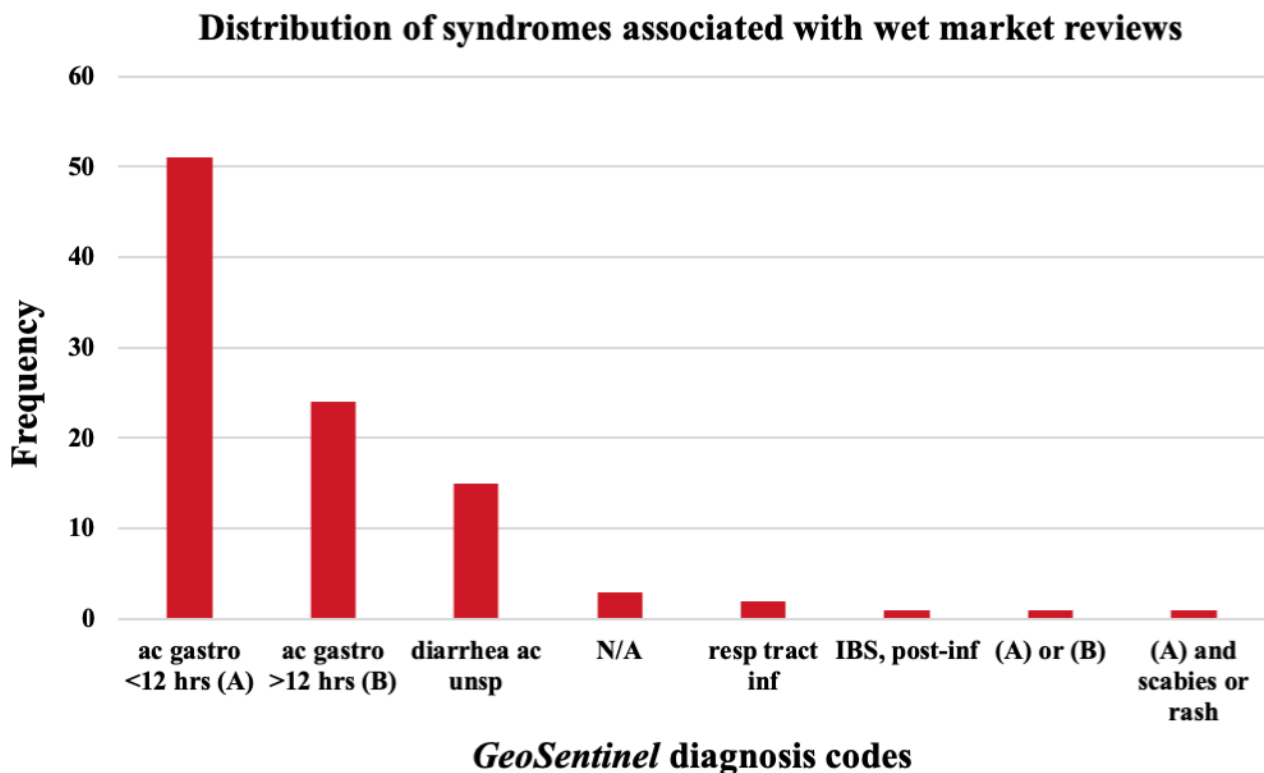


Assessment of Syndromic Information

All 393 wet markets were reviewed on *TripAdvisor*. Moreover, 14.5% (57/393) of these contained at least one review with syndromic information and were broken down regionally as 68% (39/57) Asian, 12% (7/57) European, 12% North American (7/57), 3% South American (2/57), and 3% Oceanic (2/57). In total, these yielded 98 reviews with syndromic information, of which Asian markets accounted for 74% (73/98). *Acute*

gastroenteritis <12 hrs, food poisoning was the most common diagnosis in wet markets globally 54% (51/95) syndrome references, after discarding 3 reviews corresponding to unascertainable illness; [Figure 2](#)). Asia, where the review volume was particularly noticeable, had 56% (22/39) of wet markets with this specific syndrome. The syndromes *Acute gastroenteritis > 12 hrs*, *Acute gastroenteritis <12 hrs*, and *Diarrhea, acute unspecified* comprised 95% (90/95) non-*N/A* diagnoses from the 57 wet markets.

Figure 2. Frequency distribution of 98 TripAdvisor reviewer diagnoses following GeoSentinel encoding. ac=acute; gastro=gastroenteritis; IBS=irritable bowel syndrome; N/A=symptoms present but illness unascertainable; resp tract inf=respiratory tract infection; unsp=unspecified.



Discussion

Principal Findings

This study shows the touristic importance of wet markets and provides, to our knowledge, the first global map of wet markets' locations and related syndromic information communicated on *TripAdvisor* by their visitors. Our map reveals that wet markets are heterogeneously distributed across continents. Asia is a hotspot, both in the number of markets and the absolute number of adverse associated health events. When controlling by the number of continental wet markets with syndromic information, South America and Oceania exhibited the highest proportion of *GeoSentinel acute gastroenteritis <12 hrs, food poisoning* diagnoses. However, both these regions were wet market-sparse ($n=2$ wet markets with syndromic information in each) compared with a continent like Asia.

Use of region-specific terminology for designating markets could have yielded denser mapping and should be used in future work. Furthermore, the choice to look only at *TripAdvisor* reviews in English may have impacted the wet market distribution shown in the map. An example is the density of wet markets, which seems to be in preferred destinations of English-speaking tourists; another example is that the concentration of wet markets in Africa is largely in English-speaking countries. This could prompt search on the platform using other languages to enhance the results.

Little is known about the wet market health threats to locals and travelers partly because of underreporting of adverse health events experienced by visitors [13]. Although social networking

platforms do not paint a complete picture of wet markets (eg, total number of regional wet markets and how many yearly visitors), they can provide opportunities to instantaneously report and detect these elusive cases—primarily traveler-related ones—and improve epidemiological monitoring within these settings.

Limitations and Strengths

Reviews from *Yelp*, a *TripAdvisor*-like platform connecting citizens with local businesses, have been leveraged by New York public health authorities to detect restaurant-related foodborne disease events [14]; interestingly, fewer than 2% of individuals with an alleged illness explicitly mention reporting their case to a medical professional [14]. *Twitter* has been implemented in St. Louis, Missouri, to detect food poisoning cases [15], but its potential is restricted by the short post length. The website *iwaspoisoned.com* is equally notable in its effort to crowdsource information, though it is limited in scope outside the United States. Wet markets are more challenging establishments because of their typical location in LMICs and their operation under often-limited regulations. This necessitates tapping into other Web-based communities—social platforms such as *TripAdvisor*—to glean information on visitor health.

Traveler's diarrhea is a common disorder affecting tourists visiting developing countries [16] and is generally associated with consumption of foods prepared under unhygienic conditions (common in wet markets). Our analysis shows that acute gastroenteritis (food poisoning) and diarrhea were the most frequent illnesses among wet market visitors. However, this must be carefully considered from medical and epidemiological perspectives as the result is based on rarely corroborated,

Web-based descriptions. We also cannot exclude that reported symptoms could have been caused by an event before or after (but unrelated to) a wet market visit.

Conclusions

As use of *TripAdvisor* in isolation has accounted for only a fraction of wet markets of touristic importance and their associated health risks, larger datasets are needed to confirm the results presented here and to explore others. Even so, the pipeline we present is significant for travel medicine and

epidemiology. It could ultimately contribute to predictive models for improved epidemic forecasting and to the development of diagnostic tools based on syndromic surveillance and artificial intelligence. *TripAdvisor* could crosstalk with other social networks (eg, *Yelp*) for maximal information coverage and partner with other initiatives for a more structured collection of wet market-related health information in near-real time. In this way, we can gain an improved understanding of global wet markets and their associated health risks while also ensuring their safer promotion.

Acknowledgments

NEK was supported by the MIT International Science and Technology Initiatives and by a Summer School Grant from the Zeno Karl Schindler Foundation. RRdC was supported by Fondation Louis-Jeantet, which the authors thank very much. The authors also thank Prof A Flahault for supporting this work at the Institute of Global Health. The authors thank 2 anonymous reviewers for their constructive comments.

Authors' Contributions

RRdC, IB, and NEK designed the study. NR, GA, JLFM, MMM, and SPM guided the implementation. NR and NEK developed the spatial analyses. GA paired *GeoSentinel* diagnosis codes with *TripAdvisor* reviews. NEK wrote the initial draft of the manuscript, and all coauthors reviewed and improved it.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of market types whose relevance to tourism was assessed in an initial Google Search screen.

[[DOCX File, 13KB - publichealth_v5i2e11477_app1.docx](#)]

References

1. World Health Organization. Switzerland: WHO Press; 2006. A guide to healthy food markets URL: http://www.who.int/foodsafety/publications/capacity/healthymarket_guide.pdf?ua=1 [WebCite Cache ID 70UclvVLz]
2. Wu T, Perrings C, Kinzig A, Collins JP, Minter BA, Daszak P. Economic growth, urbanization, globalization, and the risks of emerging infectious diseases in China: a review. *Ambio* 2017 Feb;46(1):18-29 [FREE Full text] [doi: [10.1007/s13280-016-0809-2](https://doi.org/10.1007/s13280-016-0809-2)] [Medline: [27492678](https://pubmed.ncbi.nlm.nih.gov/27492678/)]
3. Lee JW, McKibbin WJ. Estimating the global economic costs of SARS. In: Knobler S, Mahmoud A, Lemon S, Mack A, Sivitz L, Oberholtzer K, editors. *Learning From SARS: Preparing for the Next Disease Outbreak: Workshop Summary*. Washington DC: National Academies Press; 2004.
4. Wu J, Lu J, Faria NR, Zeng X, Song Y, Zou L, et al. Effect of live poultry market interventions on influenza A(H7N9) virus, Guangdong, China. *Emerg Infect Dis* 2016 Dec;22(12):2104-2112 [FREE Full text] [doi: [10.3201/eid2212.160450](https://doi.org/10.3201/eid2212.160450)] [Medline: [27869613](https://pubmed.ncbi.nlm.nih.gov/27869613/)]
5. Grace D. Food safety in low and middle income countries. *Int J Environ Res Public Health* 2015 Aug 27;12(9):10490-10507 [FREE Full text] [doi: [10.3390/ijerph120910490](https://doi.org/10.3390/ijerph120910490)] [Medline: [26343693](https://pubmed.ncbi.nlm.nih.gov/26343693/)]
6. Bernardo TM, Rajic A, Young I, Robiadek K, Pham MT, Funk JA. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. *J Med Internet Res* 2013 Jul 18;15(7):e147 [FREE Full text] [doi: [10.2196/jmir.2740](https://doi.org/10.2196/jmir.2740)] [Medline: [23896182](https://pubmed.ncbi.nlm.nih.gov/23896182/)]
7. Quade P. IWasPoisoned. URL: <https://iwaspoisoned.com/> [accessed 2018-06-27] [WebCite Cache ID 70Uh5XbE0]
8. Neimark J. NPR. 2017 Nov 30. 'I Was Poisoned': Can Crowdsourcing Food Illnesses Help Stop Outbreaks? URL: <https://www.npr.org/sections/thesalt/2017/11/30/565769194/i-was-poisoned-can-crowdsourcing-food-illnesses-help-stop-outbreaks> [WebCite Cache ID 70Ujfiyin]
9. Fomina A, Prelich P, Toarca S. ParseHub. URL: <https://www.parsehub.com/> [accessed 2018-07-02] [WebCite Cache ID 70bzDzEXQ]
10. Esmukov K. GitHub. 2006. geopy URL: <https://github.com/geopy/geopy> [accessed 2018-06-27] [WebCite Cache ID 70UIpq9sy]

11. National Institute of Diabetes and Digestive and Kidney Diseases. 2014. Foodborne Illnesses URL: <https://www.niddk.nih.gov/health-information/digestive-diseases/foodborne-illnesses> [accessed 2018-12-17] [WebCite Cache ID 74jmK4zPz]
12. International Society of Travel Medicine. GeoSentinel Document Center (Sites Only) URL: <http://www.istm.org/geosentinel/documentcenter2> [accessed 2018-06-29] [WebCite Cache ID 70XWu0RXn]
13. Nsoesie EO, Kluberg SA, Brownstein JS. Online reports of foodborne illness capture foods implicated in official foodborne outbreak reports. *Prev Med* 2014 Oct;67:264-269 [FREE Full text] [doi: [10.1016/j.ypmed.2014.08.003](https://doi.org/10.1016/j.ypmed.2014.08.003)] [Medline: [25124281](https://pubmed.ncbi.nlm.nih.gov/25124281/)]
14. Ostrov BF. USA Today. 2015 Oct 24. Using Yelp to track food poisoning outbreaks URL: <https://www.usatoday.com/story/news/2015/10/24/kaiser-using-yelp-to-track-food-poisoning-outbreaks/74469136/> [WebCite Cache ID 70bzbGieK]
15. Harris JK, Hinyard L, Beatty K, Hawkins JB, Nsoesie EO, Mansour R, et al. Evaluating the implementation of a Twitter-based foodborne illness reporting tool in the city of St. Louis Department of Health. *Int J Environ Res Public Health* 2018 Apr 24;15(5):- [FREE Full text] [doi: [10.3390/ijerph15050833](https://doi.org/10.3390/ijerph15050833)] [Medline: [29695038](https://pubmed.ncbi.nlm.nih.gov/29695038/)]
16. Porter CK, Olson S, Hall A, Riddle MS. Travelers' diarrhea: an update on the incidence, etiology, and risk in military deployments and similar travel populations. *Mil Med* 2017 Sep;182(S2):4-10. [doi: [10.7205/MILMED-D-17-00064](https://doi.org/10.7205/MILMED-D-17-00064)] [Medline: [28885918](https://pubmed.ncbi.nlm.nih.gov/28885918/)]

Abbreviations

LMIC: low- and middle-income country

Edited by T Sanchez; submitted 03.07.18; peer-reviewed by S Ghosh, ICH Fung; comments to author 26.11.18; revised version received 18.01.19; accepted 18.01.19; published 01.04.19.

Please cite as:

Kogan NE, Bolon I, Ray N, Alcoba G, Fernandez-Marquez JL, Müller MM, Mohanty SP, Ruiz de Castañeda R

Wet Markets and Food Safety: TripAdvisor for Improved Global Digital Surveillance

JMIR Public Health Surveill 2019;5(2):e11477

URL: <https://publichealth.jmir.org/2019/2/e11477/>

doi: [10.2196/11477](https://doi.org/10.2196/11477)

PMID: [30932867](https://pubmed.ncbi.nlm.nih.gov/30932867/)

©Nicole E Kogan, Isabelle Bolon, Nicolas Ray, Gabriel Alcoba, Jose L Fernandez-Marquez, Martin M Müller, Sharada P Mohanty, Rafael Ruiz de Castañeda. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 01.04.2019. 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

An Automated Text-Messaging Platform for Enhanced Retention and Data Collection in a Longitudinal Birth Cohort: Cohort Management Platform Analysis

Caroline M Barry¹, BA; Aditi Sabhlok¹, BS; Victoria C Saba¹, BA; Alesha D Majors¹, BA; Julia C Schechter¹, PhD; Erica L Levine², MPH; Martin Streicher², MS; Gary G Bennett², PhD; Scott H Kollins¹, PhD; Bernard F Fuemmeler³, MPH, PhD

¹Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Duke University, Durham, NC, United States

²Global Digital Health Science Center, Global Health Institute, Duke University, Durham, NC, United States

³Cancer Prevention and Control, Department of Health Behavior and Policy, Virginia Commonwealth University, Richmond, VA, United States

Corresponding Author:

Bernard F Fuemmeler, MPH, PhD

Cancer Prevention and Control

Department of Health Behavior and Policy

Virginia Commonwealth University

PO Box 980149

830 E Main St

Richmond, VA, 23219

United States

Phone: 1 8048288892

Email: bernard.fuemmeler@vcuhealth.org

Abstract

Background: Traditional methods for recruiting and maintaining contact with participants in cohort studies include print-based correspondence, which can be unidirectional, labor intensive, and slow. Leveraging technology can substantially enhance communication, maintain engagement of study participants in cohort studies, and facilitate data collection on a range of outcomes.

Objective: This paper provides an overview of the development process and design of a cohort management platform (CMP) used in the Newborn Epigenetic Study (NEST), a large longitudinal birth cohort study.

Methods: The platform uses short message service (SMS) text messaging to facilitate interactive communication with participants; it also semiautomatically performs many recruitment and retention procedures typically completed by research assistants over the course of multiple study follow-up visits.

Results: Since February 2016, 302 participants have consented to enrollment in the platform and 162 have enrolled with active engagement in the system. Daily reminders are being used to help improve adherence to the study's accelerometer wear protocol. At the time of this report, 213 participants in our follow-up study who were also registered to use the CMP were eligible for the accelerometer protocol. Preliminary data show that texters (138/213, 64.8%), when compared to nontexters (75/213, 35.2%), had significantly longer average accelerometer-wearing hours (165.6 hours, SD 56.5, vs 145.3 hours, SD 58.5, $P=.01$) when instructed to wear the devices for 1 full week.

Conclusions: This platform can serve as a model for enhancing communication and engagement with longitudinal study cohorts, especially those involved in studies assessing environmental exposures.

(*JMIR Public Health Surveill* 2019;5(2):e11666) doi:[10.2196/11666](https://doi.org/10.2196/11666)

KEYWORDS

data collection; longitudinal studies; mobile health; text messaging

Introduction

Background

Staying in contact with each other is easier today than ever before. Modern technology enables and encourages interconnectedness, overcoming the age-old obstacles of distance and time. In the past decade, cell phone use has risen globally, with the total number of mobile phone subscriptions surpassing the world population, and cloud-based apps are facilitating unprecedented, readily accessible data storage [1,2]. With these advancements saturating personal and professional networks, it follows that communication platforms would permeate health and research sectors. While studies using digital health and app-based intervention approaches to address disease self-management have boomed in recent years, there remains an apparent need to apply these approaches to clinical and population-based cohort management, particularly among large, longitudinal, community-based cohort studies [3].

Traditional methods for recruiting and maintaining contact with participants in cohort studies can be unidirectional, labor intensive, and slow. In longitudinal research studies, staff typically take various approaches to optimize study retention, including providing incentives, making repeated phone calls, sending reminder mailings, contacting family or friends, and visiting participants' homes [4,5]. These efforts seek to maximize points of meaningful contact and keep contact information up-to-date, both important contributors to retention success [6]. Automated short message service (SMS) text messaging has been used in a number of intervention studies designed to address changes in health behaviors [2,7-14]. However, the field is lacking applied demonstrations of these methods for maintaining engagement of participants in large, longitudinal cohort studies. The following is an overview of the development process and design of a cohort management platform (CMP) used in the Newborn Epigenetic Study (NEST). Broader applications, impact, and suggestions for future directions and use are discussed, with emphasis on the growing importance of integrating electronic cohort management strategies in health and research.

Objectives

A CMP was initially designed and launched for use by NEST. NEST is a federally funded, longitudinal birth cohort study based at Duke University that examines how environmental exposures and nutrition, prenatally and during childhood, affect gene expression and health outcomes in children as they develop. The cohort includes 2595 mothers recruited during pregnancy between 2005 and 2011 and any children resulting from that pregnancy. The CMP launched in February 2016 and NEST participants have been recruited to enroll in the platform during study follow-up visits, pediatrician appointments, and through study mailings. The goals of the CMP with NEST are two-fold: (1) to maximize retention in this large longitudinal cohort and (2) to assist in data collection using a multimodal platform.

Maximizing retention is vital for longitudinal studies. Attrition increases with study duration and with a large cohort and

accumulated years of participation, retention becomes more difficult. Participants move, contact information changes, connections fade, and longitudinal studies see increasing loss to follow-up. Loss to follow-up threatens internal validity because lost participants may be systematically different from those who stay enrolled [15,16], and high attrition can contribute to reduced statistical power [17]. The CMP aims to address this issue by enhancing outreach and facilitating a quick, convenient means of interactive communication. Participants are equipped with a service for immediate connection, which aims to boost not only the number of contact points between participants and staff, but also the likelihood of maintaining up-to-date contact information and overall rapport with the study.

The second objective of the CMP is to assist with data collection. Survey questions are delivered directly to participants, who can key in responses on their mobile devices in real time. The advantages to this method include convenience for both participants and staff, cost efficiency, and real-time responses. Real-time responding increases data quality by reducing recall burden [18,19]. Immediacy of data collection can also provide a more complete picture of exposures, thoughts, emotions, and other variables of interest as they are experienced throughout the day. The following provides a brief overview of the CMP's development and operation, utility for NEST, strengths and limitations of the method, and broader applications for future use.

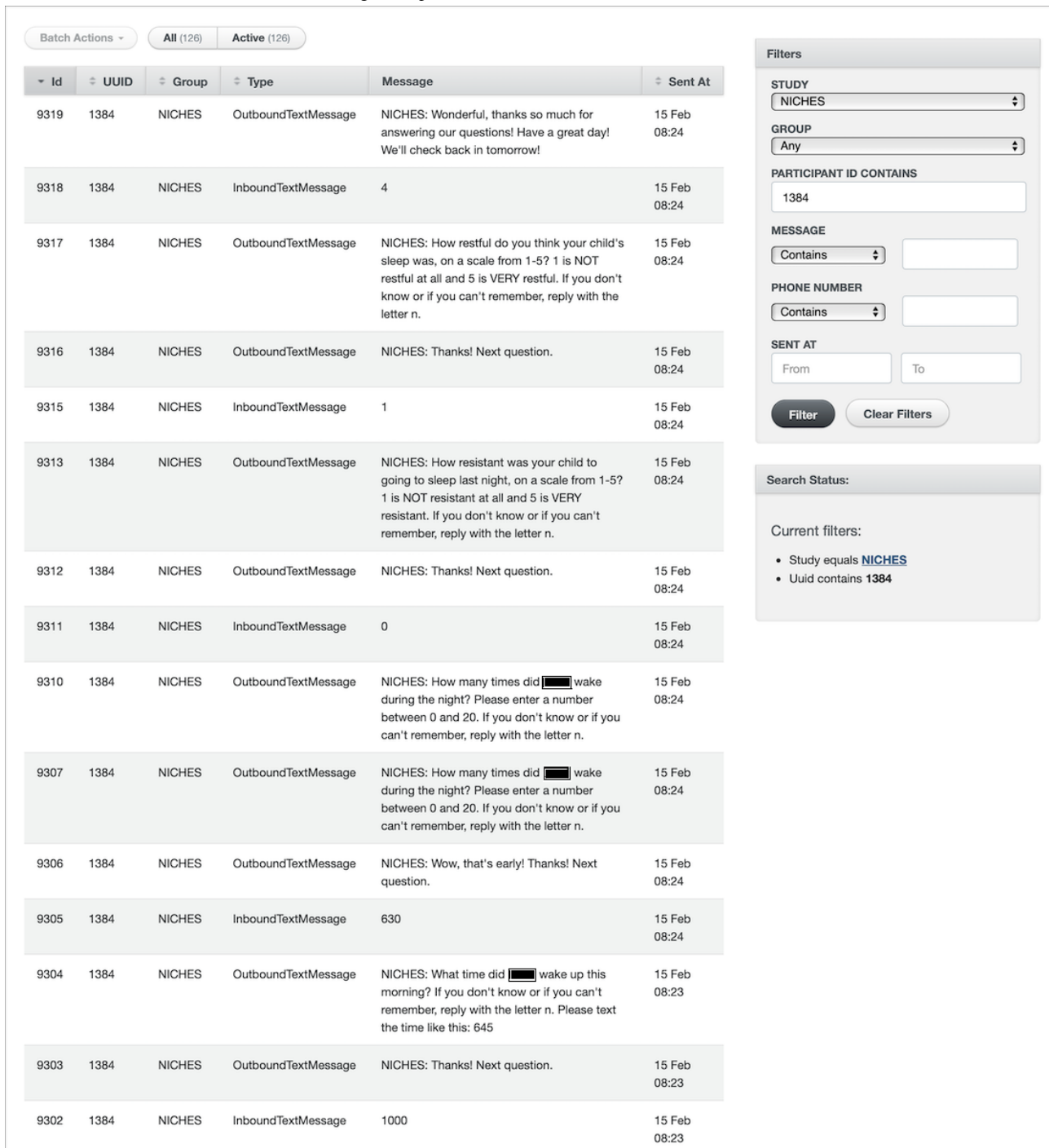
Methods

Development

The CMP was adapted and iteratively enhanced from a platform already in use by the Duke Global Digital Health Science Center. The platform is a flexible, cloud-based communication system designed to aid in behavioral interventions. Leveraging prior experience and existing code allowed for rapid development of the CMP at a lower cost than developing a system de novo. In earlier iterations of the platform, both the interactive SMS and interactive voice response (IVR) components contributed to high rates of engagement. Thus, those interactive features became central to the CMP.

The CMP runs on a custom-built, integrated, automated, interactive digital health system. Third-party vendors, such as Twilio, Heroku, Mandrill, and Amazon S3, provide relatively inexpensive infrastructure to send and receive messages from participants. Messaging within the CMP reads specific participant-tracking data fields, such as dates (eg, date of birth and date of last study visit), and delivers messages on a predetermined schedule using a customized library developed by the study team. For example, one message is sent to wish participants a happy birthday. Another message reminds participants to perform a task and yet another reminds participants about upcoming appointments. When appropriate, messages are further personalized to each participant to include the mother's and/or children's names. Daily messages sent over a 1-week period remind mothers to have their child, who is referred to by name, wear an accelerometer to record physical activity.

Figure 1. Researcher dashboard of the cohort management platform (CMP).



The platform also provides simple bidirectional messaging, where participants may respond to platform-generated text message queries. For instance, a message delivered in the morning asks participants to reply with qualitative measures of their child’s sleep quality and duration the night before.

A back-end dashboard (see Figure 1) was also built to allow the research team to review and monitor, in real time, when and what types of messages are sent and received from the participants. The study team is able to export data from the platform and merge it with central study databases for analysis. The study team also uses this back-end dashboard to register participants in the CMP once consent is given.

Consent

The CMP has been thoroughly reviewed by the Duke University’s Office of Clinical Research and Institutional Review Board for data security and integrity and to ensure ethical use. The CMP uses third-party vendors, such as Twilio, Heroku, Mandrill, and Amazon S3, to send messages, which encrypt participant information on their servers. However, if these companies or auditors opt to share these data, federal privacy protections may no longer apply. Further, it is recognized that no electronic communication channel is completely secure and confidential. The consent process involves informing participants who enroll to receive messages through the CMP of these potential risks; participants are given

the option to decline enrollment in the CMP in favor of traditional communication methods while still remaining in the study.

Results

Initial Participant Engagement

There were 302 NEST participants who were approached for CMP consent: 56.6% (171/302) were non-Hispanic black, 35.8% (108/302) were non-Hispanic white, 3.0% (9/302) were Hispanic, and 4.6% (14/302) were categorized as *other*. The mean age of 283 respondents was 33.9 years (SD 5.7). The mean annual household income of 278 respondents was US \$56,697 (SD US \$63,452). Since February 2016, of the 302 NEST participants who were approached for CMP consent, 98.7% (298/302) agreed to use the system and provided consent while 1.3% (4/302) opted out. A total of 162 participants are actively registered and using the system (ie, responded to a CMP message). Prior to registration of participants to the CMP, participants were informed of the risk and consent was obtained by study staff during an in-person study visit. Study staff then registered participants in the system using the back-end platform and shortly thereafter they received a *welcome* message. Participants then received the series of scheduled messages and were asked to interact with the system's short queries when prompted. Participants were given the option of withdrawing at any time, which they could do through the system by sending a message saying "stop." None have withdrawn thus far.

Although the CMP currently employs SMS text messaging as the primary method of communication, the study team has been asking participants, for quality improvement, to indicate how they would like to receive messages (ie, SMS, email, and/or IVR). Multiple responses are possible. Of those who have consented thus far, the majority agreed to receive text messages (266/302, 88.1%) and/or emails (278/302, 92.1%); fewer chose IVR (68/302, 22.5%). Given this initial success and that many participants chose more than one communication channel, our next step will be to roll out email channels followed by IVR to actively engage with the remaining participants for longitudinal follow-up.

Benefits to Study Management

Overall, study management and data collection for specific measures has been enhanced by using the CMP. For instance, some of the communication from study staff to participants regarding appointments can be managed through the CMP. Also, during the period when participants are asked to have their child wear their accelerometers, automated reminders are being sent and, at the same time, participants are being queried about their child's sleep duration and quality via automated text. Any unanswered text messages can be resent repeatedly on a designated schedule to improve data capture. Performing these same procedures with live study staff would result in a greater burden on the study team. For example, reminding participants

to have their child wear the accelerometer and to inquire about their child's sleep would require a staff member to make a phone contact with participants every day. For the 302 participants, one 3-minute call every day for 7 days adds up to 105.7 hours per week. Off-loading this to the CMP frees up time from staff to perform other vital tasks.

Enhancement of Data Collection and Cohort Retention

Beyond these methods of communication, the platform is used for data collection. Currently, the CMP is being used to collect data in two NEST follow-up studies. One study is designed to have two in-office appointments that are separated by no more than two weeks; during the interim, self-reported sleep data are gathered from participants using the CMP. The platform sends short text message queries asking participants to note their child's sleep duration and quality. As mentioned above, these questions are sent each morning, which gives mothers the ability to immediately reply, decreasing their burden to record or remember data from each night of sleep. In addition, the platform also sends participants daily reminders to have their child wear his or her accelerometer between their two visits. Initial data suggests that this is helping to increase adherence and maximize available data. For example, each morning at a preset time, typically 8 AM unless otherwise indicated, a reminder is sent to put the accelerometer on themselves and their child. The participant then receives questions about their child's sleep quality. The CMP waits for a valid response (ie, any 3 or 4 digits following a time pattern) and, if valid, the response is saved and the CMP sends the next question (see [Figure 2](#)). If the response is not valid, the CMP sends a prompt to try again (see [Figure 3](#)).

Preliminary data show that the CMP messaging may be helping to improve adherence to the accelerometer protocol. At the time of this report, 213 participants in our follow-up study who were also registered to use the CMP were eligible for the accelerometer protocol. This involves asking mothers to have their child wear an accelerometer on a daily basis for at least 7 days. Participants were told that they would be receiving text messages via the CMP to remind them to have their child wear his or her accelerometer. Of these 213 participants, we categorized them into *texters* and *nontexters*, based on whether they responded to a text at any point during the use of the CMP. Texters (138/213, 64.8%), when compared to nontexters (75/213, 35.2%), had significantly longer average accelerometer-wearing hours (165.6 hours, SD 56.5, vs 145.3 hours, SD 58.5, $P=.01$).

To enhance cohort retention, the CMP is being used to verify participant mailing addresses. For instance, messages are being sent after batch newsletters, and recruitment letters are mailed to confirm that they received these mailings (see [Figure 4](#)). If receipt is confirmed, the CMP thanks the participant; if not, the CMP requests updated contact information. New addresses can then be automatically fed into the study database to update records.

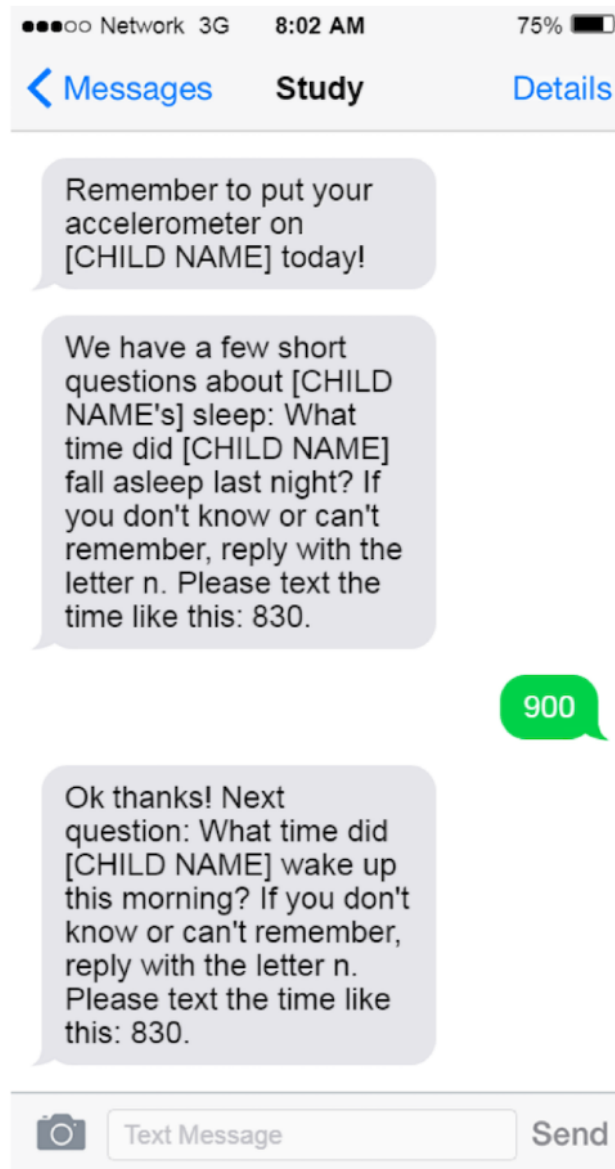
Figure 2. Successful response to the cohort management platform (CMP).

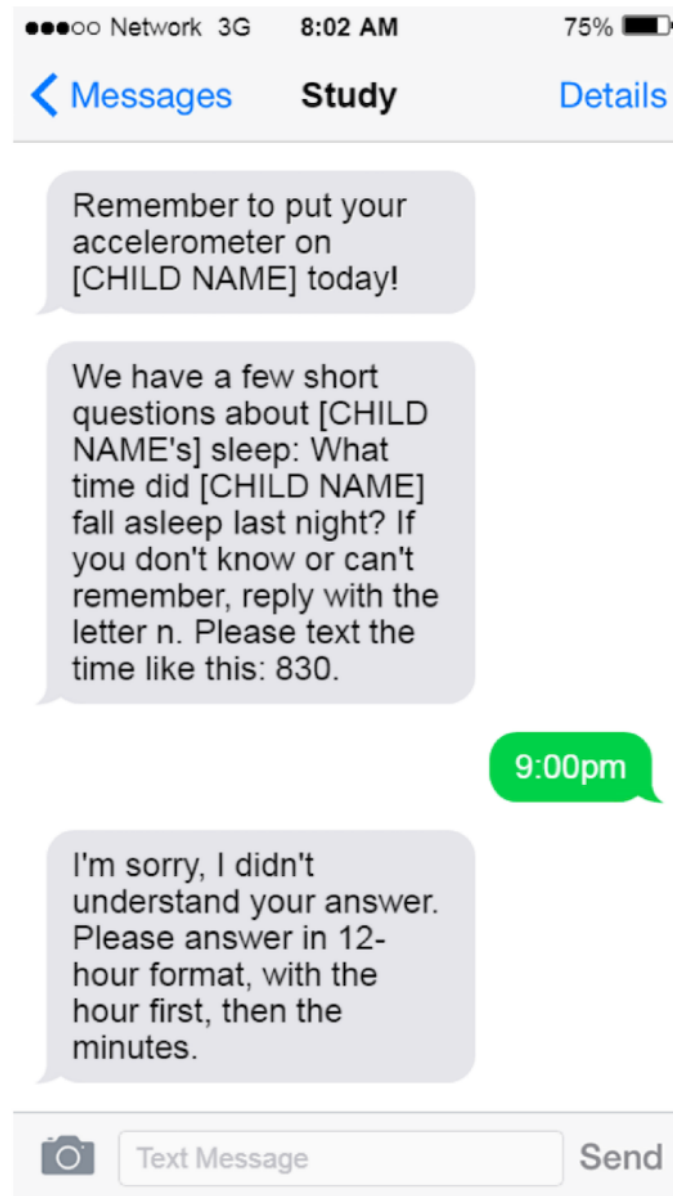
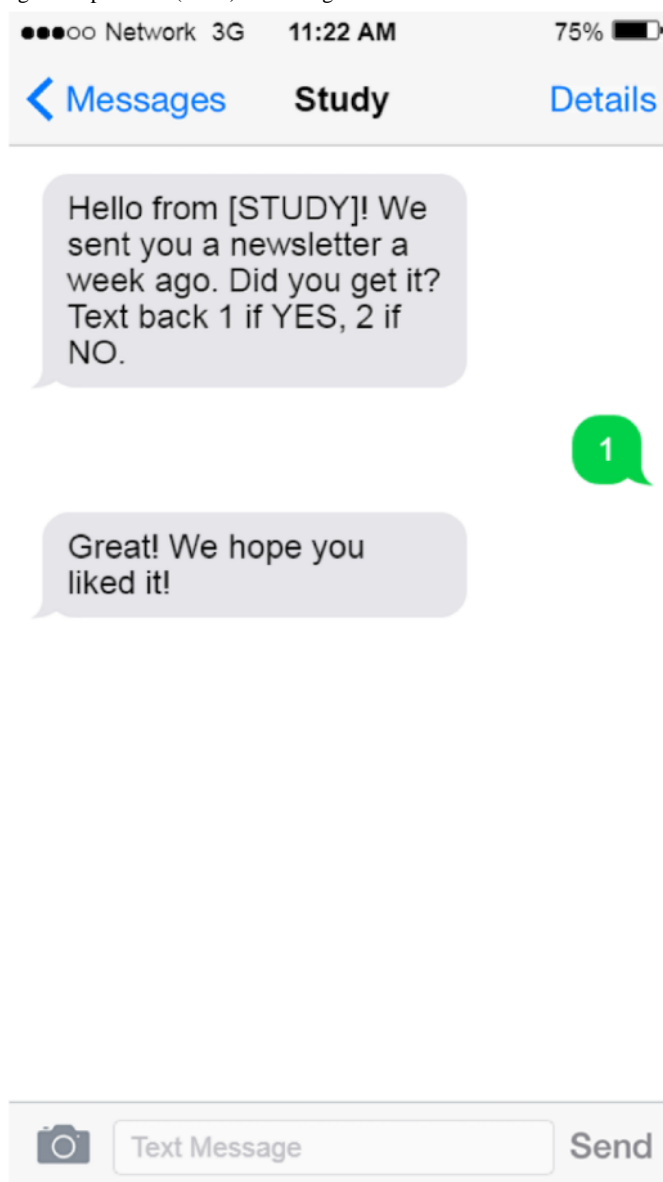
Figure 3. Unsuccessful response to the cohort management platform (CMP).

Figure 4. Response to the cohort management platform (CMP) indicating that the newsletter was received.

Discussion

Impact

To date, the CMP is showing great potential for maximizing retention and improving data collection among a large, longitudinal, community-based cohort. It is being used to address the inefficiency of traditional unidirectional participant contact protocols in numerous ways, such as the following: (1) by cutting down on staff time and effort dedicated to phone calls, mailings, and appointments; (2) by reducing resources for paper mailings and the waste of returned mail; and (3) by diminishing the burden placed on participants for measures requiring daily diaries.

Staff time and effort typically devoted to individual phone calls and single mailings can be curtailed with the aid of systems like the CMP because it capitalizes on automation while retaining personalization capacity for tailored messaging. Paper saved using the CMP is itself a compelling advantage, with immediate

resource-saving and eventual positive environmental impact. Further, the CMP allows for data collection in near real time. This can be particularly useful since measurement can be timed to be contemporaneous with key behaviors as they occur during the participant's daily life (eg, asking participants in the morning when they went to sleep the night before and when they woke up). In general, shortening the recall period for certain behaviors, such as sleep, improves accuracy [18,19]. In addition, it can be a lot easier for participants to respond quickly to a text message than to fill out a paper-based log and the data is captured automatically, thereby avoiding the need for staff to manually enter such data.

Strengths and Limitations

Strengths of the CMP include its potential cost-effectiveness, time-effectiveness, timed delivery of survey questions, real-time receipt of responses, and its accessibility through commonly used communication channels (ie, SMS text messaging). Unlike alternative data collection tools, such as REDCap or Qualtrics, the CMP is not a database platform, although it can be used in

conjunction with them. Database platforms rely primarily on manual data entry, whereas the CMP can run independently of staff to automate data collection and make real-time decisions and actions based on its collected data. Further, unlike REDCap or Qualtrics, the CMP is unique in its capacity to trigger the delivery of highly tailored messaging to perform a variety of tasks (eg, greetings, reminders, contact information changes, and data collection) over predetermined time periods. Importantly, the CMP can be used together with database platforms to achieve common goals of longitudinal cohort management. In general, the CMP can serve as a model for implementation of analogous designs.

Of note, there are some limitations of the data presented here as well as the CMP system. As it pertains to this report, we did not collect any data on why participants opted out of the CMP, which could be useful to better understand how to increase cohort management methods that are inclusive to all participants. Fortunately, for this particular sample, only around 1% of participants opted out of using the CMP, thus demonstrating that this type of system may be fairly acceptable to participants. Another limitation pertaining to the data presented here is that although we showed that compared to nontexters, texters had greater adherence to the accelerometer wear protocol, this does not confirm improved adherence in favor of use of the platform. Future studies using a rigorous randomized design are needed to fully evaluate the benefits of this type of reminder system on participant protocol adherence. A current limitation of the CMP system is that the short message queries are time sensitive. If a participant does not respond to a query within a predetermined time period, they will have missed the opportunity to provide their response and cannot respond to the query at a later time. This can be overcome by increasing the overall period of monitoring (eg, from 7 days to 14 days) to ensure sufficient data across an acceptable number of days. Finally, a limitation of using automated messaging systems in general may be the impersonal nature of the communication, which may have the unintended consequence of decreasing positive affiliation among participants to the cohort study. To guard against this, our approach has been to view the CMP as an ancillary tool to other methods we are using to make participation in the study more personal. We continue to reach out to participants via newsletters and one-on-one communication when necessary.

Future Directions

Future uses of the CMP with our particular study include increasing the number of registered participants in the system, continued recurring automated exports from the study database with participant information to the platform, regular birthday and holiday greetings, appointment reminders, integration of other data collection instruments, and promotion of study newsletters and publications. The development team is also working on integrating other modes of communication into the platform, such as building in methods that connect with Facebook and Twitter. Additionally, the team has begun to develop methods to capture data from commercial activity trackers, such as Fitbit, and linking the system with Apple's ResearchKit and Google Fit, which would allow for capturing additional health tracking data.

From an epidemiology study perspective, the CMP has the potential to cast a wider net in the collection of exposure data. Participants, equipped and comfortable with tools for immediate reporting, may selectively attend to previously overlooked or misestimated exposures. Further, the CMP could aid in research assessing physical health factors (eg, weight, physical activity, sleep, and smoking) and mental or emotional variables (eg, depressive symptoms and stress). Participants could self-report on these health outcomes to provide real-time updates and avoid the cost and time of in-person appointments, thereby boosting data quantity and quality and limiting participant burden. Although the CMP is also used for tracking diet and physical activity in obesity interventions [7,9], it could be further utilized to monitor diet and activity in observational longitudinal cohort or clinical studies. In addition, as mentioned above, future iterations will allow for capturing commercial sensor and health tracking data that could be used to enrich understanding of the myriad of factors that play a role in health and wellness.

Conclusions

The CMP can serve as a model for enhancing interactive communication and data collection with longitudinal study cohorts, as it takes advantage of modern technology and keeps pace with participants' changing communication preferences. Ease, accessibility, and points of contact increase with the help of the CMP and it facilitates more opportunities for exposure assessment. Continued implementation research is needed to advance technologies like the CMP, though its potential as a longitudinal research tool is highly promising.

Acknowledgments

Research reported in this publication was supported by the National Institute of Environmental Health Sciences of the National Institutes of Health (NIH) (award number P01ES022831) and by the United States Environmental Protection Agency (USEPA) (grant numbers RD-83523701-0 and CR-83242401-0). Content is solely a reflection of the grantee and does not necessarily represent the official views of the NIH or the USEPA.

Conflicts of Interest

GB holds equity in Coeus Health and serves on the scientific advisory board of Nutrisystem. SK has received research support and/or consulting fees from the following commercial sources in the past 12 months: Akili Interactive, Arbor, Bose, Ironshore, Jazz, KemPharm, Neos, Otsuka, Rhodes, Shire, Sunovion, and Tris. He also holds equity in Behavioral Innovations Group, LLC.

References

1. Measuring the Information: Society Report. Volume 1. Geneva, Switzerland: International Telecommunication Union; 2017. URL: https://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2017/MISR2017_Volume1.pdf [accessed 2018-12-03] [WebCite Cache ID 74OI0gYeU]
2. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. *Epidemiol Rev* 2010;32:56-69 [FREE Full text] [doi: [10.1093/epirev/mxq004](https://doi.org/10.1093/epirev/mxq004)] [Medline: [20354039](https://pubmed.ncbi.nlm.nih.gov/20354039/)]
3. Collins F. Scientific American. 2012 Jul 01. The real promise of mobile health apps URL: <https://www.scientificamerican.com/article/real-promise-mobile-health-apps/> [accessed 2018-12-03] [WebCite Cache ID 74OIkwVqf]
4. Seed M, Juarez M, Alnatour R. Improving recruitment and retention rates in preventive longitudinal research with adolescent mothers. *J Child Adolesc Psychiatr Nurs* 2009 Aug;22(3):150-153. [doi: [10.1111/j.1744-6171.2009.00193.x](https://doi.org/10.1111/j.1744-6171.2009.00193.x)] [Medline: [19702968](https://pubmed.ncbi.nlm.nih.gov/19702968/)]
5. Mitchell S, Schwartz R, Alvanzo A, Weisman M, Kyle T, Turrigiano M, et al. The use of technology in participant tracking and study retention: Lessons learned from a Clinical Trials Network study. *Subst Abuse* 2015;36(4):420-426 [FREE Full text] [doi: [10.1080/08897077.2014.992565](https://doi.org/10.1080/08897077.2014.992565)] [Medline: [25671593](https://pubmed.ncbi.nlm.nih.gov/25671593/)]
6. Nicholson LM, Schwirian PM, Klein EG, Skybo T, Murray-Johnson L, Eneli I, et al. Recruitment and retention strategies in longitudinal clinical studies with low-income populations. *Contemp Clin Trials* 2011 May;32(3):353-362 [FREE Full text] [doi: [10.1016/j.cct.2011.01.007](https://doi.org/10.1016/j.cct.2011.01.007)] [Medline: [21276876](https://pubmed.ncbi.nlm.nih.gov/21276876/)]
7. Foley P, Steinberg D, Levine E, Askew S, Batch B, Puleo EM, et al. Track: A randomized controlled trial of a digital health obesity treatment intervention for medically vulnerable primary care patients. *Contemp Clin Trials* 2016 Dec;48:12-20 [FREE Full text] [doi: [10.1016/j.cct.2016.03.006](https://doi.org/10.1016/j.cct.2016.03.006)] [Medline: [26995281](https://pubmed.ncbi.nlm.nih.gov/26995281/)]
8. Steinberg DM, Levine EL, Lane I, Askew S, Foley PB, Puleo E, et al. Adherence to self-monitoring via interactive voice response technology in an eHealth intervention targeting weight gain prevention among Black women: Randomized controlled trial. *J Med Internet Res* 2014 Apr 29;16(4):e114 [FREE Full text] [doi: [10.2196/jmir.2996](https://doi.org/10.2196/jmir.2996)] [Medline: [24780934](https://pubmed.ncbi.nlm.nih.gov/24780934/)]
9. Steinberg D, Levine E, Askew S, Foley P, Bennett G. Daily text messaging for weight control among racial and ethnic minority women: Randomized controlled pilot study. *J Med Internet Res* 2013 Nov 18;15(11):e244 [FREE Full text] [doi: [10.2196/jmir.2844](https://doi.org/10.2196/jmir.2844)] [Medline: [24246427](https://pubmed.ncbi.nlm.nih.gov/24246427/)]
10. Fuchs JD, Stojanovski K, Vittinghoff E, McMahan VM, Hosek SG, Amico KR, et al. A mobile health strategy to support adherence to antiretroviral preexposure prophylaxis. *AIDS Patient Care STDS* 2018 Mar;32(3):104-111 [FREE Full text] [doi: [10.1089/apc.2017.0255](https://doi.org/10.1089/apc.2017.0255)] [Medline: [29565183](https://pubmed.ncbi.nlm.nih.gov/29565183/)]
11. Funderburk K, Struempfer B, Parmer S, Griffin J. Body Quest Parent: A text message parent education intervention to supplement a school-based obesity prevention initiative for third-graders. *J Nutr Educ Behav* 2018 Apr;50(4):415-417.e1. [doi: [10.1016/j.jneb.2018.01.007](https://doi.org/10.1016/j.jneb.2018.01.007)] [Medline: [29456126](https://pubmed.ncbi.nlm.nih.gov/29456126/)]
12. Bauer V, Goodman N, Lapin B, Cooley C, Wang E, Craig TL, et al. Text messaging to improve disease management in patients with painful diabetic peripheral neuropathy. *Diabetes Educ* 2018 Dec;44(3):237-248. [doi: [10.1177/0145721718767400](https://doi.org/10.1177/0145721718767400)] [Medline: [29589820](https://pubmed.ncbi.nlm.nih.gov/29589820/)]
13. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text messaging and mobile phone apps as interventions to improve adherence in adolescents with chronic health conditions: A systematic review. *JMIR Mhealth Uhealth* 2017 May 15;5(5):e66 [FREE Full text] [doi: [10.2196/mhealth.7798](https://doi.org/10.2196/mhealth.7798)] [Medline: [28506955](https://pubmed.ncbi.nlm.nih.gov/28506955/)]
14. Badawy SM, Kuhns LM. Texting and mobile phone app interventions for improving adherence to preventive behavior in adolescents: A systematic review. *JMIR Mhealth Uhealth* 2017 Apr 19;5(4):e50 [FREE Full text] [doi: [10.2196/mhealth.6837](https://doi.org/10.2196/mhealth.6837)] [Medline: [28428157](https://pubmed.ncbi.nlm.nih.gov/28428157/)]
15. Akl E, Briel M, You J, Lamontagne F, Gangji A, Cukierman-Yaffe T, et al. LOST to follow-up Information in Trials (LOST-IT): A protocol on the potential impact. *Trials* 2009 Jun 11;10:40 [FREE Full text] [doi: [10.1186/1745-6215-10-40](https://doi.org/10.1186/1745-6215-10-40)] [Medline: [19519891](https://pubmed.ncbi.nlm.nih.gov/19519891/)]
16. Ioannidis J, Bassett R, Hughes M, Volberding P, Sacks H, Lau J. Predictors and impact of patients lost to follow-up in a long-term randomized trial of immediate versus deferred antiretroviral treatment. *J Acquir Immune Defic Syndr Hum Retrovirol* 1997 Sep 01;16(1):22-30. [Medline: [9377121](https://pubmed.ncbi.nlm.nih.gov/9377121/)]
17. Hansen W, Collins L. Seven ways to increase power without increasing N. *NIDA Res Monogr* 1994;142:184-195. [Medline: [9243537](https://pubmed.ncbi.nlm.nih.gov/9243537/)]
18. Gahleitner F, Legg J, Holland E, Pearson S, Roberts G. The validity and acceptability of a text-based monitoring system for pediatric asthma studies. *Pediatr Pulmonol* 2016 Jan;51(1):5-12. [doi: [10.1002/ppul.23200](https://doi.org/10.1002/ppul.23200)] [Medline: [25847724](https://pubmed.ncbi.nlm.nih.gov/25847724/)]
19. Patnaik S, Brunskill E, Thies W. Evaluating the accuracy of data collection on mobile phones: A study of forms, SMS, and voice. In: Proceedings of the International Conference on Information and Communication Technologies and Development. Washington, DC: IEEE; 2009 Apr Presented at: International Conference on Information and Communication Technologies and Development; April 17-19, 2009; Doha, Qatar. [doi: [10.1109/ICTD.2009.5426700](https://doi.org/10.1109/ICTD.2009.5426700)]

Abbreviations

CMP: cohort management platform

IVR: interactive voice response
NEST: Newborn Epigenetic Study
NIH: National Institutes of Health
SMS: short message service
USEPA: United States Environmental Protection Agency

Edited by H Wu; submitted 23.07.18; peer-reviewed by W Tian, L Huang, S Badawy; comments to author 27.09.18; revised version received 05.12.18; accepted 14.12.18; published 02.04.19.

Please cite as:

*Barry CM, Sabhlok A, Saba VC, Majors AD, Schechter JC, Levine EL, Streicher M, Bennett GG, Kollins SH, Fuemmeler BF
An Automated Text-Messaging Platform for Enhanced Retention and Data Collection in a Longitudinal Birth Cohort: Cohort
Management Platform Analysis*

JMIR Public Health Surveill 2019;5(2):e11666

URL: <https://publichealth.jmir.org/2019/2/e11666/>

doi: [10.2196/11666](https://doi.org/10.2196/11666)

PMID: [30938689](https://pubmed.ncbi.nlm.nih.gov/30938689/)

©Caroline M Barry, Aditi Sabhlok, Victoria C Saba, Alesha D Majors, Julia C Schechter, Erica L Levine, Martin Streicher, Gary G Bennett, Scott H Kollins, Bernard F Fuemmeler. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 02.04.2019. 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 Software Tool Aimed at Automating the Generation, Distribution, and Assessment of Social Media Messages for Health Promotion and Education Research

Katja Reuter^{1,2}, PhD; Alicia MacLennan², MS; NamQuyen Le², MPH; Jennifer B Unger¹, PhD; Elsi M Kaiser³, PhD; Praveen Angyan², MS

¹Department of Preventive Medicine, Keck School of Medicine of University of Southern California, Institute for Health Promotion & Disease Prevention Research, University of Southern California, Los Angeles, CA, United States

²Southern California Clinical and Translational Science Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

³Linguistics Department, Psycholinguistics Lab, University of Southern California, Los Angeles, CA, United States

Corresponding Author:

Katja Reuter, PhD

Department of Preventive Medicine, Keck School of Medicine of University of Southern California

Institute for Health Promotion & Disease Prevention Research

University of Southern California

3rd Floor, MC 9239

2001 N Soto Street

Los Angeles, CA, 90032

United States

Phone: 1 3234422046

Email: katja.reuter@usc.edu

Abstract

Background: Social media offers promise for communicating the risks and health effects of harmful products and behaviors to larger and hard-to-reach segments of the population. Nearly 70% of US adults use some social media. However, rigorous research across different social media is vital to establish successful evidence-based health communication strategies that meet the requirements of the evolving digital landscape and the needs of diverse populations.

Objective: The aim of this study was to expand and test a software tool (Trial Promoter) to support health promotion and education research by automating aspects of the generation, distribution, and assessment of large numbers of social media health messages and user comments.

Methods: The tool supports 6 functions (1) data import, (2) message generation deploying randomization techniques, (3) message distribution, (4) import and analysis of message comments, (5) collection and display of message performance data, and (6) reporting based on a predetermined data dictionary. The tool was built using 3 open-source software products: PostgreSQL, Ruby on Rails, and Semantic UI. To test the tool's utility and reliability, we developed parameterized message templates (N=102) based upon 2 government-sponsored health education campaigns, extracted images from these campaigns and a free stock photo platform (N=315), and topic-related hashtags (N=4) from Twitter. We conducted a functional correctness analysis of the generated social media messages to assess the algorithm's ability to produce the expected output for each input. We defined 100% correctness as use of the message template text and substitution of 3 message parameters (ie, image, hashtag, and destination URL) without any error. The percent correct was calculated to determine the probability with which the tool generates accurate messages.

Results: The tool generated, distributed, and assessed 1275 social media health messages over 85 days (April 19 to July 12, 2017). It correctly used the message template text and substituted the message parameters 100% (1275/1275) of the time as verified by human reviewers and a custom algorithm using text search and attribute-matching techniques.

Conclusions: A software tool can effectively support the generation, distribution, and assessment of hundreds of health promotion messages and user comments across different social media with the highest degree of functional correctness and minimal human interaction. The tool has the potential to support social media-enabled health promotion research and practice: first, by enabling the assessment of large numbers of messages to develop evidence-based health communication, and second, by providing public

health organizations with a tool to increase their output of health education messages and manage user comments. We call on readers to use and develop the tool and to contribute to evidence-based communication methods in the digital age.

(*JMIR Public Health Surveill* 2019;5(2):e11263) doi:[10.2196/11263](https://doi.org/10.2196/11263)

KEYWORDS

algorithm; automation; digital; Facebook; health communication; health promotion; Instagram; internet; online; smoking; social network; social media; tobacco; Twitter

Introduction

The use of social media (ie, social networks or social networking sites) as a health promotion and intervention tool provides new opportunities and challenges for both investigators and practitioners [1-6]. Social media includes widely accessible Web-based and mobile technologies that allow users to view, create, and share information online and to participate in social networking [7]. On the basis of previous research, these tools offer promise for communicating with larger and hard-to-reach segments of the population and for purposes as diverse as the provision of health information, delivery of behavior change interventions, disease monitoring and self-management, awareness raising, and advocacy [8-11]. Research in other fields showed that social media messaging could have a significant impact on user attitudes and behavior. In 2017, we saw experimental evidence that these tools can be deployed as engines for social manipulation and to influence voting in elections [8,9]. However, there is limited evidence as to whether social media can support the delivery of targeted and personalized behavior change interventions to improve health [10], partly owing to the challenges of implementing large-scale social media communication experiments.

Today, nearly 70% of US adults use some social media [11]. Among the most popular platforms are Facebook, YouTube, Pinterest, Instagram, Twitter, LinkedIn, and Snapchat [12]. Their user base varies by demographic characteristics such as age, gender, and race and ethnicity [11,13]. The success of digital health communication efforts might, therefore, not only depend on the type of content (eg, text, image, audio, and video) but also on variables such as the type of social media platform, organic messages versus paid (advertisements), the message date and time, and users' social networks. However, researchers and public health agencies typically lack the resources and expertise to disseminate and test the effectiveness of larger numbers of health messages on social media, and the majority of current interventions are neither evidence-based nor widely adopted [14]. Rigorous research across different social media types will be required to establish successful evidence-based health communication strategies that meet the requirements of the evolving digital landscape and the needs of diverse and vulnerable populations.

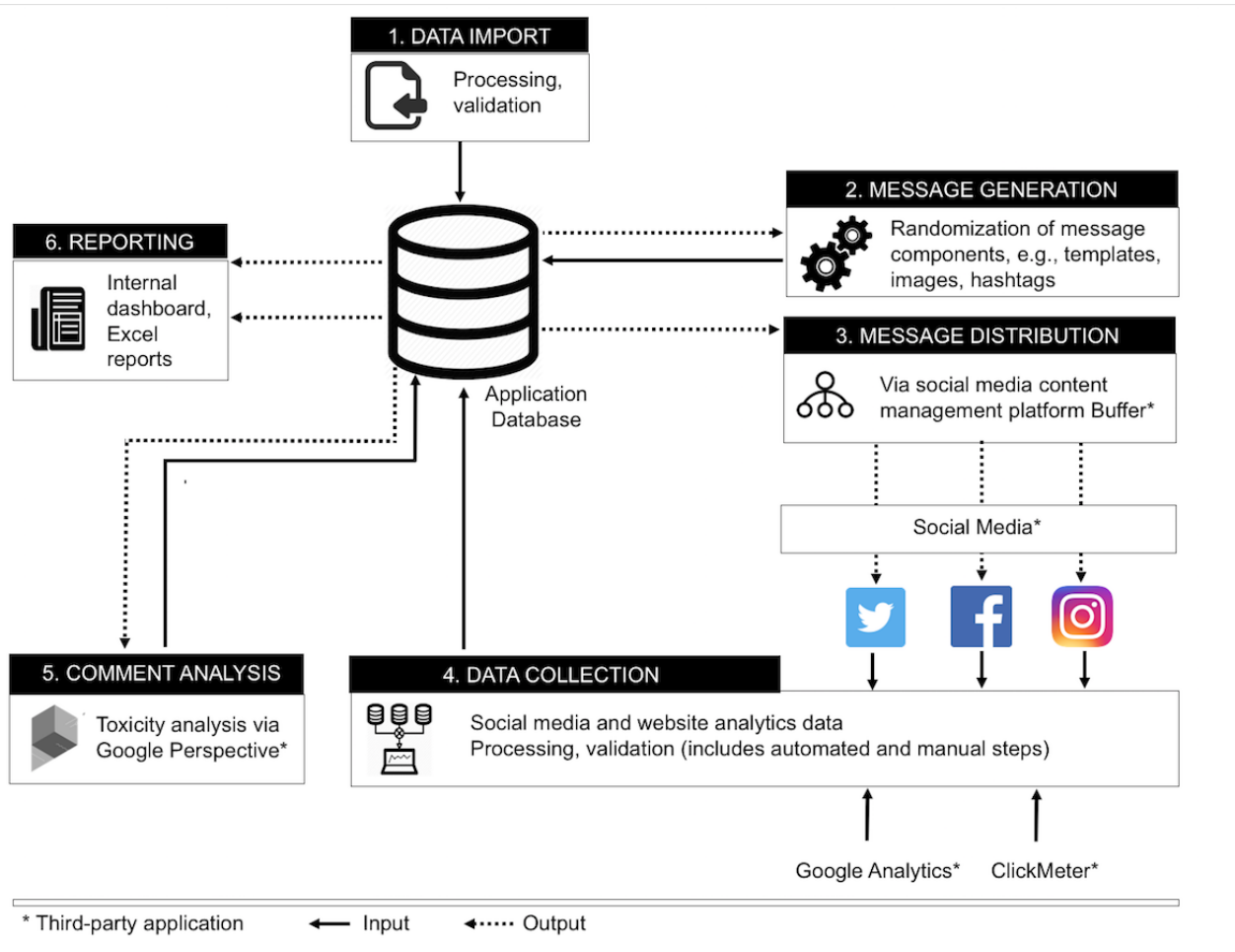
The objective of this study was to expand and test a software tool (Trial Promoter) to support health promotion and intervention research by automating aspects of the generation, distribution, and assessment of large numbers of health messages and user comments across different social media. In this study, we have described the tool including the features that support rigorous scientific study design such as randomization and the use of a data dictionary. The tool builds on our previous study where we demonstrated that a software tool has the ability to support increased output of research information on Twitter while reducing the burden of developing and distributing hundreds of individual messages [15] and that such an automated approach provides a cost-effective solution to distribute clinical trial information more efficiently [16].

Methods

Overview Description of the Technical Framework and Dataflow

The software tool, Trial Promoter, supports 6 functions (Figure 1): (1) data import (eg, parameterized message templates and images), (2) message generation deploying randomization techniques to reduce selection bias for message templates and characteristics (eg, images and hashtags), (3) message distribution across social media (the current version supports Twitter, Facebook, and Instagram), (4) collection of message and website performance data, (5) import of message comments and their toxicity score (ie, probability between 0 to 1, with higher values indicating a greater likelihood of offensive, disrespectful language), and (6) display of message performance data in the internal dashboard and output-reporting based on the predetermined data dictionary.

The tool is built using 3 open-source software products: the PostgreSQL object-relational database (version 9.3) [17], the *Ruby on Rails* Web framework (version 4.2.6) [18], and the Semantic UI frontend framework (version 2.2.11) [19]. Semantic UI supports the creation of dashboards and front-end interfaces. We further used the following infrastructure: the cloud application platform, *Heroku*, to deploy the tool quickly without the need to set up servers or install software [20] and *Amazon Web Services* to store image assets [21].

Figure 1. Application setup and data flow diagram.

Supported Data Sources, Formats, and Types

The tool is capable of importing information from different types of data sources, that is, Representational State Transfer (REST) Application Programming Interface (API), and data files. Data formats include JavaScript Object Notation (JSON), Comma Separated Values (CSV), Excel, Portable Network Graphics, and Joint Photographic Experts Group image formats. Social media messages consist of different elements such as message text, URL, hashtags, images, and videos. The current version of the tool can import all of these data types except video files.

Data Import, Processing, and Validation

The tool offers a standard template for importing data such as message templates, social media platforms to be used, hashtags, destination URLs, and experiment-specific variables such as disease terms or linguistic factors. During import, the data included in the import file are processed in 3 ways: (1) to associate imported images with message templates for randomization, (2) to associate experiment variables (eg, disease terms and linguistic factors) with a message template, and (3) to add message template parameters for the destination URL. The tool performs data validation after the import is complete. The study team can view any validation errors detected by the tool and fix message templates or images as needed. At present, the tool carries out 2 main validation steps during the setup of

the experiment: (1) it verifies that the length of the message templates for the Twitter platform fits within platform limits (ie, 280 character limit) while taking into account the length of destination URLs, image URLs, and inclusion of hashtags [22] and (2) it checks the resolution of images to meet the requirements on the social media platform, Instagram (Multimedia Appendix 1) [23]. For example, if the length of a message template does not allow for the inclusion of all hashtags from a preexisting set, the tool flags the message template as not suitable so that the study team can make the necessary adjustments.

Setting Up Experiment Parameters

The tool supports scientific study design methods by providing options for experiment customization (Multimedia Appendix 2). An experiment is defined as a set of parameters that are used to characterize a specific health communications study, for example, to examine and compare the influence of different linguistics methods used in the messages (eg, perspective taking, information packaging, and numeracy). The present version of the tool provides the following parameters for customizing an experiment, that is, name of experiment, start date, social media platforms to be used, medium (advertised messages/advertisements vs organic, nonpaid messages), image inclusion, message repetition (ie, the number of times a message should be sent), the number of messages per day, social media

accounts to be used, time schedule for message distribution, and tracking when a user clicks on a message link.

Randomization

To reduce bias in the distribution of message characteristics, the tool randomizes elements such as message templates, images, and hashtags. The message templates were shuffled into a random sequence using a Fisher-Yates shuffle [24], and the selection of hashtags and the images were randomly sampled [25].

Message Generation

Through automatic substitution of 3 message parameters (ie, images, hashtags, and destination URLs) in the message

templates, the tool generates the final messages for each social media platform (ie, Twitter, Facebook, and Instagram; Figure 2). Filled-in parameters include the destination URL to the respective Web page (ie, landing page) and a randomly chosen hashtag from a preexisting set for those messages that do not already include a hashtag (eg, #tobacco and #smoking). The message URL is tagged with Urchin Traffic Monitor parameters to track the engagement with the message on social media, that is, to track clicks on the URL that takes users to the landing page. The tool uses the REST API provided by the third-party service, Clickmeter, to generate the shortened URL. The generation of the final messages is locked once the distribution of the messages begins to prevent inadvertent changes to the messages or their deletion.

Figure 2. Screenshot of interface with the final messages the application generated for the correctness analysis described in this paper. Images shown here are samples similar to the original copyright protected campaign images and courtesy of Pixabay.com.

The screenshot displays the 'Generated messages' tab of the application. At the top, there are navigation tabs: Setup, Images, Message templates, **Generated messages**, Dashboard, Analytics file uploads, and Message comments. Below these are two main buttons: 'Generate messages' and 'Send to Buffer'. A 'Note' section contains the following text:

- Randomization: Trial Promoter uses two methods for randomization. The messages are shuffled into a random sequence using a Fisher-Yates shuffle [1]. The selection of hashtags (if applicable) and the images (if applicable) are randomized using a sampling method [2].
- Message generation: When you click on "Generate messages", all existing messages for this experiment will be deleted and then regenerated.
- Editing restriction: Set up, message templates, and generation will be locked 3 days before the experiment is scheduled to start.
- Send to Buffer: Trial Promoter sends messages to Buffer nightly (up to a week in advance). Please click on "Send to Buffer" to immediately send the messages scheduled for the next 7 days to Buffer.

Footnotes: ¹ Durstenfeld, R. "Algorithm 235: Random Permutation." Communications of the ACM. ACM, 1 July 1964. Web. 29 Mar. 2017. ² "Class: Array (Ruby 2.2.0)." Ruby Language Documentation, n.d. Web. 29 Mar. 2017.

Below the note is a pagination bar with buttons: « First, « Prev, ..., 16, 17, 18, 19, 20, 21, 22, 23, 24, ..., Next », Last ». A date selector shows 'Aug 10th' and a status bar indicates 'Displaying messages 1711 - 1800 of 3570 in total'.

The main content area shows three message cards for Instagram:

- Message 1:** On average, every cig reduces your life by 11 minutes. Even occasional #smoking can hurt you. <http://go.befreeoftobacco.org/q08>. Time: 09:00 AM. Type: Organic. Hashtag: Be Free of Tobacco. Database ID: 12212. Lin_meth_factor: 2. Lin_meth_level: 4. Includes Tracking URL and Destination URL fields.
- Message 2:** On average, every cig reduces your life by 11 minutes. Even occasional #smoking can hurt you. <http://go.befreeoftobacco.org/q07>. Time: 09:00 AM. Type: Ad. Hashtag: B Free of Tobacco. Database ID: 12211. Includes 'Add campaign id' and 'Save' buttons. Lin_meth_factor: 2. Lin_meth_level: 4. Includes Tracking URL and Destination URL fields.
- Message 3:** Even occasional #smoking can hurt you. Every #cigarette you smoke reduces your life by almost 20% of an hour. <http://go.befreeoftobacco.org/q1d>. Time: 01:00 PM. Type: Ad. Hashtag: B Free of Tobacco. Database ID: 12216. Includes 'Add campaign id' and 'Save' buttons. Lin_meth_factor: 3. Lin_meth_level: 7. Includes Tracking URL and Destination URL fields.

Message Distribution

The tool schedules and distributes the messages through the project-related social media accounts using the third-party application, *Buffer*, a social media content management Web application [26]. Each social media account that is set up in *Buffer* has a unique profile identifier assigned to it. The tool allows each experiment to specify which social media accounts are to be used. *Buffer* provides a REST API call that allows for queuing of messages directly in *Buffer*. *Buffer* then sends the

messages to multiple social media platforms at the scheduled times that were entered during the experiment setup.

Data Collection and Processing

Analytics for each message to determine the engagement among social media users with the message and on the referred to Web page were collected using a number of applications that are summarized in [Multimedia Appendix 3](#). Message comments need to be collected manually by logging into the respective social media account.

Reporting

The application provides 2 methods of reporting: (1) project-internal dashboards accessible via login and (2) reports for statistical analysis. For each experiment, the app supports the creation of a data dictionary, which centralizes the information about the data to be collected using experiment-specific data definitions (eg, value names, meaning, origin, and format) to generate customized and comprehensive

reports. See [Multimedia Appendix 4](#) for an example of a data dictionary.

In the internal dashboard, the application provides 3 types of data visualization: (1) key performance data (eg, clicks, impressions, and click rate) for each individual message by social media type ([Figure 3](#)), (2) messages with the highest click rate (number of clicks divided by number of impressions; [Figure 4](#)), and (3) comments received in response to the messages and their toxicity score ([Figure 5](#)).

Figure 3. Screenshot shows dashboard interface where the application displays key performance data for each individual message by social media type. Images shown here are samples similar to the original copyright protected campaign images. Images are courtesy of Apomares (top), Monkey Business Images (middle), Rawpixel at FreeDigitalPhotos.net (bottom).

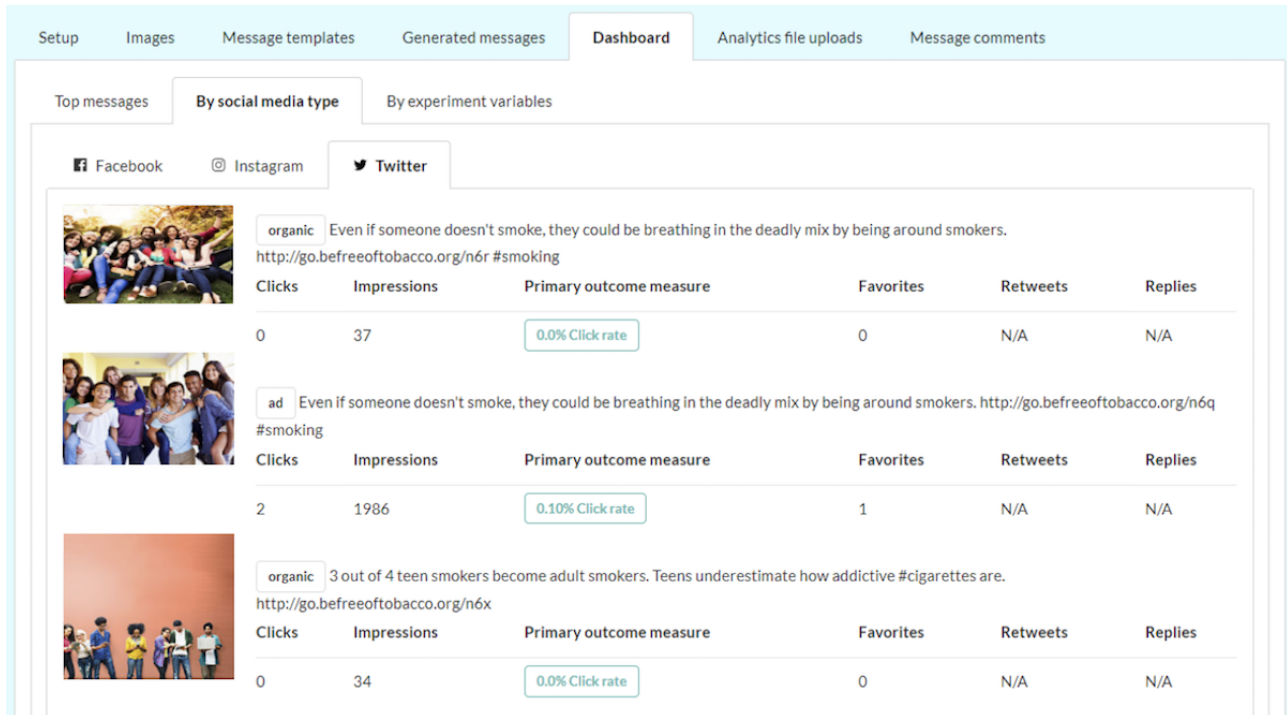


Figure 4. Screenshot shows dashboard interface where the application displays messages with the highest click rate. Images shown here are samples similar to the original copyright protected campaign images. Images are courtesy of Absolut Images (left), Pixelfit (middle), Aleksandar Georgiev at FreeDigitalPhotos.net (bottom).

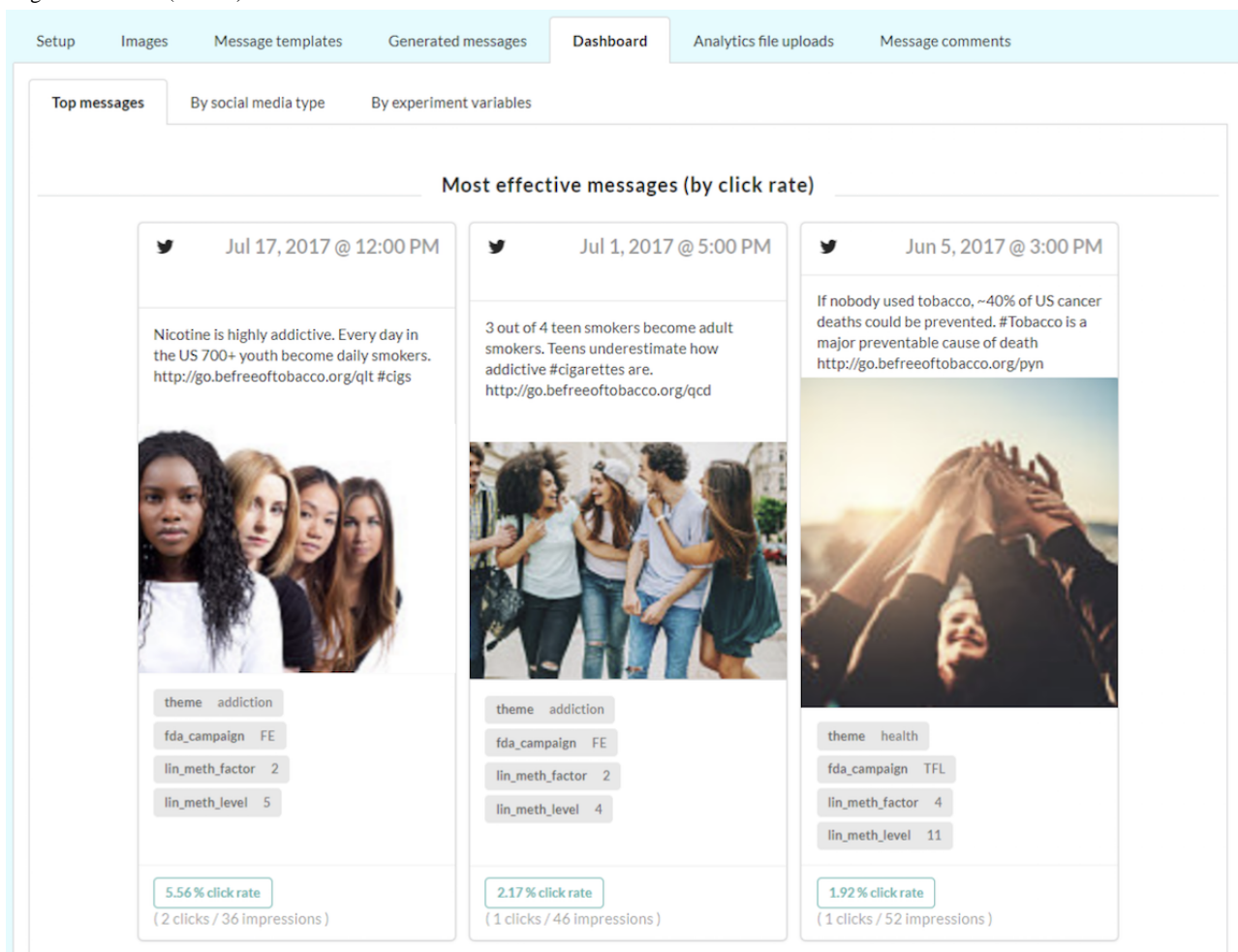
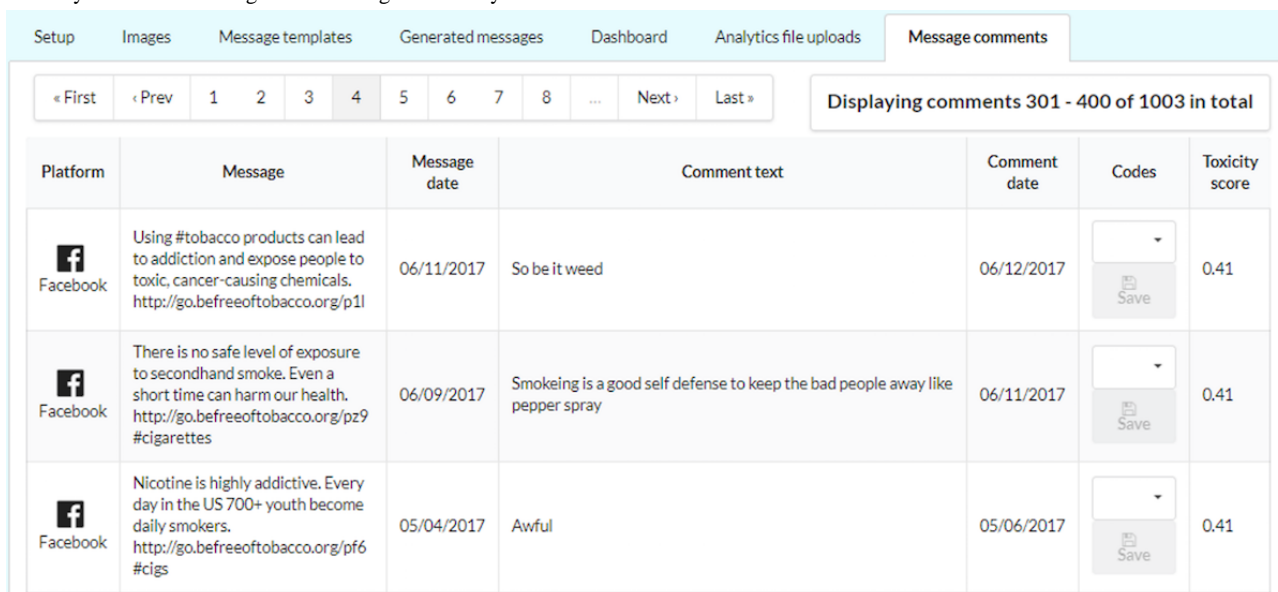


Figure 5. Screenshot shows dashboard interface where the application displays the comments received in response to the messages on Facebook and their toxicity score. The messages with the highest toxicity scores are listed first.



Comment Analysis

The tool has the capability to import data (ie, comments) into the machine learning tool, *Perspective*, developed by *Jigsaw*

under the umbrella of *Google* 's parent company, *Alphabet* [27,28]. *Perspective* then calculates the toxicity score for each imported comment and determines the probability of a comment being labeled by human moderators as *toxic*. Higher values of

a toxicity score between 0 to 1 indicate a greater likelihood of offensive, disrespectful language that could negatively impact an online conversation. The machine learning model used by *Perspective* is based on crowdsourced annotations of randomly sampled comments from the body of 63 million comments from the English Wikipedia [29]. Human annotators were given a scale for how likely an online participant would leave a conversation owing to the perceived abuse (very toxic, toxic, neither, healthy contribution, and very healthy contribution) [30]. The toxicity scores generated by *Perspective* are then returned in the JSON format and parsed, and the toxicity score is stored for each comment within the application.

Finally, the application provides reports, for example, for statistical analysis, in CSV format. The data reports are customizable and include the data specified for a particular experiment (ie, no filtering is applied to the data). The reports are tailored to reflect the data dictionary that was specified for the experiment and can be generated on the back-end of the application using a Rails console, an application that allows programmers to interact with the system from a command line interface, directly issuing commands that are interpreted and executed by the system.

Correctness Analysis

To test the tool's utility and reliability, we developed parameterized message templates (N=102) based upon 2 government-sponsored online tobacco education campaigns and extracted images from these campaigns and the free stock photo platform, Stocksnap (N=315), and extracted topic-related hashtags (N=4) from Twitter. The daily message volume per social media platform was 6 on Facebook (advertisements and organic), 6 on Twitter (advertisements and organic), and 3 on Instagram (advertisements only, owing to the fact that Instagram does not support referral URLs in organic, nonpaid messages). Advertisements and organic messages were sent to separate accounts during this experiment. The daily message volume can be customized manually during the experiment setup. It is not limited but it is recommended to stay within the social media platform-specific limit to avoid flagging or shutdown of an account, for example, the limit on Twitter is 1000 direct

messages per day and 2400 tweets per day [31]. We further recommend taking into account market research data that suggest the optimal number of posts per day, for example, on Twitter, it is 3 [32,33]. During this experiment, the tool sent 3 messages per social media account at different times. The length of the pilot project (85 days) was determined by the available budget for social media advertisements. On the basis of market research showing that messages sent at these times receive the most user engagement [32-34], the tool sent messages on Facebook at 9 am, 1 pm, and 3 pm PST; on Twitter at noon, 3 pm, and 5 pm PST; and on Instagram at 8 am, 9 am, and 5 pm PST.

We conducted a functional correctness [35] analysis of the automatically generated social media messages that were distributed across the 3 social media, Twitter, Facebook, and Instagram.

We assessed the algorithm's ability to produce the expected output for each input and defined 100% correctness as the correct use of the message template text and correct substitution of 3 message parameters (ie, image, hashtag, and destination URL). For example, an error constitutes a missing image, a missing or misspelled and therefore nonfunctional URL, or a missing or misspelled hashtag. The percent that was correct was calculated to determine the probability with which the tool generates accurate messages.

Results

Evaluation

During the 85-day experiment between April 19 and July 12, 2017, the tool successfully generated and distributed a total of 1275 messages (Twitter: N=510; Facebook: N=510; and Instagram: N=255). Figure 6 shows examples of automatically generated and distributed messages that were part of the correctness analysis described here. The software code of the application is accessible under the MIT license on GitHub [36]. The detailed analysis of the messages and influence of several variables on user attention and engagement will be discussed in a forthcoming paper.

Figure 6. Examples of automatically generated messages that were distributed by the application across Twitter, Facebook, and Instagram. Images shown here are samples similar to the original copyright protected campaign images. Images are courtesy of Pixabay.com (top), Anagoria at Wikimedia.org (middle), Mary Bates at Wikimedia.org(bottom).



Correctness Analysis

The correctness with which the application generated the social media messages during the experiment was evaluated using 3 factors for each individual message: (1) the image was randomly selected and included in the message, (2) the hashtag was randomly selected and included in the message if it did not already contain a hashtag, and (3) the URL parameter was

replaced correctly. During the experiment, the application correctly used the message template text and substituted the message parameters 100% of the time as verified by both human reviewers and a custom algorithm using text search and attribute-matching techniques (Multimedia Appendix 5). The software code used to determine the correctness of the generated messages can be found on GitHub [36].

Discussion

Principal Findings

Our findings show that a software tool can support health promotion and education research by automating aspects of the generation, distribution, and assessment of hundreds of health promotion messages and user comments across different social media types with the highest degree of functional correctness and minimal human interaction. The detailed analysis of the messages and influence of several variables on user attention and engagement will be discussed in a forthcoming paper. We chose the 3 social media types, Twitter, Facebook, and Instagram, for the experiment because they were among the most popular social media platforms used by people living in the United States at the time of the experiment [11-13]. The software code is available on GitHub for free [36]. We invite readers and developers to use and develop the tool and to contribute to the development of evidence-based health promotion and interventions for social media.

The tool that we have presented here has the potential to support research teams and public health organizations. Research teams can use the tool to generate, manage, and test larger numbers of public health messages. The tool also contributes to standardizing social media research methods through 2 features: (1) it consistently applies randomization techniques to reduce selection bias (eg, message templates and images) and (2) the tool employs a data dictionary to contribute to more consistent reporting standards for social media research metrics (including clearly defined metrics and calculations such as click rate).

By surfacing toxic comments that may include offensive and disrespectful language and could negatively impact an online health conversation, the tool also provides support for public health organizations that need to manage digital public health campaigns about controversial topics such as smoking and related regulations and policies. This became evident when an antismoking regulation campaign by the Chicago Department of Public Health resulted in significant backlash by Twitter users, sending more than 600 tweets in 1 week against the proposed regulation [37]. The tool that we have described here may support public health organizations charged with the implementation of controversial health campaigns that may require monitoring and moderation of larger volumes of comments.

Limitations of the Study and Tool

Here, we present the findings of a correctness analysis that was focused on assessing the probability with which a software tool generates and distributes correct health messages across different social media and collects message performance data and user comments. Trial Promoter focuses on social media-based campaigns. The distributed messages (organic messages and paid advertisements) would, therefore, not be viewed by individuals who do not use social media but may still be part of the targeted population of interest.

Social media also provides a method for reaching specific populations based on their characteristics (eg, age, gender, location, language, and interests). These targeting capabilities

are usually built into the social media platforms and are based upon proprietary user data that are not available to the study team. Using this version of the tool, users need to set up the targeting on the social media platform, whereas Trial Promoter integrates with the social media platform to distribute the message content.

The current version of the code only supports integration with Twitter, Facebook, and Instagram, which were selected owing to their popularity in the United States. Researchers and health organizations in other parts of the world may also want to consider integrating other social media platforms with Trial Promoter.

Research teams might also require additional functionality to answer their specific research questions, for example, different types of randomizations, other social media platforms, ability to incorporate social media monitoring data, and mentions of social media influencers—all of which would need to be developed as extensions to the current version of the application.

Furthermore, additional features could be added in the future to enhance the application such as taking into account disease and health topic awareness months, trending topics and hashtags on social media, which may affect social media user attention and engagement, as well as the automatic blocking of social media users who contribute toxic comments, or automated debiasing of social media datasets using software programs such as BotOrNot [38].

Finally, future research will need to examine in more detail the effectiveness of social media-driven health promotion efforts to communicate risks and effects of harmful products and behaviors to promote healthy lifestyles and behaviors.

Ethical and Data Privacy Considerations

As social media is designed to foster social interactions, health promotion and intervention campaigns on social media can lead to user comments that may include identifiable or personal health information, which poses privacy issues, safety risks, and dignitary violations [39]. Trial Promoter has the ability to display user comments received in response to a campaign and surface those comments that may require moderation owing to their toxicity (ie, offensive, disrespectful language that could negatively impact an online conversation). However, the tool does not directly moderate and, for example, delete specific comments or ban disruptive users on social media platforms. It is possible to disable comments on some social media, such as Facebook, as a measure to mitigate such risks. However, some social media, such as Twitter, do not allow the deletion of replies to a message. In this case, we suggest the use of disclaimer messages as suggested by Bender et al [16], for example, “Social media is not secure. Please don’t post if you are concerned about your privacy”.

In addition, the current version of the tool does not identify and moderate comments that may be nontoxic but still raise critical privacy and safety issues and may require a response. For example, a social media user may comment on a depression advertisement about their current and imminent suicidality or may leave a comment disclosing current child abuse on one of these advertisements. It is worthwhile to note that Facebook

launched a suicide alert reporting system so that Facebook users can report individuals who they believe are expressing suicidal thoughts or intent [40,41]. To manage user comments using the current version of the tool, we suggest having a moderator who monitors user comments daily and manages them on a case-by-case basis.

Furthermore, Trial Promoter uses a number of third-party applications to support specific tasks, 2 of which may raise privacy concerns that researchers should be aware of. First, for the comment analysis, Trial Promoter shares the user comments (not the username or other information) with the third-party application, *Perspective*, developed by Jigsaw under the umbrella of *Google*'s parent company, *Alphabet*. However, user comments are considered identifiable information. A recent study found that online searches of verbatim Twitter quotes found in journal articles can be traced back to individual users 84% of the time [42]. It is not possible to delete the comments after the analysis as the *Perspective* application falls under *Google*'s privacy policy [43]: "The rights you grant in this license are for the limited purpose of operating, promoting, and improving our Services, and to develop new ones." Although *Google* states that it does not share the content uploaded with third parties, which also limits potential conflict of interest, *Google* can use the comments submitted to the API to improve their machine learning model used for the analysis of the comments. We believe that as *Google* uses the data merely to improve its *Perspective* app and the dataset is not available publicly, the use of this third-party application is within the ethical and regulatory guidelines to protect users' privacy. Second, our local version of the Trial Promoter tool is hosted by the cloud-based hosting provider, Heroku, a Salesforce application. Salesforce has passed security and privacy-related audits and certifications including the EU-US Privacy Shield Framework and TRUSTe Certification [15]. Any group or institution that decides to host Trial Promoter will have to ensure the privacy and security of their preferred hosting platform.

Finally, as we expand Trial Promoter, our team intends to address the current limitations of the tool as well as ethical issues such as privacy concerns, variations in protection across different platforms, and expectations of end users and other stakeholders by incorporating the *Hippocratic Oath* for technology, that is, a greater focus on the ethics of technology design [16]. We welcome collaborators.

Comparison With Previous Work

To ensure the effective use of social media in research and to propose and assess evidence-based public health programs for social media, previous work has emphasized the need for flexible technical applications including Web-based data-gathering techniques that are readily available to research teams as well as consistent and transparent frameworks for data collection, quality assessment, debiasing techniques, and systematic reporting standards—most of which are currently lacking [15,41-44]. In addition, automated content generation and distribution for online use—in particular on social media—offers new possibilities for research and public health communities and could benefit the development and implementation of public health promotion efforts. Advanced

applications, so-called *bots*, could generate and distribute information and, in some cases, interact with messages. They are *regarded an influential but also somewhat mysterious factor in public discourse and opinion making* [45]. However, a tool that supports research efforts in this field does not exist as yet. Previous research showed that a Twitter bot sharing public health information was perceived as *credible, attractive, and competent* [46]. These data suggest that bots could potentially be utilized by research and public health organizations. Additional work demonstrated that automatically generated content by a software application is perceived as descriptive and boring but also considered to be objective and not necessarily discernible from content written by journalists" [43]. Bots have been studied in a variety of contexts (eg, prosmoking and protobacco campaigns [15,44], activism or advocacy [47], social networks and human communication decisions [48,49], social shaping [50], content pollution [51], social metric gaming [52], ranking manipulation [53], infiltration [54], political astroturfing [55], recommendation [56], scholarship dissemination [57], and journalism [58]). However, there are little data on bot-like applications that would benefit health promotion research and the development of health communication interventions.

We do not suggest that the tool described here is a bot because the content that makes up the messages needs to be selected and imported by a human actor and the application does not mimic humans and/or human behavior [59], that is, it does not act as an automated social actor similar to how humans might act in social spaces [49,51]. That said, the authors are not aware of similar research that has developed and tested a tool for automatic postings of public health messages on social media to enable better health promotion and intervention research in the digital age. Further studies on automatically generated social media content will help to better understand its role in supporting the public health agenda and health promotion research.

Conclusions

The tool (Trial Promoter) that we have presented here has the potential to influence social media-enabled health promotion and intervention research and practice. First, it enables the assessment of large numbers of messages to develop evidence-based communication approaches for social media. This is especially important as the use of social media among US adults varies by demographic characteristics such as age, gender, and race and ethnicity, and across social media [11-13]. Thus, the success of digital health communication efforts might not only depend on the type of content (eg, text, image, audio, and video) but also on other variables such as social media type, organic versus paid (advertisements) medium, the message date and time, and user's social networks. The tool presented here offers a way of assessing the influence of these variables on the effectiveness of social media-based health promotion and intervention efforts. Second, the tool can be used by public health organizations to increase their output of health education messages, for example, to potentially counteract the growing prevalence of online marketing featuring products and behaviors harmful to health, for example, tobacco products and drugs [57-59]. Finally, the tool also assists with identifying and

moderating larger volumes of user comments to the distributed messages. The tool surfaces those comments that may include offensive, disrespectful language and could negatively impact an online conversation. We call on readers and developers to

use and further develop the software code and to contribute to the development of evidence-based health communication approaches in the digital age.

Acknowledgments

This work was supported by the National Institutes of Health (NIH) National Cancer Institute (NCI), the Food and Drug Administration (FDA) Center for Tobacco Products, and the Southern California Clinical and Translational Science Institute through grant UL1TR000130 from the National Center for Advancing Translational Sciences of the NIH. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NCI, FDA, and NIH.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Example of data validation. The screenshot shows results for a resolution check of images to ensure they meet the requirements of the social media platform Instagram. The row highlighted in red indicates that the image does not meet the requirements for square pictures of a minimum of 600x600 pixels. Images shown here are samples similar to the original copyright protected campaign images. Images are courtesy of Dragqueen (top), Edwin Ortiz (middle), Alex Proimos (bottom) at commons.wikimedia.org.

[[PNG File, 125KB - publichealth_v5i2e11263_app1.png](#)]

Multimedia Appendix 2

Screenshot of experiment parameter setup form.

[[PNG File, 242KB - publichealth_v5i2e11263_app2.png](#)]

Multimedia Appendix 3

Third-party applications used for collecting key performance data.

[[PNG File, 411KB - publichealth_v5i2e11263_app3.png](#)]

Multimedia Appendix 4

Example of a data dictionary used by the application.

[[XLSX File \(Microsoft Excel File\), 16KB - publichealth_v5i2e11263_app4.xlsx](#)]

Multimedia Appendix 5

Screenshot of interface that shows the results of the automated correctness analysis performed by the application.

[[PNG File, 198KB - publichealth_v5i2e11263_app5.png](#)]

References

1. Nadarzynski T, Burton J, Henderson K, Zimmerman D, Hill O, Graham C. Targeted advertisement of chlamydia screening on social media: a mixed-methods analysis. *Digit Health* 2019;5:2055207619827193 [[FREE Full text](#)] [doi: [10.1177/2055207619827193](https://doi.org/10.1177/2055207619827193)] [Medline: [30746155](https://pubmed.ncbi.nlm.nih.gov/30746155/)]
2. Jha A, Lin L, Savoia E. The use of social media by state health departments in the US: analyzing health communication through Facebook. *J Community Health* 2016 Feb;41(1):174-179. [doi: [10.1007/s10900-015-0083-4](https://doi.org/10.1007/s10900-015-0083-4)] [Medline: [26318742](https://pubmed.ncbi.nlm.nih.gov/26318742/)]
3. Philip RK, Shapiro M, Paterson P, Glismann S, van Damme P. Is it time for vaccination to "Go Viral"? *Pediatr Infect Dis J* 2016 Dec;35(12):1343-1349. [doi: [10.1097/INF.0000000000001321](https://doi.org/10.1097/INF.0000000000001321)] [Medline: [27626913](https://pubmed.ncbi.nlm.nih.gov/27626913/)]
4. 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/)]
5. Graham JE, Moore JL, Bell RC, Miller T. Digital marketing to promote healthy weight gain among pregnant women in Alberta: an implementation study. *J Med Internet Res* 2019 Feb 1;21(2):e11534 [[FREE Full text](#)] [doi: [10.2196/11534](https://doi.org/10.2196/11534)] [Medline: [30707100](https://pubmed.ncbi.nlm.nih.gov/30707100/)]

6. Wright CJ, Schwarzman J, Dietze PM, Crockett B, Lim MS. Barriers and opportunities in the translation of mobile phone and social media interventions between research and health promotion practice in Australia: a qualitative study of expert perspectives. *Health Res Policy Syst* 2019 Jan 10;17(1):5 [FREE Full text] [doi: [10.1186/s12961-018-0406-x](https://doi.org/10.1186/s12961-018-0406-x)] [Medline: [30630497](https://pubmed.ncbi.nlm.nih.gov/30630497/)]
7. Obar J, Wildman S. Social media definition and the governance challenge: an introduction to the special issue. *Telecomm Pol* 2015 Oct 22;39(9):745-750 [FREE Full text] [doi: [10.1016/j.telpol.2015.07.014](https://doi.org/10.1016/j.telpol.2015.07.014)]
8. Matz SC, Kosinski M, Nave G, Stillwell DJ. Psychological targeting as an effective approach to digital mass persuasion. *Proc Natl Acad Sci U S A* 2017 Dec 28;114(48):12714-12719. [doi: [10.1073/pnas.1710966114](https://doi.org/10.1073/pnas.1710966114)] [Medline: [29133409](https://pubmed.ncbi.nlm.nih.gov/29133409/)]
9. Rosenberg M, Confessore N, Cadwalladr C. The New York Times. 2018 Mar 17. How Trump Consultants Exploited the Facebook Data of Millions URL: <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html> [accessed 2019-04-25] [WebCite Cache ID 77uFvy9kj]
10. Dunn AG, Mandl KD, Coiera E. Social media interventions for precision public health: promises and risks. *NPJ Digit Med* 2018;1. [doi: [10.1038/s41746-018-0054-0](https://doi.org/10.1038/s41746-018-0054-0)] [Medline: [30854472](https://pubmed.ncbi.nlm.nih.gov/30854472/)]
11. Pew Research Center. 2018. Demographics of Social Media Users and Adoption in the United States URL: <https://www.pewinternet.org/fact-sheet/social-media/> [accessed 2019-04-25] [WebCite Cache ID 77uG9ioVd]
12. Pew Research Center. 2018. Social Media Use 2018: Demographics and Statistics URL: <https://www.pewinternet.org/2018/03/01/social-media-use-in-2018/> [accessed 2019-04-25] [WebCite Cache ID 77uGC0rBb]
13. Pew Research Center. Social media preferences vary by race and ethnicity URL: <https://www.pewresearch.org/fact-tank/2015/02/03/social-media-preferences-vary-by-race-and-ethnicity/> [accessed 2019-04-25] [WebCite Cache ID 77uGDWWQE]
14. 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](https://doi.org/10.2196/jmir.3614)] [Medline: [25498835](https://pubmed.ncbi.nlm.nih.gov/25498835/)]
15. Abokhodair N, Yoo D, McDonald D. arXiv. 2016. Dissecting a Social Botnet: Growth, Content and Influence in Twitter URL: <http://arxiv.org/abs/1604.03627> [accessed 2019-04-25] [WebCite Cache ID 77uLWUK7M]
16. Bender JL, Cyr AB, Arbuckle L, Ferris LE. Ethics and privacy implications of using the internet and social media to recruit participants for health research: a privacy-by-design framework for online recruitment. *J Med Internet Res* 2017 Dec 6;19(4):e104 [FREE Full text] [doi: [10.2196/jmir.7029](https://doi.org/10.2196/jmir.7029)] [Medline: [28385682](https://pubmed.ncbi.nlm.nih.gov/28385682/)]
17. PostgreSQL. URL: <https://www.postgresql.org/> [accessed 2019-04-25] [WebCite Cache ID 77uGN7u0E]
18. Ruby on Rails Guides. Ruby on Rails 4.2 Release Notes URL: https://guides.rubyonrails.org/4_2_release_notes.html [accessed 2019-04-25] [WebCite Cache ID 77uGWTdnD]
19. Semantic UI. URL: <https://semantic-ui.com/> [accessed 2019-04-26] [WebCite Cache ID 77uh8nyL4]
20. Heroku. 2019. URL: <https://www.heroku.com/> [accessed 2019-04-26] [WebCite Cache ID 77u6WK7c]
21. Amazon Web Services (AWS). 2019. URL: <https://aws.amazon.com/> [accessed 2019-04-25] [WebCite Cache ID 77uPJ8BNf]
22. Twitter. Counting characters URL: <https://developer.twitter.com/en/docs/basics/counting-characters.html> [accessed 2019-04-25] [WebCite Cache ID 77uHbBsM8]
23. Facebook Ads Help Center. Design Requirements for Instagram Feed Ads URL: <https://www.facebook.com/business/help/430958953753149> [accessed 2019-04-25] [WebCite Cache ID 77uHTdIA1]
24. Durstenfeld R. Algorithm 235: Random Permutation. USA: ACM; 1964.
25. Ruby-Doc.org. Array URL: <https://ruby-doc.org/core-2.2.0/Array.html> [accessed 2019-04-25] [WebCite Cache ID 77uHG5qb7]
26. Buffer: Social Media Management Platform. URL: <https://buffer.com/> [accessed 2019-04-25] [WebCite Cache ID 77uHQg27Y]
27. Perspective. URL: <https://www.perspectiveapi.com/> [accessed 2019-04-25] [WebCite Cache ID 77uHn7pgX]
28. GitHub. 2019. Perspective Comment Analyzer API URL: <https://github.com/conversationai/perspectiveapi> [accessed 2019-04-25] [WebCite Cache ID 77uHpMDAo]
29. Wulczyn E, Thain N, Dixon L. Ex Machina: Personal Attacks Seen at Scale. In: Proceedings of the 26th International Conference on World Wide Web. Ex Machina; 2017 Presented at: WWW'17; April 3-7, 2017; Perth, Australia p. 1391-1399. [doi: [10.1145/3038912.3052591](https://doi.org/10.1145/3038912.3052591)]
30. GitHub.: GitHub; 2019. wiki-detox URL: <https://github.com/ewulczyn/wiki-detox> [accessed 2019-04-26] [WebCite Cache ID 77uLsnFW]
31. Twitter. About Twitter limits URL: <https://help.twitter.com/en/rules-and-policies/twitter-limits> [accessed 2019-05-01] [WebCite Cache ID 782JV12GA]
32. CoSchedule. The Best Times to Post on Social Media in 2018 Based on Research URL: <https://coschedule.com/blog/best-times-to-post-on-social-media/> [accessed 2019-05-01] [WebCite Cache ID 782Jd9FR1]
33. Lee K. Buffer. Infographic: How Often Should You Post on Social Media? See the Most Popular Research and Tips URL: <https://buffer.com/library/how-often-post-social-media> [accessed 2019-05-01] [WebCite Cache ID 782K0YmXI]
34. Spark Staff. Adobe Spark. How Often to Post to Twitter, Facebook, Instagram, and Pinterest URL: <https://blog.adobespark.com/2016/05/06/how-often-to-post-to-twitter-facebook-instagram-and-pinterest/> [accessed 2019-05-01] [WebCite Cache ID 782KEwpFp]

35. Dunlop D, Basili V. A comparative analysis of functional correctness. *ACM Comput Surv* 1982;14(2):229-244. [doi: [10.1145/356876.356881](https://doi.org/10.1145/356876.356881)]
36. GitHub. 2019. trial-promoter-app URL: <https://github.com/scctsi/trial-promoter-app> [accessed 2019-04-26] [WebCite Cache ID 77ugTZ6LU]
37. Harris JK, Moreland-Russell S, Choucair B, Mansour R, Staub M, Simmons K. Tweeting for and against public health policy: response to the Chicago Department of Public Health's electronic cigarette Twitter campaign. *J Med Internet Res* 2014 Oct 16;16(10):e238 [FREE Full text] [doi: [10.2196/jmir.3622](https://doi.org/10.2196/jmir.3622)] [Medline: [25320863](https://pubmed.ncbi.nlm.nih.gov/25320863/)]
38. Davis C, Varol O, Ferrara E, Flammini A, Menczer F. BotOrNot: A System to Evaluate Social Bots. In: Proceedings of the 25th International Conference Companion on World Wide Web. 2016 Presented at: WWW'16 Companion; April 11-15, 2016; Geneva, Switzerland.
39. Hammack CM. Ethical use of social media data: beyond the clinical context. *Hastings Cent Rep* 2019 Jan;49(1):40-42. [doi: [10.1002/hast.979](https://doi.org/10.1002/hast.979)] [Medline: [30790308](https://pubmed.ncbi.nlm.nih.gov/30790308/)]
40. Clerwall C. Enter the robot journalist. *Journalism Pract* 2014 Feb 25;8(5):519-531. [doi: [10.1080/17512786.2014.883116](https://doi.org/10.1080/17512786.2014.883116)]
41. Allem J, Ramanujam J, Lerman K, Chu K, Boley CT, Unger JB. Identifying sentiment of hookah-related posts on Twitter. *JMIR Public Health Surveill* 2017 Oct 18;3(4):e74 [FREE Full text] [doi: [10.2196/publichealth.8133](https://doi.org/10.2196/publichealth.8133)] [Medline: [29046267](https://pubmed.ncbi.nlm.nih.gov/29046267/)]
42. Allem J, Ferrara E, Uppu SP, Cruz TB, Unger JB. E-cigarette surveillance with social media data: social bots, emerging topics, and trends. *JMIR Public Health Surveill* 2017 Dec 20;3(4):e98 [FREE Full text] [doi: [10.2196/publichealth.8641](https://doi.org/10.2196/publichealth.8641)] [Medline: [29263018](https://pubmed.ncbi.nlm.nih.gov/29263018/)]
43. Wilkie A, Michael M, Plummer-Fernandez M. *Speculative Method and Twitter: Bots, Energy and Three Conceptual Characters*. Thousand Oaks, California: Sage Publications; 2015:79.
44. Tavares G, Faisal A. Scaling-laws of human broadcast communication enable distinction between human, corporate and robot Twitter users. *PLoS One* 2013;8(7):e65774 [FREE Full text] [doi: [10.1371/journal.pone.0065774](https://doi.org/10.1371/journal.pone.0065774)] [Medline: [23843945](https://pubmed.ncbi.nlm.nih.gov/23843945/)]
45. Grimme C, Preuss M, Adam L, Trautmann H. arXiv. 2017. Social Bots: Human-Like by Means of Human Control? URL: <https://arxiv.org/pdf/1706.07624.pdf> [accessed 2019-04-25] [WebCite Cache ID 77uP3kWoC]
46. Edwards C, Edwards A, Spence P, Shelton A. Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter. *Comput Human Behav* 2014 Apr;33:372-376. [doi: [10.1016/j.chb.2013.08.013](https://doi.org/10.1016/j.chb.2013.08.013)]
47. Hwang T, Pearce I, Nanis M. *Socialbots: Voices from the Fronts*. USA: ACM; 2012.
48. Lee K, Eoff B, Caverlee J. Seven Months with the Devils: A Long-Term Study of Content Polluters on Twitter. 2011 Presented at: Fifth International AAAI Conference on Weblogs and Social Media; July 17-21, 2011; Barcelona (Spain) URL: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewPaper/2780>
49. Berkeley EP. Works Bepress. 2014. Twitter Deception and Influence: Issues of Identity, Slacktivism, and Puppetry URL: https://works.bepress.com/david_cook/15/ [accessed 2019-04-25] [WebCite Cache ID 77uJZDD9H]
50. Mustafaraj E, Metaxas P. Wellesley College. 2010. From obscurity to prominence in minutes: political speech and real-time search URL: <http://cs.wellesley.edu/~pmetaxas/Metaxas-Obscurity-to-prominence.pdf> [accessed 2019-04-25] [WebCite Cache ID 77uJu3EZA]
51. Boshmaf Y, Muslukhov I, Beznosov K, Ripeanu M. The Socialbot Network: When Bots Socialize for Fame and Money. In: Proceedings of the 27th Annual Computer Security Applications Conference. USA: ACM; 2011 Presented at: 27th Annual Computer Security Applications Conference; December 3-7, 2011; New York, NY, USA p. 93-102.
52. Ratkiewicz J, Conover M, Meiss M. Association for the Advancement of Artificial Intelligence. 2011. Detecting and tracking political abuse in social media URL: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewPaper/2850> [accessed 2019-04-25] [WebCite Cache ID 77uH8W1mD]
53. Aiello L, Deplano M, Schifanella R, Ruffo G. People Are Strange When You're a Stranger: Impact and Influence of Bots on Social Networks. 2012 Presented at: Sixth International AAAI Conference on Weblogs and Social Media; June 4-7, 2012; Dublin, Ireland URL: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/viewPaper/4523>
54. Haustein S, Bowman T, Holmberg K, Tsou A, Sugimoto C, Larivière V. arXiv. 2014. Tweets as impact indicators examining the implications of automated bot accounts on Twitter URL: <https://arxiv.org/abs/1410.4139> [accessed 2019-04-25] [WebCite Cache ID 77uGvsUm2]
55. Lokot T, Diakopoulos N. News Bots: automating news and information dissemination on Twitter. *Digit Journal* 2016 Aug 17;4(6):682-699. [doi: [10.1080/21670811.2015.1081822](https://doi.org/10.1080/21670811.2015.1081822)]
56. Wagner C, Mitter S, Körner C, Strohmaier M. When Social Bots Attack: Modeling Susceptibility of Users in Online Social Networks. 2012 Presented at: MSM2012; April 12, 2012; Lyon p. 48 URL: http://ceur-ws.org/Vol-838/paper_11.pdf
57. Salimian P, Chunara R, Weitzman E. Averting the perfect storm: addressing youth substance use risk from social media use. *Pediatr Ann* 2014 Oct;43(10):411. [doi: [10.3928/00904481-20140924-08](https://doi.org/10.3928/00904481-20140924-08)] [Medline: [25290130](https://pubmed.ncbi.nlm.nih.gov/25290130/)]
58. Kavuluru R, Sabbir AK. Toward automated e-cigarette surveillance: spotting e-cigarette proponents on Twitter. *J Biomed Inform* 2016 Dec;61:19-26 [FREE Full text] [doi: [10.1016/j.jbi.2016.03.006](https://doi.org/10.1016/j.jbi.2016.03.006)] [Medline: [26975599](https://pubmed.ncbi.nlm.nih.gov/26975599/)]
59. Collins L, Glasser A, Abudayyeh H, Pearson J, Villanti A. E-cigarette marketing and communication: how e-cigarette companies market e-cigarettes and the public engages with e-cigarette information. *Nicotine Tob Res* 2019 Jan 1;21(1):14-24. [doi: [10.1093/ntr/ntx284](https://doi.org/10.1093/ntr/ntx284)] [Medline: [29315420](https://pubmed.ncbi.nlm.nih.gov/29315420/)]

Abbreviations

API: Application Programming Interface
CSV: Comma Separated Values
FDA: Food and Drug Administration
JSON: JavaScript Object Notation
NCI: National Cancer Institute
NIH: National Institutes of Health
REST: Representational State Transfer

Edited by G Eysenbach; submitted 15.06.18; peer-reviewed by JP Allem, M McVay, C Rubanovich; comments to author 19.07.18; revised version received 19.03.19; accepted 02.04.19; published 07.05.19.

Please cite as:

Reuter K, MacLennan A, Le N, Unger JB, Kaiser EM, Angyan P

A Software Tool Aimed at Automating the Generation, Distribution, and Assessment of Social Media Messages for Health Promotion and Education Research

JMIR Public Health Surveill 2019;5(2):e11263

URL: <http://publichealth.jmir.org/2019/2/e11263/>

doi: [10.2196/11263](https://doi.org/10.2196/11263)

PMID: [31066708](https://pubmed.ncbi.nlm.nih.gov/31066708/)

©Katja Reuter, Alicia MacLennan, NamQuyen Le, Jennifer B Unger, Elsi M Kaiser, Praveen Angyan. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 07.05.2019. 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

Exploring Australian Hajj Tour Operators' Knowledge and Practices Regarding Pilgrims' Health Risks: A Qualitative Study

Amani S Alqahtani¹, BSc, MPH, PhD; Mohamed Tashani², BSc, MPH, PhD; Anita E Heywood³, BSc Biomed, MPH, PhD; Robert Booy², MD; Harunor Rashid², MD; Kerrie E Wiley⁴, BSc, MScMed, PhD

¹Saudi Food and Drug Authority, Riyadh, Saudi Arabia

²Child and Adolescent Health, Children's Hospital at Westmead Clinical School, University of Sydney, Sydney, Australia

³University of New South Wales, Sydney, Australia

⁴University of Sydney, Sydney, Australia

Corresponding Author:

Amani S Alqahtani, BSc, MPH, PhD

Saudi Food and Drug Authority

Northern Ring Branch Rd, An Nafal

Riyadh, 12232

Saudi Arabia

Phone: 966 545099789

Email: amani.shelwa@gmail.com

Abstract

Background: Travel agents are known to be one of the main sources of health information for pilgrims, and their advice is associated with positive health behaviors.

Objective: This study aimed to investigate travel agents' health knowledge, what health advice they provide to the pilgrims, and their sources of health information.

Methods: In-depth interviews were conducted among specialist Hajj travel agents in Sydney, Australia. Thematic analysis was undertaken.

Results: Of the 13 accredited Hajj travel agents, 9 (69%) were interviewed. A high level of awareness regarding gastrointestinal infections, standard hygiene methods, and the risk of injury was noted among the participants and was included in advice provided to pilgrims. However, very limited knowledge and provision of advice about the risk of respiratory infections was identified. Knowledge of the compulsory meningococcal vaccine was high, and all participated travel agents reported influenza vaccine (a recommended vaccine) as a second "compulsory" vaccine for Hajj visas. Conversely, participants reported very limited knowledge about other recommended vaccines for Hajj. The Ministry of Hajj website and personal Hajj experience were the main sources of information.

Conclusions: This study identifies a potential path for novel health promotion strategies to improve health knowledge among Hajj travel agents and subsequently among Hajj pilgrims.

(*JMIR Public Health Surveill* 2019;5(2):e10960) doi:[10.2196/10960](https://doi.org/10.2196/10960)

KEYWORDS

Hajj; health advice; mass gathering; travel agents; travel; respiratory infections

Introduction

As the world's largest annual mass gathering, Hajj presents a major challenge to disease control in Saudi Arabia, as more than 2 million people travel each year to commemorate the pilgrimage [1]. This challenge is not only for Saudi Arabia but is also international with the risk of pilgrims importing infectious diseases into their home countries when they return [2].

Respiratory tract infections, including influenza and pneumonia, are the leading health risks at Hajj [3-6]. Emerging infections like Middle East respiratory syndrome coronavirus (MERS-CoV) pose additional threats to disease control [7,8]. The Saudi Arabian Ministry of Health (MoH) and World Health Organization have introduced and recommended various preventive health measures for pilgrims including vaccines and other disease control measures (Table 1 and Textbox 1) [9].

Table 1. Summary of the recommended vaccines for the prevention of infectious disease in travellers to Saudi Arabia for Hajj.

Preventive measures	Instructions	Remarks
Compulsory vaccines		
Quadrivalent meningococcal vaccine (ACWY)	Compulsory for all pilgrims	Administered not less than 10 days before arrival
Oral polio (OPV) or inactivated poliovirus (IPV)	Compulsory for pilgrims from endemic countries	Administered at least 4 weeks before arrival Other pilgrims should remain up to date
Yellow fever vaccine	Compulsory for pilgrims from endemic countries or those transiting through endemic countries	Administered at least 10 days before arrival
Recommended vaccines		
Seasonal influenza vaccine	Recommended for all, in particular at-risk pilgrims	N/A ^a
Vaccines against diphtheria, tetanus, pertussis, measles, and mumps	Remaining up to date	N/A

^aN/A: not applicable.

Textbox 1. Summary of the recommended non-pharmaceutical measures for the prevention of infectious disease in travellers to Saudi Arabia for Hajj.

<p>Non-pharmaceutical measures</p> <ul style="list-style-type: none"> • Wash hands with soap and water or disinfectant, especially after coughing and sneezing, after using the toilet, before handling and consuming food, and after touching animals • Use disposable tissues when coughing or sneezing and dispose of them afterwards in waste baskets • Avoid hand contact with the eyes, nose, and mouth • Wear facemasks, especially in crowded places • Avoid direct contact with persons who appear to be ill with coughing, sneezing, expectorating, vomiting, diarrhea, and do not share personal belongings • Maintain good personal hygiene • Avoid contact with sick animals • Avoid drinking raw camel milk or camel urine or eating meat that has not been properly cooked • Take insect bite avoidance measures during daytime and night time hours to reduce the risk of infection with dengue and other mosquito-borne diseases <p>Health education</p> <ul style="list-style-type: none"> • Health authorities in countries of origin are required to provide health information to pilgrims on infectious disease symptoms, transmission mode, and measures to prevent
--

Travel to Hajj is unique compared to other types of travel, including travel to other mass gathering events, in that pilgrims must travel in groups by enrolling with a travel agent authorized by the Saudi Arabian authority to run Hajj tourism [10]. These tour operators play an important role in preparing pilgrims for their journey, including securing Hajj visas, conducting pre-travel briefing sessions about the ritual steps (popularly known as “Hajj seminars”), and providing some safety advice to pilgrims to enable a safe journey for their clients [10].

The provision of pre-travel health advice plays an important role in raising awareness among the travellers of the health risks and preventive measures. Seeking health advice from travel agents has been reported previously among travellers including Hajj pilgrims [11-15]. Recent quantitative studies have found that Hajj pilgrims report travel agents as one of their main

sources of health information, and their advice was associated with increased likelihood of positive travel health behaviors [14,16-21]. However, no studies have attempted to investigate the provision of pre-travel health advice from the travel agents’ perspective. We conducted a qualitative study aimed at addressing this knowledge gap through better understanding travel agents’ knowledge about the health risks at Hajj, what health advice they provide to the pilgrims, and their sources of health knowledge.

Methods

Study Design

A qualitative study was conducted to enable exploration of unknown or unanticipated issues not usually possible with

quantitative methods. Qualitative methods provide detailed, in-depth insights into why people act the way they do. A grounded theory approach was used [22], as it provided the ideal methodological framework in which to explore the pre-travel health advice Hajj travel agents provide to pilgrims in a way that allows for unexpected findings to be further explored. Grounded theory uses an inductive approach, whereby a cycle of data collection (interviews) and analysis is followed by subsequent interview and analysis cycles. Themes emerge from earlier interviews that inform subsequent interview questions to allow further exploration of emergent themes. This cycle continues until no new themes emerge and theoretical saturation is reached.

Sampling Strategy and Participants

A purposive sampling strategy was employed. This study was conducted in Sydney, New South Wales (NSW), Australia. Our research group previously conducted quantitative studies in 2014 and 2015 among Sydney-based Hajj pilgrims [14,16-19]. Sampling Sydney-based travel agents enabled us to further examine the quantitative survey findings. Agents were invited to participate in the study through email or telephone. Once the agent responded and agreed to participate, a consent form was sent via email or post, and verbal consent was obtained from all participants prior to interview.

Interviews

Between February and September 2016, in-depth interviews were conducted. All interviews were audio recorded with the participants' verbal consent. Participants were assured that interviews would be strictly confidential and any published results would contain non-identifying data only. Each of the interviews took between 30 and 60 minutes to complete. While English was the primary language used in the interviews, participants were given the option to be interviewed in Arabic if they preferred.

A semistructured interview guide was used that evolved with each iteration of the grounded theory data collection and analysis cycle [22]. Participants were encouraged to narrate events and situations they experienced and to describe the circumstances around their behaviors and choices. Participants' actual behaviors and experiences were the focus, rather than their opinions. The interviews explored the pre-travel health services and information that tour group leaders provide to the pilgrims in regard to the risk of health hazards at Hajj and the preventive measures they promote to protect the pilgrims against those health risks. The interviews also explored pre-travel health advice issues such as access, utilization and barriers to health services, and current information sources and their cultural appropriateness.

Analysis

Throughout the process, the first author kept a research journal detailing experiences and thoughts in relation to the collection and analysis of the data. This method is known to promote reflexivity on the part of the researcher during data collection, analysis, and reporting. The interviewer independently constructed a code list of major themes emerging from the data using line-by-line coding of the transcripts, followed by focused

coding to synthesize emergent themes from the data. Finally, axial coding was used to draw relationships between the themes. The coding of transcripts and the compilation and analysis of code files were done using NVivo 11 software.

Ethical Approval

This study was reviewed and approved by the Human Research Ethics Committee (HREC) at the University of Sydney (Project No. 2014/599).

Results

Participant Characteristics

There are 13 accredited Hajj tour group leaders operating in Sydney, Australia. All were approached, and 9 (69%) agreed to participate in the study. Eight of nine (89%) were male, all were aged between 38 and 65 (median 45) years, and the number of years of Hajj travel business experience among participants ranged between 3 and 30 (median 9) years.

Thematic analysis of interview responses identified two main themes: (1) the type of advice given to pilgrims by these travel agents and (2) the sources of health information for travel agents. Each key theme can be divided into subthemes. In the following sections, the main findings will be described and links between themes highlighted.

Major Theme 1: The Type of Health Advice Provided to Hajj Pilgrims by Travel Agents

Generally, advice provided by travel agents typically included the compulsory vaccination requirements for securing the Hajj visa, standard hygiene methods such as hand hygiene, and some safety advice to prevent injury or trauma during Hajj. A typical response to the question of what kinds of general advice they provide to pilgrims was:

We usually provide them information regarding the required vaccines for the visa and general hygiene practice such as keeping their hand clean most of the time. We also ask them to follow the group leader all the time so they can perform the Hajj in the best way.
[Male, 65 years old]

Most of the agents additionally requested at-risk pilgrims (the elderly and/or those with chronic conditions) to provide a letter from their local doctor about their ability to perform Hajj. Also, they advised pilgrims to bring adequate quantities of their regular medications along with prescriptions for these medications. For example:

We also ask the pilgrims to bring a medical certificate from their doctors to approve that he or she is fit physically and able to do the Hajj pilgrimage, and this only for pilgrims who are elderly or have any medical conditions. [Male, 37 years old]

Yes, I ask them to bring adequate medication and have a certificate from the doctor that they are able to do the Hajj. [Male, 42 years old]

Specific Advice About Health Risks During Hajj

When asked what specific advice they give to prospective pilgrims about health risks, the majority of travel agents spontaneously responded with their advice about non-infectious ailments such as sunstroke, trauma and injuries, and food-borne infectious diseases. None mentioned the risk of communicable diseases, with the exception of influenza:

Usually in Hajj, the major things they can face in there is the crowded (sic) and the dehydration. So this is a very major issue because it's really hot weather and the hot climate (sic) especially for the elderly, and also the food is really, really important because they need to eat healthy, clean food, and need to know where to buy it from, and all that sort of things. [Male, 48 years old]

The risks mainly were about the diarrhea, indigestion, and flu. That's what we usually advise, plus heat stroke. [Male, 37 years old]

Advice on the Risk of MERS-CoV During Hajj

Few agents were aware of the emerging infection, MERS-CoV, in Saudi Arabia. Of those who knew about it, there was a demonstrated lack of knowledge or incorrect beliefs about issues such as mode of transmission and appropriate preventive measures.

No, we were not aware of MERS-CoV in particular, but we only advised them about the problem with the poor health condition of some third world countries, which causes some serious diseases that do not exist in Australia. So we usually ask them to avoid some crowded places. [Male, 44 years old]

Yes, like in 2009 we told them to take care of their selves from H1N1, so in MERS we told them that there is more severe infection now. [Male, 65 years old]

Yeah, we heard about it and it's been, like, there for about two...three years now. So, as I said, we told them that this infection can be transferred, it's contagious, and could be by blood or by water and also by breath. [Male, 46 years old]

Postponement of Hajj travel for at-risk groups was the only preventive measure for MERS-CoV that was reported (by two travel agents). No other advice or recommendations about MERS-CoV preventive measures such as facemasks was reported. Travel agents reported a lack of regular information about MERS-CoV from the Saudi Arabian Embassy as a factor in not providing advice.

In Hajj 2014, the Saudi Embassy sent us a message and advised to visit some links which has some information regarding MERS and advised the elderly to not perform the Hajj, but not last year (2015) or this year (2016). Actually, this was very good and helped us to provide the pilgrims proper information. But this year we did not give them any advice regarding MERS-CoV as we did not receive any message from the Saudi Embassy. [Male, 48 years old]

Vaccination Advice

The meningococcal vaccine (ACWY) is a compulsory vaccine for obtaining the Hajj visa, and most travel agents described a high level of knowledge regarding meningococcal disease severity and vaccine effectiveness in protecting pilgrims. This knowledge was also shared with the pilgrims as a further incentive to get the vaccine.

We described to them that this disease is fatal, and I also give them examples about myself that I always get this vaccine and renew the certificate when it's expired. In Australia, this infection is rare, but in some countries this infection maybe common and as Hajj involved many pilgrims from different countries, this may increase the risk to get this disease during Hajj. [Male, 65 years old]

Yes, we do. Because we know that it's a dangerous disease and if they catch it they will be in grave danger of life and so forth. So yes, we do advise them that it's important that you need to take it. [Male, 37 years old]

Other agents described the importance of the meningococcal vaccine only to get the visa and relied on doctors to provide any further medical information when the pilgrims visited them to get the vaccine.

I told them that without this vaccine you cannot get the visa, but medically, we deal with some specific doctors and advise the pilgrims to visit them and they provide them the medical information about the vaccine. [Male, 46 years old]

When the travel agents were asked if they recommended any other vaccine to the pilgrims besides meningococcal vaccine, interestingly, all travel agents reported seasonal influenza vaccine as a second "compulsory" vaccine for the visa, despite its being "recommended" and not "required" according to the Saudi Arabian Ministry of Health website.

Yes, I also asked them to have flu vaccine because it is compulsory too. [Male, 55 years old]

The flu vaccine is compulsory too. [Male, 39 years old]

Yes ACWY, that's meningococcal and then they need to have a seasonal flu vaccine, so these two are required before they can even be given visa and the medical certificate are attached to the back of the passport. [Male, 44 years old]

None of the travel agents recommended any vaccines other than the compulsory one or those they considered compulsory. They considered this to be the doctors' job as they are medically qualified and have a better understanding of the pilgrim's health history.

I don't really recommend any other vaccine. I'm not a doctor myself. But I recommended them to speak to their local doctor and if they would like them to take extra vaccines. [Male, 40 years old]

Advice on Nonpharmaceutical Measures

Constant attention to hand hygiene was strongly advised among the non-pharmaceutical measures. Advice to use facemasks was also reported among some travel agents, especially in crowded places and peak times. No other recommendations for preventive measures were spontaneously reported among the travel agents.

I always advised them to wash their hands, especially before and after eating and when using the toilet. We also advised them to use facemask, especially in crowded places and peak hours. [Male, 65 years old]

Mainly the advice was washing the hands, always keeping washing the hands at certain times like when you go out, come back to your room, before you eat, wash your hands. [Male, 39 years old]

Some agents stated that they continually follow their pilgrims during the Hajj pilgrimage and remind them to use hand hygiene and facemasks.

Yes, we usually visit them in the tent especially in Arafat and Mina and, for example, if the weather is hot we advise them to stay inside the tent. And if we will move and if we expect any crowds, we asked them to put on the mask. Also I keep remind them about hand hygiene and food hygiene. [Male, 55 years old]

Dissemination of Health Advice by Travel Agents

All Hajj travel agents conducted pre-Hajj seminars several weeks before travel to Hajj. While some discussion was included, there was no specific health session conducted as a part of these seminars.

We organize Hajj seminars to teach them the Hajj worship; sometime we shared some health advice as well. [Male, 42 years old]

Well, in our general booklet, we do talk about some issues of health, but it's not a particular book in just specifically for health. [Male, 37 years old]

Major Theme 2: Sources of Health Information for Travel Agents

Travel agents described obtaining health information from different sources. The Ministry of Hajj and Umrah website was the main cited source.

Mostly I sought the information from Ministry of Hajj website (Saudi Arabia), I do not use any other sources; I only give information of what Ministry of Hajj advises. [Male, 48 years old]

Yes, Ministry of Hajj Saudi Arabia, we just do not use any other source, because we have no other information to be verified. So we just only give information of what Ministry of Hajj advises. [Male, 44 years old]

Moreover, they also relied on their previous Hajj experiences as a source of their health information.

My health information is mostly from my Hajj experience. I have been organizing the Hajj travel for about 30 years. [Male, 65 years old]

Only one agent mentioned the Australian government website, Smartraveller. This travel agent found this website to be disappointing because it does not provide enough information regarding Hajj, so they do not visit it frequently.

Smartraveller has no information apart from threatening the people to not go there, do not talk to this, and do not do this; it's useless source. They are not culturally aware and they just only give negative information. So it's not sufficient at all. [Male, 37 years old]

Interestingly, none of the participants mentioned the Saudi Arabian MoH website (different to the Ministry of Hajj site mentioned above) as a source of their health information. When asked if they referred their pilgrims to visit any specific Hajj website, some agents encouraged their clients to visit the general Hajj websites but not the Ministry of Health website.

Usually yes, and once we send them information via email there's a link down there of some websites about Hajj, so they can retrieve it at any time. [Male, 44 years old]

Yes, we advised them if they need any further information to Google some Hajj websites. [Female, 45 years old]

None of the travel agents undertook any special pre-Hajj training program either as a requirement for approval from the Saudi Arabian authority or just to update their knowledge of health regulations regarding Hajj.

No, we aren't required to have those sorts of programs and I think it must be included. So for me being a younger Hajj group leader, I feel it's a duty of care in all this so I always seek the information from different sources. But there's a lot of older Hajj group organizers that pretty much do not have the resources that I have. I'm not aware if the Ministry of Health, or Ministry of Hajj have such programs for this. [Male, 44 years old]

Type of Health Information Hajj Pilgrims Usually Seek From the Travel Agent

Travel agents stated that, in their experience, information regarding the visa requirements, weather, and crowding during Hajj was the most sought after by pilgrims. Among those who take regular medications, most inquiries were about the way they can carry their medications and keep them in good condition during their travel.

They were asking about vaccines for the visa, the crowd, or the general conditions and the temperature and also the travelling route, the toilet, and these are just general basic handling that they need to look at what they require during their travel. That's all they're usually worried about. [Male, 37 years old]

Mostly those who have diabetes and asthma, they're concerned and always asked about their medication, how they can bring it with them and keep it safe in the journey. As you know, specially the diabetes medication it should be kept cool, so we provide the

cold container when we travel from Australia to Saudi Arabia and from Madinah to Makkah. [Male, 55 years old]

Barriers to and Facilitators of the Availability of Health Information

All respondents reported that there were no barriers regarding health information access because much of the information is widely available through the Internet. The travel agents assumed the information they provide to their clients (Hajj pilgrims) is adequate to make the pilgrims' travel safe.

Well, this generation is more aware and knowledgeable; they always visit the Internet to get all the information they need. [Male, 65 years old]

I don't think there is a lack of information. It's alright. We've given them all the information they need. [Male, 48 years old]

Travel agents suggested some methods to better deliver health information to Hajj pilgrims. This included providing health lectures for Australian pilgrims before travel, conducted by Muslim doctors; disseminating health booklets organized by the Saudi health authority at entry points to Saudi Arabia or through the Saudi embassies in pilgrims' original countries; and supplying the travel agents with adequate, factual health information so they can pass it on to pilgrims.

All the information is accessible; however, if the Saudi Ministry of Health or Ministry of Hajj provide us a booklet that have (sic) all the health information regarding the vaccinations and any other measures that the pilgrims need, this will be a better way to deliver the information, so the pilgrim can keep it with them during Hajj journey. Like what happened in 2014, they provided us booklets which were very helpful. Or provide it in the entry point in Mecca or Medina but better to be before Hajj so they can well prepare. [Male, 65 years old]

The Relationships Between the Identified Themes

The policy system that develops Hajj health regulations is under the "umbrella" of the Saudi Ministry of Health and Ministry of Hajj, with the involvement of local health authorities in the original country of pilgrims [9,10]. However, personal Hajj experience and the Ministry of Hajj website were the most direct influencers of health knowledge among the interviewed Hajj tour operators. Other indirect factors such as Hajj experiences of their previous clients and advice from the Saudi Arabian Embassy were also influential. Health information was distributed through the pre-Hajj seminars run by the tour operators, either verbally or in written form ([Multimedia Appendix 1](#)).

Discussion

Principal Findings

This study provides insight into Hajj travel agents' knowledge and awareness of health issues and the health advice they provide to their Hajj travellers. It demonstrates that advice from travel agents was the main source of knowledge about vaccines

and significantly improved vaccine uptake, compared to advice from other sources such as general practitioners, in Hajj 2014 and 2015 [14]. It also illustrates that, with the exception of influenza, the agents reported very limited knowledge and advice about the risk of respiratory infectious diseases, which are the most common health hazards at Hajj [3,23].

This study found a high level of knowledge among travel agents regarding non-infectious illnesses such as sunstroke, trauma and injuries, and food-borne infectious diseases. These results agree with the findings of our previous quantitative studies where we found that Australian Hajj pilgrims in both years (2014 and 2015) were very concerned about the risk of food-borne diseases, diarrhea, influenza, and trauma [14], yet showed a lower level of concern about other common respiratory infections at Hajj, including pneumonia. Interestingly, in this study, while few tour operators were aware of the ongoing MERS-CoV risks in Saudi Arabia and had limited knowledge about its transmission mode and preventive measures, some participants mentioned that the advice about MERS-CoV was distributed only in 2014, but not in 2015 and 2016. Information from the Saudi Arabian Embassy in 2014 (that we deem appropriate) was the main source of their knowledge. These results align with the pilgrim knowledge base examined in our quantitative surveys, which found a decline in pilgrims' awareness about MERS-CoV in 2015 (117/421, 27.8%) compared with the previous survey in 2014 (179/350, 51.1%) [16,17]. Given that no cases were detected at these Hajj events, education regarding MERS-CoV was arguably adequate.

Unsurprisingly, in this study we found that knowledge among travel agents about the compulsory meningococcal vaccine was high, including the severity of the disease and the transmission mode. Remarkably, all participating travel agents also reported seasonal influenza vaccine as a second compulsory vaccine for obtaining the Hajj visa. This helps explain the finding from the quantitative surveys of pilgrims, which reported a comparatively high rate of influenza vaccine coverage among Australian Hajj pilgrims in several years (80% in 2014 [14] and 76% in 2015), which is much higher than for other countries (eg, 7.1% among Turkish pilgrims in 2015 [24] and 20% among Egyptian pilgrims between 2012 and 2015 [25]). In addition, our qualitative data revealed no knowledge or advice about the other recommended vaccines from the sampled Hajj travel agents. This result aligns with the quantitative findings relating to the low uptake of other recommended vaccines such as pneumococcal and pertussis vaccines among Australian Hajj pilgrims [14]. Similarly, low uptake has also been reported among international Hajj pilgrims [26].

While our study found that no pre-Hajj training programs have been offered to the Hajj tour operators, this study showed that the Ministry of Hajj and Umrah website and personal Hajj experience were the main sources of information for the advice travel agents provided. This is an important finding given the fact that the Ministry of Hajj and Umrah is mostly responsible for the management of administrative procedures of Hajj and Umrah, not for ensuring the health of Hajj pilgrims. The website provides information for authorized Hajj coordinators about related services such as issuing visas and organizing accommodation. Although the Saudi Arabian MoH is the official

source of information about the health conditions for Hajj travellers, none of the sampled travel agents reported it as a source of their health knowledge. This also aligns with findings from the quantitative survey, which found more than half of Australian pilgrims were not aware of the official health recommendations issued by the Saudi Arabian MoH [14]. This may indicate that published official guidelines do not uniformly reach the travel agents, let alone reach the Hajj pilgrims.

Another important stakeholder identified from this study is the Saudi Arabian Embassy in Australia. This study found that information distributed from the embassy to the authorized Hajj travel agents helped increase the agents' knowledge regarding MERS-CoV in 2014. The Saudi Arabian Embassy is considered the first contact point for organizing Hajj travel including providing the approval for travel agents to conduct Hajj travel and receiving the applications for Hajj visas [10]. This unique finding identifies a potential path for a new health promotion strategy to help improve the health knowledge among Hajj travel agents and, as a result, among Hajj pilgrims. This can be achieved by encouraging the Saudi embassies around the world to disseminate accurate and adequate health information from the MoH to all authorised Hajj travel operators, and provide annual pre-Hajj training and educational programs to the tour operators, encouraging them to distribute the information to their pilgrims. It is noteworthy to highlight that the current health guidelines for Hajj travellers are published in medical journals or other platforms primarily targeted for health care

providers, using technical/medical language that would be difficult for lay tour operator or pilgrims to comprehend. There is a need for a guideline published in simple and understandable language for lay pilgrims and tour operators [27].

Limitations

To our knowledge, this is the first qualitative study involving Hajj tour operators. It provides valuable insight into travel agents' knowledge, attitudes, and advice related to health risks and preventive measures during Hajj. Nevertheless, this study has some limitations. While our purposive sample included 70% of the Hajj travel agents in NSW, who service the Hajj travel needs of 49% of Australia's Muslims, the inclusion of Hajj travel agents from other states and territories might have yielded themes not identified here. Time and resource constraints resulted in recruitment ceasing at nine participants, meaning that most, but not all, themes met theoretical saturation. Another limitation was that only the first author was involved in coding the transcripts. This may limit the credibility of the coding, although other senior authors were involved in the discussion and development of the thematic analysis throughout the process.

Conclusion

This study identifies important opportunities to use Hajj travel agents and Saudi embassies as conduits for health promotion for Hajj pilgrims and can inform travel health policy and practice.

Acknowledgments

We thank the United Muslims Association, Al-Bayan institute, Lebanese Muslim Association, Al-Rawdhawh Hajj travel, BMA travel agency, Daar Ibn Abbas organization, Hasan Hajj tours, and Rooti Hill Hajj group for their help and support to facilitate attending the pre-Hajj seminars. The authors also would like to acknowledge the help and the support of the Custodian of the two Holy Mosques Institute for Hajj and Umrah Research (Prof Atif Asghar), Mecca, Saudi Arabia.

Conflicts of Interest

RB has received funding from Baxter, CSL, GSK, Merck, Novartis, Pfizer, Roche, Romark, and Sanofi Pasteur for conducting this research, travel to conferences or consultancy work; all funding received is directed to research accounts at The Children's Hospital at Westmead. AEH has received grant funding from GSK and Sanofi Pasteur for investigator-driven research. HR has received fees from Pfizer and Novartis for consulting or serving on an advisory board. KEW has received travel support from Fondation Mérieux for conference attendance.

Multimedia Appendix 1

Factors influencing the health knowledge and practices of Hajj tour operators.

[DOC File, 516KB - [publichealth_v5i2e10960_app1.doc](#)]

References

1. Memish ZA, Zumla A, Alhakeem RF, Assiri A, Turkestani A, Al Harby KD, et al. Hajj: infectious disease surveillance and control. *Lancet* 2014 Jun 14;383(9934):2073-2082. [doi: [10.1016/S0140-6736\(14\)60381-0](#)] [Medline: [24857703](#)]
2. Memish Z, Assiri A, Turkestani A, Yezli S, Al Masri M, Charrel R, et al. Mass gathering and globalization of respiratory pathogens during the 2013 Hajj. *Clin Microbiol Infect* 2015 Jun;21(6):571.e1-571578 [FREE Full text] [doi: [10.1016/j.cmi.2015.02.008](#)] [Medline: [25700892](#)]
3. Al-Tawfiq J, Gautret P, Benkouiten S, Memish Z. Mass Gatherings and the Spread of Respiratory Infections. Lessons from the Hajj. *Ann Am Thorac Soc* 2016 Dec;13(6):759-765. [doi: [10.1513/AnnalsATS.201511-772FR](#)] [Medline: [27088298](#)]
4. Alfalali M, Rashid H. Prevalence of influenza at Hajj: is it correlated with vaccine uptake? *Infect Disord Drug Targets* 2014;14(3):213-218. [Medline: [25809627](#)]

5. Memish Z, Assiri A, Almasri M, Alhakeem R, Turkestani A, Al Rabeeah AA, et al. Impact of the Hajj on pneumococcal transmission. *Clin Microbiol Infect* 2015 Jan;21(1):77.e11-77.e78 [FREE Full text] [doi: [10.1016/j.cmi.2014.07.005](https://doi.org/10.1016/j.cmi.2014.07.005)] [Medline: [25636939](https://pubmed.ncbi.nlm.nih.gov/25636939/)]
6. Alqahtani A, Alfelali M, Arbon P, Booy R, Rashid H. Burden of vaccine preventable diseases at large events. *Vaccine* 2015 Nov 27;33(48):6552-6563. [doi: [10.1016/j.vaccine.2015.09.076](https://doi.org/10.1016/j.vaccine.2015.09.076)] [Medline: [26437018](https://pubmed.ncbi.nlm.nih.gov/26437018/)]
7. Azhar E, Lanini S, Ippolito G, Zumla A. The Middle East Respiratory Syndrome Coronavirus - A Continuing Risk to Global Health Security. *Adv Exp Med Biol* 2017;972:49-60. [doi: [10.1007/5584_2016_133](https://doi.org/10.1007/5584_2016_133)] [Medline: [27966107](https://pubmed.ncbi.nlm.nih.gov/27966107/)]
8. Gardner L, Rey D, Heywood A, Toms R, Wood J, Travis Waller S, et al. A scenario-based evaluation of the Middle East respiratory syndrome coronavirus and the Hajj. *Risk Anal* 2014 Aug;34(8):1391-1400 [FREE Full text] [doi: [10.1111/risa.12253](https://doi.org/10.1111/risa.12253)] [Medline: [25041625](https://pubmed.ncbi.nlm.nih.gov/25041625/)]
9. World Health Organization. Health conditions for travellers to Saudi Arabia for the pilgrimage to Mecca (Hajj). *Wkly Epidemiol Rec* 2010 Oct 22;85(43):425-428 [FREE Full text] [Medline: [21038714](https://pubmed.ncbi.nlm.nih.gov/21038714/)]
10. Aleeban M, Mackey T. Global Health and Visa Policy Reform to Address Dangers of Hajj during Summer Seasons. *Front Public Health* 2016;4:280 [FREE Full text] [doi: [10.3389/fpubh.2016.00280](https://doi.org/10.3389/fpubh.2016.00280)] [Medline: [28066758](https://pubmed.ncbi.nlm.nih.gov/28066758/)]
11. Al-Abri S, Abdel-Hady D, Al-Abaidani I. Knowledge, attitudes, and practices regarding travel health among Muscat International Airport travelers in Oman: Identifying the gaps and addressing the challenges. *J Epidemiol Glob Health* 2016 Dec;6(2):67-75 [FREE Full text] [doi: [10.1016/j.jegh.2016.02.003](https://doi.org/10.1016/j.jegh.2016.02.003)] [Medline: [26948720](https://pubmed.ncbi.nlm.nih.gov/26948720/)]
12. Heywood A, Watkins R, Iamsirithaworn S, Nilvarangkul K, MacIntyre C. A cross-sectional study of pre-travel health-seeking practices among travelers departing Sydney and Bangkok airports. *BMC Public Health* 2012 May 02;12:321 [FREE Full text] [doi: [10.1186/1471-2458-12-321](https://doi.org/10.1186/1471-2458-12-321)] [Medline: [22550996](https://pubmed.ncbi.nlm.nih.gov/22550996/)]
13. Leggat P, Zwar N, Hudson B, Travel Health Advisory Group, Australia. Hepatitis B risks and immunisation coverage amongst Australians travelling to southeast Asia and east Asia. *Travel Med Infect Dis* 2009 Nov;7(6):344-349. [doi: [10.1016/j.tmaid.2009.03.008](https://doi.org/10.1016/j.tmaid.2009.03.008)] [Medline: [19945011](https://pubmed.ncbi.nlm.nih.gov/19945011/)]
14. Alqahtani A, Wiley K, Tashani M, Willaby H, Heywood A, BinDhim NF, et al. Exploring barriers to and facilitators of preventive measures against infectious diseases among Australian Hajj pilgrims: cross-sectional studies before and after Hajj. *Int J Infect Dis* 2016 Jun;47:53-59 [FREE Full text] [doi: [10.1016/j.ijid.2016.02.005](https://doi.org/10.1016/j.ijid.2016.02.005)] [Medline: [26875699](https://pubmed.ncbi.nlm.nih.gov/26875699/)]
15. Wilder-Smith A, Khairullah N, Song J, Chen C, Torresi J. Travel health knowledge, attitudes and practices among Australasian travelers. *J Travel Med* 2004;11(1):9-15 [FREE Full text] [Medline: [14769281](https://pubmed.ncbi.nlm.nih.gov/14769281/)]
16. Alqahtani A, Wiley K, Mushta S, Yamazaki K, BinDhim N, Heywood AE, et al. Association between Australian Hajj Pilgrims' awareness of MERS-CoV, and their compliance with preventive measures and exposure to camels. *J Travel Med* 2016 May;23(5). [doi: [10.1093/jtm/taw046](https://doi.org/10.1093/jtm/taw046)] [Medline: [27432904](https://pubmed.ncbi.nlm.nih.gov/27432904/)]
17. Alqahtani A, Wiley K, Tashani M, Heywood A, Willaby H, BinDhim NF, et al. Camel exposure and knowledge about MERS-CoV among Australian Hajj pilgrims in 2014. *Viol Sin* 2016 Feb;31(1):89-93. [doi: [10.1007/s12250-015-3669-1](https://doi.org/10.1007/s12250-015-3669-1)] [Medline: [26810660](https://pubmed.ncbi.nlm.nih.gov/26810660/)]
18. Alqahtani A, Wiley K, Willaby H, BinDhim N, Tashani M, Heywood AE, et al. Australian Hajj pilgrims' knowledge, attitude and perception about Ebola, November 2014 to February 2015. *Euro Surveill* 2015 Mar 26;20(12) [FREE Full text] [Medline: [25846489](https://pubmed.ncbi.nlm.nih.gov/25846489/)]
19. Alqahtani A, Yamazaki K, Alqahtani W, Tashani M, Heywood A, Booy R, et al. Australian Hajj pilgrims' perception about mass casualty incidents versus emerging infections at Hajj. *Travel Med Infect Dis* 2017;15:81-83. [doi: [10.1016/j.tmaid.2016.11.002](https://doi.org/10.1016/j.tmaid.2016.11.002)] [Medline: [27856351](https://pubmed.ncbi.nlm.nih.gov/27856351/)]
20. Al-Zahrani I, Chaudhry A, Alhamdan N. Sources of health education for international Arab pilgrims and the effect of this education on their practices towards health hazards in Hajj, 1427 H (2006). *Saudi Epidemiol Bull* 2007;14(4):25-29.
21. Aljoudi A, Nooh N, Choudhry A. Effect of health education advice on Saudi Hajj Is, Hajj 1423 H (2003 G). *Saudi Epidemiol Bull* 2004;11(2):11-12.
22. Charmaz K. Constructing grounded theory. London, UK: Sage; 2014.
23. Khan I, Khan S, Asima B, Hussaini S, Zakiuddin M, Faisal F. Morbidity and mortality amongst Indian Hajj pilgrims: A 3-year experience of Indian Hajj medical mission in mass-gathering medicine. *J Infect Public Health* 2018;11(2):165-170 [FREE Full text] [doi: [10.1016/j.jiph.2017.06.004](https://doi.org/10.1016/j.jiph.2017.06.004)] [Medline: [28668659](https://pubmed.ncbi.nlm.nih.gov/28668659/)]
24. Sahin M, Aker S, Kaynar Tuncel E. Knowledge, attitudes and practices concerning Middle East respiratory syndrome among Umrah and Hajj pilgrims in Samsun, Turkey, 2015. *Euro Surveill* 2015;20(38) [FREE Full text] [doi: [10.2807/1560-7917.ES.2015.20.38.30023](https://doi.org/10.2807/1560-7917.ES.2015.20.38.30023)] [Medline: [26535787](https://pubmed.ncbi.nlm.nih.gov/26535787/)]
25. Refaey S, Amin M, Roguski K, Azziz-Baumgartner E, Uyeki T, Labib M, et al. Cross-sectional survey and surveillance for influenza viruses and MERS-CoV among Egyptian pilgrims returning from Hajj during 2012-2015. *Influenza Other Respir Viruses* 2017 Dec;11(1):57-60 [FREE Full text] [doi: [10.1111/irv.12429](https://doi.org/10.1111/irv.12429)] [Medline: [27603034](https://pubmed.ncbi.nlm.nih.gov/27603034/)]
26. Alqahtani A, Rashid H, Heywood A. Vaccinations against respiratory tract infections at Hajj. *Clin Microbiol Infect* 2015 Feb;21(2):115-127 [FREE Full text] [doi: [10.1016/j.cmi.2014.11.026](https://doi.org/10.1016/j.cmi.2014.11.026)] [Medline: [25682277](https://pubmed.ncbi.nlm.nih.gov/25682277/)]
27. Greenhalgh T. Time for patient leaflets on the Hajj. *BMJ* 2011 Nov 01;343:d6609. [doi: [10.1136/bmj.d6609](https://doi.org/10.1136/bmj.d6609)] [Medline: [22046006](https://pubmed.ncbi.nlm.nih.gov/22046006/)]

Abbreviations

ACWY: quadrivalent meningococcal vaccine
MERS-CoV: Middle East respiratory syndrome coronavirus
MoH: Ministry of Health (Saudi Arabia)
NSW: New South Wales

Edited by M Libel, Y Khader; submitted 03.05.18; peer-reviewed by P Gautret, F Lami, Y Khader; comments to author 21.12.18; revised version received 02.02.19; accepted 07.02.19; published 23.05.19.

Please cite as:

Alqahtani AS, Tashani M, Heywood AE, Booy R, Rashid H, Wiley KE
Exploring Australian Hajj Tour Operators' Knowledge and Practices Regarding Pilgrims' Health Risks: A Qualitative Study
JMIR Public Health Surveill 2019;5(2):e10960
URL: <http://publichealth.jmir.org/2019/2/e10960/>
doi: [10.2196/10960](https://doi.org/10.2196/10960)
PMID: [31124464](https://pubmed.ncbi.nlm.nih.gov/31124464/)

©Amani S Alqahtani, Mohamed Tashani, Anita E Heywood, Robert Booy, Harunor Rashid, Kerrie E Wiley. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 23.05.2019. 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/>