



JMIR Public Health and Surveillance

Impact Factor (2023): 3.5
Volume 3 (2017), Issue 2 ISSN 2369-2960 Editor in Chief: Travis Sanchez, PhD, MPH

Contents

Viewpoints

Experiences From Developing and Upgrading a Web-Based Surveillance System for Malaria Elimination in Cambodia (e30) Clementine Fu, Sérgio Lopes, Steve Mellor, Siddhi Aryal, Siv Sovannaroeth, Arantxa Roca-Feltrer.	4
Measuring Sexual Behavior Stigma to Inform Effective HIV Prevention and Treatment Programs for Key Populations (e23) Shauna Stahlman, James Hargreaves, Laurel Sprague, Anne Stangl, Stefan Baral.	138
Ensuring Inclusion of Adolescent Key Populations at Higher Risk of HIV Exposure: Recommendations for Conducting Biological Behavioral Surveillance Surveys (e40) Lisa Johnston, Justine Sass, Jeffry Acaba, Wing-Sie Cheng, Shirley Mark Prabhu.	283

Original Papers

Associations Between Maternal Depressive Symptoms and Nonresponsive Feeding Styles and Practices in Mothers of Young Children: A Systematic Review (e29) Ana Lindsay, Tatiana Mesa, Mary Greaney, Sherrie Wallington, Julie Wright.	18
Using Google Trends Data to Study Public Interest in Breast Cancer Screening in Brazil: Why Not a Pink February? (e17) Paulo Vasconcellos-Silva, Dárlinton Carvalho, Valéria Trajano, Lucia de La Rocque, Anunciata Sawada, Leidjaira Juvanhol.	30
How Do You #relax When You're #stressed? A Content Analysis and Infodemiology Study of Stress-Related Tweets (e35) Son Doan, Amanda Ritchart, Nicholas Perry, Juan Chaparro, Mike Conway.	38
Health Care Professionals' Evidence-Based Medicine Internet Searches Closely Mimic the Known Seasonal Variation of Lyme Borreliosis: A Register-Based Study (e19) Samuli Pesälä, Mikko Virtanen, Jussi Sane, Jukkapekka Jousimaa, Outi Lyytikäinen, Satu Murtopuro, Pekka Mustonen, Minna Kaila, Otto Helve.	59
Public Response to Scientific Misconduct: Assessing Changes in Public Sentiment Toward the Stimulus-Triggered Acquisition of Pluripotency (STAP) Cell Case via Twitter (e21) Alberto Gayle, Motomu Shimaoka.	66

Zika in Twitter: Temporal Variations of Locations, Actors, and Concepts (e22)
 Anthony Stefanidis, Emily Vraga, Georgios Lamprianidis, Jacek Radzikowski, Paul Delamater, Kathryn Jacobsen, Dieter Pfoser, Arie Croitoru, Andrew Crooks. 80

What Are People Tweeting About Zika? An Exploratory Study Concerning Its Symptoms, Treatment, Transmission, and Prevention (e38)
 Michele Miller, Tanvi Banerjee, Roopteja Muppalla, William Romine, Amit Sheth. 97

Measuring Audience Engagement for Public Health Twitter Chats: Insights From #LiveFitNOLA (e34)
 Kristina Rabarison, Merriah Croston, Naomi Englar, Connie Bish, Shelbi Flynn, Carolyn Johnson. 111

The Use of Intervention Mapping to Develop a Tailored Web-Based Intervention, Condom-HIM (e20)
 Joyal Miranda, José Côté. 123

Physician's Perceptions of Telemedicine in HIV Care Provision: A Cross-Sectional Web-Based Survey (e31)
 Kelly Anderson, Troy Francis, Francisco Ibanez-Carrasco, Jason Globerman. 160

Recruiting Young Gay and Bisexual Men for a Human Papillomavirus Vaccination Intervention Through Social Media: The Effects of Advertisement Content (e33)
 Paul Reiter, Mira Katz, Jose Bauermeister, Abigail Shoben, Electra Paskett, Annie-Laurie McRee. 167

Perceptions of HIV Seriousness, Risk, and Threat Among Online Samples of HIV-Negative Men Who Have Sex With Men in Seven Countries (e37)
 Anna Chard, Nicholas Metheny, Rob Stephenson. 177

Twimed: Twitter and PubMed Comparable Corpus of Drugs, Diseases, Symptoms, and Their Relations (e24)
 Nestor Alvaro, Yusuke Miyao, Nigel Collier. 188

Filtering Entities to Optimize Identification of Adverse Drug Reaction From Social Media: How Can the Number of Words Between Entities in the Messages Help? (e36)
 Redhouane Abdellaoui, Stéphane Schück, Nathalie Texier, Anita Burgun. 203

Social Media and Sexual Behavior Among Adolescents: Is there a link? (e28)
 Megan Landry, Monique Turner, Amita Vyas, Susan Wood. 216

Understanding Health Information Seeking on the Internet Among Sexual Minority People: Cross-Sectional Analysis From the Health Information National Trends Survey (e39)
 Jennifer Jabson, Joanne Patterson, Charles Kamen. 226

GapMap: Enabling Comprehensive Autism Resource Epidemiology (e27)
 Nikhila Albert, Jena Daniels, Jessey Schwartz, Michael Du, Dennis Wall. 235

Making Air Pollution Visible: A Tool for Promoting Environmental Health Literacy (e16)
 Ekaterina Cleary, Allison Patton, Hsin-Ching Wu, Alan Xie, Joseph Stubblefield, William Mass, Georges Grinstein, Susan Koch-Weser, Doug Brugge, Carolyn Wong. 245

Determinants of Participants' Follow-Up and Characterization of Representativeness in Flu Near You, A Participatory Disease Surveillance System (e18)
 Kristin Baltrusaitis, Mauricio Santillana, Adam Crawley, Rumi Chunara, Mark Smolinski, John Brownstein. 252

Saúde na Copa: The World's First Application of Participatory Surveillance for a Mass Gathering at FIFA World Cup 2014, Brazil (e26)

Onicio Leal Neto, George Dimech, Marlo Libel, Wayner de Souza, Eduarda Cesse, Mark Smolinski, Wanderson Oliveira, Jones Albuquerque.
2 6 3

Understanding Environmental and Contextual Influences of Physical Activity During First-Year University: The Feasibility of Using Ecological Momentary Assessment in the MovingU Study (e32)

Chloe Bedard, Sara King-Dowling, Madeline McDonald, Genevieve Dunton, John Cairney, Matthew Kwan. 273

HIV Strategic Information in Non-European Union Countries in the World Health Organization European Region: Capacity Development Needs (e41)

Ivana Bozicevic, Senad Handanagic, Jurja Cakalo, Annemarie Rinder Stengaard, George Rutherford. 294

Short Paper

The RSVP Project: Factors Related to Disengagement From Human Immunodeficiency Virus Care Among Persons in San Francisco (e25)

Susan Scheer, Miao-Jung Chen, Maree Parisi, Maya Yoshida-Cervantes, Erin Antunez, Viva Delgado, Nicholas Moss, Kate Buchacz. 151

Viewpoint

Experiences From Developing and Upgrading a Web-Based Surveillance System for Malaria Elimination in Cambodia

Clementine Fu^{1*}, MSc; Sérgio Lopes^{1*}, MSc; Steve Mellor², MSc; Siddhi Aryal³, PhD; Siv Sovannaroth⁴, MSc; Arantxa Roca-Feltrer⁵, PhD

¹Malaria Consortium Cambodia, Phnom Penh, Cambodia

²Independent Consultant, Phnom Penh, Cambodia

³Malaria Consortium Asia Regional Office, Bangkok, Thailand

⁴The National Center for Parasitology, Entomology and Malaria Control, Phnom Penh, Cambodia

⁵Malaria Consortium Global Office, London, United Kingdom

*these authors contributed equally

Corresponding Author:

Sérgio Lopes, MSc

Malaria Consortium Cambodia

House # 91, St. 95, Boeung Trabek, Chamkar Morn

Phnom Penh, 12305

Cambodia

Phone: 855 017753032

Fax: 855 017753032

Email: s.lopes@malariaconsortium.org

Abstract

Strengthening the surveillance component is key toward achieving country-wide malaria elimination in Cambodia. A Web-based upgraded malaria information system (MIS) was deemed to essentially act as the central component for surveillance strengthening. New functionality (eg, data visualization) and operational (eg, data quality) attributes of the system received particular attention. However, building from the lessons learned in previous systems' developments, other aspects unique to Cambodia were considered to be equally important; for instance, feasibility issues, particularly at the field level (eg, user acceptability at various health levels), and sustainability needs (eg, long-term system flexibility). The Cambodian process of identifying the essential changes and critical attributes for this new information system can provide a model for other countries at various stages of the disease control and elimination continuum. Sharing these experiences not only facilitates the establishment of "best practices" but also accelerates global and regional malaria elimination efforts. In this article, Cambodia's experience in developing and upgrading its MIS to remain responsive to country-specific needs demonstrates the necessity for considering functionality, operationalization, feasibility, and sustainability of an information system in the context of malaria elimination.

(*JMIR Public Health Surveill* 2017;3(2):e30) doi:[10.2196/publichealth.6942](https://doi.org/10.2196/publichealth.6942)

KEYWORDS

malaria elimination; surveillance system; Cambodia

Introduction

Within the Greater Mekong Subregion (GMS), Cambodia has seen dramatic shifts in malaria incidence during the last 10-15 years. Since 2000, confirmed malaria cases have decreased by over 75%, partly as a result of targeted treatment and management activities outlined in the National Strategic Plan 2011-2025 [1,2]. Accordingly, in line with the regional shift toward malaria elimination and the country's waning malaria burden, an update to this plan was developed in 2015—the malaria elimination action framework (MEAF) 2016-2020 [3].

This framework sets out an ambitious plan to eliminate malaria by 2025, with a particular emphasis on the surveillance system and accompanying activities. It reflects new strategic updates based on changes in the country's epidemiological and programmatic contexts, while taking into account recommendations from previous containment experiences and other regional and global policy guidelines. Strengthening the surveillance component is a necessary marker toward achieving country-wide elimination. Consequently, considering the need to incorporate and facilitate the capture of elimination-focused data, a new and upgraded information system was deemed

essential to act as the central component for the surveillance plans. The Cambodian experience of identifying the essential changes and critical attributes for this new information system can provide a model for other countries at various stages of the disease control and elimination continuum. In this viewpoint paper, we highlight the main steps needed to ensure a smooth and progressive shift from an offline information system toward an integrated Web-based system.

Cambodia Malaria Information System in Cambodia: A System Upgrade to Meet Elimination Needs

The first national malaria information system (MIS) was constructed as a Microsoft Access database (Microsoft Corporation, Redmond, USA) in 2009. It aimed to collect and analyze data reported from various levels of the health system to the centrally based National Centre for Parasitology, Entomology, and Malaria Control (CNM) [4].

The MIS evolved over time, and by the end of 2013, several modules and functionalities existed, namely, (1) individual case-based line listing reported from village malaria workers (VMWs), district-level health facilities (HFs), and private providers; (2) provincial and national data reports, including epidemiological summaries and basic visual outputs; and (3) rudimentary data flow for real-time reporting mechanisms.

Cambodia Malaria Information System—Initial Development, Objectives, and Functionalities

The system was originally developed to primarily capture monthly individual-level malaria case data from both the national community-level VMW program and government-supported health facilities in malaria-endemic provinces. The prototype system also added a level of granularity (eg, geographical and functional detail) deemed necessary to shift toward elimination [5]. Such level of detail is not present in the national Ministry of Health (MOH)-sponsored health management information system (HMIS) database, which only captures aggregate case data at the provincial and district level.

The MIS has since expanded to also capture data directly from private providers. Further functionalities were also added to the MIS to support and document other CNM-related malaria activities, including intervention coverage like bed net distribution data.

Other components were created as operational and analytical needs arose; specifically, efforts to analyze data at the village level for risk stratification purposes resulted in a new, separate module [6].

Additionally, a quarterly malaria bulletin was introduced to improve feedback and promote data-driven decisions at both peripheral and central levels.

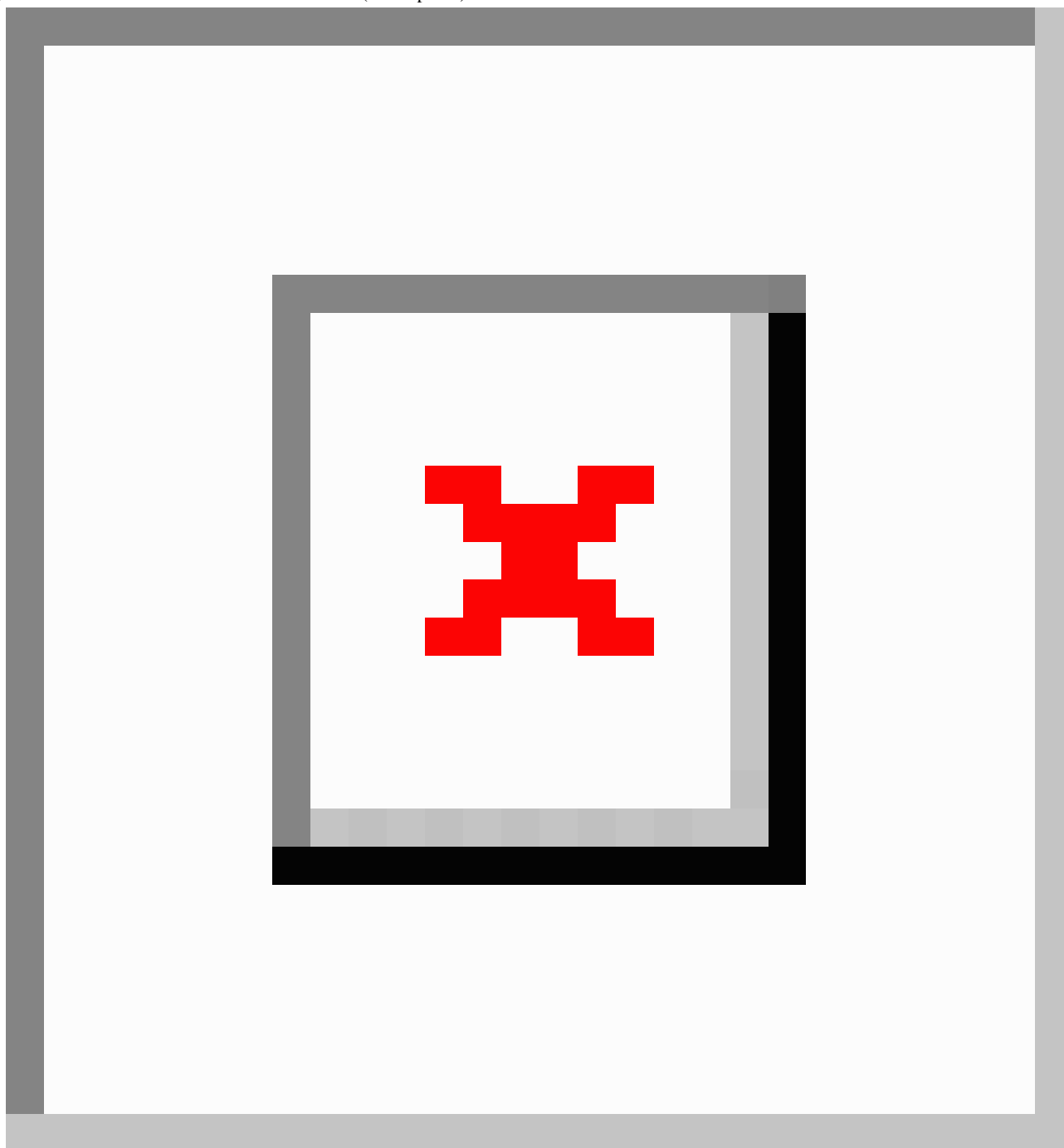
Initial Real-Time Case Reporting Activities

During the initial development phase of the MIS, an attempt at utilizing the system for real-time reporting activities was implemented in the form of day 0 or day 3 SMS (short text message, SMS) alert systems [4,6]. The day 3 system was created to notify relevant health facility or operational district (OD) officials about post-day 3 parasitaemic patients, indicating the potential need for an active response. This system was then later expanded to include day 0 case notifications, allowing for further field level responses. A separate server was set up to capture this SMS information, which was then pushed passively into the MIS. Hence, although the MIS could receive and gather this data immediately, it did not have the ability to process the information and communicate directly back to the SMS server. At that time, since no operational field follow-up was done that required supplementary SMS data entry from other sources (eg, health facility report for interventional treatment of an initially reported day 3 positive case), the MIS structure did not need to have the interoperability and flexibility needed for an active response.

Malaria Information System (MIS) Usage and Data Flow

All cases (tested, confirmed, and treated) localized to the village level are recorded via paper-based forms from VMWs, health facilities or health centers, and private providers (public private mix) in CNM-supported (malaria-endemic) districts. This data is collated monthly by each source's respective OD and then entered electronically into the MIS Access file to be sent to CNM and compiled into the comprehensive database (see [Figure 1](#)). Data from all MOH-maintained health facilities throughout the country is also entered in aggregate form to the HMIS. As a result, the lowest level of localization for HMIS-captured data is at the health facility level. Data from the HMIS is extracted regularly by CNM and entered manually into the MIS. For those areas not captured by the MIS directly (eg, district and provincial hospitals, and nonendemic regions), the data helps to complete the full national malaria context. For areas already covered by the MIS, a cross-checking analysis is done to monitor the comparative data quality. Accordingly, end users of the full MIS are located at the OD and central level only.

Data flow mimics ground-level operations, functioning primarily in a uni-directional manner toward the central level without a specific operational response. This conceptually embodies the classical model of a malaria surveillance system [5]. The mixed paper-based and electronic reporting system is still currently in use until a full transition to electronic reporting is piloted, evaluated, and fully operational.

Figure 1. Malaria surveillance data flow into MIS (Initial phase).

The Need for Upgrade—New Nationally Integrated Surveillance and Elimination Activities

According to the malaria elimination action framework 2016-2020 strategy [3], Cambodia's plans for elimination include a phased approach wherein elimination-targeted ODs are planned to scale up progressively from 18 in 2016 to 42 in 2018. This categorization is based on individual districts' risk stratification.

As this phased approach takes place, new and accompanying activities, including reactive case investigations, proactive detection, and real-time responses, become crucial components to a successful elimination strategy [7].

Functionally, the MIS needs to respond to an implementation plan with different ODs in different risk strata and needs to be able to process incoming data from multiple sources, activities, and geographical regions. The necessity for mobile platform integration is also critical, as technological tools have been proven to effectively facilitate field-level elimination responses [8].

System Components Necessary for Elimination

Defining the key features of the system in a participatory way is essential prior to the development and roll out of any surveillance system [5,9]. Toward that end, CNM and its key technical partners agreed upon several critical data elements, indicators, and surveillance activities for malaria elimination

through a series of working groups. The corresponding components needed in the complementary information system mirror the main points agreed in this group (Ministry of Health, Kingdom of Cambodia, Surveillance for Elimination: Operational Manual, unpublished data). The necessity for integrated and interrelated data justifies the transition to other type of platforms and database structure responsive to these requirements.

Real-Time Reporting and Case Notification

As caseload continues to drop, the need to respond to every case quickly for appropriate treatment and follow-up becomes more critical. Ideally, every positive case reported from an elimination OD must be registered in real time, regardless of which point of care source it originated at [10]. Therefore, all sources must have a direct notification link (eg, mobile phone, SMS, and so on) to the central information system, allowing the subsequent outwards data flow to occur. This communication is a critical piece of the elimination strategy, especially considering the potential mobility and nonpermanent status of migrant populations at highest risk for malaria.

Case Investigation

Once a case is reported, an investigation should be initiated in a timely manner to obtain further patient information, establish transmission potential, and determine need for additional response activities [11]. Operationally, the investigation may involve environmental inspection and in-depth travel history questions to be captured on a mobile platform. Additional activities or interventions may be recommended, either based on the team's manual assessment or an automatic criteria-defined trigger from the central database. From a systems perspective, this also suggests that an initial case captured by the database from a particular source will need to be linked to additional follow-up information from a different source.

Foci Investigation

In elimination settings, additional follow-up on a passively detected index case is insufficient; foci investigation is necessary to determine where the infection occurred, whether active vectors may be present to transmit the parasite, and potentially identify geographical centers of interest [7,11]. Based on results from the initial case investigation, an operational decision can be made about the need for foci information. This data can additionally inform geographical risk and vector surveillance initiatives. Although it is sufficient for foci information in the information system to be linked to the case through its reporting source, additional structural considerations are necessary to analyze risk and interventions.

Case Response and Intervention Coverage

Once the case investigation is complete, additional measures may be taken at the point of identification. Reactive case detection (RACD) around the index case and intervention coverage may be necessary, and need to be documented as a result of the initial case notification [7]. These activities can also be mapped and identified via geolocation. The operational response may be different for each case, as determined by multiple factors (eg, the OD in which the case is located, the

village-level risk, and so on)—therefore, the system should be able to automatically identify the defined course of action based on captured information and trigger resultant notifications to the responsible field-level actors.

Risk Stratification

Initial efforts at stratifying malaria risk at the village level proved to be useful for directing interventions [6]. However, this exercise was only partially automated—the risk indicator was limited and could not incorporate other nonincidence related data. Additionally, the analyses suffered from lack of data completeness and obligatory usage of time-delayed malaria data, decreasing the practical benefits. Currently, under the elimination context, a more fine-grained classification is necessary to guide the classification and transition of new MEAF risk categories. Regular analyses will assist field-level facilities to direct intervention and financial coverage [5,10]. Both monthly passive data from burden reduction areas and real-time data from elimination areas can be combined with other key factors to create a consistently updated picture of malaria risk within the country. The core system therefore needs to be able to actively interact with and integrate both types of data to create an automated, systematic output.

Data Visualization

Quick, regularly updated data visualization and utilization is a key component to decentralizing malaria control and elimination efforts [5,9,10,12]. Provincial and field-level staff can benefit from understanding and interpreting infection trends to better direct their resources. In the Access system, although the OD users could produce reports based on their own district-level data, the full scope of updated data is only captured and stored locally at the central level. Data usage is purely passive through the production of quarterly dashboards. For field staff, data application must be useful, immediate, and precise; indicators and visualizations should be predefined as much as possible to facilitate ease of operations. Bidirectionality of data flow is critical to empower those at points of care to understand the importance of their role in the surveillance system and take more ownership in these processes [5,8,13].

Moving Toward a Web-Based Malaria Information System: Key Requirements

By 2011, all 45 malaria-targeted ODs had received training and incorporated the original MIS, evidencing the rapid uptake and strong ownership of the system by the national program [6].

As Cambodia approached national and regional goals for malaria elimination, new functionality development (eg, data visualization) and operational aspects (eg, data quality) were the primary broad issues for improving the information system. However, other context-specific aspects are equally important in defining the framework, particularly system acceptability and sustainability [14]. For instance, feasibility issues at the field level (eg, user acceptability at various health levels) and sustainability needs (eg, long-term system flexibility) were considered key points to be addressed.

Corresponding System Development Timeline With Malaria Elimination Action Framework (MEAF) Elimination Plans

To respond to the ambitious phasing and objectives set by the MEAF, not only is it necessary to improve or reconstruct the current passive reporting system, but also to develop and roll out all newly required malaria elimination modules. Thus, it is crucial to follow a phased approach to ensure that each component can operate successfully both in isolation and in coordination with the entire system.

The first key step was to develop and transition to a Web-based platform. This new system, completed and piloted in 2016, moves passive data reporting from a local database to a centrally accessible server. Not only will this change facilitate speed of data entry and immediate feedback to the field level, but it also expands the necessary framework for subsequent development (see [Table 1](#)). In the following phase, previously described malaria elimination modules will be added to the MIS in a layered manner, ensuring sufficient flexibility to account for further additions or changes identified during the implementation phase (see Multimedia [Appendix 1](#)).

Table 1. Characteristics of local versus Web-based malaria information system (MIS) in Cambodia.

Characteristics	Microsoft Access: local-based MIS ^a	Web-based MIS
Modality	Offline—Fill in and send to CNM ^b monthly	Online—Web-based, “client-server”
Access	Local—Software and app must be downloaded on each device	Web—Accessible from any device with Internet access
Users	OD ^c , CNM Smaller number of users enabled	OD, CNM, HF ^d , HCs ^e , VMW ^f Can handle many users
Importation	Manual from HMIS ^g Cannot be automated	Has capacity to be automated from HMIS and other individual databases (eg, private providers)
Size	2GB maximum	10 GB+
Other	Limited capacity for capturing and communicating with mobile app data	Greater flexibility for data input and outputs from other sources, including mobile

^aMIS: malaria information system.

^bCNM: National Centre for Parasitology, Entomology, and Malaria Control.

^cOD: operational district.

^dHF: health facility.

^eHC: health center.

^fVMW: village malaria workers.

^gHMIS: health management information system.

Despite the time pressure to roll out elimination activities (18 ODs were targeted for elimination in 2016), it is essential to guarantee that the system upgrade and implementation is responsive to functional and operational needs. Although the phased process may not be the most rapid approach to reach elimination, the continuous “learning by doing” approach responds to country needs as planning between stakeholders progresses. Considering the added functions of elimination-focused field activities, concentrated technical discussions on the surveillance system are necessary to clearly establish country-specific procedures. Conceptualization and development of the core information system are tied to these decisions, as the technical components directly reflect and respond to surveillance definitions. Additionally, it is crucial to ensure that the system incorporates expansive flexibility to allow for future changes in policy or operation. Despite initial guidance on the key surveillance activities for malaria elimination, discussions about the malaria surveillance system in Cambodia are ongoing and may change based on program prioritization.

Required Functionalities for MIS Upgrade

Functional Attributes: Database Structure and Web Attributes

A significant change in the back-end data structure of the upgraded MIS is crucial to ensure system adaptability and functionality. The relational tables in the original Access-based MIS were constructed by separating data sources, mimicking the functional data flow needs at the time. Case data detected and reported by a VMW were stored independently from case data detected by a health facility, even though individual case-level information captured by each provider was similar. Hence, cases were linked to their respective data sources rather than pooled together as an aggregate collection of cases. However, as the system evolves to respond to elimination needs, a source-based structure is no longer the most efficient solution for data storage and systems processing. A secondary case investigation or case follow-up activity might be completed by a different field source than the initial reporting source. The data captured from all subsequent activities need to be linked to the original report, an effort which would be inefficient and impractical to implement in a source-based structure.

Considering the lack of a country-systematized unique patient identifier, the information system must adopt a case-based

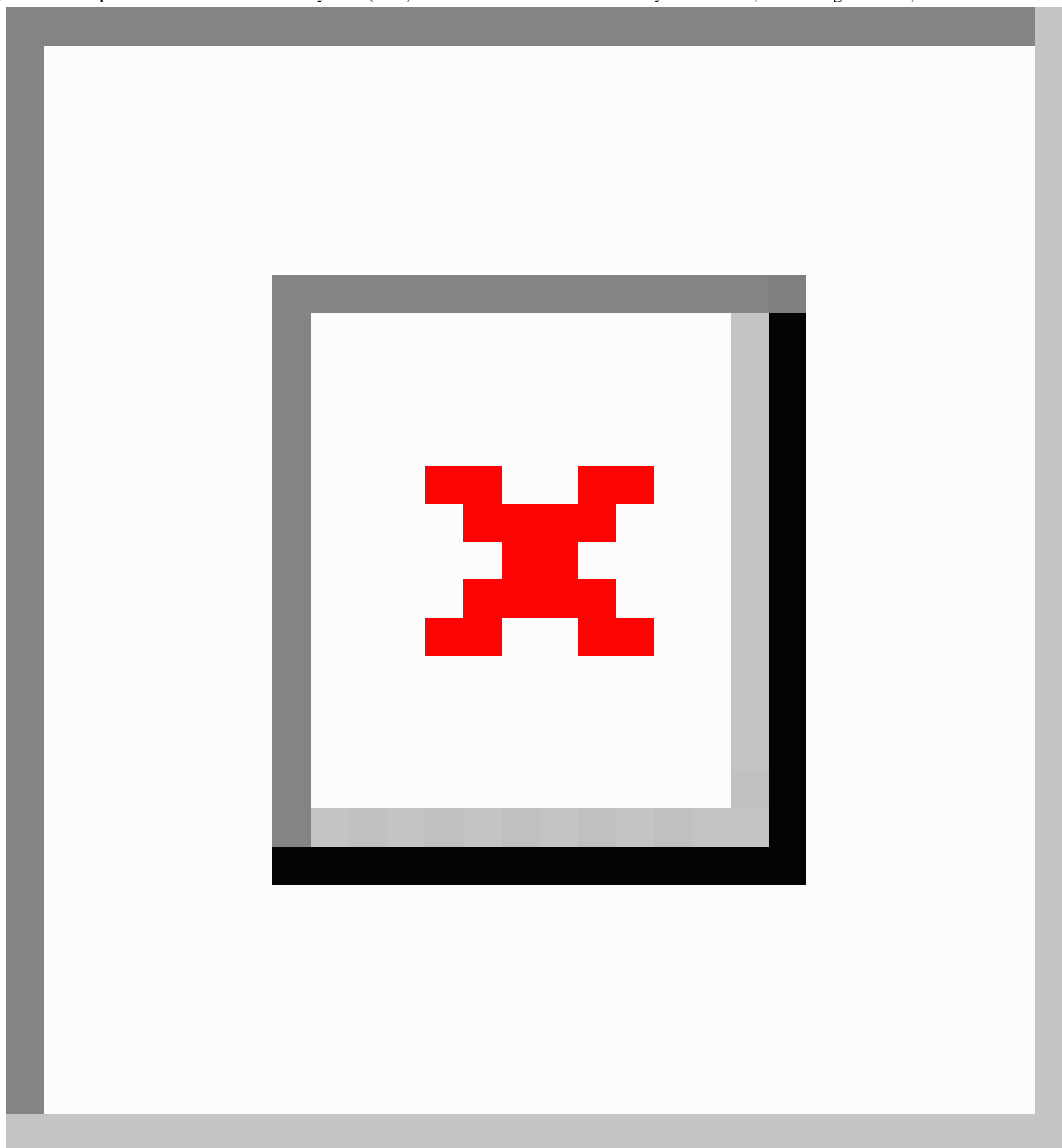
structure, wherein all cases are stored in the same table regardless of reporting source. This produces a system-derived case identification, and allows data from all initial and follow-up activities to be linked without data replication and system inefficiency. Each case's initial reporting source is still recorded to allow for programmatic monitoring. This change in data structure prepares the system to be more responsive in accumulating and analyzing information from elimination activities. Follow-up investigations and interventions can be linked to their initial case reports, enabling capacity to monitor program performance. It also allows for increased ease of query from the system database, eliminating the need to extract information from multiple separate sources. Long-term future possibility also exists for additional data sources or elimination-context activities, changing the types of incoming data and intersecting relationships. A case-based structure enables far greater system flexibility for currently unknown changes.

A main feature of the MIS upgrade is its Web-based functionality. Since the original MIS is a locally stored database, the most updated and complete dataset is only accessible by the CNM central-level data unit. Any additions or changes made at this level are not reflected in field-level office datasets unless newly updated files are sent or the Access package is reinstalled locally. Under that system, regular, timely data visualization and analysis is not possible at the peripheral levels where data is entered; rather, the primary field role is simply to report passively and wait for feedback from the CNM epidemiology unit.

As the country moves toward elimination, immediate data visualization at lower levels of the system is key to ensure appropriate, targeted responses [5,7,8]. With a Web-based functionality, automated dashboards can be produced and integrated within the MIS user interface, enabling rapid up-to-date data visualization at the lower operational levels. The increased operational capacity of a client-server relationship creates greater potential for both the system and users. This system gains an enhanced ability to communicate and acquire data through different platforms, particularly for the usage of Internet-based devices. Intersystem exchanges are also greatly enhanced, particularly for data analysis and feedback.

Additionally, any changes made to the core data reporting can be reflected immediately—for instance, if treatment guidelines are changed to validate administration of a new drug, CNM can add this option to the central administrative reporting database which updates instantaneously. Field-based end users can view and select the new option during data entry immediately rather than having to spend time and incur resource costs to implement another database configuration (see [Figure 2](#)).

Finally, utilizing a Microsoft structured query language (SQL) database rather than an Access one increases the capacity to store larger quantities of data. As elimination activities involve more operational resources and complex data analysis, they will generate more information and require greater systems processing. SQL offers greater flexibility in size for future expansions while still retaining retrospective data from all years of implementation.

Figure 2. Examples of malaria information system (MIS) attributes and their immediate system effect (treatment guidelines).

During the initial stage of the MIS upgrade, it was critical to ensure that the existing Access-based MIS modules and functionalities were replicated for continuity. The user-viewed front end design remained relatively similar to the existing MIS in order to minimize the impact of the system upgrade to its main end users (periphery staff). Data entry and core modules were visually maintained as similar as possible to ease the transition, guarantee maximum user acceptance and adherence to the new system, and minimize training needs (see [Figures 3-6](#)).

A new key functionality is the tiered user administration access. All users are assigned a level based on their function within the system (eg, OD, HF, and VMW), each of which accords a

different type of functionality and data access. Peripheral users can only view their type's operational modules and can only enter data for cases which occur in their geographical coverage area. Access to additional functions, including customizable data dashboards, are also determined by the user category. This allows for the system to be developed as a single front-end structure with different viewing permissions, rather than necessitating multiple structures to be built for different user groups. More importantly, the central server platform also allows multiple users to access the same system components, increasing the potential for user efficiency in data entry. This tiered access was designed to be managed by a central level user (CNM) who can administer changes according to the needs.

Figure 3. Microsoft Access malaria information system (MIS) main page.

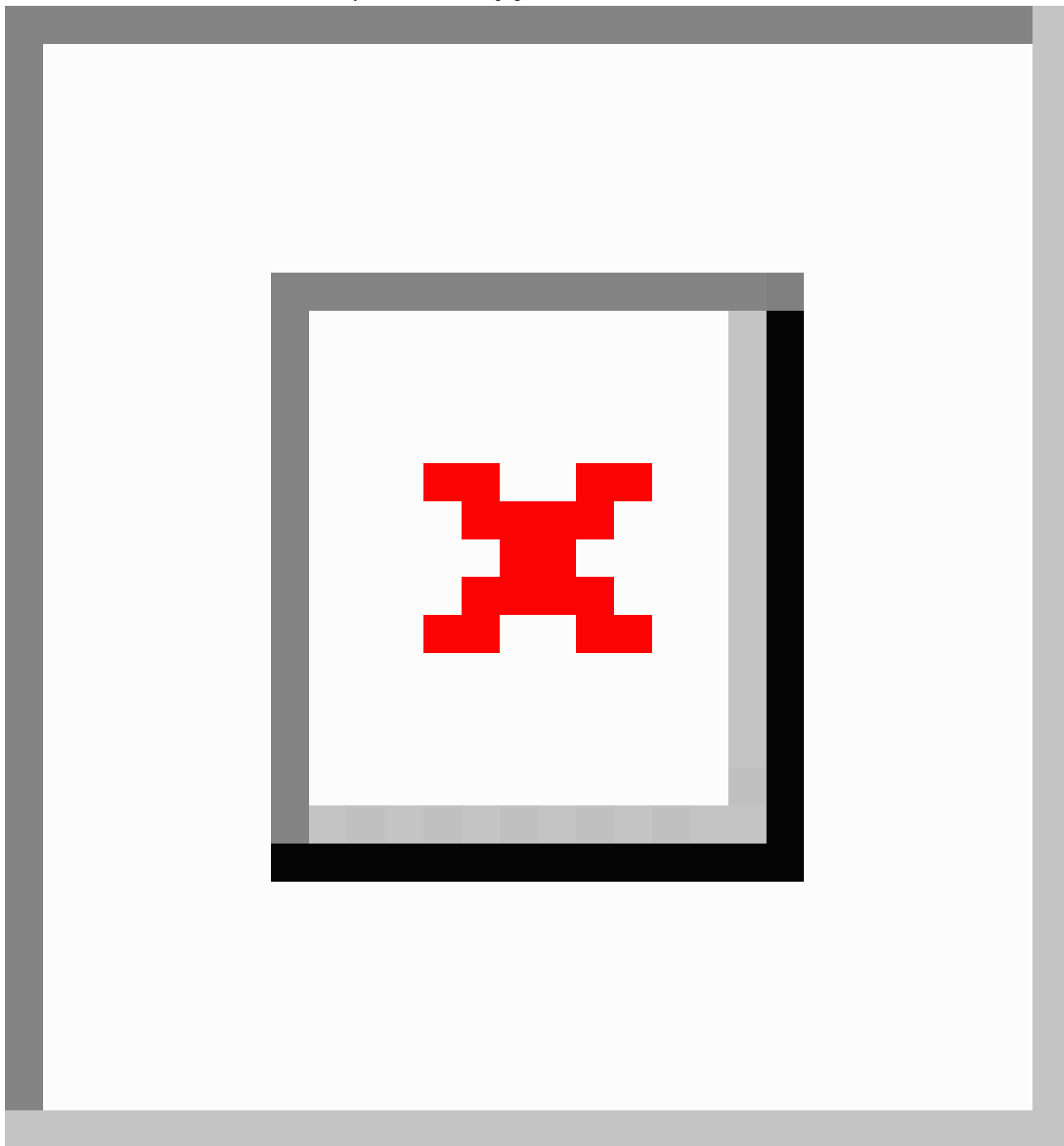


Figure 4. Web-based malaria information system (MIS) main page.

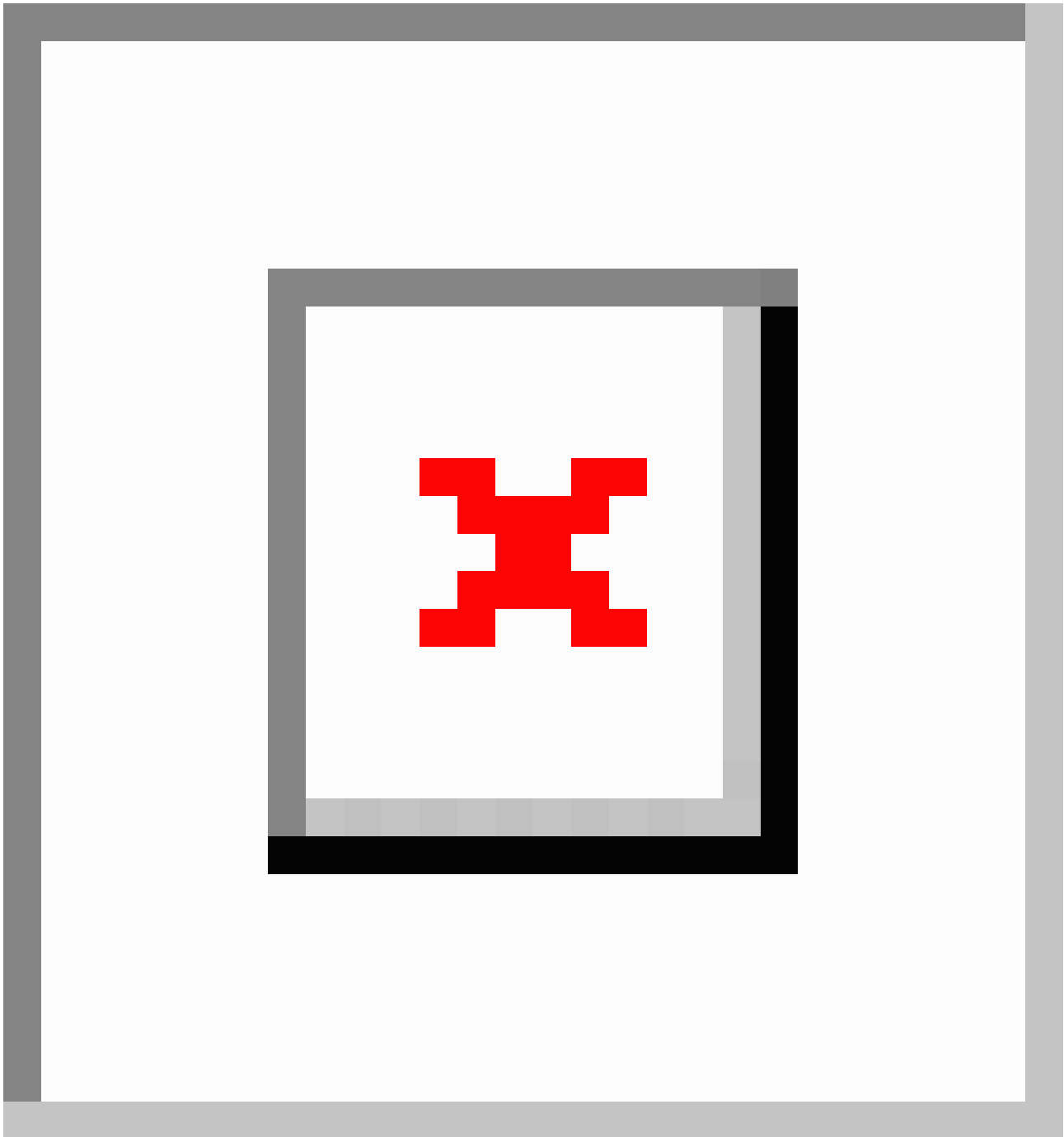


Figure 5. Microsoft Access malaria information system (MIS) village malaria worker (VMW) data entry page.

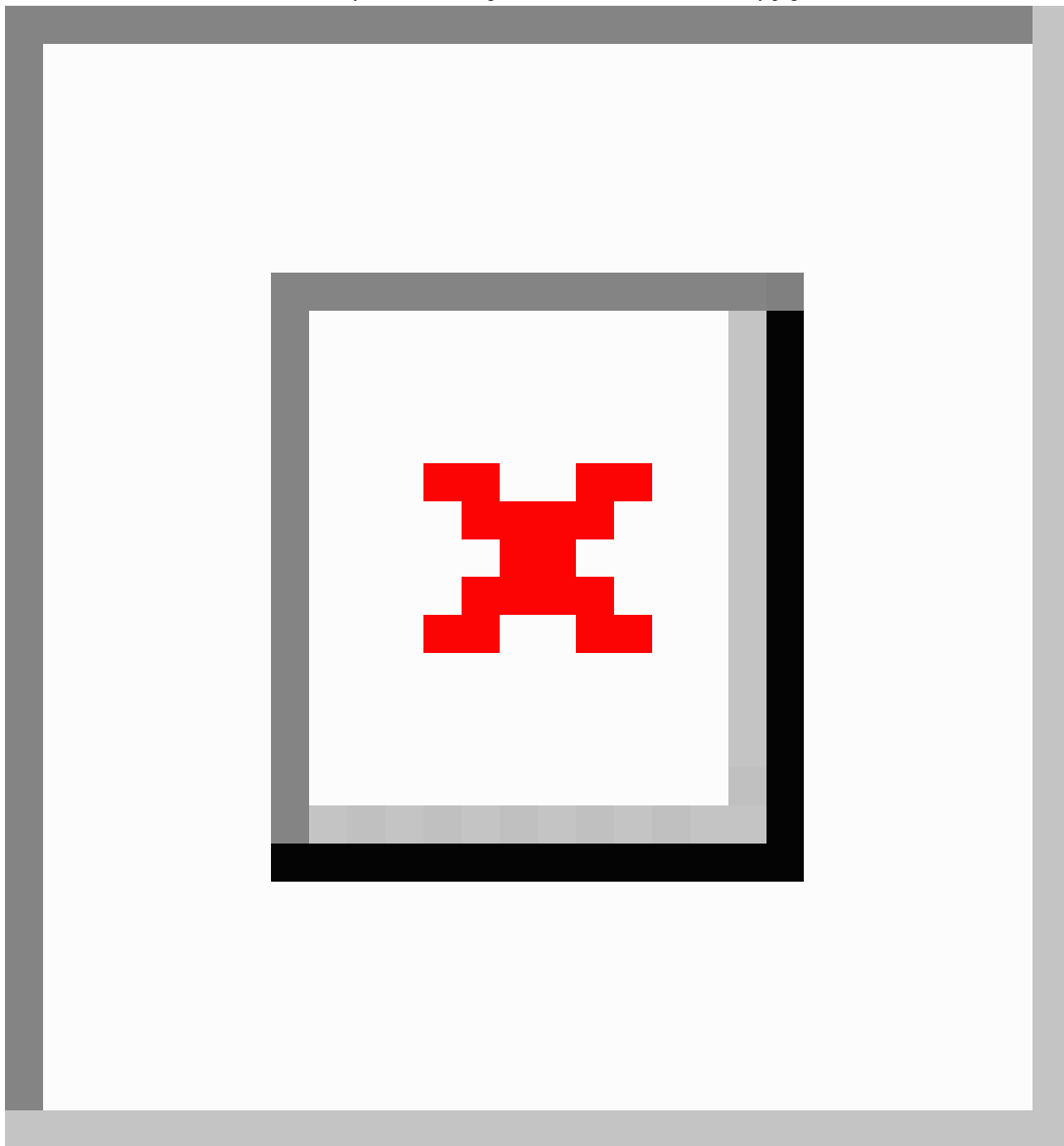
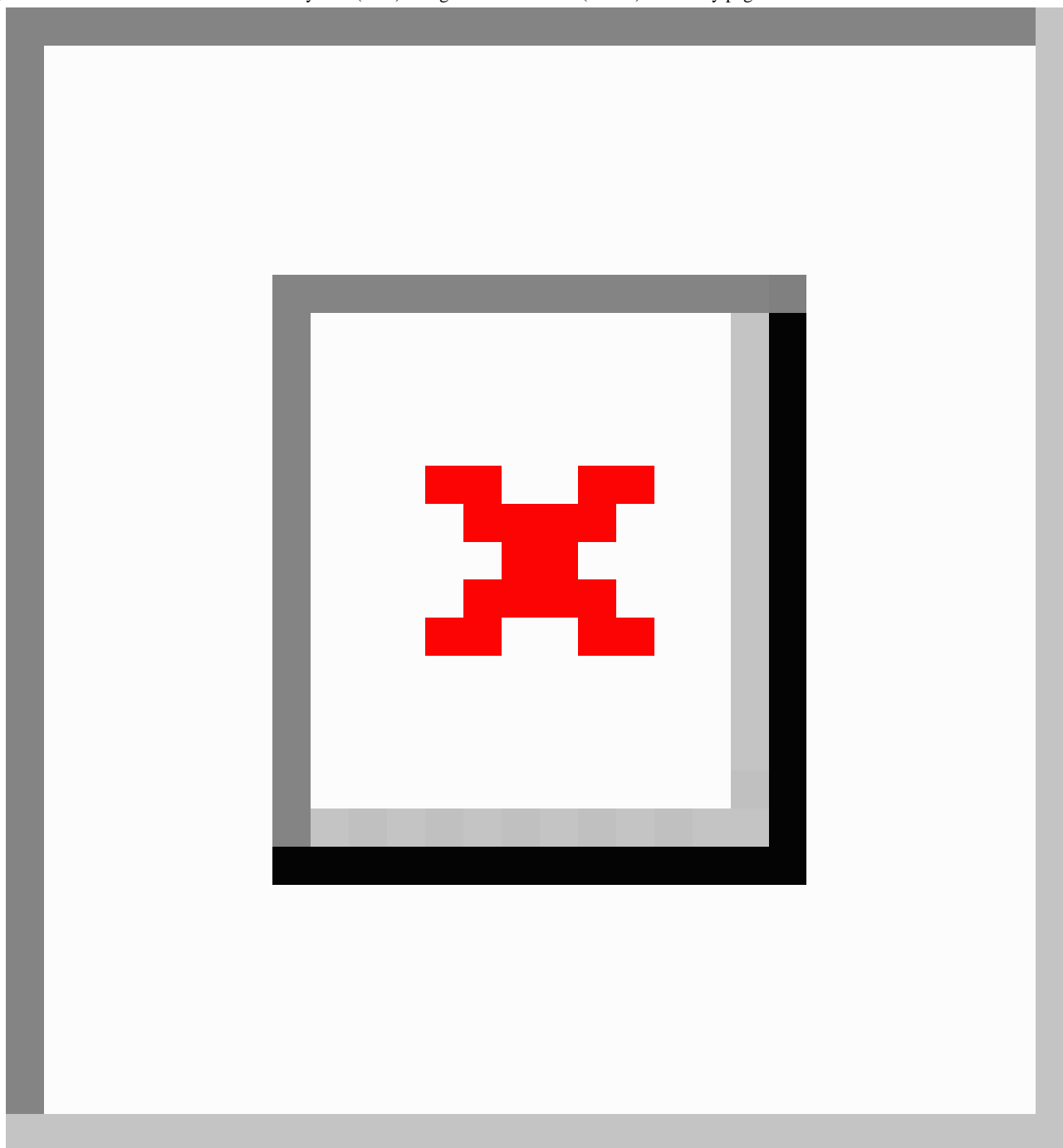


Figure 6. Web-based malaria information system (MIS) village malaria worker (VMW) data entry page.



Operational Attributes

System Piloting and Feedback

The phased approach used in the MIS development will enable health staff at District and Provincial (PHD) levels to be gradually exposed to the system as it develops. This is particularly important with new or unfamiliar features like data visualization through dashboards. It also encourages system ownership by giving additional access to immediately updated information, supporting decision making at lower administrative units. Initially, OD and PHD staff can become familiarized with simplified dashboards designed to display the most relevant information for activities and implementation. As the system develops and malaria elimination activities are scaled up, the

dashboards are expected to integrate more data with greater detail, as well as add customization to understand how their raw data is utilized.

Initial feedback from first-stage pilots at the OD level have already indicated increased levels of systems convenience and ease of usage; additionally, introducing features in stages allows time and response to incorporate suggestions from end users' experience. Although these pilot activities were not comprehensive, they were intended as low-scale feasibility assessments by actual field representatives to obtain critical practical feedback for further improvements [9].

Data Checks and Quality

Ensuring data quality, completeness, and accuracy is critical not only for malaria elimination activities, but also builds crucial aspects of a surveillance system database [5,10]. The enhanced MIS was designed with a built-in capacity to alert and correct data entry errors, reduce potential user mistakes, and improve variable organization. Minimizing data problems at the point of entry reduces central-level time and resources for database cleaning and field monitoring. All modules, including those on mobile platforms have been designed with similar preliminary data quality checks to reduce reporting errors.

The MIS upgrade also automates the capacity to identify missing data or reports depending on the submission timelines defined by CNM. Specific field evaluation indicators are generated for follow-up. Reporting completeness is critical to ensure proper surveillance coverage, and is considered a major priority area for improvement. Previous risk stratification exercises were not automated due to gaps in data completeness, necessitating expensive and time consuming field operations to actively fill them retrospectively through field visits. This functionality is expected to have a wider impact on the system and operations, as it can generate automatic updates to better inform planning, execution, and oversight.

Lastly, the new MIS is expected to automatically integrate data from at least two other non-CNM operated sources (private sector outlets and the national health management information system), as well as potentially others in the future. This regularized data flow within the MIS will not only ease the burden of potentially duplicative data entry from periphery users required to report to multiple sources, but also streamlines data processing within the surveillance system. For these connections to happen, the database must be built on a Web platform to allow data transfer and communications.

Feasibility and Sustainability Attributes

As described previously, replicating MIS modules and front-end functionalities was initially a critical aspect of the Web-based MIS upgrade due to the well-understood challenges of ensuring high quality data from data entry staff. This approach ensured minimal impact to familiarize users with a new system and stayed within the already existing time and financial constraints. In addition, it ensured the MIS upgrade was perceived as a continuation of the previous system rather than a new activity burdening already overstretched workers. These aspects have a significant impact on maintaining sustainability and operation of an already highly accepted system.

In addition, a key principle was to ensure data use and collection at the central level was not interrupted due to a system transition. Hence, the MIS upgrade includes planning for a one-time data migration to be conducted in parallel with the operational capacity for roll-out. This will ensure that data flow will not be diverted into two separate streams on different platforms and databases.

As with the initial MIS development, systems modifications and development are conducted under the leadership of CNM

to ensure that new functionalities respond to country needs. This process is particularly important as Cambodia builds from its previous MIS experience and looks to preserve full ownership of the system. Upgrading a preexisting malaria surveillance system while maintaining characteristics specific to the country context is critical to ensure system acceptance and sustainability.

In countries where a functional case-based surveillance system is already developed and operational, with all levels of the health system invested in its usage, building and adapting that existing system to address the technical upgrade needs is the optimal approach. Rather than set up a brand new system on a different platform, modifying the established system also ensures that ownership is maintained with minimal additional training needs. Other neighboring countries in the region that used case-based reporting systems pre-elimination have followed this phased approach to great success [13,14].

Countries in the malaria control phase without a previously functional surveillance or information system typically opt to incorporate a system following a predeveloped data warehouse model. In this scenario, the central database and most core pre-elimination and elimination functionalities are already created and readily available for minor modifications. However, customization and programming remain a challenge to address country-specific contextual factors such as local language support, incorporation of the existing health system structure, and the need for elimination module adaptation to the national program's operational and procedural decisions. Regardless of the pre-elimination situation, it is still essential to carefully plan a transition phase to ensure that the existing system, whatever its form, can migrate into the new phase with minimal disruption. "One size does not fit all" clearly applies to surveillance and information systems. Despite the pressing global and regional agendas involved with elimination efforts, it is critical that countries undertake a comprehensive needs assessment exercise capturing dimensions for a well-functioning system that go beyond functionality and operationalization, while including feasibility and sustainability attributes to guide country-specific decision making.

Conclusions

As countries move toward malaria elimination, national programs gather experiences and lessons learned in the process of strengthening their surveillance systems and strategies throughout the control-pre-elimination-elimination continuum. Sharing these experiences not only facilitates the establishment of "best practices" but also accelerates worldwide and regional malaria elimination efforts. Countries at different transitioning stages have the benefit of learning from and incorporating or modifying methods implemented in other countries which have already passed through the same phases. Cambodia's experience in developing and upgrading its MIS to remain responsive to country-specific needs demonstrates the necessity for considering functionality, operationalization, feasibility, and sustainability of an information system in the context of malaria elimination.

Acknowledgments

The authors acknowledge the continuing assistance, support, and cooperation from the National Centre for Parasitology, Entomology and Malaria Control (CNM) in Cambodia.

These activities are supported by funding from the Global Fund for AIDS, TB, and Malaria (GFATM). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of its sponsors.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Malaria information system upgrade phasing.

[[PDF File \(Adobe PDF File\), 34KB - publichealth_v3i2e30_app1.pdf](#)]

References

1. World Health Organization. World Malaria Report 2015. Geneva: WHO; 2015.
2. National Center for Entomology, Parasitology and Malaria Control (CNM). The National Strategic Plan for Elimination of Malaria in the Kingdom of Cambodia 2011-2025. Phnom Penh, Cambodia: Ministry of Health of Kingdom of Cambodia; 2011.
3. National Center for Entomology, Parasitology and Malaria Control (CNM). Cambodia Malaria Elimination Action Framework 2016-2020. Phnom Penh, Cambodia: Ministry of Health, Kingdom of Cambodia; 2016.
4. Mellor S. Malaria Consortium. 2013. Moving towards malaria elimination: developing innovative tools for malaria surveillance in Cambodia URL: <http://www.malariaconsortium.org/pages/learning-papers.htm> [accessed 2017-05-11] [WebCite Cache ID 6qNVAzGDZ]
5. Ohrt C, Roberts KW, Sturrock HJ, Wegbreit J, Lee BY, Gosling RD. Information systems to support surveillance for malaria elimination. *Am J Trop Med Hyg* 2015 Jul;93(1):145-152 [FREE Full text] [doi: [10.4269/ajtmh.14-0257](https://doi.org/10.4269/ajtmh.14-0257)] [Medline: [26013378](https://pubmed.ncbi.nlm.nih.gov/26013378/)]
6. Cox J, Sovannaroth S, Dy SL, Ngor P, Mellor S, Roca-Feltrer A. Novel approaches to risk stratification to support malaria elimination: an example from Cambodia. *Malar J* 2014 Sep 19;13:371. [doi: [10.1186/1475-2875-13-371](https://doi.org/10.1186/1475-2875-13-371)] [Medline: [25233886](https://pubmed.ncbi.nlm.nih.gov/25233886/)]
7. World Health Organization. Disease Surveillance for Malaria Elimination: an operational manual. Geneva: WHO; 2012:1-68.
8. Khamsiriwatchara A, Sudathip P, Sawang S, Vijakadge S, Potithavoranan T, Sangvichean A, et al. Artemisinin resistance containment project in Thailand. (I): Implementation of electronic-based malaria information system for early case detection and individual case management in provinces along the Thai-Cambodian border. *Malar J* 2012 Jul 29;11:247 [FREE Full text] [doi: [10.1186/1475-2875-11-247](https://doi.org/10.1186/1475-2875-11-247)] [Medline: [22839508](https://pubmed.ncbi.nlm.nih.gov/22839508/)]
9. Barclay VC, Smith RA, Findeis JL. Surveillance considerations for malaria elimination. *Malar J* 2012 Aug 31;11:304 [FREE Full text] [doi: [10.1186/1475-2875-11-304](https://doi.org/10.1186/1475-2875-11-304)] [Medline: [22938625](https://pubmed.ncbi.nlm.nih.gov/22938625/)]
10. malERA Consultative Group on Basic Science Enabling Technologies. A research agenda for malaria eradication: basic science and enabling technologies. *PLoS Med* 2011 Jan 25;8(1):e1000399 [FREE Full text] [doi: [10.1371/journal.pmed.1000399](https://doi.org/10.1371/journal.pmed.1000399)] [Medline: [21311584](https://pubmed.ncbi.nlm.nih.gov/21311584/)]
11. Robert B, World Health Organization. Manual of epidemiology and epidemiological services in malaria programmes. Geneva: WHO; 1968:1-223.
12. Chisha Z, Larsen, DA, Burns M, Miller JM, Chirwa J, Mbwili C, et al. RESEARCH open access enhanced surveillance and data feedback loop associated with improved malaria data in Lusaka, Zambia. *Malaria Journal* 2015;14:222. [doi: [10.1186/s12936-015-073](https://doi.org/10.1186/s12936-015-073)]
13. Ma S, Lawpoolsri S, Soonthornworasiri N, Khamsiriwatchara A, Jandee K, Taweeseeneepitch K, et al. Effectiveness of implementation of electronic malaria information system as the national malaria surveillance system in Thailand. *JMIR Public Health Surveill* 2016;2(1):e20 [FREE Full text] [doi: [10.2196/publichealth.5347](https://doi.org/10.2196/publichealth.5347)] [Medline: [27227156](https://pubmed.ncbi.nlm.nih.gov/27227156/)]
14. Cao J, Sturrock Hugh J W, Cotter C, Zhou S, Zhou H, Liu Y, et al. Communicating and monitoring surveillance and response activities for malaria elimination: China's "1-3-7" strategy. *PLoS Med* 2014 May;11(5):e1001642 [FREE Full text] [doi: [10.1371/journal.pmed.1001642](https://doi.org/10.1371/journal.pmed.1001642)] [Medline: [24824170](https://pubmed.ncbi.nlm.nih.gov/24824170/)]

Abbreviations

CNM: National Centre for Parasitology, Entomology, and Malaria Control

GMS: Greater Mekong Subregion

HC: health center

HF: health facility
HMIS: health management information system
MEAF: malaria elimination action framework 2016-2020
MIS: malaria information system
MOH: ministry of health
OD: operational district
PHD: provincial health department
PPM: public private mix
RACD: reactive case detection
VMW: village malaria worker

Edited by M Focsa; submitted 04.11.16; peer-reviewed by S Ursoniu, S Lawpoolsri, J Kaewkungwal; comments to author 06.12.16; revised version received 01.03.17; accepted 29.03.17; published 14.06.17.

Please cite as:

*Fu C, Lopes S, Mellor S, Aryal S, Sovannaroeth S, Roca-Feltrer A
Experiences From Developing and Upgrading a Web-Based Surveillance System for Malaria Elimination in Cambodia
JMIR Public Health Surveill 2017;3(2):e30
URL: <http://publichealth.jmir.org/2017/2/e30/>
doi: [10.2196/publichealth.6942](https://doi.org/10.2196/publichealth.6942)
PMID: [28615155](https://pubmed.ncbi.nlm.nih.gov/28615155/)*

©Clementine Fu, Sérgio Lopes, Steve Mellor, Siddhi Aryal, Siv Sovannaroeth, Arantxa Roca-Feltrer. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 14.06.2017. 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

Associations Between Maternal Depressive Symptoms and Nonresponsive Feeding Styles and Practices in Mothers of Young Children: A Systematic Review

Ana Cristina Lindsay^{1*}, DDS, MPH, DrPH; Tatiana Mesa^{1*}, BSc, MS; Mary L Greaney^{2*}, AB, MPH, PhD; Sherrie F Wallington^{3*}, BSHE, MA, PhD; Julie A Wright^{1*}, BS, MS, PhD

¹Exercise and Health Sciences, College of Nursing and Health Sciences, University of Massachusetts Boston, Boston, MA, United States

²Department of Kinesiology, Kinesiology/Health Studies, University of Rhode Island, Kingston, RI, United States

³Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Georgetown University, Washington, DC, United States

*all authors contributed equally

Corresponding Author:

Ana Cristina Lindsay, DDS, MPH, DrPH

Exercise and Health Sciences

College of Nursing and Health Sciences

University of Massachusetts Boston

100 Morrissey Boulevard

Boston, MA, 02125

United States

Phone: 1 617287579

Fax: 1 6172877500

Email: Ana.Lindsay@umb.edu

Abstract

Background: Childhood obesity is a significant global public health problem due to increasing rates worldwide. Growing evidence suggests that nonresponsive parental feeding styles and practices are important influences on children's eating behaviors and weight status, especially during early childhood. Therefore, understanding parental factors that may influence nonresponsive parental feeding styles and practices is significant for the development of interventions to prevent childhood obesity.

Objective: The objectives of this systematic review were to (1) identify and review existing research examining the associations between maternal depressive symptoms and use of nonresponsive feeding styles and practices among mothers of young children (2-8 years of age), (2) highlight the limitations of reviewed studies, and (3) generate suggestions for future research.

Methods: Using the PRISMA (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) guidelines, six electronic academic databases were searched for peer-reviewed, full-text papers published in English between January 2000 and June 2016. Only studies with mothers 18+ years old of normally developing children between 2 and 8 years of age were included. Of the 297 citations identified, 35 full-text papers were retrieved and 8 were reviewed.

Results: The reviewed studies provided mixed evidence for associations between maternal depressive symptoms and nonresponsive feeding styles and practices. Two out of three studies reported positive associations with nonresponsive feeding styles, in that mothers with elevated depressive symptoms were more likely than mothers without those symptoms to exhibit uninvolved and permissive or indulgent feeding styles. Furthermore, results of reviewed studies provide good evidence for association between maternal depressive symptoms and instrumental feeding (3 of 3 reviewed studies) and nonresponsive family mealtime practices (3/3), but mixed evidence for pressuring children to eat (3/6) and emotional feeding (1/3). In addition, evidence for the association between maternal depressive symptoms and restricting child food intake was mixed: one study (1/6) found a positive association; two studies (2/6) found a negative association; whereas one study (1/6) found no association.

Conclusions: This review indicates that the results of studies examining the associations between maternal depressive symptoms and parental feeding styles and practices are mixed. Limitations of studies included in this review should be noted: (1) the use of a diverse set of self-report questionnaires to assess parental feeding practices is problematic due to potential misclassification and makes it difficult to compare these outcomes across studies, thus caution must be taken in drawing conclusions; and (2) the majority of included studies (6/8) were cross-sectional. There is a need for additional longitudinal studies to disentangle the influence of depression on parental feeding styles and practices. Nevertheless, given that depressive symptoms and feeding styles

and practices are potentially modifiable, it is important to understand their relationship to inform obesity prevention interventions and programs.

(*JMIR Public Health Surveill* 2017;3(2):e29) doi:[10.2196/publichealth.6492](https://doi.org/10.2196/publichealth.6492)

KEYWORDS

maternal depression; child; feeding behavior; practices; feeding styles; obesity

Introduction

Childhood obesity is an important global public health issue due to existing prevalence and increasing rates worldwide [1,2]. The increasing prevalence of childhood obesity in young children is particularly concerning, given the evidence that children's weight status is associated with weight status in adulthood, making early childhood a critical period for prevention of overweight and obesity [2-4]. Consequently, identifying modifiable factors associated with increased risk of early childhood obesity is a priority [1-4].

Early childhood is an important period of growth and development that influences one's health during childhood and beyond [3-7]. It is when the foundations for healthful eating habits that have long-lasting implications for weight status and related comorbidities are established [4,5,7-9]. Several parental characteristics are associated with children's risk of overweight and obesity including parents' weight status [1-4], sociodemographic and economic characteristics (eg, income, education) [1-4], and mental health status (eg, depression) [10-19]. Parents, especially mothers, influence their children's development and maintenance of eating habits and food preferences [2,4-9,10]. Parental feeding styles and parental feeding practices have been identified as particularly important influences on children's eating behaviors during early childhood [5,7,8,9-16,20-24].

Parental feeding style, the overarching feeding strategy parents adopt during feeding situations [9,11,20,21], has been conceptualized as having two main dimensions: demandingness (also defined as control) and responsiveness (also defined as warmth). Within these two dimensions, there are four parental feeding styles typologies: (1) authoritative (high level of demandingness and high level of responsiveness), (2) authoritarian (high level of demandingness and low level of responsiveness), (3) indulgent or permissive (low demandingness and high responsiveness), and (4) uninvolved or neglectful (low demandingness and low responsiveness). Parental feeding practices are specific behaviors that parents use to influence the amount and/or type of food a child eats and include monitoring and controlling food intake, pressuring to eat, instrumental and emotional feeding, and so on [9,11,21-25].

Family meals and family mealtime practices are key family routines relevant to obesity prevention [26]. Family mealtime practices encompass habits and processes that a family engages in around eating together [27]. Family mealtime may offer several benefits to children's health and development such as helping children develop healthful eating patterns and weight status [27-29].

Understanding factors that may be associated with parental feeding styles and practices, and family mealtime practices that are unintentionally detrimental to children's development of healthful eating habits is of great importance to the development of interventions to prevent child obesity. Providing parents with guidance on healthful feeding styles and practices will help children develop healthful eating habits and, ultimately, maintain a healthy weight status [2-4,8,30,31].

Research suggests that mental health status of the parents may influence the weight status of their child through parental feeding styles and practices [7,10-13]. Mental health conditions (eg, depression and depressive symptoms) among mothers of young children are increasingly recognized as an important public health concern [32-36]. According to the National Institute of Mental Health one in seven women of reproductive age are affected by depression, and 15% of women in the United States experience postpartum depression [32,36]. Depressive symptoms can affect mothers' sensitivity and responsiveness to their children and can contribute to less engaged or responsive mother-child interactions as well as a higher use of disengaged (eg, uninvolved and permissive/indulgent) or controlling (eg, authoritarian) parenting behaviors [10-14]. In addition, elevated depressive symptoms such as low energy and diminished pleasure in activities may contribute to decreased maternal-child involvement [10-13], with mothers choosing strategies for coping that require less cognitive effort [11-13].

Given the high prevalence (15%-38%) of depression and depressive symptoms among women of childbearing age [32-36] and increasing evidence linking maternal depressive symptoms to nonresponsive feeding styles and practices related to the risk of childhood obesity [11-14,30,37], the objectives of this systematic literature review were to (1) identify and review existing research examining the associations between maternal depressive symptoms and use of nonresponsive feeding styles and practices among mothers of young children (2-8 years of age), (2) highlight the limitations of reviewed studies, and (3) generate suggestions for future research.

Methods

Study Design

We conducted this review by following the reporting guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement [38]. These guidelines include a four-phase flow diagram to guide the inclusion and exclusion of research papers [38]. In addition, the guidelines provide a 27-item checklist outlining standards per review section (eg, title, abstract, introduction, methods, results, discussions, funding) to ensure that reviews are systematically conducted and reported [38].

Search Strategy

We searched six electronic databases: Science Direct, PubMed, PsycINFO, PsycARTICLES, Medline, and Cumulative Index to Nursing and Allied Health Literature (CINAHL) between April and June 21, 2016. The search was limited to full-text, peer-reviewed articles published in English between January 2000 and June 2016. Key search terms included: (1) child* OR pediatric*; (2) maternal depress* OR maternal depress* symptoms OR maternal “depressive symptoms”; and (3) “feeding practices” OR “feeding behavior” OR feeding strateg* OR feeding style (see [Figure 1](#)). Two authors (ACL, TM) independently examined the titles and abstracts of all citations, and the citations were excluded when both authors determined that the study did not meet the inclusion criteria. The same two authors independently reviewed the retrieved articles and identified studies to be included in this systematic review. They also searched the reference lists [39-46] and other review studies focusing on infants and/or children older than 8 years of age [7,10-16,25,31,47-49].

Study Selection

This systematic review was limited to peer-reviewed studies that included mothers 18+ years old of normally developing children (ie, not born preterm, not diagnosed with physical or mental complications) between 2 and 8 years of ages (ie, early childhood). We identified studies that (1) examined the association between maternal depression and/or maternal depressive symptoms (independent variable) and parental feeding styles, parental feeding practices, and/or family mealtime practices (outcome variables), and (2) measured maternal depression or depressive symptoms with a validated questionnaire or scale at any period prior to or following childbirth. Studies that focused on special groups (eg, teen mothers, children born pre-term or low birth weight, or special needs such as cerebral palsy) or populations with health concerns (eg, mothers diagnosed with HIV) were excluded. Studies that used qualitative methods exclusively were also excluded to simplify comparison of findings across studies. Additionally, studies that focused exclusively on breastfeeding and/or complementary feeding practices were excluded as previous review papers have examined the association between maternal depressive symptoms and infant feeding practices [15,25,31,47,49].

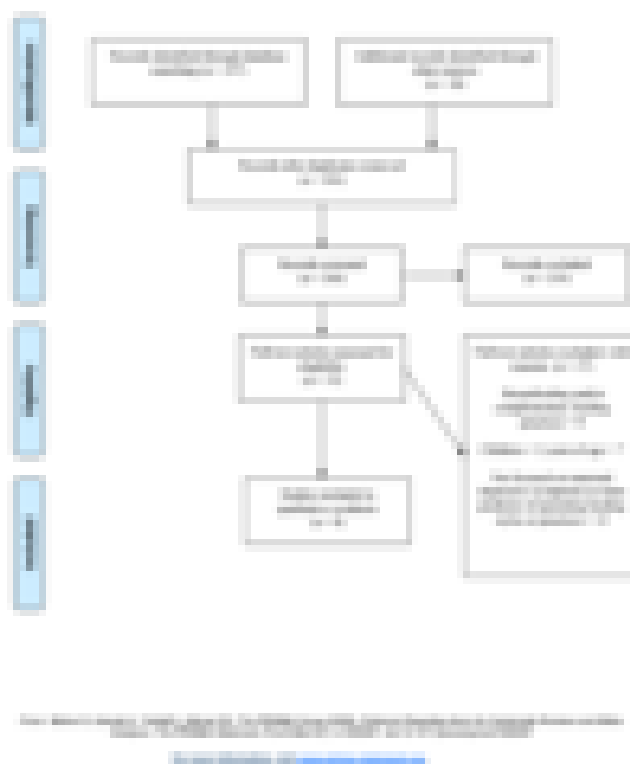
Data Extraction and Data Synthesis

Using the search strategy illustrated in [Figure 1](#) (PRISMA flow diagram), we identified eight observational studies meeting eligibility requirements [39-46]. Two authors (ACL, TM) independently read and completed an article extraction form for all articles. The data extraction form gathered the following information: (1) authors, (2) study setting, (3) sample size, (4) participant characteristics, (5) study design, (6) study aim(s), (7) measure (s) of maternal depressive symptoms, (8) measure(s) of parental feeding styles and practices and family mealtime practices, and (9) study results. The two authors who completed the data extraction forms compared their results and discussed discrepancies, which were resolved with feedback from a third author.

This review extracted data on associations between maternal depressive symptoms (exposure) and parental feeding styles, parental feeding practices, and/or family mealtime practices (outcomes) and summarized the findings. Due to the range of study designs, assessment of exposure, and outcomes, conducting a meta-analysis of the data was not appropriate, and results of this review are presented as a narrative summary.

Quality Assessment of Included Studies

The first author (ACL) assessed the quality of reviewed studies using a modified version of the Strengthening in the Reporting of Observational Studies in Epidemiology (STROBE) statement [50], which were confirmed by two authors (MLG, JAW). STROBE is an international, collaborative initiative of epidemiologists, methodologists, statisticians, researchers and journal editors who have a common goal of strengthening reporting of observational studies in epidemiology. The combined STROBE checklist for cohort, case-control, and cross-sectional studies includes 22 items [50]. For this review, we included 11 of these items to (1) identify potential sources of bias, and (2) identify possible methodological areas that were insufficiently addressed (see [Multimedia Appendix 1](#)). Each question was designed to be answered either “yes” or “no,” with a score of 1 assigned to “yes” response, and a score of 0 to “no” response (range of scores 0-11). Total scores were then used to assign a rating of the study as strong (score > 8), moderate (score between 8-6), or weak (score between 5-0).

Figure 1. PRISMA Flow Diagram.

Results

Search Results

We identified 297 unique citations and two authors (ACL, TM) independently examined the titles and abstracts of the identified citations. We excluded 272 citations that did not meet the eligibility criteria, and 35 full-text articles were selected for detailed review and assessed by two authors independently. Eight studies that met the eligibility criteria were included in this systematic review (see [Figure 1](#)). Studies included in this review fell into three categories, and this review is organized by the study's purpose. The three purpose categories were to examine the associations between: (1) maternal depressive symptoms and nonresponsive parental feeding styles, (2) maternal depressive symptoms and nonresponsive parental feeding practices, and (3) maternal depressive symptoms and nonresponsive family mealtime practices.

Study Characteristics

Eligible studies examined maternal depressive symptoms [51-56] and parental feeding styles [57] and feeding practices [58-64] using a wide variety of tools. Of the 8 included studies, 3 focused on feeding styles [39,41,46], 6 on nonresponsive feeding practices [40,41,43-46] and 3 on nonresponsive family mealtime practices [41,42,44]. Summary characteristics of included studies are presented in [Multimedia Appendix 2](#),

whereas synthesized information on methodology and main findings are presented in [Multimedia Appendix 3](#).

Study Quality

The quality of the reviewed studies varied (See [Multimedia Appendix 1](#)). Out of the 8 included studies, 6 were assigned a rating of > 8 (strong), and 2 were rated between 8-6 (moderate). Most of the reviewed studies presented methodological limitations: (1) nearly all (6/8) used cross-sectional study designs that prohibit determining causal inferences between maternal depressive symptoms and child-feeding practices and styles [39,41,42,44-46], (2) all studies (8/8) relied on self-reporting of maternal depressive symptoms [39-46], and (3) the majority (6/8) relied on self-reporting of parental feeding styles, parental feeding practices and family mealtime practices [39-42,44,46]. Furthermore, it was not always possible to determine whether analyses controlled for potential confounding factors and analyses were not stratified by child gender.

Moreover, maternal depressive symptoms [51-56] and parental feeding practices [58-64] were assessed by an array of instruments, which made it difficult to compare results across studies. For example, across the 8 reviewed studies, 6 different instruments were used to measure maternal depressive symptoms [51-56], whereas 7 instruments were used to assess parental feeding practices [58-64], 2 instruments were used to assess family mealtime practices [59,61], and 1 to assess parental feeding styles [57].

Associations Between Maternal Depressive Symptoms and Nonresponsive Feeding Styles

Out of the 8 included studies [39-46], 3 studies [39,41,46] examined the association between maternal depressive symptoms and nonresponsive feeding styles. Two [39,41] of these studies used the validated Center for Epidemiological Study Depression (CES-D) instrument [51], whereas 1 used the Depression Anxiety Stress Scale 21-item (DASS-21) [53] to assess maternal depressive symptoms. All three studies were cross-sectional [39,41,46] and used the validated Caregiver's Feeding Styles Questionnaire (CFSQ) [57] to assess parental feeding styles. The CFSQ is a 19-item valid, reliable measure developed for use with caregiver's of preschool-age children [57].

Of the 3 studies examining associations between maternal depressive symptoms and nonresponsive parental feeding styles in adjusted analysis, only 2 studies found positive associations [39,41], with mothers with elevated depressive symptoms being more likely to exhibit uninvolved [39,41] and permissive styles [41] than those without these symptoms.

A cross-sectional study [41] conducted in the United States with a sample of low-income mothers of whom approximately 30% were Hispanics, found that mothers reporting elevated depressive symptoms reported using more demandingness (eg, encouragement or discouragement of child's eating behaviors) and permissive feeding styles (eg, fewer authority narratives about feeding) than mothers without elevated depressive symptoms after adjusting for potential confounders. One cross-sectional study [39] conducted in the United States with a sample of low-income mothers, of whom approximately 55% were Hispanic and 45% were African-American, found that after adjusting for potential confounders, mothers employing an uninvolved feeding style (a permissive feeding style) reported less positive affect and more parenting stress than mothers who used authoritative, authoritarian, or indulgent/permissive feeding styles. In addition, mothers with elevated depressive symptoms were more likely to present low authority in child feeding [39].

One cross-sectional study [46] conducted in Australia found that mothers who reported experiencing higher levels of depressive symptoms also reported using higher levels of the authoritarian feeding style. In adjusted analysis, however, none of the maternal psychosocial well-being variables independently contributed to the prediction of authoritarian parental feeding style.

Associations Between Maternal Depressive Symptoms and Nonresponsive Feeding Practices

Out of the 8 reviewed studies, 6 [40,41,43-46] examined the association between maternal depressive symptoms and nonresponsive feeding practices. Four studies [41,44-46] employed cross-sectional designs and two used longitudinal designs [40,43]. The 6 studies used five different instruments [51,53-56] to assess maternal depressive symptoms, with the Edinburgh Postnatal Depression Scale (EPDS) [53] being used in 2 [40,43] and the Depression Anxiety Stress Scales 21-item (DASS-21) [54] also being used in 2 studies [43,46]. The validated Child Feeding Questionnaire (CFQ) [58] was the

instrument most used to assess parental feeding practices (5/6). In addition, five other validated instruments were used across the 6 studies [59-64]. All 6 reviewed studies [40,41,43-46] provided information on the reliability and validity of the instruments used to assess both maternal depressive symptoms and parental feeding practices (see [Multimedia Appendix 3](#)).

In summary, 3 studies reported positive associations between maternal depressive symptoms and use of instrumental feeding (eg, using food as a reward) (3/3) [40,44,45], and pressure to eat (3/6) [40,41,45]. One study reported positive associations between maternal depressive symptoms and restriction of child's food intake (1/6) [40], and emotional (eg, using food to manage child's mood) feeding (1/3) [40]. In addition, 2 studies (2/6), one with a cross-sectional design [44] and one using a longitudinal design [43], reported negative associations between maternal depressive symptoms and restriction of child food intake. In contrast, a cross-sectional study (1/6) [45] found that elevated depressive symptoms were not associated with restriction of child food intake. Moreover, 1 study (1/4) reported that elevated depressive symptoms were negatively associated with monitoring of child food intake [43].

Pressure to Eat

Six of the reviewed studies examined the association between maternal depressive symptoms and mothers' use of pressure to get their children to eat [40,41,43-46], and all found positive associations between elevated maternal depressive symptoms and pressure to eat in unadjusted analyses. However, only three studies (3/6), one using a longitudinal design [40] and two using cross-sectional designs [41,45], reported significant positive associations between maternal depressive symptoms and pressure to eat after adjusting for several key child (eg, age, gender, child body mass index) and maternal characteristics (eg, age, BMI, race, income, educational level). Results of these 3 studies [40,41,45] indicated that mothers reporting elevated depressive symptoms were more likely to report pressuring their children to eat than mothers without elevated depressive symptoms after adjusting for potential confounding factors.

Restriction of Child's Food Intake

Six of the reviewed studies [40,41,43-46] examined associations between maternal depressive symptoms and restrictions in child food intake. Three of these studies [40,41,46] reported positive associations between maternal general depressive symptoms and restriction in feeding in unadjusted analyses. However, after adjusting for key maternal (age, education, BMI) and child covariates (age, gender, BMI at 4 months and feeding mode at 4 months), only one longitudinal study [40] found that maternal general depressive symptoms were associated with the restriction of children's food intake. In contrast, a longitudinal study [43] found that high depressive symptoms predicted less maternal use of restriction. Likewise, a study [44] using a cross-sectional design found a negative association between maternal depressive symptoms and use of restriction of child's food intake, with mothers reporting mild and moderate to severe symptoms were less likely to restrict their child's intake than mothers not reporting depressive symptoms. Moreover, a cross-sectional study [45] determined that maternal depressive symptoms were not predictive of mothers' restrictive feeding practices.

Monitoring of Child Food Intake

Four of the reviewed studies, two employing longitudinal study design [40,43] and two cross-sectional designs [41,44], examined associations between maternal depressive symptoms and monitoring of child food intake. Of the 4 studies, 1 longitudinal follow-up study found that maternal depressive symptoms partially negatively predicted monitoring of child food intake [43].

Instrumental and Emotional Feeding Practices

Three [40,44,45] studies examined the relationship between maternal depressive symptoms and instrumental feeding practices (eg, using food as a reward, increased use of incentives) and emotional feeding (eg, using food to manage child mood). All 3 studies found a positive association between maternal depressive symptoms and instrumental feeding [40,44,45], whereas one (1/3) found a positive association between maternal depressive symptoms and emotional feeding [40]. One longitudinal study found that mothers with elevated depressive symptoms were more likely to employ both instrumental and emotional feeding practices adjusting for maternal and child covariates [40] than mothers with low or without depressive symptoms. Additionally, 2 studies using cross-sectional designs [44,45], one conducted in the United States [44], and one in England [45], found positive associations between maternal depressive symptoms and the use of instrumental feeding practices (eg, use of food as reward, or use of incentive and conditions to get child to eat). In adjusted analysis, higher maternal depressive symptoms were significantly associated with use of food as a reward [44] and with greater use of incentives or conditions to eat [45].

Associations Between Maternal Depressive Symptoms and Nonresponsive Family Mealtime Practices

Out of the 8 reviewed studies [39-46] 3 [41,42,44], all of which employed cross-sectional designs and were conducted in the United States with low-income mothers, examined the association between maternal depressive symptoms and a number of nonresponsive family mealtime practices. Two [41,42] of these studies used the Center for Epidemiological Study Depression (CES-D) instrument [51], whereas one [44] used the Patient Health Questionnaire-9 (PHQ-9) [55] to assess maternal depressive symptoms. Two different instruments [59,61] were used to assess family mealtime practices by 2 [42,44] of the 3 studies, and 1 study [41] used both semistructured narrative interview and videotaped observations of mother-child feeding situations.

All 3 studies determined that children in households with mothers with elevated depressive symptoms were more likely to be exposed to less optimal mealtime practices and routines than children in households with mothers with low or without any depressive symptoms [41,42,44]. Moreover, all 3 studies found that mothers reporting elevated depressive symptoms were more likely to report nonresponsive feeding practices which were associated with both uninvolved (eg, mother not being present during meals, child skipping breakfast, child eating while watching television) and permissive (eg, lower levels of

maternal control over child eating routines, greater child choice over snacking) feeding styles.

One study [41] found that in households of mothers with elevated depressive symptoms, children were less likely to eat at the kitchen or dining table, the television was more likely to be audible during meals, and children were less likely to eat with their mothers. Similarly, 1 study [42] found that maternal depression was significantly associated with lower maternal presence when the child ate, lower levels of maternal control over child eating routines, greater child choice over snacking, and fewer optimal mealtime practices than in homes of mothers without higher depression scores [42]. Likewise, 1 study [44] found in adjusted analyses that mothers reporting mild depressive symptoms were more likely to have children who consumed sweetened drinks daily, who did not eat breakfast regularly, and who ate out in restaurants 3 or more times per week than mothers without depressive symptoms.

Discussion

Principal Findings

The aim of this systematic review was to identify and review existing research examining associations between maternal depressive symptoms and nonresponsive parental feeding styles and parental feeding practices in mothers of young children. The 8 reviewed studies provide mixed support for associations between maternal depressive symptoms and nonresponsive feeding styles, feeding practices, and family mealtime practices. Uninvolved and permissive feeding styles, and feeding practices use of instrumental feeding (eg, use of food as reward) and pressuring children to eat were the most consistently associated with depressive symptoms among studies included in this review. In addition, maternal depressive symptoms were associated with uninvolved and permissive family mealtime practices.

Across the reviewed studies, elevated maternal depressive symptoms were most often associated with uninvolved and permissive parental feeding styles [39,41]. Two of the three studies examining associations between maternal depressive symptoms and nonresponsive feeding styles (uninvolved and permissive) found a positive association [39,41]. These findings concur with results of studies [11,12] and review papers [10,15,25,47,49,65] conducted among mothers with infants. Furthermore, available evidence from the extant literature on maternal mental health and parenting suggests that maternal mental health issues may impair mothers' responsiveness to, and interactions with, their children. The reduced interaction may manifest in nonresponsive, more controlling, and less-sensitive parenting [10,29,37,47,49,60]. In addition, research suggest that elevated depressive symptoms may contribute to decreased maternal-child interactions [30,66], with mothers being less responsive to their children and choosing strategies for coping that require less cognitive effort [11-13].

Evidence for associations between maternal depressive symptoms and nonresponsive maternal feeding practices was mixed across reviewed studies [40,41,43-46]. Instrumental

feeding (eg, use of food as reward; 3/3) [40,44,45] and pressure to eat (3/6) [40,41,45] were the most consistently nonresponsive feeding practices associated with elevated depressive symptoms across the studies included in this review. Evidence from studies [11,12,16] and systematic reviews [25,47] with mothers of infants and toddlers suggests that mothers experiencing elevated depressive symptoms are more likely to use restrictive and controlling feeding practices than mothers without elevated depressive symptoms. Moreover, evidence suggest that mothers experiencing elevated depressive symptoms are less likely to be responsive to their children's cues of hunger and satiety and less likely to respect their child's ability to self-regulate food intake [10-13,66]. Previous studies indicate that nonresponsive feeding practices interfere with a child's natural ability to self-regulate food intake based on hunger and satiety cues [14,19,30,67-69]. Furthermore, research suggests that both parental pressure to eat and feeding restrictions are associated with unrestrained eating and disinhibited eating in later life, excessive weight gain, and increased risk of child obesity in children [5,20,21,67-71].

Studies included in this review provide consistent evidence for the association between maternal depressive symptoms and nonresponsive family mealtime practices [41,42,44]. Children in households with mothers having elevated depressive symptoms were more likely to be exposed to less optimal mealtime practices and routines than children in households with mothers having low or no depressive symptoms [41,42,44]. Elevated depressive symptoms such as low energy and diminished pleasure in activities may contribute to decreased maternal involvement with the child [5,21,67-69,30], resulting in mothers being less responsive to their children and choosing strategies for coping that require less cognitive effort [5]. Suboptimal family mealtime practices have been reported to be associated with children's unhealthy eating habits [72-77], which in turn have been linked to risk of overweight and obesity [72-77].

Limitations and Strengths

Our evaluation of the methodologies of studies included in this systematic review suggests some limitations, and therefore caution in the interpretation of study findings. The majority (6/8) of studies used cross-sectional study designs precluding a causal assessment of associations between maternal depressive symptoms and feeding styles and practices [39,41,42,44-46]. Additional longitudinal studies are necessary to understand whether mothers' depressive symptoms influence their feeding styles and practices. Furthermore, nearly all examined (6/8) studies used an array of self-reported questionnaires for assessments of maternal depressive symptoms and parental feeding practices (7/8), which is potentially problematic due to possible misclassification of depressive symptoms (exposure) and parental feeding practices (outcome). Finally, variability in the assessment of maternal depressive symptoms (eg, CES-D, DASS, BSI) and parental feeding practices (eg, Child Feeding Styles Questionnaire [CFSQ], CFQ, Family Mealtime Coding System [FMCS]) make it difficult to compare findings across

studies and indicate that caution must be taken in drawing conclusive assertions.

Strengths of this review include the use of systematic criteria (ie, PRISMA) to identify and select studies and a quality assessment tool for the critical appraisals of studies. Nonetheless, this review may be incomplete given limitation to studies published in English. Another possible limitation of this review is the variability in the studies' location. Multiple countries (United States, Australia, and England) were represented, which may limit cross-study comparisons. Finally, publication bias should also be considered, as should the fact that this review is limited to full-text studies published in English and may have excluded studies published in other formats and/or languages.

Future Directions

Additional research is needed to further examine the relationships between maternal depressive symptoms and nonresponsive parental feeding styles and practices. Specifically, longitudinal studies and additional studies including low-income and racial/ethnic minority populations at increased risk of depressive symptoms are needed. Future studies should explore the associations between maternal depressive symptoms, food insecurity and maternal feeding styles and practices. This is required especially due to documented evidence of greater prevalence of obesity among racial/ethnic minority populations [78]. In 2 examined studies [39,42], authors suggest that food insecurity may interact with maternal depressive symptoms such as stress to increase the risk of unintentionally detrimental feeding practices such as pressuring child to eat and/or restricting child food intake. Therefore, studies that assess the potential interactions of food security status and maternal depressive symptoms on maternal feeding styles and practices are needed. Moreover, given the inconsistencies in results across studies included in this review, future research should also consider the potential influence of additional factors such as mother's socioeconomic status, acculturation level, social support, as well as contextual factors such as work strain, access to healthful foods, and so on. Finally, future research may benefit from examining differentials of depressive symptoms and parental feeding styles and practices according to the gender of the parent and the child.

Conclusions

In summary, studies identified and synthesized in this review provided mixed evidence for associations between maternal depressive symptoms and nonresponsive maternal feeding styles and practices. Nevertheless, given the high prevalence of maternal depressive symptoms among women of reproductive age [32-35], the indication from some studies of associations between maternal depressive symptoms and nonresponsive feeding styles and practices, and the fact that both maternal depressive symptoms and that nonresponsive feeding styles and practices are potentially modifiable, further understanding of these associations are likely to provide important insights for the development of interventions to prevent and control childhood obesity.

Acknowledgments

The authors are grateful for library assistance provided by Ms Teresa Maceira, Reference Librarian at the University of Massachusetts, Boston.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Quality assessment of 8 included studies using an adapted version of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE).

[[PDF File \(Adobe PDF File\), 49KB - publichealth_v3i2e29_app1.pdf](#)]

Multimedia Appendix 2

Description of the 8 studies included in systematic review.

[[PDF File \(Adobe PDF File\), 85KB - publichealth_v3i2e29_app2.pdf](#)]

Multimedia Appendix 3

Characteristics of 8 included studies included in systematic review.

[[PDF File \(Adobe PDF File\), 56KB - publichealth_v3i2e29_app3.pdf](#)]

References

1. WHO. Report of the first meeting of the ad hoc working group on science and evidence for ending childhood obesity. Geneva: World Health Organization; 2014 Jun 18 Presented at: World Health Organization; June 18-20, 2014; Geneva, Switzerland p. 1-49 URL: <http://www.who.int/end-childhood-obesity/echo-final-report-august-2014.pdf>
2. Institute OM. Examining a Developmental Approach to Childhood Obesity: The Fetal and Early Childhood Years: Workshop Summary (2015). Washington, DC: National Academies Press; 2015 Presented at: Institute of Medicine; Food and Nutrition Board; 2015; Washington, DC.
3. Institute OM. Early childhood obesity prevention policies. Washington, DC: National Academies Press; 2011.
4. Lumeng JC, Taveras EM, Birch L, Yanovski SZ. Prevention of obesity in infancy and early childhood: a National Institutes of Health workshop. *JAMA Pediatr* 2015 May;169(5):484-490. [doi: [10.1001/jamapediatrics.2014.3554](https://doi.org/10.1001/jamapediatrics.2014.3554)] [Medline: [25775180](https://pubmed.ncbi.nlm.nih.gov/25775180/)]
5. Anzman SL, Rollins BY, Birch LL. Parental influence on children's early eating environments and obesity risk: implications for prevention. *Int J Obes Relat Metab Disord* 2010 Mar 02;34(7):1116-1124. [doi: [10.1038/ijo.2010.43](https://doi.org/10.1038/ijo.2010.43)] [Medline: [20195285](https://pubmed.ncbi.nlm.nih.gov/20195285/)]
6. Lindsay AC, Sussner KM, Kim J, Gortmaker S. The role of parents in preventing childhood obesity. *Future Child* 2006;16(1):169-186. [Medline: [16532663](https://pubmed.ncbi.nlm.nih.gov/16532663/)]
7. Birch LL. Learning to eat: behavioral and psychological aspects. *Nestle Nutr Inst Workshop Ser* 2016;85:125-134. [doi: [10.1159/000439503](https://doi.org/10.1159/000439503)] [Medline: [27088340](https://pubmed.ncbi.nlm.nih.gov/27088340/)]
8. Patrick H, Hennessy E, McSpadden K, Oh A. Parenting styles and practices in children's obesogenic behaviors: scientific gaps and future research directions. *Child Obes* 2013 Aug;9 Suppl:S73-S86 [FREE Full text] [doi: [10.1089/chi.2013.0039](https://doi.org/10.1089/chi.2013.0039)] [Medline: [23944926](https://pubmed.ncbi.nlm.nih.gov/23944926/)]
9. Ventura AK, Birch LL. Does parenting affect children's eating and weight status? *Int J Behav Nutr Phys Act* 2008 Mar 17;5:15 [FREE Full text] [doi: [10.1186/1479-5868-5-15](https://doi.org/10.1186/1479-5868-5-15)] [Medline: [18346282](https://pubmed.ncbi.nlm.nih.gov/18346282/)]
10. El-Behadli AF, Sharp C, Hughes SO, Obasi EM, Nicklas TA. Maternal depression, stress and feeding styles: towards a framework for theory and research in child obesity. *Br J Nutr* 2015 Jan 15;113(S1):S55-S71. [doi: [10.1017/S000711451400333X](https://doi.org/10.1017/S000711451400333X)] [Medline: [25588385](https://pubmed.ncbi.nlm.nih.gov/25588385/)]
11. Hurley KM, Black MM, Papas MA, Caulfield LE, Caulfield LE. Maternal symptoms of stress, depression, and anxiety are related to nonresponsive feeding styles in a statewide sample of WIC participants. *J Nutr* 2008 Apr;138(4):799-805 [FREE Full text] [Medline: [18356338](https://pubmed.ncbi.nlm.nih.gov/18356338/)]
12. Savage JS, Birch LL. WIC mothers' depressive symptoms are associated with greater use of feeding to soothe, regardless of perceived child negativity. *Pediatr Obes* 2016 Feb 29. [doi: [10.1111/ijpo.12122](https://doi.org/10.1111/ijpo.12122)] [Medline: [26923811](https://pubmed.ncbi.nlm.nih.gov/26923811/)]
13. Lovejoy M, Graczyk P, O'Hare E, Neuman G. Maternal depression and parenting behavior. *Clin Psychol Rev* 2000 Aug;20(5):561-592. [doi: [10.1016/S0272-7358\(98\)00100-7](https://doi.org/10.1016/S0272-7358(98)00100-7)] [Medline: [10860167](https://pubmed.ncbi.nlm.nih.gov/10860167/)]
14. Milgrom J, Skouteris H, Worotniuk T, Henwood A, Bruce L. The association between ante- and postnatal depressive symptoms and obesity in both mother and child: a systematic review of the literature. *Women's Health Issues* 2012 May;22(3):e319-e328. [doi: [10.1016/j.whi.2011.12.001](https://doi.org/10.1016/j.whi.2011.12.001)] [Medline: [22341777](https://pubmed.ncbi.nlm.nih.gov/22341777/)]

15. Weng SF, Redsell SA, Swift JA, Yang M, Glazebrook CP. Systematic review and meta-analyses of risk factors for childhood overweight identifiable during infancy. *Arch Dis Child* 2012 Dec;97(12):1019-1026 [FREE Full text] [doi: [10.1136/archdischild-2012-302263](https://doi.org/10.1136/archdischild-2012-302263)] [Medline: [23109090](https://pubmed.ncbi.nlm.nih.gov/23109090/)]
16. Barrett KJ, Thompson AL, Bentley ME. The influence of maternal psychosocial characteristics on infant feeding styles. *Appetite* 2016 Aug 1;103:396-402. [doi: [10.1016/j.appet.2016.04.042](https://doi.org/10.1016/j.appet.2016.04.042)] [Medline: [27174251](https://pubmed.ncbi.nlm.nih.gov/27174251/)]
17. Lampard AM, Franckle RL, Davison KK. Maternal depression and childhood obesity: a systematic review. *Prev Med* 2014 Feb;59:60-67 [FREE Full text] [doi: [10.1016/j.ypmed.2013.11.020](https://doi.org/10.1016/j.ypmed.2013.11.020)] [Medline: [24291685](https://pubmed.ncbi.nlm.nih.gov/24291685/)]
18. Tate EB, Wood W, Liao Y, Dunton GF. Do stressed mothers have heavier children? a meta-analysis on the relationship between maternal stress and child body mass index. *Obes Rev* 2015 May;16(5):351-361 [FREE Full text] [doi: [10.1111/obr.12262](https://doi.org/10.1111/obr.12262)] [Medline: [25879393](https://pubmed.ncbi.nlm.nih.gov/25879393/)]
19. Duarte CS, Shen S, Wu P, Must A. Maternal depression and child BMI: longitudinal findings from a US sample. *Pediatr Obes* 2012 Apr;7(2):124-133 [FREE Full text] [doi: [10.1111/j.2047-6310.2011.00012.x](https://doi.org/10.1111/j.2047-6310.2011.00012.x)] [Medline: [22434752](https://pubmed.ncbi.nlm.nih.gov/22434752/)]
20. Vollmer RL, Mobley AR. Parenting styles, feeding styles, and their influence on child obesogenic behaviors and body weight. a review. *Appetite* 2013 Dec;71:232-241. [doi: [10.1016/j.appet.2013.08.015](https://doi.org/10.1016/j.appet.2013.08.015)]
21. Shloim N, Edelson LR, Martin N, Hetherington MM. Parenting styles, feeding styles, feeding practices, and weight status in 4-12 year-old children: a systematic review of the literature. *Front Psychol* 2015;14(6):1849. [doi: [10.3389/fpsyg.2015.01849](https://doi.org/10.3389/fpsyg.2015.01849)] [Medline: [26696920](https://pubmed.ncbi.nlm.nih.gov/26696920/)]
22. Hubbs-Tait L, Kennedy TS, Page MC, Topham GL, Harrist AW. Parental feeding practices predict authoritative, authoritarian, and permissive parenting styles. *J Am Diet Assoc* 2008 Jul;108(7):1154-61; discussion 1161. [doi: [10.1016/j.jada.2008.04.008](https://doi.org/10.1016/j.jada.2008.04.008)] [Medline: [18589022](https://pubmed.ncbi.nlm.nih.gov/18589022/)]
23. Kiefner-Burmeister A, Hoffmann D, Zbur S, Musher-Eizenman D. Implementation of parental feeding practices: does parenting style matter? *Public Health Nutr* 2016 Sep;19(13):2410-2414. [doi: [10.1017/S1368980016000446](https://doi.org/10.1017/S1368980016000446)] [Medline: [26975423](https://pubmed.ncbi.nlm.nih.gov/26975423/)]
24. Hurley KM, Cross MB, Hughes SO. A systematic review of responsive feeding and child obesity in high-income countries. *J Nutr* 2011 Mar;141(3):495-501 [FREE Full text] [doi: [10.3945/jn.110.130047](https://doi.org/10.3945/jn.110.130047)] [Medline: [21270360](https://pubmed.ncbi.nlm.nih.gov/21270360/)]
25. Dennis C, McQueen K. The relationship between infant-feeding outcomes and postpartum depression: a qualitative systematic review. *Pediatrics* 2009 Apr;123(4):e736-e751. [doi: [10.1542/peds.2008-1629](https://doi.org/10.1542/peds.2008-1629)] [Medline: [19336362](https://pubmed.ncbi.nlm.nih.gov/19336362/)]
26. Fiese BH, Hammons A, Grigsby-Toussaint D. Family mealtimes: a contextual approach to understanding childhood obesity. *Econ Hum Biol* 2012 Dec;10(4):365-374. [doi: [10.1016/j.ehb.2012.04.004](https://doi.org/10.1016/j.ehb.2012.04.004)] [Medline: [22652025](https://pubmed.ncbi.nlm.nih.gov/22652025/)]
27. Black M, Hurley K. Child-encyclopedia. 2007. Helping children develop health eating habits URL: <http://www.child-encyclopedia.com/sites/default/files/textes-experts/en/540/helping-children-develop-healthy-eating-habits.pdf> [accessed 2017-05-15] [WebCite Cache ID 6qTOFXmEL]
28. Fiese B, Foley K, Spagnola M. Routine and ritual elements in family mealtimes: contexts for child well-being and family identity. *New Dir Child Adolesc Dev* 2006;Spring(111):67-89. [Medline: [16646500](https://pubmed.ncbi.nlm.nih.gov/16646500/)]
29. Franko DL, Thompson D, Affenito SG, Barton BA, Striegel-Moore RH. What mediates the relationship between family meals and adolescent health issues. *Health Psychol* 2008;27(2, Suppl):S109-S117. [doi: [10.1037/0278-6133.27.2\(Suppl.\).S109](https://doi.org/10.1037/0278-6133.27.2(Suppl.).S109)]
30. Barlow J, Coren E. Parent-training programmes for improving maternal psychosocial health. *Cochrane Database Syst Rev* 2004(1):CD002020. [doi: [10.1002/14651858.CD002020.pub2](https://doi.org/10.1002/14651858.CD002020.pub2)] [Medline: [14973981](https://pubmed.ncbi.nlm.nih.gov/14973981/)]
31. Hurley KM, Black MM. Introduction to a supplement on responsive feeding: promoting healthy growth and development for infants and toddlers. *J Nutr* 2011 Mar;141(3):489 [FREE Full text] [doi: [10.3945/jn.110.130070](https://doi.org/10.3945/jn.110.130070)] [Medline: [21270358](https://pubmed.ncbi.nlm.nih.gov/21270358/)]
32. Ertel KA, Rich-Edwards JW, Koenen KC. Maternal depression in the United States: nationally representative rates and risks. *J Womens Health (Larchmt)* 2011 Nov;20(11):1609-1617 [FREE Full text] [doi: [10.1089/jwh.2010.2657](https://doi.org/10.1089/jwh.2010.2657)] [Medline: [21877915](https://pubmed.ncbi.nlm.nih.gov/21877915/)]
33. Ko JY, Farr SL, Dietz PM, Robbins CL. Depression and treatment among U.S. pregnant and nonpregnant women of reproductive age, 2005-2009. *J Womens Health (Larchmt)* 2012 Aug;21(8):830-836 [FREE Full text] [doi: [10.1089/jwh.2011.3466](https://doi.org/10.1089/jwh.2011.3466)] [Medline: [22691031](https://pubmed.ncbi.nlm.nih.gov/22691031/)]
34. Farr S, Bitsko R, Hayes D, Dietz P. Mental health and access to services among US women of reproductive age. *Am J Obstet Gynecol* 2010 Dec;203(6):542.e1-542.e9. [doi: [10.1016/j.ajog.2010.07.007](https://doi.org/10.1016/j.ajog.2010.07.007)] [Medline: [20817143](https://pubmed.ncbi.nlm.nih.gov/20817143/)]
35. Le Strat SY, Dubertret C, Le FB. Prevalence and correlates of major depressive episode in pregnant and postpartum women in the United States. *J Affect Disord* 2011 Dec;135(1-3):128-138. [doi: [10.1016/j.jad.2011.07.004](https://doi.org/10.1016/j.jad.2011.07.004)] [Medline: [21802737](https://pubmed.ncbi.nlm.nih.gov/21802737/)]
36. National Institute of Mental Health. Depression: What You Need to Know. 2015. URL: https://www.nimh.nih.gov/health/publications/depression-what-you-need-to-know/depression-what-you-need-to-know-pdf_151827.pdf [accessed 2017-05-15] [WebCite Cache ID 6qTOdbTpI]
37. Kendall-Tackett K. The new paradigm for depression in new mothers: current findings on maternal depression, breastfeeding and resiliency across the lifespan. *Breastfeed Rev* 2015 Mar;23(1):7-10. [Medline: [25906491](https://pubmed.ncbi.nlm.nih.gov/25906491/)]
38. 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/)]

39. Hughes SO, Power TG, Liu Y, Sharp C, Nicklas TA. Parent emotional distress and feeding styles in low-income families. the role of parent depression and parenting stress. *Appetite* 2015 Sep;92:337-342. [doi: [10.1016/j.appet.2015.06.002](https://doi.org/10.1016/j.appet.2015.06.002)] [Medline: [26050915](https://pubmed.ncbi.nlm.nih.gov/26050915/)]
40. Mallan KM, Daniels LA, Wilson JL, Jansen E, Nicholson JM. Association between maternal depressive symptoms in the early post-natal period and responsiveness in feeding at child age 2 years. *Matern Child Nutr* 2015 Oct;11(4):926-935. [doi: [10.1111/mcn.12116](https://doi.org/10.1111/mcn.12116)] [Medline: [24784325](https://pubmed.ncbi.nlm.nih.gov/24784325/)]
41. Gouling AN, Rosenblum KL, Miller AL, Peterson KE, Chen Y, Kaciroti N, et al. Associations between maternal depressive symptoms and child feeding practices in a cross-sectional study of low-income mothers and their young children. *Int J Behav Nutr Phys Act* 2014;11:75 [FREE Full text] [doi: [10.1186/1479-5868-11-75](https://doi.org/10.1186/1479-5868-11-75)] [Medline: [24935753](https://pubmed.ncbi.nlm.nih.gov/24935753/)]
42. McCurdy K, Gorman KS, Kisler T, Metallinos-Katsaras E. Associations between family food behaviors, maternal depression, and child weight among low-income children. *Appetite* 2014 Aug;79:97-105 [FREE Full text] [doi: [10.1016/j.appet.2014.04.015](https://doi.org/10.1016/j.appet.2014.04.015)] [Medline: [24768937](https://pubmed.ncbi.nlm.nih.gov/24768937/)]
43. Gemmill AW, Worotniuk T, Holt CJ, Skouteris H, Milgrom J. Maternal psychological factors and controlled child feeding practices in relation to child body mass index. *Child Obes* 2013 Aug;9(4):326-337. [doi: [10.1089/chi.2012.0135](https://doi.org/10.1089/chi.2012.0135)] [Medline: [23782306](https://pubmed.ncbi.nlm.nih.gov/23782306/)]
44. Gross RS, Velazco NK, Briggs RD, Racine AD. Maternal depressive symptoms and child obesity in low-income urban families. *Acad Pediatr* 2013 Jul;13(4):356-363. [doi: [10.1016/j.acap.2013.04.002](https://doi.org/10.1016/j.acap.2013.04.002)] [Medline: [23830021](https://pubmed.ncbi.nlm.nih.gov/23830021/)]
45. Haycraft E, Farrow C, Blissett J. Maternal symptoms of depression are related to observations of controlling feeding practices in mothers of young children. *J Fam Psychol* 2013 Feb;27(1):159-164. [doi: [10.1037/a0031110](https://doi.org/10.1037/a0031110)] [Medline: [23421843](https://pubmed.ncbi.nlm.nih.gov/23421843/)]
46. Mitchell S, Brennan L, Hayes L, Miles CL. Maternal psychosocial predictors of controlling parental feeding styles and practices. *Appetite* 2009 Dec;53(3):384-389. [doi: [10.1016/j.appet.2009.08.001](https://doi.org/10.1016/j.appet.2009.08.001)] [Medline: [19666066](https://pubmed.ncbi.nlm.nih.gov/19666066/)]
47. DiSantis KI, Hodges EA, Johnson SL, Fisher JO. The role of responsive feeding in overweight during infancy and toddlerhood: a systematic review. *Int J Obes (Lond)* 2011 Apr;35(4):480-492. [doi: [10.1038/ijo.2011.3](https://doi.org/10.1038/ijo.2011.3)] [Medline: [21427696](https://pubmed.ncbi.nlm.nih.gov/21427696/)]
48. Bascom EM, Napolitano MA. Breastfeeding duration and primary reasons for breastfeeding cessation among women with postpartum depressive symptoms. *J Hum Lact* 2016 May;32(2):282-291. [doi: [10.1177/0890334415619908](https://doi.org/10.1177/0890334415619908)] [Medline: [26644420](https://pubmed.ncbi.nlm.nih.gov/26644420/)]
49. McLearn KT, Minkovitz CS, Strobino DM, Marks E, Hou W. The timing of maternal depressive symptoms and mothers' parenting practices with young children: implications for pediatric practice. *Pediatrics* 2006 Jul;118(1):e174-e182. [doi: [10.1542/peds.2005-1551](https://doi.org/10.1542/peds.2005-1551)] [Medline: [16818531](https://pubmed.ncbi.nlm.nih.gov/16818531/)]
50. von EE, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007 Oct 20;335(7624):806-808 [FREE Full text] [doi: [10.1136/bmj.39335.541782.AD](https://doi.org/10.1136/bmj.39335.541782.AD)] [Medline: [17947786](https://pubmed.ncbi.nlm.nih.gov/17947786/)]
51. Radloff L. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977;1(3):385-401 [FREE Full text]
52. Abidin R. Parenting stress index: professional manual. 3rd edition. Odessa, FL: Psychological Assessment Resources; 1995.
53. Cox J, Holden J, Sagovsky R. Detection of postnatal depression. development of the 10-item Edinburgh postnatal depression scale. *Br J Psychiatry* 1987 Jun 01;150(6):782-786. [doi: [10.1192/bjp.150.6.782](https://doi.org/10.1192/bjp.150.6.782)] [Medline: [3651732](https://pubmed.ncbi.nlm.nih.gov/3651732/)]
54. Antony M, Bieling P, Cox BJ, Enns M, Swinson R. Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment* 1998;10(2):176-181. [doi: [10.1037/1040-3590.10.2.176](https://doi.org/10.1037/1040-3590.10.2.176)]
55. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
56. Zigmond A, Snaith R. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983 Jun;67(6):361-370. [Medline: [6880820](https://pubmed.ncbi.nlm.nih.gov/6880820/)]
57. Hughes SO, Power TG, Orlet FJ, Mueller S, Nicklas TA. Revisiting a neglected construct: parenting styles in a child-feeding context. *Appetite* 2005 Feb;44(1):83-92. [doi: [10.1016/j.appet.2004.08.007](https://doi.org/10.1016/j.appet.2004.08.007)] [Medline: [15604035](https://pubmed.ncbi.nlm.nih.gov/15604035/)]
58. Birch L, Fisher J, Grimm-Thomas K, Markey C, Sawyer R, Johnson S. Confirmatory factor analysis of the child feeding questionnaire: a measure of parental attitudes, beliefs and practices about child feeding and obesity proneness. *Appetite* 2001 Jun;36(3):201-210. [doi: [10.1006/appe.2001.0398](https://doi.org/10.1006/appe.2001.0398)] [Medline: [11358344](https://pubmed.ncbi.nlm.nih.gov/11358344/)]
59. McCurdy K, Gorman KS. Measuring family food environments in diverse families with young children. *Appetite* 2010 Jun;54(3):615-618. [doi: [10.1016/j.appet.2010.03.004](https://doi.org/10.1016/j.appet.2010.03.004)] [Medline: [20227449](https://pubmed.ncbi.nlm.nih.gov/20227449/)]
60. Musher-Eizenman D, Holub S. Comprehensive feeding practices questionnaire: validation of a new measure of parental feeding practices. *J Pediatr Psychol* 2007 Sep;32(8):960-972. [doi: [10.1093/jpepsy/jsm037](https://doi.org/10.1093/jpepsy/jsm037)] [Medline: [17535817](https://pubmed.ncbi.nlm.nih.gov/17535817/)]
61. Larios SE, Ayala GX, Arredondo EM, Baquero B, Elder JP. Development and validation of a scale to measure Latino parenting strategies related to children's obesigenic behaviors. the parenting strategies for eating and activity scale (PEAS). *Appetite* 2009 Feb;52(1):166-172 [FREE Full text] [doi: [10.1016/j.appet.2008.09.011](https://doi.org/10.1016/j.appet.2008.09.011)] [Medline: [18845197](https://pubmed.ncbi.nlm.nih.gov/18845197/)]
62. Wardle J, Sanderson S, Guthrie CA, Rapoport L, Plomin R. Parental feeding style and the inter-generational transmission of obesity risk. *Obes Res* 2002 Jun;10(6):453-462 [FREE Full text] [doi: [10.1038/oby.2002.63](https://doi.org/10.1038/oby.2002.63)] [Medline: [12055321](https://pubmed.ncbi.nlm.nih.gov/12055321/)]

63. Ogden J, Reynolds R, Smith A. Expanding the concept of parental control: a role for overt and covert control in children's snacking behaviour? *Appetite* 2006 Jul;47(1):100-106. [doi: [10.1016/j.appet.2006.03.330](https://doi.org/10.1016/j.appet.2006.03.330)] [Medline: [16682098](https://pubmed.ncbi.nlm.nih.gov/16682098/)]
64. Haycraft EL, Blissett JM. Maternal and paternal controlling feeding practices: reliability and relationships with BMI. *Obesity (Silver Spring)* 2008 Jul;16(7):1552-1558 [FREE Full text] [doi: [10.1038/oby.2008.238](https://doi.org/10.1038/oby.2008.238)] [Medline: [18421263](https://pubmed.ncbi.nlm.nih.gov/18421263/)]
65. Fildes A, van Jaarsveld CH, Llewellyn C, Wardle J, Fisher A. Parental control over feeding in infancy. influence of infant weight, appetite and feeding method. *Appetite* 2015 Aug;91:101-106. [doi: [10.1016/j.appet.2015.04.004](https://doi.org/10.1016/j.appet.2015.04.004)]
66. Paulson JF, Dauber S, Leiferman JA. Individual and combined effects of postpartum depression in mothers and fathers on parenting behavior. *Pediatrics* 2006 Aug;118(2):659-668. [doi: [10.1542/peds.2005-2948](https://doi.org/10.1542/peds.2005-2948)] [Medline: [16882821](https://pubmed.ncbi.nlm.nih.gov/16882821/)]
67. Tschann JM, Martinez SM, Penilla C, Gregorich SE, Pasch LA, de Groat CL, et al. Parental feeding practices and child weight status in Mexican American families: a longitudinal analysis. *Int J Behav Nutr Phys Act* 2015 May;12:66 [FREE Full text] [doi: [10.1186/s12966-015-0224-2](https://doi.org/10.1186/s12966-015-0224-2)] [Medline: [25986057](https://pubmed.ncbi.nlm.nih.gov/25986057/)]
68. Rodgers RF, Paxton SJ, Massey R, Campbell KJ, Wertheim EH, Skouteris H, et al. Maternal feeding practices predict weight gain and obesogenic eating behaviors in young children: a prospective study. *Int J Behav Nutr Phys Act* 2013;10:24 [FREE Full text] [doi: [10.1186/1479-5868-10-24](https://doi.org/10.1186/1479-5868-10-24)] [Medline: [23414332](https://pubmed.ncbi.nlm.nih.gov/23414332/)]
69. Hennessy E, Hughes SO, Goldberg JP, Hyatt RR, Economos CD. Permissive parental feeding behavior is associated with an increase in intake of low-nutrient-dense foods among American children living in rural communities. *J Acad Nutr Diet* 2012 Jan;112(1):142-148. [doi: [10.1016/j.jada.2011.08.030](https://doi.org/10.1016/j.jada.2011.08.030)] [Medline: [22709645](https://pubmed.ncbi.nlm.nih.gov/22709645/)]
70. Gross RS, Fierman AH, Mendelsohn AL, Chiasson MA, Rosenberg TJ, Scheinmann R, et al. Maternal perceptions of infant hunger, satiety, and pressuring feeding styles in an urban Latina WIC population. *Acad Pediatr* 2010 Jan;10(1):29-35. [doi: [10.1016/j.acap.2009.08.001](https://doi.org/10.1016/j.acap.2009.08.001)] [Medline: [20004633](https://pubmed.ncbi.nlm.nih.gov/20004633/)]
71. Lumeng JC, Ozbeki TN, Appugliese DP, Kaciroti N, Corwyn RF, Bradley RH. Observed assertive and intrusive maternal feeding behaviors increase child adiposity. *Am J Clin Nutr* 2012 Jan 25;95(3):640-647. [doi: [10.3945/ajcn.111.024851](https://doi.org/10.3945/ajcn.111.024851)]
72. Horodyski MA, Stommel M, Brophy-Herb HE, Weatherspoon L. Mealtime television viewing and dietary quality in low-income African American and Caucasian mother-toddler dyads. *Matern Child Health J* 2010 Jul;14(4):548-556. [doi: [10.1007/s10995-009-0501-2](https://doi.org/10.1007/s10995-009-0501-2)] [Medline: [19629662](https://pubmed.ncbi.nlm.nih.gov/19629662/)]
73. Domoff SE, Lumeng JC, Kaciroti N, Miller AL. Early childhood risk factors for mealtime TV exposure and engagement in low-income families. *Acad Pediatr* 2016 Dec 13;17(4):411-415. [doi: [10.1016/j.acap.2016.12.003](https://doi.org/10.1016/j.acap.2016.12.003)] [Medline: [27979749](https://pubmed.ncbi.nlm.nih.gov/27979749/)]
74. Trofholz AC, Tate AD, Draxten ML, Neumark-Sztainer D, Berge JM. Home food environment factors associated with the presence of fruit and vegetables at dinner: a direct observational study. *Appetite* 2016 Jan 01;96:526-532 [FREE Full text] [doi: [10.1016/j.appet.2015.10.019](https://doi.org/10.1016/j.appet.2015.10.019)] [Medline: [26527254](https://pubmed.ncbi.nlm.nih.gov/26527254/)]
75. Martin-Biggers J, Spaccarotella K, Berhaupt-Glickstein A, Hongu N, Worobey J, Byrd-Bredbenner C. Come and get it! a discussion of family mealtime literature and factors affecting obesity risk. *Adv Nutr* 2014 May;5(3):235-247 [FREE Full text] [doi: [10.3945/an.113.005116](https://doi.org/10.3945/an.113.005116)] [Medline: [24829470](https://pubmed.ncbi.nlm.nih.gov/24829470/)]
76. Wyse R, Campbell E, Nathan N, Wolfenden L. Associations between characteristics of the home food environment and fruit and vegetable intake in preschool children: a cross-sectional study. *BMC Public Health* 2011 Dec 16;11:938 [FREE Full text] [doi: [10.1186/1471-2458-11-938](https://doi.org/10.1186/1471-2458-11-938)] [Medline: [22177136](https://pubmed.ncbi.nlm.nih.gov/22177136/)]
77. Østbye T, Malhotra R, Stroo M, Lovelady C, Brouwer R, Zucker N, et al. The effect of the home environment on physical activity and dietary intake in preschool children. *Int J Obes (Lond)* 2013 Oct;37(10):1314-1321 [FREE Full text] [doi: [10.1038/ijo.2013.76](https://doi.org/10.1038/ijo.2013.76)] [Medline: [23736357](https://pubmed.ncbi.nlm.nih.gov/23736357/)]
78. Ogden CL, Carroll MD, Lawman HG, Fryar CD, Kruszon-Moran D, Kit BK, et al. Trends in obesity prevalence among children and adolescents in the United States, 1988-1994 through 2013-2014. *JAMA* 2016 Jun 07;315(21):2292-2299. [doi: [10.1001/jama.2016.6361](https://doi.org/10.1001/jama.2016.6361)] [Medline: [27272581](https://pubmed.ncbi.nlm.nih.gov/27272581/)]

Abbreviations

BMI: body mass index

BSI: Brief Symptom Inventory

CES-D: Center for Epidemiologic Studies Depression Scale

CFSQ: Child Feeding Styles Questionnaire

CFQ: Child Feeding Questionnaire

CINAHL: Cumulative Index to Nursing and Allied Health Literature

DASS: Depression Anxiety Stress Scales

EPDS: Edinburgh Postnatal Depression Scale

FMCS: Family Mealtime Coding System

HIV: human immunodeficiency virus

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis

STROBE: Strengthening in the Reporting of Observational Studies in Epidemiology

Edited by G Eysenbach; submitted 13.08.16; peer-reviewed by F Aboud, R Rodgers; comments to author 08.09.16; revised version received 16.12.16; accepted 25.03.17; published 26.05.17.

Please cite as:

Lindsay AC, Mesa T, Greaney ML, Wallington SF, Wright JA

Associations Between Maternal Depressive Symptoms and Nonresponsive Feeding Styles and Practices in Mothers of Young Children: A Systematic Review

JMIR Public Health Surveill 2017;3(2):e29

URL: <http://publichealth.jmir.org/2017/2/e29/>

doi: [10.2196/publichealth.6492](https://doi.org/10.2196/publichealth.6492)

PMID: [28550007](https://pubmed.ncbi.nlm.nih.gov/28550007/)

©Ana Cristina Lindsay, Tatiana Mesa, Mary L. Greaney, Sherrie F. Wallington, Julie A. Wright. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 26.05.2017. 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

Using Google Trends Data to Study Public Interest in Breast Cancer Screening in Brazil: Why Not a Pink February?

Paulo Roberto Vasconcellos-Silva^{1,2*}, MD, PhD; Dárlinton Barbosa Feres Carvalho^{3*}, PhD; Valéria Trajano^{1*}, PhD; Lucia Rodriguez de La Rocque^{1,4*}, PhD; Anunciata Cristina Marins Braz Sawada^{1*}, MSc; Leidjaira Lopes Juvanhol^{5*}, RN, PhD

¹Laboratory of Innovation in Therapies, Teaching and Bioproducts /LITEB, Oswaldo Cruz Institute/IOC, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

²Research Coordination, National Cancer Institute, Rio de Janeiro, Brazil

³Department of Computer Science, Federal University of São João del-Rei, São João del-Rei, Brazil

⁴Institute of Letters. Sector of English Literature, Department of Germanic Languages, State University of Rio de Janeiro/UERJ, Rio de Janeiro, Brazil

⁵National School of Public Health, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

* all authors contributed equally

Corresponding Author:

Paulo Roberto Vasconcellos-Silva, MD, PhD

Laboratory of Innovation in Therapies, Teaching and Bioproducts /LITEB

Oswaldo Cruz Institute/IOC

Oswaldo Cruz Foundation

Manguinhos

Avenida Brasil 4365

Rio de Janeiro, CEP21040-360

Brazil

Phone: 55 21 25621477

Fax: 55 21 2562 1477

Email: bioeticaunirio@yahoo.com.br

Abstract

Background: One of the major challenges of the Brazilian Ministry of Health is to foster interest in breast cancer screening (BCS), especially among women at high risk. Strategies have been developed to promote the early identification of breast cancer mainly by Pink October campaigns. The massive number of queries conducted through Google creates traffic data that can be analyzed to show unrevealed interest cycles and their seasonalities.

Objectives: Using Google Trends, we studied cycles of public interest in queries toward *mammography* and *breast cancer* along the last 5 years. We hypothesize that these data may be correlated with collective interest cycles leveraged by national BCS campaigns such as Pink October.

Methods: Google Trends was employed to normalize traffic data on a scale from 0 (<1% of the peak volume) to 100 (peak of traffic) presented as weekly relative search volume (RSV) concerning mammography and breast cancer as search terms. A time series covered the last 261 weeks (November 2011 to October 2016), and RSV of both terms were compared with their respective annual means. Polynomial trendlines (second order) were employed to estimate overall trends.

Results: We found an upward trend for both terms over the 5 years, with almost parallel trendlines. Remarkable peaks were found along Pink October months—mammography and breast cancer searches were leveraged up reaching, respectively, 119.1% (2016) and 196.8% (2015) above annual means. Short downward RSVs along December-January months were also noteworthy along all the studied period. These trends traced an N-shaped pattern with higher peaks in Pink October months and sharp falls along subsequent December and January.

Conclusions: Considering these findings, it would be reasonable to bring Pink October to the beginning of each year, thereby extending the beneficial effect of the campaigns. It would be more appropriate to start screening campaigns at the beginning of the year, when new resolutions are taken and new projects are added to everyday routines. Our work raises attention to the study of traffic data to encourage health campaign analysts to undertake better analysis based on marketing practices.

KEYWORDS

Internet; cancer information seeking; breast cancer; mass screening; health communication; early detection of cancer; infoveillance; infodemiology

Introduction

The increasing number of searches for health-related issues generates “big data,” providing meaningful research in infodemiology, which is the study of patterns and determinants of information on the Web with the purpose to inform public health and public policy [1,2]. This new concept is grounded in two approaches: supply-based infodemiology (studies the dynamics, quantity, and quality of information available on websites, media reports, blogs, tweets, etc) and demand-based (health information seeking, eg, what people are searching for on the Internet) [1]. Information-seeking behavior (measured, eg, by the frequency with which the public enters specified search terms) has been used successfully to show unexpected interest cycles and regional seasonalities. Numerous conditions and situations have been studied, from influenza epidemics [3] to seasonality of headache [4]. Previous studies suggest that infoveillance can measure the success of a campaign in driving information-seeking behaviors [5], showing significant relationships between health campaigns and information-seeking behaviors [6]. Web data surveillance holds a strong potential to lead to overlooked phenomena [2,7,8] and might increase knowledge on campaign strategies centered on timing of interest cycles. This paper works under the demand-based approach, examining patterns of public seeking for information on breast cancer early identification, like similar works that have become increasingly frequent in recent literature [6,8-10].

Breast cancer is the most common form of cancer among women all over the world, whether in high-income or in poor countries [11], accounting for 22% of the 4.7 million new cases occurring annually among females worldwide [12,13]. There is plenty of evidence that early diagnosis initiatives of breast cancer save far more lives and are much more cost-effective than treatment of late stages [14,15]. From the perception of countries like Brazil, the efficacy and adherence to breast cancer screening (BCS) is still a problematic issue from a public health policy perspective [16]. Brazilian mortality rates are increasing with striking variations between geographic regions, and several factors may account for the disparities, including delays in diagnosis due to low education level [17], low adherence to screening programs, and gaps in their implementation [18,19]. Surveillance systems databases have been used to assess self-reported cancer screening utilization. Although invaluable in identifying determinants of screening use and describing trends, these database systems are too complex and costly, and remain a challenge for the largest country in Latin America. On the other hand, a massive number of queries conducted through Google create data that can be analyzed with Google Trends, a

publicly available tool used to compare the volume of Web search queries in different periods [20].

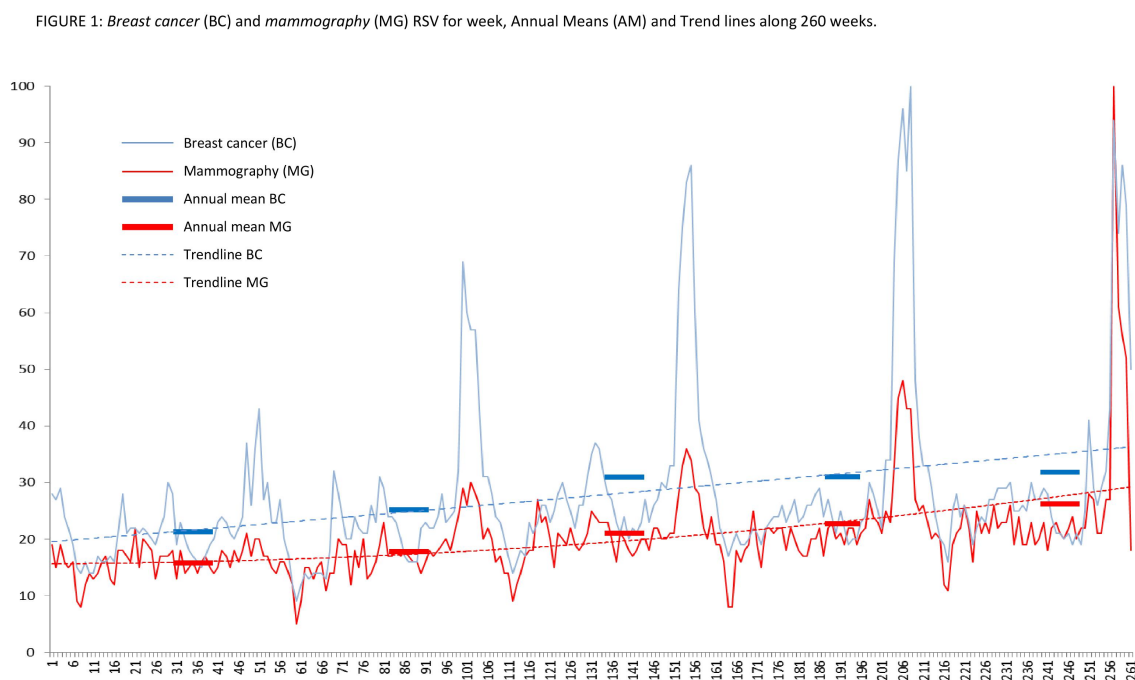
The use of search volume for predicting real-world events may have less to do with their superiority over other data systems than with matters of low cost, transparency, simplicity, and reproducibility across a variety of domains. Among other free access tools available [19], Google Trends provides essential data to public health planners as weekly reports on the volume of queries related to pertinent issues. Google Trends shows oscillations whenever a particular search term is searched for, relative to the top number of searches [20]. We hypothesize that these query data may be correlated with collective interest cycles affected by campaigns and, thus, may be suitable in “predicting the present” in terms of “BCS attitudes.” If that is the case, Google Trends would be a low-cost support for screening campaigns planners, providing feedback information almost immediately after interventions. In this paper, we studied oscillations of public interest in queries toward mammography and breast cancer along the last 5 years.

Methods

Google Trends is a Web-based free tracking system of Google search volumes. Google Trends algorithmics normalize data for the overall number of searches on a scale from 0 (search volume <1% of the peak volume) to 100 (peak of popularity), presenting them as a weekly relative search volume (RSV). RSV values are by definition, as presented in the y-axis (Figure 1), always less than 100, and display a proportion compared with the highest search volume. This approach corrects results for population size and Internet access, both of which increased during the study period.

Mammography and breast cancer (“mamografia” and “cancer de mama” in Portuguese) were used as search terms to produce separate time series (put together in Figure 1) in the last 260 weeks (November 2011 to October 2016), with the filter “Brazil” (country) in category “Health.” We selected these search terms based on their face validity, excluding their plural forms or any other unusual forms, which resulted in low weekly RSV.

The results were analyzed considering the data export through comma separated value (CSV) files. The weekly and monthly RSV values were compared with annual means, and a graph was plotted adding up annual means to highlight differences between weekly RSV series for both terms. Polynomial trendlines (second order) were added to the weekly RSV to estimate trends over the 261 weeks.

Figure 1. Breast cancer and mammography relative search volume for week, annual means, and trendlines along 260 weeks.

1

Results

Results show an upward trend for both breast cancer and mammography searches over the 5 years, with almost parallel trend curves (Figure 1). The annual means on breast cancer queries show a slight decline between 2011 and 2012 (not so relevant, considering just the last two months, without Pink October) followed by a raise from there, with a “jump” in 2015 annual means, as shown in Table 1. Annual means of mammography searches rose steadily along the 5 years (Table 1). Interest in breast cancer seems to significantly increase in Pink October months with remarkable higher means (reaching

up to 196.8% above the 2015 annual means). There were several minor peaks throughout the years (without impacts comparable with Pink October months and no obvious seasonality). Likewise, there were remarkable growing peaks for mammography searches along Pink October months (reaching 119.1% above the 2015 annual means), though not so “instable” throughout the years when compared with breast cancer. A short downward trend along December-January months was also noteworthy along the 261 weeks—mammography reached 27.1% and breast cancer 36.6% below the annual means. These oscillations traced an N-shaped curve with higher peaks in Pink October months and sharp falls along the two subsequent months (Figure 1).

Table 1. Breast cancer and mammography relative search volume along Pink October and December-January (means) related to annual means.

Relative search volume	2011	2012	2013	2014	2015	2016 ^a
Mammography						
Annual means		15.8	17.8	21	22.7	26.2
(std)		(2.9)	(4.5)	(5.2)	(7.4)	(14.6)
Pink October		18.5	27.5	33	44.8	57.4
(annual means)		(↑17.1%)	(↑54.5%)	(↑57.1%)	(↑97.4%)	(↑119.1%)
December-January means	13	12.7	14.8	15.3	19.3	
(annual means)	(↓11%)	(↓19.6%)	(↓16.9%)	(↓27.1%)	(↓15%)	
Breast cancer						
Annual means		21.3	25.2	30.9	31	31.8
(std)		(6.3)	(12)	(15.4)	(19.7)	(17.6)
Pink October		33	54.3	76	92	76.6
(annual means)		(↑54.9)	(↑115.5%)	(↑146%)	(↑196.8%)	(↑140.9%)
December-January means	16.3	13.9	18.8	19.6	22.8	
(annual means)	(↓26.9%)	(↓34.7%)	(↓25.4%)	(↓36.6%)	(↓26.5%)	

^aFrom January to October.

Discussion

Principal Findings

In this study, we examined the utility of Google Trends in identifying cycles of public interest in breast cancer and BCS. Although Internet access is still concentrated in metropolitan areas in Brazil, limiting Google Trends' use in areas with a low search volume, several studies seem to support the assumption that queries are sensible to foresee collective movements in real life. This is a well-known truism in marketing sciences grounded on studies on Google Trends' power in "predicting the present" [21], meaning that search volume correlates with contemporaneous events. The same rationale is being employed in several health research fields, and it has been useful to elucidate a wide range of questions from vaccination compliance [22], to protection against ultraviolet exposure during summer season [23], and interest in cancer issues after prevention campaigns [24].

There are two points to be highlighted in the results: (1) the growing interest in the early identification of a major public health problem and (2) the short collapse of this interest cycle at the end of each year (Table 1). In short, our results showed N-shaped RSV curves both in mammography and breast cancer (Figure 1), with higher peaks along Pink October months and sharp declines along December and January. This "cancer screening vortex" has been also described by Schootman and colleagues—the highest RSV along breast, colorectal, cervical, prostate, and lung cancer screening campaigns and the lowest during December-January [10]. In this case, this gap may be due to the Brazilian cultural aspects concerning summer vacations, Christmas, and New Year's celebrations. People tend to disregard issues related to illness and death, typically postponing some health decisions for the next year.

In Brazil, Pink October's strategy has been planned to promote collective interest in BCS in the context of cultural taboos and misconceptions. In fact, interest in breast cancer seems to significantly increase in October, although with several peaks throughout the year, with no evident seasonality. Likewise, in recent years, RSV concerning the early diagnosis of breast cancer has been significantly higher along Pink October months. It seems to be growing almost exponentially, and perhaps will go beyond searches on breast cancer in the next few years. It is consistent, with several works describing the use of Internet (boosted by higher educational level and the worldwide widespread of mobile phones) as a resource to self-care [25,26]. There is also a close correlation between the level of education—which has grown in Brazil in the last decades [27]—and accesses by Google to issues concerning science and health [28]. Interest in breast cancer always outperformed (in absolute and relative terms) mammography, but showed erratic patterns over the years and irregular growth in annual means. This may be consequent to events without seasonality, linked to the high incidence of new cases and constant media coverage—especially among celebrities who seem to boost the number of hits [29,30].

Surveillance systems databases have been very useful to assess self-reported cancer screening utilization. These data have been invaluable in identifying determinants of screening practices and describing trends and regional inequalities over time [31,32]. Unfortunately, due to the need of massive survey interviews for data collection, these database systems are too costly for low-income countries [33]. The complexity of a suitable survey structure required to aggregate reliable data, requiring the participation of a large study population, is also a huge obstacle for the largest country in Latin America. Methodological problems are also involved—public health planners must consider accuracy problems caused by self-report questionnaires and selection bias [32]. As a result, the Brazilian

population-based prevalence of cancer screening methods are not precise and the cultural impact of Pink October campaigns concerning BCS behavior is still unknown. Schootman and colleagues [10] examined the utility of Google Trends relative to a surveillance system focused on cancer screening (behavioral risk factor surveillance system). Social interest in learning about cancer screening exams was compared with surveillance systems based on self-reported use of these tests. In the same manner, the present results are eloquent to point out that attention has been increasingly drawn to the means of early identification of cancer. It is not clear if these findings may be taken as a plain evidence of well-succeeded campaigns supported by huge Brazilian government investments in access to screening [34] leveraged by the raise in educational levels [27] and widespread use of Google in mobile phones, tablets, notebooks, and desktops [17]. It is not possible to be sure if women moved forward from curiosity in Google queries to effective action. Nonetheless, the number of mammograms performed in the Brazilian Public Health System has jumped to just over 2.5 million (61.9% growth) in the period studied [34].

Timing-Based Strategies and the “Cancer Vortex”

There are several reports in the literature concerning campaigns and health interventions based on “what” and “how” (selection of qualified information and proper vehicles to deliver messages) [35,36], “where” (environments in which campaigns would be more effective) [37,38], and “who” (who are the best counselors) [30,39,40]. Nonetheless, reports based on “when”—the ideal timing for intervention—are not so frequent. Given the described findings and considering that the effectiveness of campaigns may be influenced by their impact and persistency in everyday life (measured in terms of “intensity,” “duration,” and convergence with relevant facts), it would be reasonable to consider some changes in Pink October timing. It would be reasonable to assume that, in Brazil, the anticipation of “Pink October” to the beginning of each year could extend the beneficial effect of the campaigns. Considering that both RSV curves decline sharply in December and January of each year (consistently with other authors in other continents [10]), would it be reasonable to expand the beneficial effects of Pink October by adding some months between its interest peak and the “December-January cancer vortex”? If we go further in this perspective and change to “Pink February,” would more people be interested in BCS for a longer period of time? Following this reasoning, it would be more appropriate to have screening campaigns at the beginning of the year, when new resolutions are taken and new projects are added to everyday routines.

Limitations

In Brazil, Web access is still concentrated in (but not limited to) metropolitan areas, which would limit the use of Google Trends in rural areas or regions with a low search volume. In fact, specific subpopulations and their cultural disparities may not be reachable by RSV algorithmics. In addition, Google Trends data only represent searches performed in Google. It is also important to consider that, although it represents a simple and low-cost alternative to nationwide screening database, Google Trends is still insufficient to describe screening behavior peculiarities at a global level. Nonetheless, as mentioned before, several works have described information-seeking behavior as a proxy of self-care attitudes. The potential of Google Trends to generate hypotheses about public awareness and interest in multiple aspects of cancer is also well documented [6,9,10].

Future Work

Future studies based on algorithmics sensible to interest cycles among small community groups should be useful to plan interventions tailored to the local needs. Study designs and analytic tools more appropriate to estimate the effects of media coverage on screening behavior would also be of invaluable help.

Conclusions

The leading goal of this study is to raise attention to forecasting methods using massive data to encourage health policy makers to undertake more sophisticated analyses based on classic marketing practices. Timely evaluations after campaigns may inform policy makers about awareness and interest seasonal cycles, which would leverage further interventions. Transparency of methods, simplicity, and reproducibility make the use of these new approaches an important alternative for low-income and huge-dimension countries. Timing-based strategies and Google Trends evaluations after campaigns may inform policy makers about seasonal cycles of attention and interest, which would leverage further interventions. We believe that patterns described here can be useful as baselines to help campaign analysts get started with specialized techniques that can be subsequently employed in more effective campaigns. The understanding and proper use of Google Trends oscillations, even being common sense for marketing researchers, are challenging for disciplines like public health, where government agencies work with a different concept of timing and public health demands. However, RSV trends should be clear for public communication planners with broad perspectives and committed in a timely fashion with users' demands.

Conflicts of Interest

None declared.

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 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/)]
2. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med* 2011 May;40(5):S154-S158. [doi: [10.1016/j.amepre.2011.02.006](https://doi.org/10.1016/j.amepre.2011.02.006)] [Medline: [21521589](https://pubmed.ncbi.nlm.nih.gov/21521589/)]

3. 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/)]
4. Cjuno J, Taype-Rondan A. [Headache seasonality in the Northern and Southern hemispheres: an approach using Google Trends]. *Rev Med Chil* 2016 Jul;144(7):947 [FREE Full text] [doi: [10.4067/S0034-98872016000700019](https://doi.org/10.4067/S0034-98872016000700019)] [Medline: [27661561](https://pubmed.ncbi.nlm.nih.gov/27661561/)]
5. 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;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/)]
6. Ling R, Lee J. Disease monitoring and health campaign evaluation using Google search activities for HIV and AIDS, stroke, colorectal cancer, and marijuana use in Canada: a retrospective observational study. *JMIR Public Health Surveill* 2016;2(2):e156 [FREE Full text] [doi: [10.2196/publichealth.6504](https://doi.org/10.2196/publichealth.6504)] [Medline: [27733330](https://pubmed.ncbi.nlm.nih.gov/27733330/)]
7. Goel S, Hofman JM, Lahaie S, Pennock DM, Watts DJ. Predicting consumer behavior with Web search. *Proc Natl Acad Sci USA* 2010 Oct 12;107(41):17486-17490. [doi: [10.1073/pnas.1005962107](https://doi.org/10.1073/pnas.1005962107)] [Medline: [20876140](https://pubmed.ncbi.nlm.nih.gov/20876140/)]
8. Anker AE, Reinhart AM, Feeley TH. Health information seeking: a review of measures and methods. *Patient Educ Couns* 2011;82(3):346-354. [doi: [10.1016/j.pec.2010.12.008](https://doi.org/10.1016/j.pec.2010.12.008)] [Medline: [21239134](https://pubmed.ncbi.nlm.nih.gov/21239134/)]
9. Glynn RW, Kelly JC, Coffey N, Sweeney KJ, Kerin MJ. The effect of breast cancer awareness month on internet search activity--a comparison with awareness campaigns for lung and prostate cancer. *BMC Cancer* 2011 Oct 12;11:442 [FREE Full text] [doi: [10.1186/1471-2407-11-442](https://doi.org/10.1186/1471-2407-11-442)] [Medline: [21993136](https://pubmed.ncbi.nlm.nih.gov/21993136/)]
10. Schootman M, Toor A, Cavazos-Rehg P, Jeffe DB, McQueen A, Eberth J, et al. The utility of Google Trends data to examine interest in cancer screening. *BMJ Open* 2015 Jun 08;5(6):e006678 [FREE Full text] [doi: [10.1136/bmjopen-2014-006678](https://doi.org/10.1136/bmjopen-2014-006678)] [Medline: [26056120](https://pubmed.ncbi.nlm.nih.gov/26056120/)]
11. Benson JR, Jatoti I. The global breast cancer burden. *Future Oncol* 2012 Jun;8(6):697-702. [doi: [10.2217/fon.12.61](https://doi.org/10.2217/fon.12.61)] [Medline: [22764767](https://pubmed.ncbi.nlm.nih.gov/22764767/)]
12. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005;55(2):74-108 [FREE Full text] [Medline: [15761078](https://pubmed.ncbi.nlm.nih.gov/15761078/)]
13. Parkin DM, Fernández LM. Use of statistics to assess the global burden of breast cancer. *Breast J* 2006;12(Suppl 1):S70-S80. [doi: [10.1111/j.1075-122X.2006.00205.x](https://doi.org/10.1111/j.1075-122X.2006.00205.x)] [Medline: [16430400](https://pubmed.ncbi.nlm.nih.gov/16430400/)]
14. Tabár L, Vitak B, Chen TH, Yen AM, Cohen A, Tot T, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology* 2011 Sep;260(3):658-663. [doi: [10.1148/radiol.11110469](https://doi.org/10.1148/radiol.11110469)] [Medline: [21712474](https://pubmed.ncbi.nlm.nih.gov/21712474/)]
15. Walters S, Maringe C, Butler J, Rachet B, Barrett-Lee P, Bergh J, ICBP Module 1 Working Group. Breast cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK, 2000-2007: a population-based study. *Br J Cancer* 2013 Mar 19;108(5):1195-1208 [FREE Full text] [doi: [10.1038/bjc.2013.6](https://doi.org/10.1038/bjc.2013.6)] [Medline: [23449362](https://pubmed.ncbi.nlm.nih.gov/23449362/)]
16. Amorim V, Barros M, César C, Carandina L, Goldbaum M. Factors associated with lack of mammograms and clinical breast examination by women: a population-based study in Campinas, São Paulo State, Brazil. *Cad Saude Publica* 2008;24(11):2623-2632. [Medline: [19009142](https://pubmed.ncbi.nlm.nih.gov/19009142/)]
17. Damiani G, Basso D, Acampora A, Bianchi C, Silvestrini G, Frisicale E, et al. The impact of level of education on adherence to breast and cervical cancer screening: evidence from a systematic review and meta-analysis. *Prev Med* 2015;81:28. [doi: [10.1016/j.yjmed.2015.09.011](https://doi.org/10.1016/j.yjmed.2015.09.011)] [Medline: [26408405](https://pubmed.ncbi.nlm.nih.gov/26408405/)]
18. Lee BL, Liedke PE, Barrios CH, Simon SD, Finkelstein DM, Goss PE. Breast cancer in Brazil: present status and future goals. *Lancet Oncol* 2012 Mar;13(3):e95-e102. [doi: [10.1016/S1470-2045\(11\)70323-0](https://doi.org/10.1016/S1470-2045(11)70323-0)] [Medline: [22381937](https://pubmed.ncbi.nlm.nih.gov/22381937/)]
19. Instituto Nacional de Câncer. INCA. 2016. Estimativa 2016 URL: <http://www.inca.gov.br/estimativa/2016/sintese-de-resultados-comentarios.asp> [accessed 2017-03-20] [WebCite Cache ID 6p6h9Pp0r]
20. Nuti S, Wayda B, Ranasinghe I, Wang S, Dreyer R, Chen S, et al. The use of Google Trends in health care research: a systematic review. *PLoS One* 2014;9(10):e109583. [doi: [10.1371/journal.pone.0109583](https://doi.org/10.1371/journal.pone.0109583)] [Medline: [25337815](https://pubmed.ncbi.nlm.nih.gov/25337815/)]
21. Choi H, Varian H. Google. 2009. Predicting the present with Google Trends: Technical report URL: http://google.com/googleblogs/pdfs/google_predicting_the_present.pdf [accessed 2017-03-20] [WebCite Cache ID 6p6hIMVhE]
22. Barak-Corren Y, Reis BY. Internet activity as a proxy for vaccination compliance. *Vaccine* 2015 May 15;33(21):2395-2398 [FREE Full text] [doi: [10.1016/j.vaccine.2015.03.100](https://doi.org/10.1016/j.vaccine.2015.03.100)] [Medline: [25869888](https://pubmed.ncbi.nlm.nih.gov/25869888/)]
23. Vasconcellos-Silva PR, Castiel LD, Griep RH. Patterns of access to information on protection against UV during the Brazilian summer: is there such a thing as the "summer effect"? *Cien Saude Colet* 2015;20(8):2533-2538. [doi: [10.1590/1413-81232015208.18932014](https://doi.org/10.1590/1413-81232015208.18932014)]
24. Vasconcellos-Silva PR, Castiel LD, Griep RH, Zanchetta M. Cancer prevention campaigns and Internet access: promoting health or disease? *J Epidemiol Community Health* 2008 Oct;62(10):876-881. [doi: [10.1136/jech.2007.069054](https://doi.org/10.1136/jech.2007.069054)] [Medline: [18791045](https://pubmed.ncbi.nlm.nih.gov/18791045/)]
25. Jiang S, Street R. Pathway linking Internet health information seeking to better health: a moderated mediation study. *Health Commun* 2016:1-8 Epub ahead of print. [doi: [10.1080/10410236.2016.1196514](https://doi.org/10.1080/10410236.2016.1196514)] [Medline: [27464036](https://pubmed.ncbi.nlm.nih.gov/27464036/)]

26. Cao W, Zhang X, Xu K, Wang Y. Modeling online health information-seeking behavior in China: the roles of source characteristics, reward assessment, and Internet self-efficacy. *Health Commun* 2016 Sep;31(9):1105-1114. [doi: [10.1080/10410236.2015.1045236](https://doi.org/10.1080/10410236.2015.1045236)] [Medline: [26861963](https://pubmed.ncbi.nlm.nih.gov/26861963/)]
27. Portal B. Brasil. 2014 Nov 05. Women's education increases relative to men URL: <http://www.brasil.gov.br/cidadania-e-justica/2014/11/escolaridade-das-mulheres-aumenta-em-relacao-a-dos-homens> [accessed 2017-03-20] [WebCite Cache ID 6p6hvblyj]
28. Segev E, Baram-Tsabari A. Seeking science information online: data mining Google to better understand the roles of the media and the education system. *Public Underst Sci* 2012 Oct;21(7):813-829. [doi: [10.1177/0963662510387560](https://doi.org/10.1177/0963662510387560)] [Medline: [23832560](https://pubmed.ncbi.nlm.nih.gov/23832560/)]
29. Cooper CP, Gelb CA, Lobb K. Celebrity appeal: reaching women to promote colorectal cancer screening. *J Womens Health (Larchmt)* 2015 Mar;24(3):169-173 [FREE Full text] [doi: [10.1089/jwh.2014.5084](https://doi.org/10.1089/jwh.2014.5084)] [Medline: [25521047](https://pubmed.ncbi.nlm.nih.gov/25521047/)]
30. Vasconcellos-Silva PR, Sormunen T, Craftman AG. Evolution of accesses to information on breast cancer and screening on the Brazilian national cancer institute website: an exploratory study. *Cien Saude Colet* 2017;21(8):25-28 (forthcoming).
31. Joseph DA, King JB, Miller JW, Richardson LC, Centers for Disease Control and Prevention (CDC). Prevalence of colorectal cancer screening among adults--Behavioral Risk Factor Surveillance System, United States, 2010. *MMWR Suppl* 2012 Jun 15;61(2):51-56. [Medline: [22695464](https://pubmed.ncbi.nlm.nih.gov/22695464/)]
32. Hiatt RA, Klabunde C, Breen N, Swan J, Ballard-Barbash R. Cancer screening practices from National Health Interview Surveys: past, present, and future. *J Natl Cancer Inst* 2002 Dec 18;94(24):1837-1846. [Medline: [12488477](https://pubmed.ncbi.nlm.nih.gov/12488477/)]
33. Dey S. Preventing breast cancer in LMICs via screening and/or early detection: the real and the surreal. *World J Clin Oncol* 2014 Aug 10;5(3):509-519 [FREE Full text] [doi: [10.5306/wjco.v5.i3.509](https://doi.org/10.5306/wjco.v5.i3.509)] [Medline: [25114864](https://pubmed.ncbi.nlm.nih.gov/25114864/)]
34. Portal B. Brasil. 2017. Number of mammograms performed by SUS in four years grows by 62% URL: <http://www.brasil.gov.br/saude/2015/10/mamografias-realizadas-no-sus-cresceram-62-em-quatro-anos> [accessed 2017-03-20] [WebCite Cache ID 6p6jBmNsr]
35. Airhihenbuwa CO, Obregon R. A critical assessment of theories/models used in health communication for HIV/AIDS. *J Health Commun* 2000;5 Suppl:5-15. [doi: [10.1080/10810730050019528](https://doi.org/10.1080/10810730050019528)] [Medline: [11010357](https://pubmed.ncbi.nlm.nih.gov/11010357/)]
36. Cook PF, Carrington JM, Schmiege SJ, Starr W, Reeder B. A counselor in your pocket: feasibility of mobile health tailored messages to support HIV medication adherence. *Patient Prefer Adherence* 2015;9:1353-1366 [FREE Full text] [doi: [10.2147/PPA.S88222](https://doi.org/10.2147/PPA.S88222)] [Medline: [26491263](https://pubmed.ncbi.nlm.nih.gov/26491263/)]
37. Paskett ED, Tatum CM, D'Agostino Jr R, Rushing J, Velez R, Michielutte R, et al. Community-based interventions to improve breast and cervical cancer screening: results of the Forsyth County Cancer Screening (FoCaS) Project. *Cancer Epidemiol Biomarkers Prev* 1999 May;8(5):453-459 [FREE Full text] [Medline: [10350442](https://pubmed.ncbi.nlm.nih.gov/10350442/)]
38. Paskett ED, McMahan K, Tatum C, Velez R, Shelton B, Case LD, et al. Clinic-based interventions to promote breast and cervical cancer screening. *Prev Med* 1998;27(1):120-128. [doi: [10.1006/pmed.1997.0254](https://doi.org/10.1006/pmed.1997.0254)] [Medline: [9465362](https://pubmed.ncbi.nlm.nih.gov/9465362/)]
39. Noar SM, Willoughby JF, Myrick JG, Brown J. Public figure announcements about cancer and opportunities for cancer communication: a review and research agenda. *Health Commun* 2014;29(5):445-461. [Medline: [23845155](https://pubmed.ncbi.nlm.nih.gov/23845155/)]
40. Goodenberger ML, Thomas BC, Wain KE. The utilization of counseling skills by the laboratory genetic counselor. *J Genet Couns* 2015 Feb;24(1):6-17. [doi: [10.1007/s10897-014-9749-9](https://doi.org/10.1007/s10897-014-9749-9)] [Medline: [25138081](https://pubmed.ncbi.nlm.nih.gov/25138081/)]

Abbreviations

- BCS:** breast cancer screening
CSV: comma separated values
RSV: relative search volume

Edited by G Eysenbach; submitted 20.11.16; peer-reviewed by R Hawkins, E Davies; comments to author 04.01.17; revised version received 18.01.17; accepted 25.01.17; published 06.04.17.

Please cite as:

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;3(2):e17
URL: <http://publichealth.jmir.org/2017/2/e17/>
doi: [10.2196/publichealth.7015](https://doi.org/10.2196/publichealth.7015)
PMID: [28385679](https://pubmed.ncbi.nlm.nih.gov/28385679/)

©Paulo Roberto Vasconcellos-Silva, Dárlinton Barbosa Feres Carvalho, Valéria Trajano, Lucia Rodriguez de La Rocque, Anunciata Cristina Marins Braz Sawada, Leidjaira Lopes Juvanhol. Originally published in *JMIR Public Health and Surveillance*

(<http://publichealth.jmir.org>), 06.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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

How Do You #relax When You're #stressed? A Content Analysis and Infodemiology Study of Stress-Related Tweets

Son Doan¹, PhD; Amanda Ritchart², CPhil; Nicholas Perry³, MS; Juan D Chaparro¹, MD; Mike Conway⁴, PhD

¹Department of Biomedical Informatics, University of California, San Diego, La Jolla, CA, United States

²Linguistics Department, University of California, San Diego, La Jolla, CA, United States

³Department of Psychology, University of Utah, Salt Lake City, UT, United States

⁴Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States

Corresponding Author:

Son Doan, PhD

Department of Biomedical Informatics

University of California, San Diego

9500 Gilman Dr

La Jolla, CA, 92093

United States

Phone: 1 858 534 2230

Email: sondoan@gmail.com

Abstract

Background: Stress is a contributing factor to many major health problems in the United States, such as heart disease, depression, and autoimmune diseases. Relaxation is often recommended in mental health treatment as a frontline strategy to reduce stress, thereby improving health conditions. Twitter is a microblog platform that allows users to post their own personal messages (tweets), including their expressions about feelings and actions related to stress and stress management (eg, relaxing). While Twitter is increasingly used as a source of data for understanding mental health from a population perspective, the specific issue of stress—as manifested on Twitter—has not yet been the focus of any systematic study.

Objective: The objective of our study was to understand how people express their feelings of stress and relaxation through Twitter messages. In addition, we aimed at investigating automated natural language processing methods to (1) classify stress versus nonstress and relaxation versus nonrelaxation tweets, and (2) identify first-hand experience—that is, who is the experimenter—in stress and relaxation tweets.

Methods: We first performed a qualitative content analysis of 1326 and 781 tweets containing the keywords “stress” and “relax,” respectively. We then investigated the use of machine learning algorithms—in particular naive Bayes and support vector machines—to automatically classify tweets as stress versus nonstress and relaxation versus nonrelaxation. Finally, we applied these classifiers to sample datasets drawn from 4 cities in the United States (Los Angeles, New York, San Diego, and San Francisco) obtained from Twitter’s streaming application programming interface, with the goal of evaluating the extent of any correlation between our automatic classification of tweets and results from public stress surveys.

Results: Content analysis showed that the most frequent topic of stress tweets was education, followed by work and social relationships. The most frequent topic of relaxation tweets was rest & vacation, followed by nature and water. When we applied the classifiers to the cities dataset, the proportion of stress tweets in New York and San Diego was substantially higher than that in Los Angeles and San Francisco. In addition, we found that characteristic expressions of stress and relaxation varied for each city based on its geolocation.

Conclusions: This content analysis and infodemiology study revealed that Twitter, when used in conjunction with natural language processing techniques, is a useful data source for understanding stress and stress management strategies, and can potentially supplement infrequently collected survey-based stress data.

(*JMIR Public Health Surveill* 2017;3(2):e35) doi:[10.2196/publichealth.5939](https://doi.org/10.2196/publichealth.5939)

KEYWORDS

social media; Twitter; stress; relaxation; natural language processing; machine learning

Introduction

Psychological stress has been linked to multiple health conditions, including depression [1], heart disease [2], autoimmune disease [3], and general all-cause mortality [4]. Stress has also been associated with worse health outcomes among those living with chronic illness [5], suggesting that stress may exacerbate preexisting health conditions, as well as contribute to the development of new health problems. Stress not only contributes to physical and mental health problems, such as heart disease, depression, and autoimmune diseases [6], but also has negative impacts on family life and work, significantly impairing quality of life [7,8]. Accordingly, stress is an important concern for public health prevention initiatives [7,8].

Health surveys have demonstrated that stress negatively affects a large proportion of the US population [9]. Underscoring the magnitude of the problem, a study conducted by the Harvard School of Public Health found that 49% of the American public reported being stressed within the last year, and also found that 60% of those who reported being in poor health also reported experiencing a substantial amount of stress within the last month [7]. Further, levels of stress appear to be unequally distributed throughout the population [10]. National surveys have documented that higher levels of stress are reported among those who have lower income, are less educated, and are younger [11]. Theorists have suggested that geographic clustering of psychological characteristics may be driven by selective migration (in this case, people more vulnerable to stress seek out others like themselves), social influence (ie, people with attitudes and beliefs that lead to greater stress cluster together geographically), or environmental influence (ie, features of the physical environment, such as neighborhoods, increase stress among those who live close to one another) [12]. In short, large-scale studies have documented both the high prevalence of stress within the United States and geographic clustering of psychological distress, suggesting that symptoms of stress should ideally be tracked at both the national and local levels.

Relaxation is considered a key component of frontline stress management techniques, such as cognitive-behavioral stress management [13]. General stress management can include adaptive coping (eg, distraction), physical relaxation strategies (eg, diaphragmatic breathing), cognitive reappraisal (eg, reconsidering the stressor from a different perspective), and mindfulness (ie, increasing awareness of the present moment). These stress management strategies are intended to reduce psychological and physiological arousal related to stress, promote healthier coping alternatives, and, in turn, reduce some of the negative health impacts of stress. Indeed, these strategies have been found to be effective for improving health outcomes among those living with chronic illness [14-16], as well as for improving general mental health and quality of life [17,18].

Understanding what the major causes of stress are and how people negatively or positively manage their stress (eg, through stress management techniques such as cultivating relaxation) is important [7,19]. Population health surveys often use telephone interviews or questionnaires from samples of the

population, such as the US Centers for Disease Control and Prevention's (CDC) Behavioral Risk Factor Surveillance System [20]. These methods, although reliable, are conducted relatively infrequently due to cost and may be less effective at reaching certain populations, such as those without a dedicated landline telephone. With the rapid growth of online social networks today, social media data can serve as a useful additional resource to understand aspects of stress that are difficult to assess in general surveys or clinical care. For example, social media provide a means to rapidly and dynamically address new and evolving research questions with a degree of flexibility not possible with surveys. Social media may also provide insights into populations that may be underrepresented in surveys (depending on the demographics of the particular social media platform used). Thus, social media can potentially serve as a beneficial supplement to detailed surveys when trying to understand public health concerns.

Twitter—one of the most popular social media platforms—is a microblog service that allows users to post their own personal messages (a “tweet” with a 140-character limit). As of May 2016, it had 310 million active users with 1 billion unique visits monthly to sites with embedded tweets [21]. The utility of Twitter as a data source has been investigated in numerous applications such as election prediction [22], stock market prediction [23], oil price changes [22], and earthquake and disasters [24].

Twitter has also been used in public health for tracking influenza [25-27], and for studying breast cancer prevention [28], childhood obesity [29], issues related to general health [30], tobacco and e-cigarette use [31], dental pain [32,33], general pain [34], sexually transmitted diseases [35], and weight loss [36]. There has also been research regarding the general well-being of people in different geographical locations using Twitter messages [37], and correlation studies of Twitter messages with depression [38] and with heart disease mortality [39]. However, to our knowledge, no studies specifically focused on stress and stress management have been conducted until now.

In this study, we investigated how people express their own stress and relaxation through an in-depth content analysis of Twitter messages. In addition, we investigated automated methods to classify stress and relaxation tweets using machine learning techniques. Furthermore, we ranked stress and relaxation levels based on the relative proportions of stress- and relaxation-related tweets (as identified by our natural language processing classifiers) originating in 4 US cities: New York, Los Angeles, San Diego, and San Francisco. We then compared these results with public surveys reported by Forbes and CNN [40,41]. Using easily acquired, naturalistic Twitter data, and complementing existing survey-based epidemiological methods, this study provides another perspective on how people think about and cope with stress.

Methods

Data Collection

Dataset 1

To begin our investigation of stress and relaxation (stress management) tweets, we first collected tweets with user-defined stress and relaxation topics using the Twitter REST application programming interface (API) [42]. The user-defined topics included the hashtagged topics #stress and #relax, as well as

variations of these words. [Textbox 1](#) lists the full search list we used. We collected tweets between July 9 and July 14, 2014. We supplemented this seed dataset with tweets from the random sample stream Twitter streaming API [43] (1% sample rate) in order to have better representation of “everyday” tweets that did not necessarily contain stress- and relaxation-related hashtags, but that still contained the keywords “stress” or “relax.” This dataset consisted of 1326 stress-related and 781 relaxation-related tweets. We referred to this dataset as dataset 1.

Textbox 1. List of hashtags related to stress and relaxation to create dataset 1.

Stress-related hashtags
#stress
#stressed
#stressful
#stressin
#stressing
#sostressful
#sostressed
#stressinout
#stressingout
Relaxation-related hashtags
#relaxed
#relaxin
#relaxing
#sorelaxin
#sorelaxing

Dataset 2

We further investigated the characteristics of stress and stress management by geographical location (4 US cities) and compared the locations against each other using dataset 2. This dataset—much larger than dataset 1—consisted of geotagged tweets obtained from the Twitter streaming API [43] in 1 of 4 possible cities: Los Angeles, New York, San Diego, and San Francisco. We chose these cities because they are densely populated and major metropolitan areas on the east and west coasts of the United States. Tweets were collected between September 30, 2013 and February 10, 2014. The number of tweets for each city for this time period was 8.2 million for New York, 6.6 million for Los Angeles, 3 million for San Diego, and 4.4 million for San Francisco. Note that the most populous cities—that is, New York and Los Angeles—generated the greatest number of tweets during the study period. We referred to this dataset as dataset 2.

Criterion Standard and Manual Analysis of Tweets

Since our primary goal in this study was to understand how people express stress and relaxation through Twitter, we developed annotation guidelines for both stress and relaxation tweets based on reports from the American Psychological Association [7], CDC [8,44], and medical websites [6,45,46].

Following these guidelines, we classified tweets by both genre and theme. Genre reflects the format of the tweet (eg, personal experience), and theme reflects the domain of the actual content conveyed (including such categories as stress symptoms and stress topics).

Details for each genre and theme for stress and relaxation tweets were as follows.

Genre

We categorized tweets as being first-hand experience versus other genres. We defined first-hand experience as a direct personal experience, or an experience directly related to the user writing the tweet. Other genres were second-hand experience, advertisements, news articles, etc. This genre classification was based on previous work on classifying health-related tweets [31]. After classifying a tweet as first-hand experience, we assigned its content into 2 themes: stress and relaxation.

Stress Themes

Content analysis focused on 3 main questions: (1) What kind of stress was being experienced? (2) What was the cause of the stress? and (3) What kind of actions, if any, were being taken regarding the stress? Based on these questions, we categorized

the theme into 3 categories: stress symptoms, topics, and action(s) taken.

Symptoms fell into 3 classes: (1) psychological and emotional, (2) physical, and (3) behavioral. These categories were based on guidelines for stress symptoms [47-49].

Topics referred to the general topic of a tweet: (1) work, (2) education, (3) finances, (4) social relationships, (5) travel, (6) temporal, and (7) other. These topics were identified based on an analysis of data from dataset 1.

The action taken theme indicated the action that people reported taking when they were stressed. The action could be either negative or positive. An example of a negative action is “I need a drink tonight. #sostressed.” An example of a positive action is “I need a nap, and a hug. #stressingout #tired.”

The nonspecific theme was for users who simply tweeted without any symptom, topic, or action; for example, “#stressed!!!,” “Bad Night :(,” and “#SoStressed.”

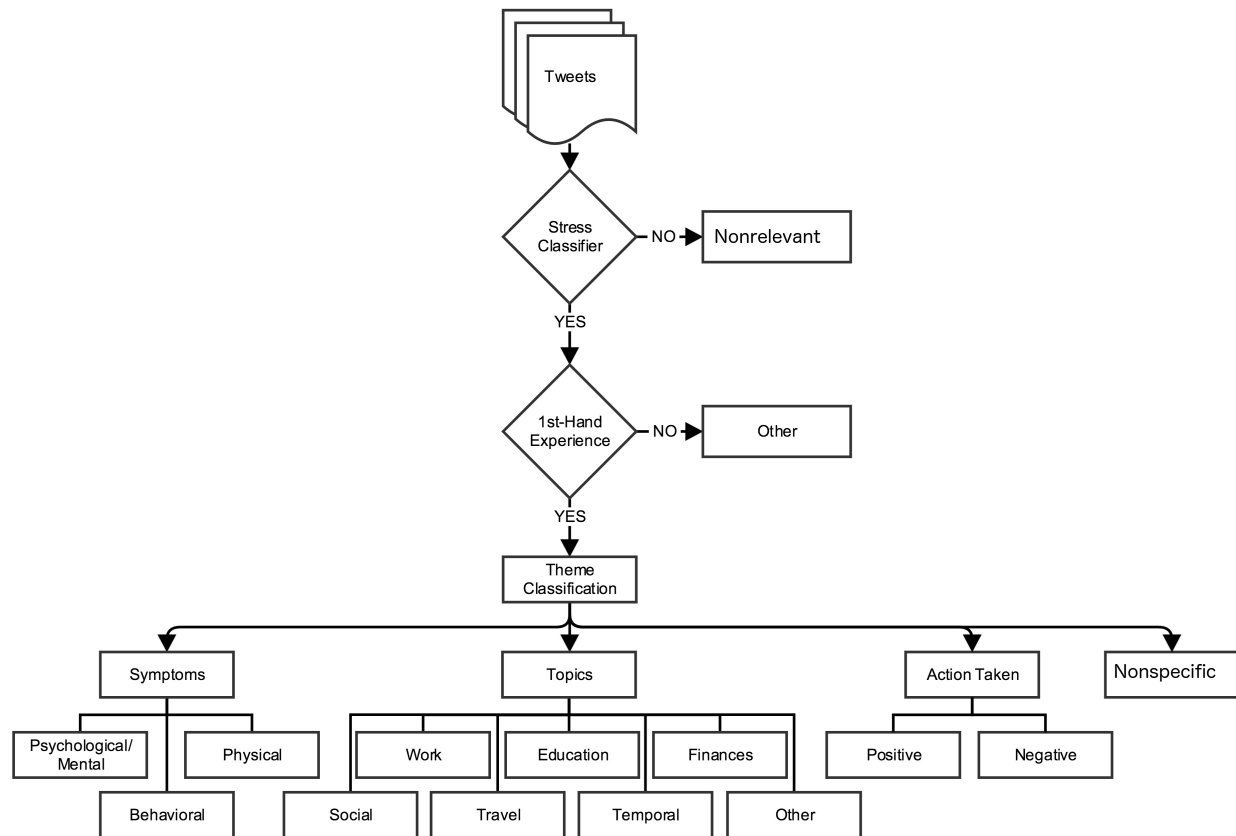
Relaxation Themes

We categorized first-hand experience relaxation tweets by the following topics (themes), which referred to the action reported being taken by the user to relax, such as exercising or listening to music. We created 11 topics based on data from dataset 1: (1) physical, (2) water, (3) self-care, (4) alcohol & drugs, (5) entertainment & hobbies, (6) food & drink, (7) nature, (8) rest & vacation, (9) social relationships, (10) other, and (11) nonspecific.

Figure 1 depicts the schema for stress tweets and Figure 2 depicts the schema for relaxation tweets. Definitions and examples of each category of first-hand experience tweets and its themes for stress and relaxation tweets are listed in Multimedia Appendix 1 and Multimedia Appendix 2, respectively.

One author (AR) annotated stress and relaxation tweets from dataset 1 and another (SD) annotated and verified the dataset to ensure that all tweets were annotated correctly. Any disagreements were resolved by meetings or exchanging emails. Dataset 1 contained a total of 664 stress and 662 nonstress tweets among the 1326 stress-related tweets, and a total of 391 relaxation and 390 nonrelaxation tweets among the 781 relaxation-related tweets. For each stress or relaxation tweet, 2 authors (AR, SD) discussed and manually annotated tweets based on the guidelines as described above. After annotation, there were a total of 479 stress tweets and 335 relaxation tweets related to first-hand experience in dataset 1. Figure 3 depicts the details of dataset 1.

Since the prevalences of some of the stress themes (eg, finances, work) and relaxation themes (eg, food & drink, social) in dataset 1 were very low (ie, too infrequent to train a machine learning classifier), we developed an automatic keyword-based theme classifier using a manually crafted lexicon of stress and relaxation keywords associated with each category. We first generated unigrams and bigrams from dataset 1, and one author (AR) manually reviewed and selected the highest-frequency unigram and bigram keywords. We then manually added corresponding synonyms into each theme to increase the coverage of the classifier. For example, the topic “education” in the stress schema contained the unigrams “school,” “college,” and “classes” and the bigram “high school” in dataset 1. We manually added synonyms of those terms, such as “exams” and “studying” as unigram keywords and “college life,” “my tuition,” and “on finals” into bigram keywords. The list was iteratively reviewed and confirmed by another author (SD). There was an average of 20 unigram and 20 bigram terms for each theme. We created only unigram and bigram keywords, since tweet messages are short in nature. Bigram keywords were necessary to include idiomatic expressions like “vicious cycle” and “hate feeling,” and they also added more specificity, such as “my heart” and “my sanity,” which helped to increase the accuracy of the classifiers.

Figure 1. Schema used to classify stress tweets.

Machine Learning Algorithms

Leveraging the annotated data derived from our content analysis of dataset 1, we applied and evaluated machine learning algorithms for classification of stress versus nonstress tweets and relaxation versus nonrelaxation tweets (on dataset 1). To apply the classifier trained on dataset 1 to the unseen, much larger dataset 2 (cities dataset), we first filtered tweets by keeping only the tweets that contained stress- or

relaxation-related hashtags in [Textbox 1](#) or the keywords “stress” or “relax” for each city in dataset 2. After this step, dataset 2 contained only tweets with stress- or relaxation-related keywords or hashtags. To calculate the proportion of stress or relaxation tweets at the city level, we used the stress or relaxation classifier trained on dataset 1 to filter stress or relaxation tweets, and then applied the classifier for first-hand experienter to tweets from each city in dataset 2. [Figure 4](#) shows a flowchart describing our machine learning design.

Figure 2. Schema used to classify relaxation tweets.

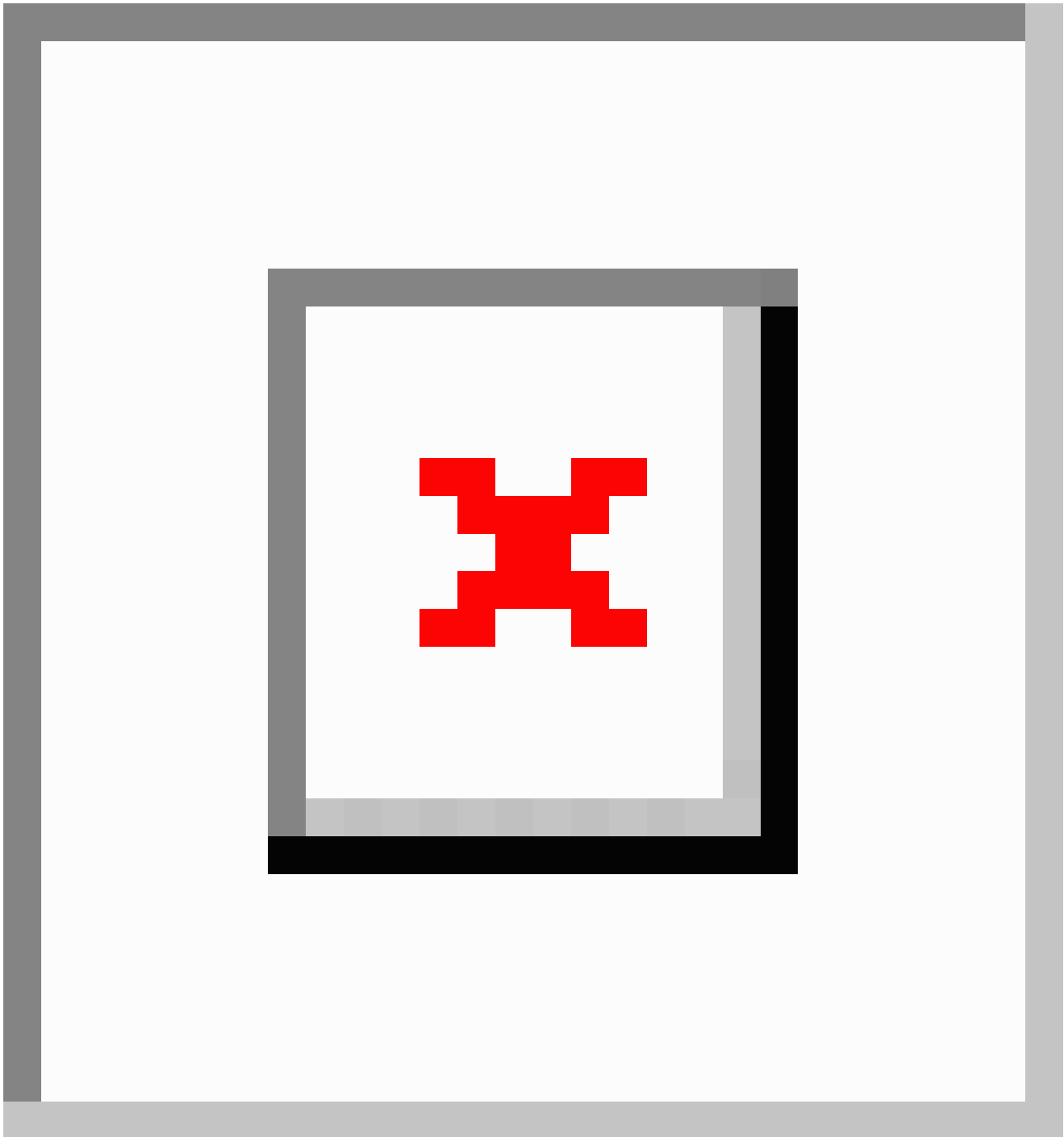
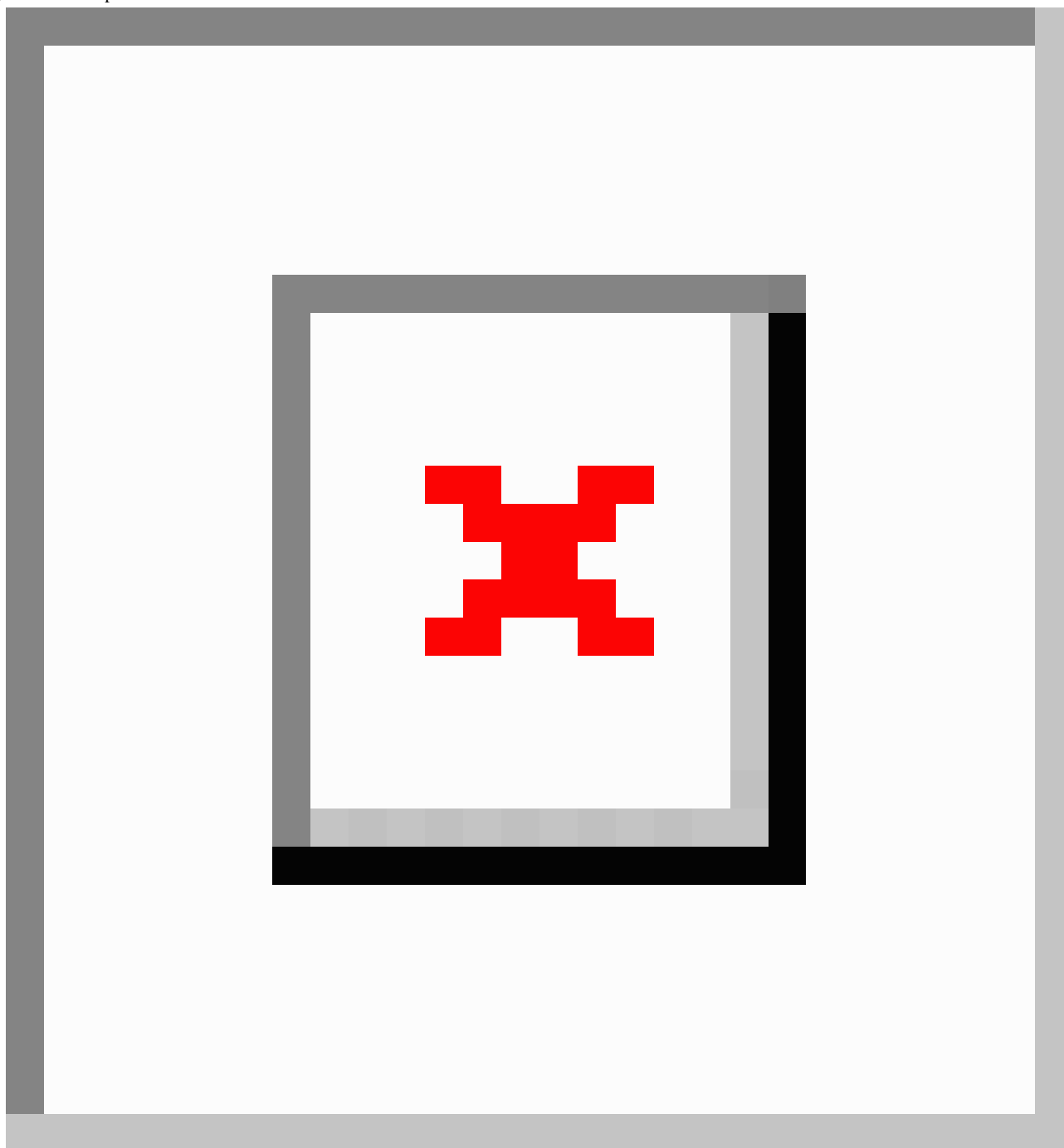
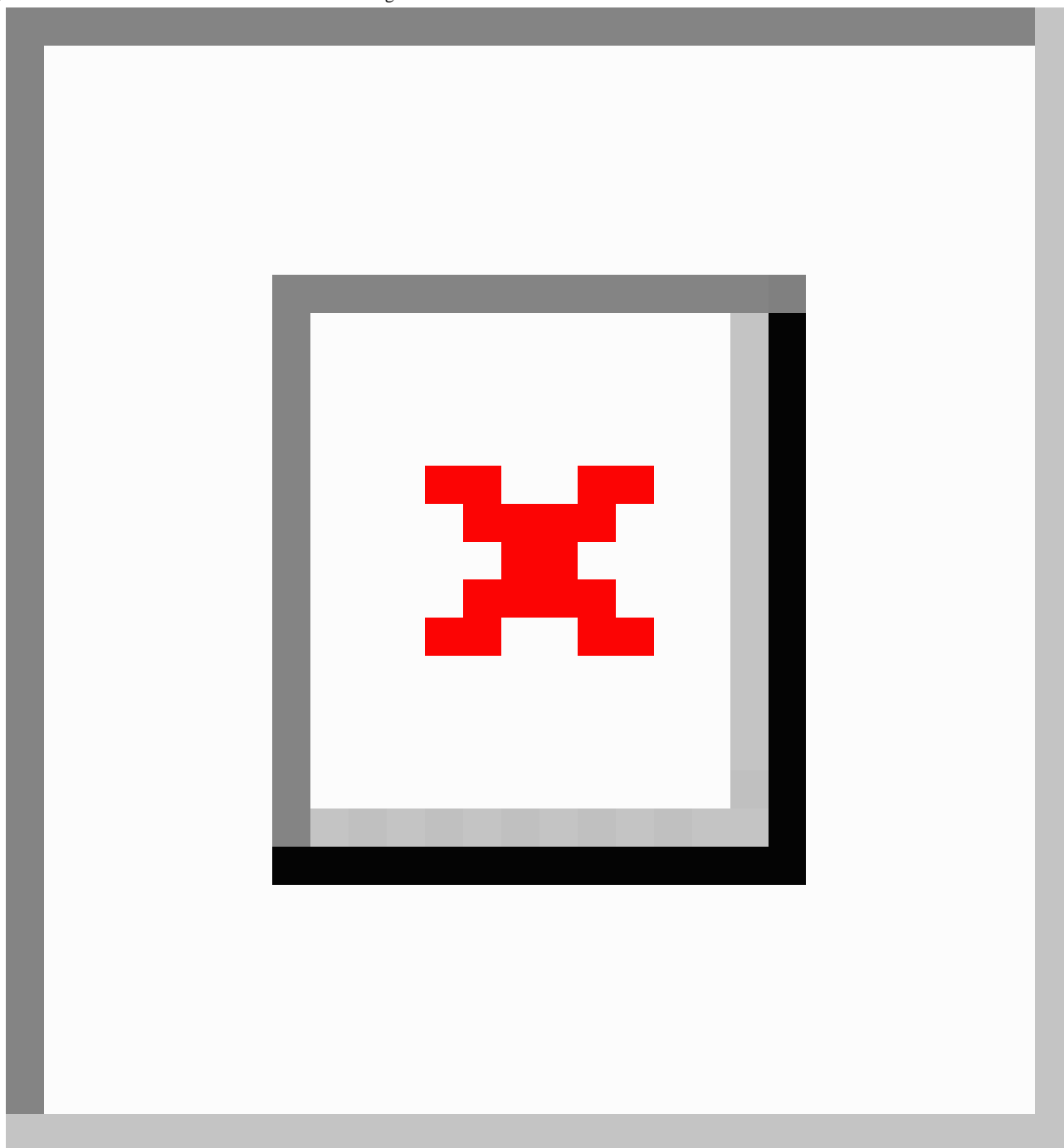


Figure 3. Description of dataset 1.

Our study focused on 2 machine learning-based classification tasks. First, tweets were classified into the appropriate stress and relaxation category (ie, is it stress or relaxation related?). Second, first-hand experience tweets versus nonfirst-hand experience tweets were classified. We used 2 machine learning algorithms: naive Bayes and support vector machines (SVMs), which were implemented on dataset 1 using 10-fold cross-validation. We used both the naive Bayes and SVM algorithms, as both these algorithms have been used extensively

for text classification tasks [50-52]. We used the *Rainbow* package [51] for implementing both naive Bayes and SVMs (linear kernel). We used “bag-of-words” as feature sets for both algorithms. The reason we used the bag-of-words representation is that this feature representation is considered as a baseline and the most common text representation in text classification in general [50-52]. To the best of our knowledge, this is the first study on classifying tweets on stress and relaxation tweets.

Figure 4. Datasets and tasks used for machine learning.

Calculating the Proportion of Stress and Relaxation Tweets at the City Level

We applied the 2-step classification to each city in dataset 2 to automatically identify stress and relaxation tweets. We calculated the proportions of stress and relaxation tweets to the total number of tweets in each city.

Measurements and Statistical Analysis

For both stress or relaxation and first-hand experience classifications, we used accuracy, sensitivity, specificity, and positive predictive values (PPVs) as metrics [53-55]. They were defined as follows: sensitivity = $TP/(TP + FN)$; PPV = $TP/(TP + FP)$; specificity = $TN/(FP + TN)$; and accuracy = $(TP +$

$TN)/(TP + TN + FP + FN)$, where TP is the number of tweets that are correctly classified as true, FP is the number of tweets that are incorrectly classified as true, FN is the number of tweets that are true but incorrectly classified as false, and TN is the number of tweets that are correctly classified as false.

To compare data among cities, we used Pearson chi-square test and reported significance if the *P* value was less than .05 [56]. Statistical analyses were performed using the publicly available R package software version 3.2.3 (R Foundation). Note that, to preserve the anonymity of Twitter users, all example tweets reported in this paper are paraphrases of original tweets.

Results

Content Analysis in Stress and Relaxation Tweets (Dataset 1)

Figure 5 shows the distribution of themes in first-hand experience stress tweets. The highest-frequency theme in stress tweets was topic, followed by nonspecific (eg, “#stressed!!!”), action taken (eg, “I need a drink #sostressed”), and symptoms (eg, “Not sure what to do...#stressed #worried #lost”). This suggests that Twitter users who posted about stress usually posted more about the cause or topic of their stress and less about actions and symptoms associated with stress.

Among the total number of stress-related tweets, as Figure 5 shows, the most frequent topic was education, followed by other topic, work, and social relationships. This is interesting because many of Twitter’s users are young people who attend school [57,58]. It seems that education and issues related to education, such as exams and finals, were of the utmost concern for Twitter

users. Examples of the education topic are “Never doing a session B math course ever again #sostressful” and “my exam in less than a month?! #stressing.” Figure 6 shows the topic distribution of first-hand experience stress tweets.

Relaxation-related tweets encompassed a wider range of topics than stress-related tweets. The most frequent topic of relaxation tweets was rest & vacation, followed by nature and water. Figure 7 shows topic distribution of first-hand experience of relaxation tweets.

Automatic Classification of Stress and Relaxation Tweets (Dataset 1)

Table 1 shows cross-validated classification results. Our results indicated that both algorithms achieved high accuracy (range 78.08%-85.64%), sensitivity (range 90.26%-99.09%), and PPV (range 70.68%-89.32%). Specificity was rather lower, especially with first-hand relaxation classification (naive Bayes: 11.67%, SVM: 18.33%).

Figure 5. Distribution by theme of first-hand experience stress tweets in dataset 1.

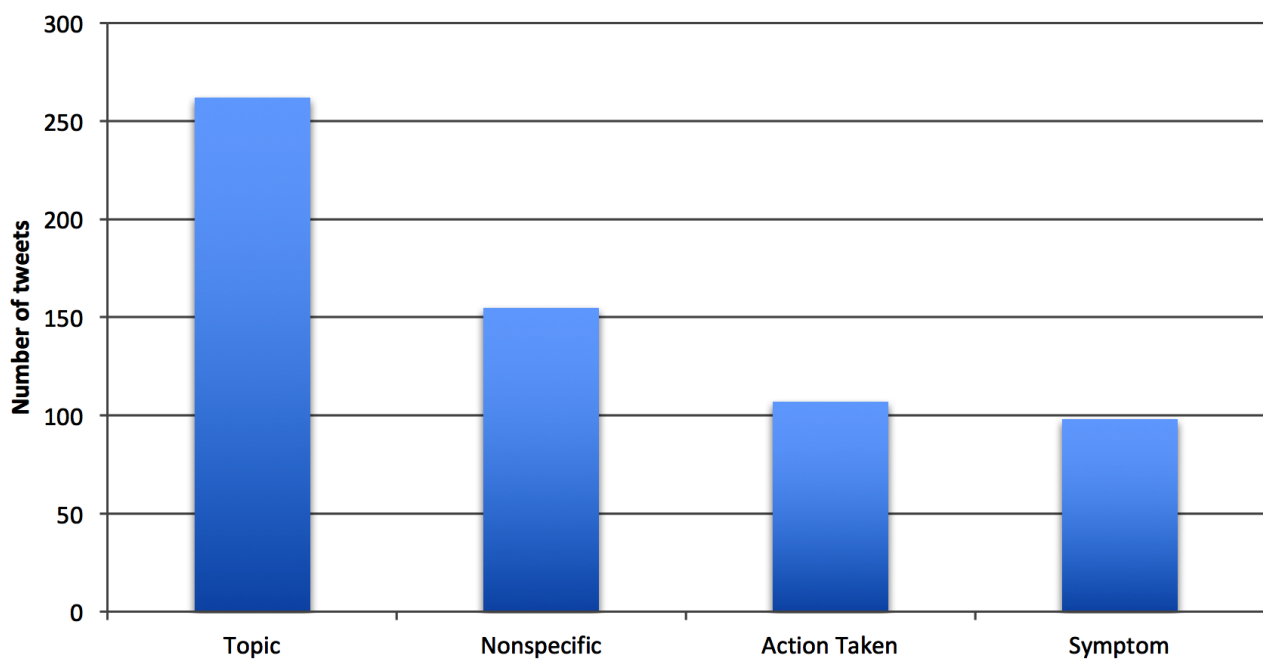


Figure 6. Distribution by topic of first-hand experience stress tweets in dataset 1.

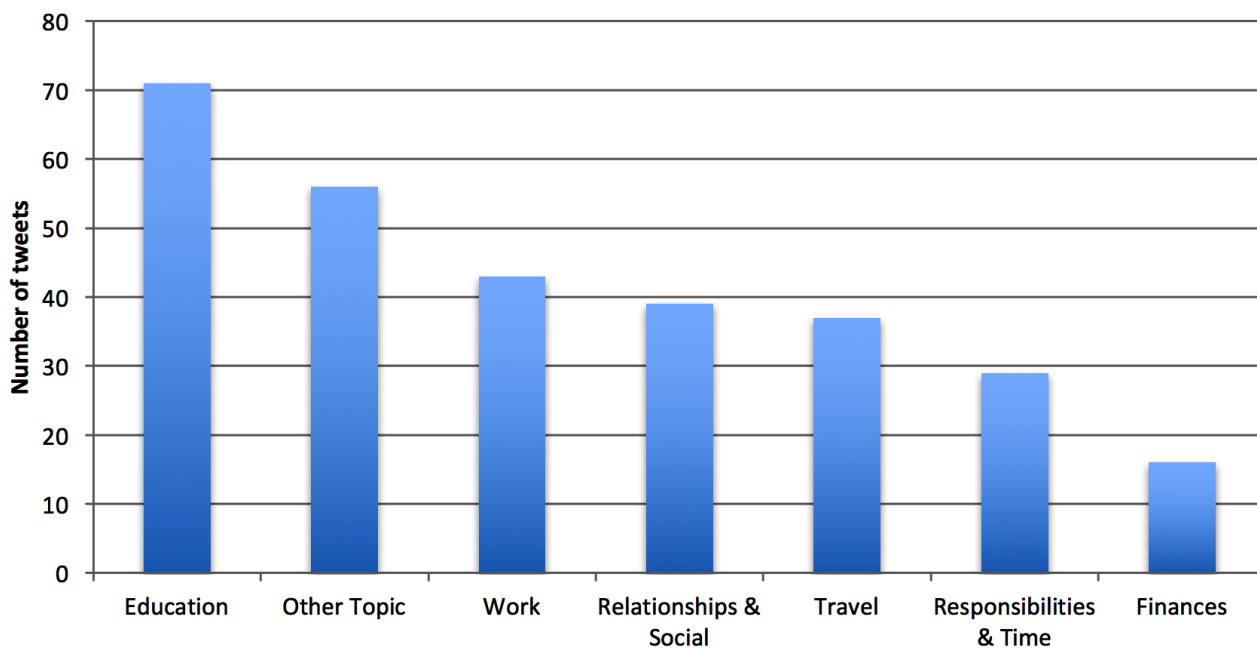


Figure 7. Distribution by topic of first-hand experience relaxation tweets in dataset 1.

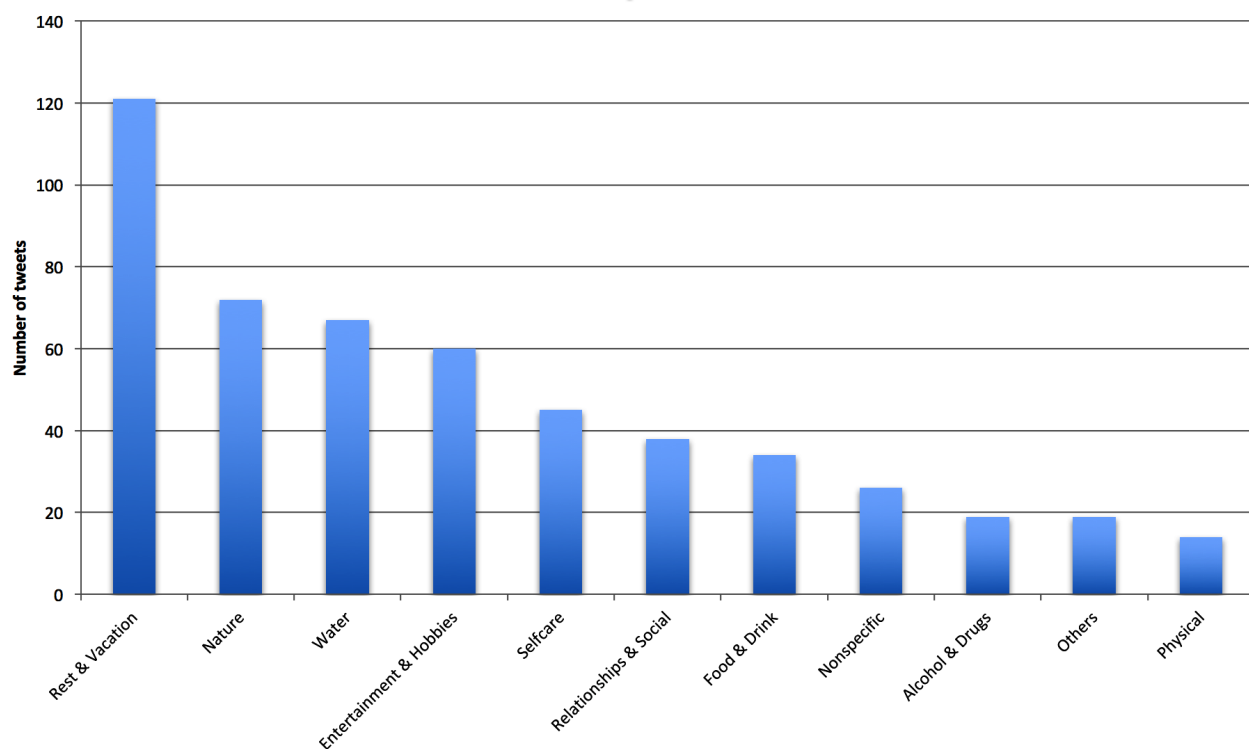


Table 1. Classification evaluation using 10-fold cross-validation on dataset 1.

Classification	Machine learning algorithm							
	Naive Bayes				Support vector machine (linear kernel)			
	Acc ^a (%)	Sen ^b (%)	Spec ^c (%)	PPV ^d (%)	Acc (%)	Sen (%)	Spec (%)	PPV (%)
Stress vs nonstress	78.64	91.97	65.30	72.69	81.66	92.73	70.61	76.07
Relaxation vs nonrelaxation	78.08	96.15	60.00	70.68	83.72	90.26	77.18	79.86
First-hand vs nonfirst-hand experience stress	87.58	95.53	67.89	88.14	85.61	90.64	73.16	89.32
First-hand vs nonfirst-hand experience relaxation	85.64	99.09	11.67	86.07	83.85	95.76	18.33	86.56

^aAcc: accuracy.

^bSen: sensitivity.

^cSpec: specificity.

^dPPV: positive predictive value.

Of the 2 machine learning algorithms used, SVM (with linear kernel) performed better than naive Bayes in classifying stress versus nonstress tweets (81.66% vs 78.64% accuracy, 92.73% vs 91.97% sensitivity, 70.61% vs 65.30% specificity, 76.07% vs 72.69% PPV). SVM was also better than naive Bayes in classifying relaxation versus nonrelaxation tweets in accuracy (83.72% vs 78.08%), specificity (77.18% vs 60.00%), and PPV (79.86% vs 70.68%) but slightly lower in sensitivity (90.26% vs 96.15%).

Table 1 also indicates that naive Bayes had better accuracy and sensitivity than SVM in identifying first-hand experience stress and relaxation tweets: 87.58% versus 85.61% (accuracy) and

95.53% versus 90.64% (sensitivity) for stress; 85.64% versus 83.85% (accuracy) and 99.09% versus 95.76% (sensitivity) for relaxation tweets. In contrast, SVM performed better in specificity and PPV in classifying first-hand experience stress and relaxation tweets.

Table 2 shows the terms that had the highest information gain for stress and relaxation classification. Interestingly, we found that most terms characteristic of the stress class were related to the term “stress,” such as “stressed” or “stressin.” In contrast, the terms most characteristic of the relaxation class were “vacation,” “water,” or “beach,” which are related to the topics as categorized in our relaxation schema.

Table 2. Top 30 keywords ranked by information gain in stress and relaxation classification in dataset 1.

Stress vs nonstress	First-hand stress vs nonstress	First-hand relaxation vs nonrelaxation	Relaxation vs nonrelaxation
stressed	http	rt	rt
stress	rt	relaxing	relaxing
rt	stressed	relaxin	relaxin
mistress	stressing	sorelaxing	sorelaxing
stressful	stressful	relaxed	relaxed
stressing	mistress	work	time
http	stressingout	night	work
stressingout	sostressed	time	night
cashnewvideo	stressin	day	day
camerondallas	cashnewvideo	shower	cashnewvideo
burdenofstress	school	cashnewvideo	relax
tiger	ly	camerondallas	shower
stressin	stress	finally	camerondallas
sostressed	camerondallas	bath	relaxa
day	day	relax	video
nashgrier	love	listening	finally
distressed	sostressful	beach	bath
school	college	relaxa	home
anxiety	packing	video	vacation
life	life	home	listening
busy	twitter	vacation	beach
learn	tiger	pool	nashgrier
woods	hours	sitting	relaxar
bitch	big	enjoying	pool
hours	nashgrier	watching	enjoying
packing	distressed	rain	rain
twitter	hate	give	long
haha	long	nashgrier	sitting
college	weeks	long	watching
love	figure	bed	nice

Automatic Classification of Stress and Relaxation Tweets at the City Level (Dataset 2)

Using an SVM algorithm trained on our annotated data (dataset 1), we automatically classified the much larger dataset 2 (cities dataset). We used a 3-step classification process. First, we filtered by the keywords “stress” and “relax.” Second, we

applied the stress or relaxation classifier to these filtered data. Third, we used the first-hand classifier to identify first-hand stress and relaxation tweets. In both steps, we used SVM (linear kernel) trained on dataset 1 as the classifier. We used SVM because it had advantages in stress and relaxation classification in comparison with naive Bayes in dataset 1. [Table 3](#) shows the number of tweets after each step.

Table 3. Number of tweets remaining after automatic classification.

Cities	Stress rank 2011 (2014) ^a	No. of tweets	No. of tweets containing "relax"	No. of tweets containing "stress"	No. of relaxation tweets	No. of stress tweets	No. of relaxation tweets (first-hand)	No. of stress tweets (first-hand)
Los Angeles	1 (3)	6,627,969	5061	7925	3216	5914	2788	2386
New York	2 (1)	8,229,442	6992	11,789	4412	8245	3766	3278
San Diego	5 (38)	2,908,774	2178	3769	1449	2830	1275	1193
San Francisco	7 (39)	4,372,966	2554	4558	1682	3384	1471	1389

^aStress ranking is based on 2011 Forbes [40] and 2014 CNN studies [41]. Statistical tests between cities showed there are differences between cities ($P < .001$), except San Diego and New York (stress: $P = .18$, relaxation: $P = .02$). P values of relaxation and stress tweets between San Diego and Los Angeles are .41 and $< .001$, respectively. Ranks based on stress tweets are New York=San Diego, Los Angeles, and San Francisco.

To evaluate the performance of stress and relaxation classification in dataset 2, we randomly sampled 2 sets of 100 tweets, with each set consisting of 100 tweets containing either the keyword "stress" (set 1) or "relax" (set 2) from a city in dataset 2. We chose New York for evaluation, since New York had the greatest number of tweets. Then 100 tweets from set 1 were manually annotated (conducted by author SD) as stress or nonstress and first-hand experience stress or nonfirst-hand experience stress class. Similarly, 100 tweets from set 2 were also manually annotated as relaxation or nonrelaxation and

first-hand relaxation experience or nonfirst-hand experience relaxation class.

Table 4 shows the results of classification of set 1 and set 2 using the SVM algorithm. It indicated fair accuracy (66.0%-92.0%) and high PPV (84.6%-100.0%); however, it had lower sensitivity in first-hand stress classification (44.0%) and specificity in relaxation classification (57.1%). The results of the SVM algorithm in dataset 2 were different from those in dataset 1, perhaps due to different data distribution. Figure 8 shows the descriptions of manual annotation of 100 random tweets of set 1 and set 2.

Table 4. Classification evaluation using a random sample of 200 tweets (100 containing the keyword "stress" and 100 containing the keyword "relax") from New York in dataset 2.

Classification	SVM (linear kernel)			
	Acc ^a (%)	Sen ^b (%)	Spec ^c (%)	PPV ^d (%)
Stress vs nonstress	75.0	76.7	70.4	87.5
Relaxation vs nonrelaxation	66.0	67.4	57.1	90.6
First-hand vs nonfirst-hand experience stress	68.0	44.0	92.0	84.6
First-hand vs nonfirst-hand experience relaxation	92.0	87.5	100.0	100.0

^aAcc: accuracy.

^bSen: sensitivity.

^cSpec: specificity.

^dPPV: positive predictive value.

Figure 9 shows the proportion of stress and relaxation tweets out of all tweets by city in dataset 2. The number of stress tweets was twice that of the number of relaxation tweets, indicating that Twitter users were more likely to tweet about stress than relaxation.

To evaluate theme classification by keyword matching, we randomly sampled 50 classified tweets for each theme from New York. Manual review showed that keyword classification achieved a PPV from 60% to 90% for relaxation tweets and 40% to 80% for stress tweets. Themes that had high PPV in relaxation tweets were alcohol & drugs (94%), entertainment & hobbies (94%), and water (92%). Themes having lower PPV were nature (60%) and food & drink (78%). For stress tweets, themes having high PPV are finances (84%), education (82%), and behavioral (82%), while travel (50%) and temporal (62%) had lower PPV. Multimedia Appendix 3 shows the numbers of

classified first-hand stress and relaxation tweets by theme for each city.

First-hand classification results from dataset 2 showed that cities manifested a uniform pattern of stress and relaxation tweets. We found that the singular first-person pronoun "I" was consistently used the most across all cities when expressing stress, found in approximately 4% of all stress tweets, while in relaxation tweets "I" was used less often (ranked 7), at around 2.4%. Multimedia Appendix 4 shows details of the 30 highest-frequency keywords in first-hand experience stress and relaxation tweets for Los Angeles, New York, San Diego, and San Francisco.

We also found that linguistic expressions of negation such as "not," "but," and "don't" or quantifying words such as "much" were among the 30 unigrams most characteristic of stress-related tweets. In addition, users often used emotionally laden

swearwords when expressing stress. It is important to note, however, that the affective polarity of certain swearwords can be highly context dependent (“it’s shit” vs “it’s the shit”) [59]. Relaxation tweets, on the other hand, tended to contain words indicating relaxation and time, such as “relax,” “home,” “time,”

“day,” and “now.” We found that “home” was among the highest-frequency terms in relaxation tweets, as was “weekend.” [Multimedia Appendix 5](#) depicts tag clouds of stress and relaxation tweets for each city.

Figure 8. Description of manual annotation of 100 random tweets containing the keywords “stress” and “relax” from dataset 2.

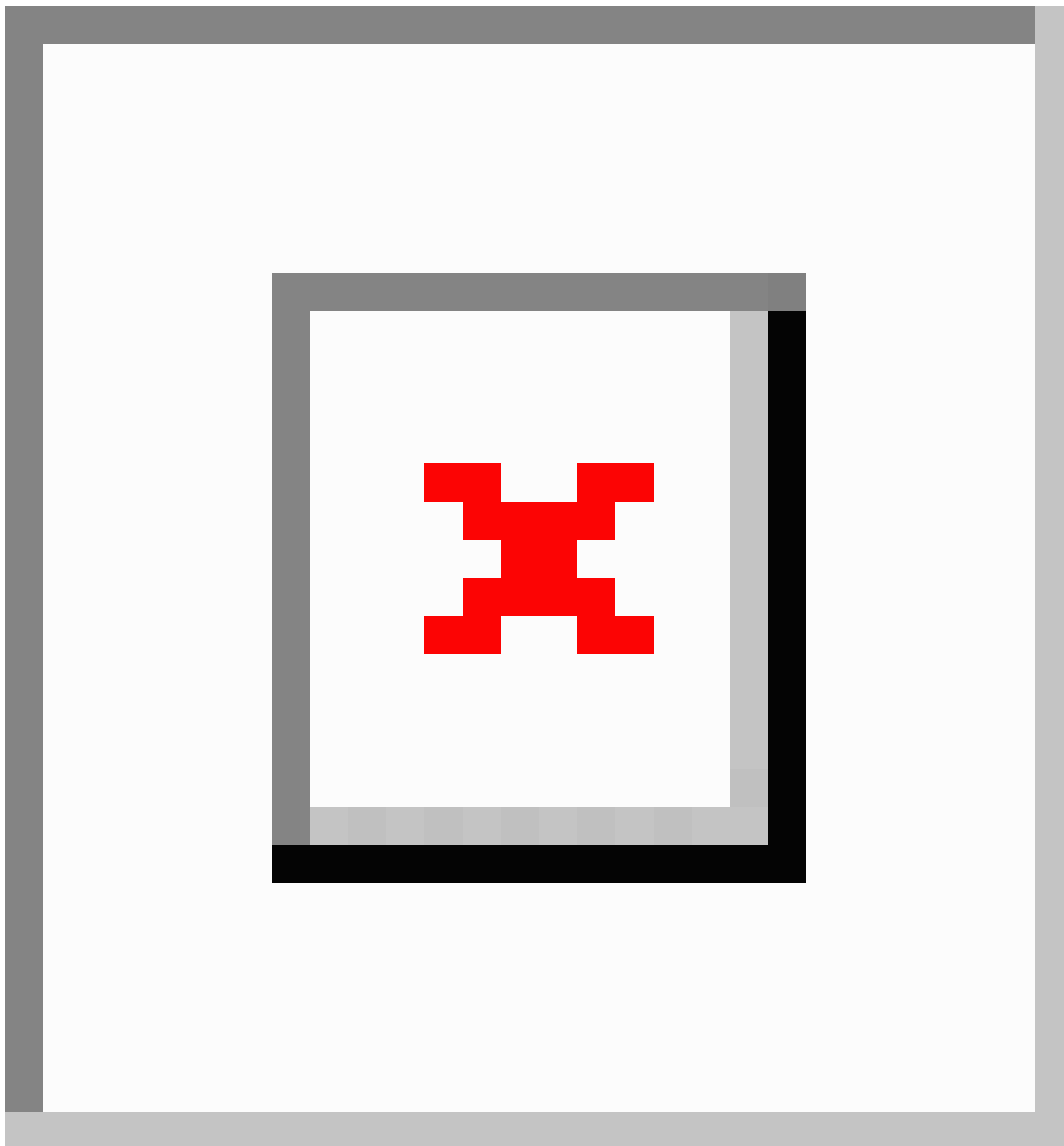
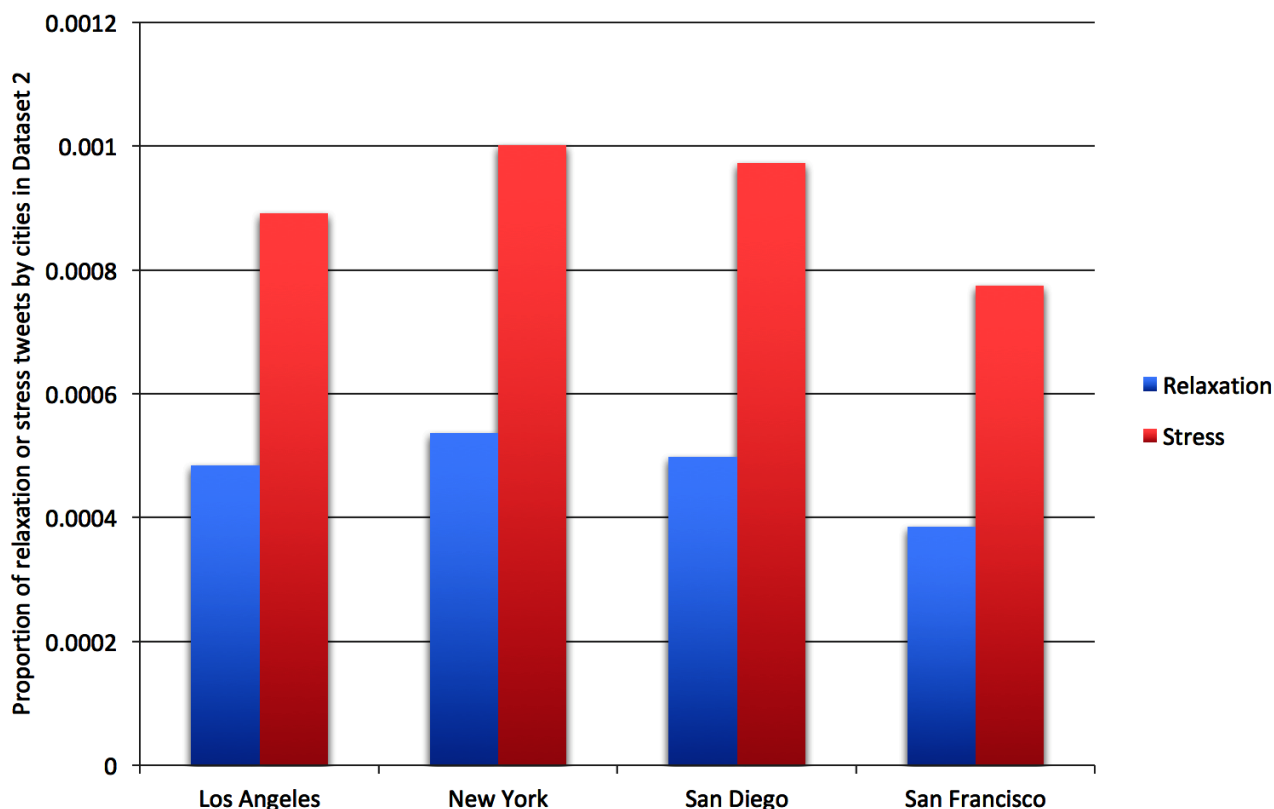


Figure 9. Proportion of relaxation and stress tweets by city in dataset 2.

Theme Distributions of Tweets at the City Level (Dataset 2)

Figure 10 shows the theme distributions of stress tweets among the 4 cities. Education was the highest-frequency topic (12%-14%), followed by work (4%-5%) and travel (4%) (data presented in [Multimedia Appendix 3](#)). Interestingly, we found that tweets describing action taken and psychological and emotional symptoms also had relatively high frequencies (8%-10%). This indicates that, besides topic, people often posted about their emotional state and reaction to stress.

The topic distributions of relaxation tweets were also consistent across cities. Figure 11 shows that rest & vacation was the highest-frequency topic (27%-31%), followed by entertainment & hobbies (13%-14%), food & drink (9%-10%), and nature

(9%-10%). [Multimedia Appendix 3](#) shows detailed numbers of stress and relaxation tweets for each city.

Although we did not find statistically significant differences in theme distributions among cities for stress tweets, there were significant differences between New York and the other cities in the topics of nature and water in relaxation tweets. This may indicate the different activities taken for relaxation between the east coast (New York) and the west coast (Los Angeles, San Diego, and San Francisco). We found that high-frequency terms for relaxation tweets in New York included “watching,” while in San Diego “beach” was more common. This intuitively suggests that San Diegans more often relaxed by going to the beach, while New Yorkers relaxed by enjoying indoor (or spectator) entertainment (“watching,” “listening”).

Figure 10. Stress theme distribution by each of the 4 cities in dataset 2. There are no significant differences between cities ($P>.05$). Neg: negative; Pos: positive; S: symptoms; T: topics.

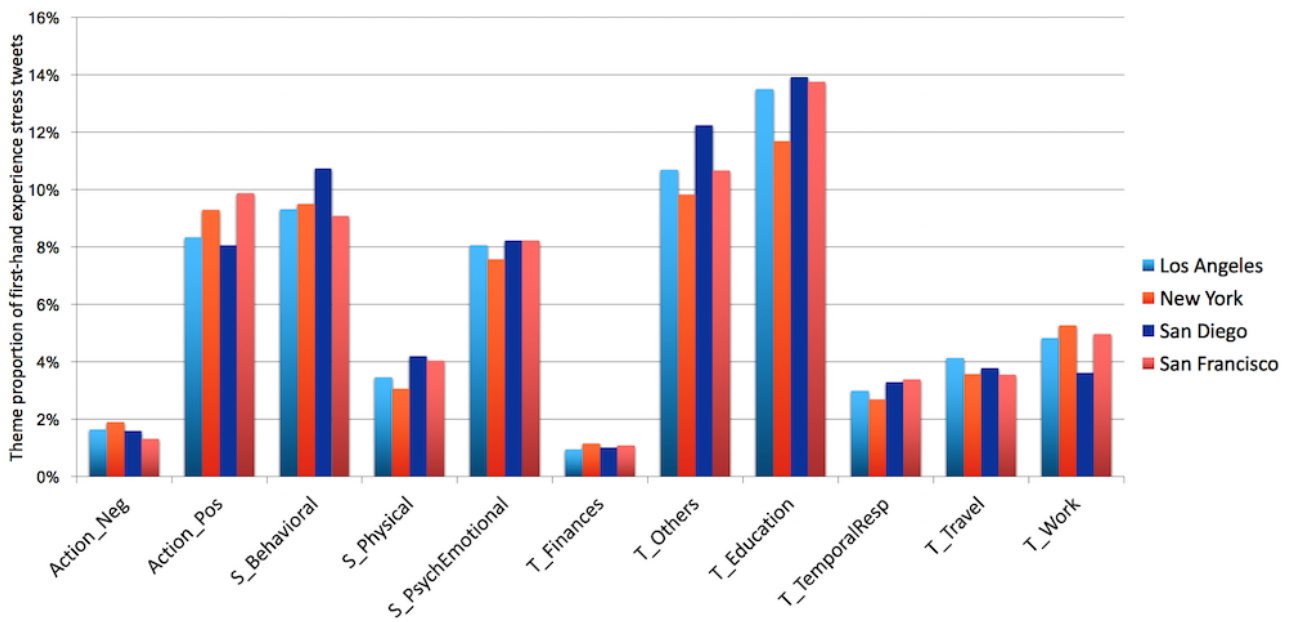
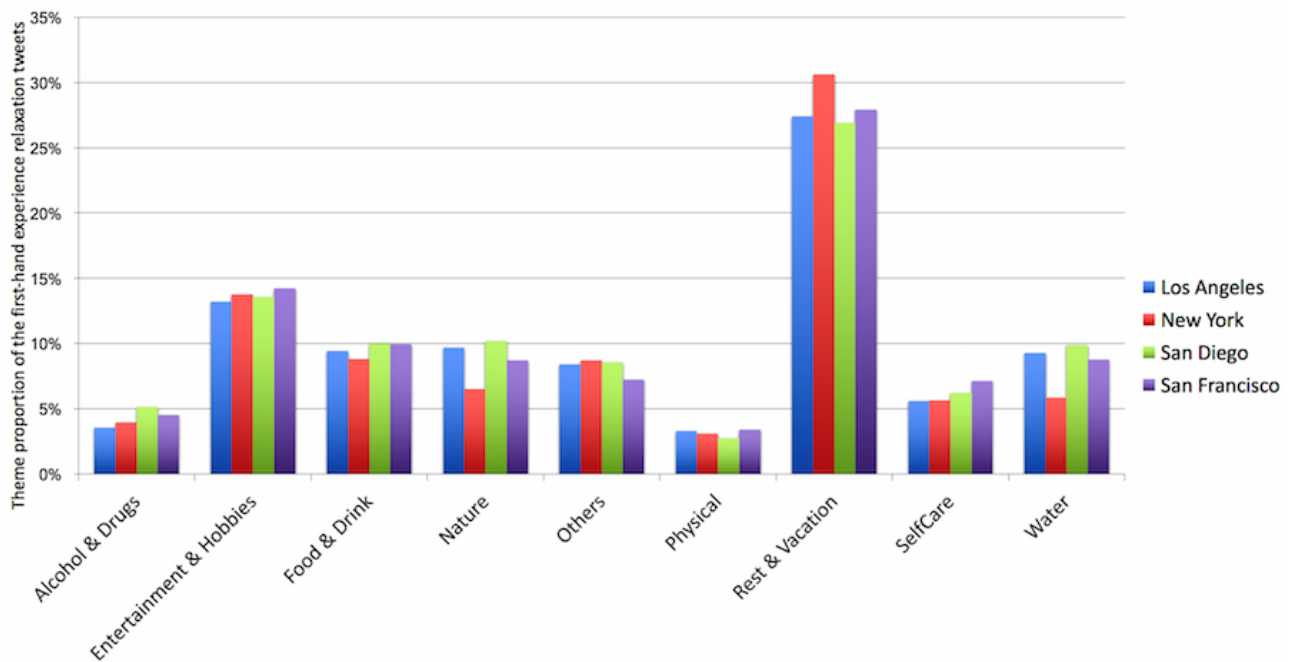


Figure 11. Relaxation theme distribution by each of the 4 cities in dataset 2. There are significant differences between New York and the other cities in the topics of nature and water.



Correlations Between Tweets Data Analysis and Public Surveys

Compared with 2 public surveys on the most stressful cities in the United States by Forbes [40] in 2011 and CNN [41] in 2014, the proportion of stress tweets found here were different. Both surveys ranked New York and Los Angeles among the most stressful cities in the country, while San Diego and San

Francisco were categorized as less stressful. Our city ranking based on the proportion of first-hand experience stress tweets was New York followed by San Diego, Los Angeles, and San Francisco (Table 5 and Figure 9). While we found no significant difference between New York and San Diego, we did find significant differences ($P<.001$) in pairwise comparisons between San Diego, Los Angeles, and San Francisco (Table 5).

Table 5. *P* values of pairwise comparisons of the proportion of stress and relaxation tweets between the 4 studied cities.

Cities	Los Angeles	New York	San Francisco
San Diego			
Stress	<.001	.18	<.001
Relaxation	.41	.02	<.001
San Francisco			
Stress	<.001	<.001	N/A ^a
Relaxation	<.001	<.001	N/A
New York			
Stress	<.001	N/A	<.001
Relaxation	<.001	N/A	<.001

^aN/A: not applicable.

Differences between results found in public stress surveys and our automatic classification of Twitter messages could be due to differences in methodology and population when collecting data. Public surveys collect data using telephones and paper-based reports, while Twitter messages are user generated, are naturalistic, and reflect personal thoughts.

Stress Relief by Relaxation in Tweets

The distribution of stress topics across cities shows an interesting finding: peoples' reactions to stress were more positive than negative. [Figure 10](#) shows that, for all cities, 8%-10% of tweets reported positive action taken in response to stress, while only 1%-2% reported negative action (see [Multimedia Appendix 3](#) for details). This suggests that people may react to stress positively, or that people are more likely to publicly report positive rather than negative actions. Examples of positive reaction in stress tweets are rest ("Rest is best when you are stressed") and exercising ("I'm so stressed, thank god I'm heading to yoga now").

Relaxation can be considered a stress management activity. [Figure 9](#) shows that the numbers of relaxation tweets were consistently proportional across all cities to those of stress tweets, indicating that Twitter users were consistently more inclined to post about stressful life events or experiences than about relaxing experiences. Examples of stress relief from relaxation tweets are personal contact ("I don't need anything but a hug..."), exercising ("Went for a run, feel awesome, now time to relax"), shopping ("Last day in #SanDiego Just relaxing, shopping and say bye to friends"), and entertainment ("Relaxing watching a movie:-) :-)"). [Figure 7](#) and [Figure 10](#) also indicate that rest & vacation was the highest-frequency topic within relaxation tweets, followed by entertainment & hobbies, nature, and water. These topics can be considered common activities for stress relief.

Discussion

Principal Results

Our research addressed several aspects of the use of Twitter as a medium of expression of stress and relaxation by users. First, we created schema for categorizing stress- and relaxation-related tweets based on previously published psychological guidelines.

By categorizing first-hand experience tweets into the primary themes of content topics, symptoms, and actions taken, we gained further insight into the common patterns of expressions of stress.

Second, we analyzed in detail the contents of tweets based on our annotation scheme and found both similarities and differences in the prevalence and characteristics of stress and relaxation tweets across cities on the east and west coasts of the United States. The most frequent topic of stress tweets in our datasets was education, which likely reflects the younger demographic of Twitter users [57,58], but work and travel were also common topics. It is notable that, despite poverty rates, unemployment rates, and cost of living being significant factors in the methodology of CNN's and Forbes's stress ranking systems of the most stressful cities, finances were not a major content topic of the stress tweets in any city in our studies. Although this result could be partially attributable to the need for either computer or mobile phone access in order to use Twitter and may cause underrepresentation in lower-income groups, it may also indicate that certain topics, such as personal finances, still remain relatively taboo in social media settings. Regarding positive and negative actions regarding stress, positive actions far outnumbered more destructive behavior. The use of Twitter in itself to discuss feelings of stress and stress management can be seen as a constructive manner of dealing with stress by expressing these feelings and using the support of "followers" and friends. Social media platforms are increasingly being used as support networks in the management of chronic health conditions as varied as cancer, depression, and obesity. A recent systematic review by Patel et al found that the impact of social media use on those experiencing chronic disease was positive in 48% of studies reviewed, neutral in 45%, and harmful in only 7% [60].

Third, our study indicated that words most associated with relaxation strategies (see [Table 2](#)) fell into 3 main groups: (1) bathing and personal care (eg, "bath," "shower"), (2) vacationing ("vacation," "pool," "beach"), and (3) watching sports or television ("videos," "sitting," "watching"), indicating that relaxation strategies involved purposefully taking time away from work-based activities and daily responsibilities. A further key theme that emerged from a qualitative analysis of the data

was the idea of nature—in this case, particularly water (eg, “pool,” “beach,” “rain”)—as being of key importance for relaxation. This result is consistent with recent research demonstrating the link between stress reduction and exposure to the natural environment (eg, [61]).

Finally, we showed that machine learning algorithms could be employed to achieve good accuracy for the automatic classification of stress and relaxation tweets.

Limitations

This study has several limitations. First, we obtained dataset 2 from the Twitter API’s 1% sample. Second, the annotation scheme we developed, although well suited for our purpose, could benefit from further refinement. For example, we found that many tweets were categorized as topic “other.” Third, it is

likely that classification results could be improved given the availability of additional training data, in particular for first-hand experience classification of stress and relaxation tweets. Furthermore, using additional feature sets, such as ngrams, emotions, and negations, could help improve accuracy. Fourth, Twitter reports of stress and relaxation may be influenced by self-presentation issues (eg, stress related to excessive workload can be used as a status indicator in some contexts). Finally, as with all social media-based research, the population studied is unlikely to be a representative sample of the general population.

Conclusions

This research showed that Twitter can be a useful tool for the analysis of stress and relaxation levels in the community, and has the potential to provide a valuable supplement to social and psychological studies of stress and stress management.

Acknowledgments

SD and AR were partially supported by NIH grant U54HL108460. NP and MC were partially supported by NIH grant R00LM011393. JDC were partially supported by the NLM Medical Informatics Training Grant 5T15LM011271-04. We would like to thank Mr Gregory Stoddard, MPH, MBA at the University of Utah’s Division of Epidemiology for his valuable comments on an earlier version of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Examples of each category of first-hand experience stress tweets with its themes.

[[PDF File \(Adobe PDF File\), 34KB - publichealth_v3i2e35_app1.pdf](#)]

Multimedia Appendix 2

Examples of each category of first-hand experience relaxation tweets with its themes.

[[PDF File \(Adobe PDF File\), 32KB - publichealth_v3i2e35_app2.pdf](#)]

Multimedia Appendix 3

Number of classified first-hand stress tweets by theme and first-hand relaxation tweets in each city.

[[PDF File \(Adobe PDF File\), 29KB - publichealth_v3i2e35_app3.pdf](#)]

Multimedia Appendix 4

Top 30 highest-frequency keywords in first-hand experience stress and relaxation tweets for Los Angeles, New York, San Diego, and San Francisco.

[[PDF File \(Adobe PDF File\), 30KB - publichealth_v3i2e35_app4.pdf](#)]

Multimedia Appendix 5

Tag clouds of stress and relaxation tweets in New York, Los Angeles, San Diego, and San Francisco.

[[PDF File \(Adobe PDF File\), 4MB - publichealth_v3i2e35_app5.pdf](#)]

References

1. Hammen C. Stress and depression. *Annu Rev Clin Psychol* 2005;1:293-319. [doi: [10.1146/annurev.clinpsy.1.102803.143938](https://doi.org/10.1146/annurev.clinpsy.1.102803.143938)] [Medline: [17716090](https://pubmed.ncbi.nlm.nih.gov/17716090/)]

2. Stansfeld S, Marmot M. *Stress and the Heart: Psychosocial Pathways to Coronary Heart Disease*. London, UK: BMJ Books; 2002.
3. McEwen BS, Stellar E. Stress and the individual. Mechanisms leading to disease. *Arch Intern Med* 1993 Sep 27;153(18):2093-2101. [Medline: [8379800](#)]
4. Nielsen NR, Kristensen TS, Schnohr P, Grønbaek M. Perceived stress and cause-specific mortality among men and women: results from a prospective cohort study. *Am J Epidemiol* 2008 Sep 01;168(5):481-491. [doi: [10.1093/aje/kwn157](#)] [Medline: [18611955](#)]
5. Andersen BL, Kiecolt-Glaser JK, Glaser R. A biobehavioral model of cancer stress and disease course. *Am Psychol* 1994 May;49(5):389-404 [FREE Full text] [Medline: [8024167](#)]
6. Segal J, Smith M, Segal R, Robinson L. *Stress symptoms, signs, and causes.*: HelpGuide.org International; 2016 Apr. URL: <http://www.helpguide.org/articles/stress/stress-symptoms-causes-and-effects.htm> [accessed 2016-05-03] [WebCite Cache ID [6hDuhAkA8](#)]
7. NPR, Robert Wood Johnson Foundation, Harvard School of Health. *The burden of stress in America*. 2014. URL: http://www.rwjf.org/content/dam/farm/reports/surveys_and_polls/2014/rwjf414295 [accessed 2016-05-03] [WebCite Cache ID [6hDur5MgJ](#)]
8. The National Institute for Occupational Safety Health. *STRESS...at work*. 2014 Jun 06. URL: <http://www.cdc.gov/niosh/docs/99-101/> [accessed 2016-05-03] [WebCite Cache ID [6hDuxw1X6](#)]
9. American Psychological Association. *Stress in America: paying with our health*. 2015 Feb 4. URL: <http://www.apa.org/news/press/releases/stress/2014/stress-report.pdf> [accessed 2016-05-03] [WebCite Cache ID [6hDv3vfHK](#)]
10. Moriarty DG, Zack MM, Holt JB, Chapman DP, Safran MA. Geographic patterns of frequent mental distress: U.S. adults, 1993-2001 and 2003-2006. *Am J Prev Med* 2009 Jun;36(6):497-505. [doi: [10.1016/j.amepre.2009.01.038](#)] [Medline: [19460657](#)]
11. Cohen S, Janicki-Deverts D. Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. *J Appl Soc Psychol* 2012 Jun;42(6):1320-1334. [doi: [10.1111/j.1559-1816.2012.00900.x](#)]
12. Rentfrow PJ, Gosling SD, Jokela M, Stillwell DJ, Kosinski M, Potter J. Divided we stand: three psychological regions of the United States and their political, economic, social, and health correlates. *J Pers Soc Psychol* 2013 Dec;105(6):996-1012. [doi: [10.1037/a0034434](#)] [Medline: [24128185](#)]
13. Antoni M, Schneiderman N, Ironson G. *Stress Management for HIV: Clinical Validation and Intervention Manual*. Mahwah, NJ: Lawrence Erlbaum Associates; 2007.
14. Antoni MH, Baggett L, Ironson G, LaPerriere A, August S, Klimas N, et al. Cognitive-behavioral stress management intervention buffers distress responses and immunologic changes following notification of HIV-1 seropositivity. *J Consult Clin Psychol* 1991 Dec;59(6):906-915. [Medline: [1774375](#)]
15. Brown JL, Vanable PA. Cognitive-behavioral stress management interventions for persons living with HIV: a review and critique of the literature. *Ann Behav Med* 2008 Feb;35(1):26-40 [FREE Full text] [doi: [10.1007/s12160-007-9010-y](#)] [Medline: [18347902](#)]
16. Cruess DG, Antoni MH, McGregor BA, Kilbourn KM, Boyers AE, Alferi SM, et al. Cognitive-behavioral stress management reduces serum cortisol by enhancing benefit finding among women being treated for early stage breast cancer. *Psychosom Med* 2000;62(3):304-308. [Medline: [10845343](#)]
17. Carlson LE, Speca M, Patel KD, Goodey E. Mindfulness-based stress reduction in relation to quality of life, mood, symptoms of stress and levels of cortisol, dehydroepiandrosterone sulfate (DHEAS) and melatonin in breast and prostate cancer outpatients. *Psychoneuroendocrinology* 2004 May;29(4):448-474. [Medline: [14749092](#)]
18. Chiesa A, Serretti A. Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. *J Altern Complement Med* 2009 May;15(5):593-600. [doi: [10.1089/acm.2008.0495](#)] [Medline: [19432513](#)]
19. Contrada R, Baum A. *The Handbook of Stress Science: Biology, Psychology, and Health*. New York, NY: Springer Publishing Inc; 2010.
20. Centers for Disease Control and Prevention. *Behavioral Risk Factor Surveillance System*. Atlanta, GA: CDC; 2016 Feb 01. URL: <http://www.cdc.gov/brfss/> [accessed 2016-05-03] [WebCite Cache ID [6hDv7Dtg9](#)]
21. Twitter. *Twitter usage: company facts*. San Francisco, CA: Twitter, Inc; 2016. URL: <https://about.twitter.com/company> [accessed 2016-05-03] [WebCite Cache ID [6hDvCMq5e](#)]
22. O'Connor B, Balasubramanyan R, Routledge B, Smith N. *From tweets to polls: linking text sentiment to public opinion time series*. Palo Alto, CA: AAI Press; 2010 Presented at: The 4th International AAI Conference on Weblogs and Social Media; May 23-26, 2010; Washington, DC.
23. Bollen J, Mao H, Zeng X. Twitter mood predicts the stock market. *J Comput Sci* 2011 Mar;2(1):1-8.
24. Doan S, Vo B, Collier N. *An analysis of Twitter messages in the 2011 Tohoku Earthquake*. Berlin, Germany: Springer; 2011 Presented at: 4th ICST International Conference on eHealth; Nov 21-23, 2011; Malaga, Spain p. 58-66.
25. Doan S, Ohno-Machado L, Collier N. *Enhancing Twitter data analysis with simple semantic filtering: example in tracking influenza-like illnesses*. 2012 Presented at: IEEE Second International Conference on Healthcare Informatics, Imaging, and Systems Biology; Sept 27-28, 2012; La Jolla, CA, USA p. 62-71.

26. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PLoS One* 2010 Nov 29;5(11):e14118 [FREE Full text] [doi: [10.1371/journal.pone.0014118](https://doi.org/10.1371/journal.pone.0014118)] [Medline: [21124761](https://pubmed.ncbi.nlm.nih.gov/21124761/)]
27. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS One* 2011;6(5):e19467 [FREE Full text] [doi: [10.1371/journal.pone.0019467](https://doi.org/10.1371/journal.pone.0019467)] [Medline: [21573238](https://pubmed.ncbi.nlm.nih.gov/21573238/)]
28. Thackeray R, Burton SH, Giraud-Carrier C, Rollins S, Draper CR. Using Twitter for breast cancer prevention: an analysis of breast cancer awareness month. *BMC Cancer* 2013;13(1):508 [FREE Full text] [doi: [10.1186/1471-2407-13-508](https://doi.org/10.1186/1471-2407-13-508)] [Medline: [24168075](https://pubmed.ncbi.nlm.nih.gov/24168075/)]
29. Harris JK, Moreland-Russell S, Tabak RG, Ruhr LR, Maier RC. Communication about childhood obesity on Twitter. *Am J Public Health* 2014 Jul;104(7):e62-e69 [FREE Full text] [doi: [10.2105/AJPH.2013.301860](https://doi.org/10.2105/AJPH.2013.301860)] [Medline: [24832138](https://pubmed.ncbi.nlm.nih.gov/24832138/)]
30. Lee JL, DeCamp M, Dredze M, Chisolm MS, Berger ZD. What are health-related users tweeting? A qualitative content analysis of health-related users and their messages on twitter. *J Med Internet Res* 2014;16(10):e237 [FREE Full text] [doi: [10.2196/jmir.3765](https://doi.org/10.2196/jmir.3765)] [Medline: [25591063](https://pubmed.ncbi.nlm.nih.gov/25591063/)]
31. Myslín M, Zhu S, Chapman W, Conway M. Using Twitter to examine smoking behavior and perceptions of emerging tobacco products. *J Med Internet Res* 2013;15(8):e174 [FREE Full text] [doi: [10.2196/jmir.2534](https://doi.org/10.2196/jmir.2534)] [Medline: [23989137](https://pubmed.ncbi.nlm.nih.gov/23989137/)]
32. Heavilin N, Gerbert B, Page JE, Gibbs JL. Public health surveillance of dental pain via Twitter. *J Dent Res* 2011 Sep;90(9):1047-1051 [FREE Full text] [doi: [10.1177/0022034511415273](https://doi.org/10.1177/0022034511415273)] [Medline: [21768306](https://pubmed.ncbi.nlm.nih.gov/21768306/)]
33. Ahlwardt K, Heavilin N, Gibbs J, Page J, Gerbert B, Tsoh J. Tweeting about pain: comparing self-reported toothache experiences with those of backaches, earaches and headaches. *J Am Dent Assoc* 2014;145(7):737-743. [doi: [10.14219/jada.2014.30](https://doi.org/10.14219/jada.2014.30)] [Medline: [24982280](https://pubmed.ncbi.nlm.nih.gov/24982280/)]
34. Tighe PJ, Goldsmith RC, Gravenstein M, Bernard HR, Fillingim RB. The painful tweet: text, sentiment, and community structure analyses of tweets pertaining to pain. *J Med Internet Res* 2015;17(4):e84 [FREE Full text] [doi: [10.2196/jmir.3769](https://doi.org/10.2196/jmir.3769)] [Medline: [25843553](https://pubmed.ncbi.nlm.nih.gov/25843553/)]
35. Gabarron E, Serrano JA, Wynn R, Lau AY. Tweet content related to sexually transmitted diseases: no joking matter. *J Med Internet Res* 2014 Oct 06;16(10):e228 [FREE Full text] [doi: [10.2196/jmir.3259](https://doi.org/10.2196/jmir.3259)] [Medline: [25289463](https://pubmed.ncbi.nlm.nih.gov/25289463/)]
36. Turner-McGrievy GM, Beets MW. Tweet for health: using an online social network to examine temporal trends in weight loss-related posts. *Transl Behav Med* 2015 Jun;5(2):160-166 [FREE Full text] [doi: [10.1007/s13142-015-0308-1](https://doi.org/10.1007/s13142-015-0308-1)] [Medline: [26029278](https://pubmed.ncbi.nlm.nih.gov/26029278/)]
37. Schwartz H, Eichstaedt J. Characterizing geographic variation in well-being using tweets. 2013 Presented at: The 7th International AAAI Conference on Weblogs and Social Media; Jul 8-11, 2013; Cambridge, MA, USA p. 583-591.
38. De Choudhury M, Counts S, Horvitz E. Social media as a measurement tool of depression in populations. New York, NY: ACM Press; 2013 Presented at: The 5th Annual ACM Web Science Conference; May 2-4, 2013; Paris, France p. 47-56.
39. Eichstaedt JC, Schwartz HA, Kern ML, Park G, Labarthe DR, Merchant RM, et al. Psychological language on Twitter predicts county-level heart disease mortality. *Psychol Sci* 2015 Feb;26(2):159-169 [FREE Full text] [doi: [10.1177/0956797614557867](https://doi.org/10.1177/0956797614557867)] [Medline: [25605707](https://pubmed.ncbi.nlm.nih.gov/25605707/)]
40. Greenfield B. America's most stressful cities. New York, NY: Forbes Media LLC; 2011 Sep 23. URL: <http://www.forbes.com/sites/bethgreenfield/2011/09/23/americas-most-stressful-cities/> [accessed 2016-05-03] [WebCite Cache ID 6hDvdoSqu]
41. CNN Money. Stressed out cities. Atlanta, GA: Cable News Network; 2014. URL: http://money.cnn.com/pf/stressed-cities/2014/full_list/ [accessed 2016-05-03] [WebCite Cache ID 6hDvhNv6y]
42. Twitter. REST APIs. 2016. URL: <https://dev.twitter.com/rest/public> [accessed 2016-05-03] [WebCite Cache ID 6hDv15qNA]
43. Twitter. Streaming APIs. 2016. URL: <https://dev.twitter.com/streaming/public> [accessed 2016-05-03] [WebCite Cache ID 6hDvniaql]
44. Centers for Disease Control and Prevention. Managing stress. Atlanta, GA: National Center for Injury Prevention and Control; 2012 Dec 19. URL: <http://www.cdc.gov/features/handlingstress/> [accessed 2016-05-03] [WebCite Cache ID 6hDvqNo39]
45. Healthline Editorial Team. Stress and anxiety.: Healthline Media; 2016. URL: <http://www.healthline.com/health/stress-and-anxiety> [accessed 2016-05-03] [WebCite Cache ID 6hDvttDbj]
46. Statistic Brain. Stress statistics. Los Angeles, CA: Statistic Brain Research Institute; 2016. URL: <http://www.statisticbrain.com/stress-statistics/> [accessed 2016-05-03] [WebCite Cache ID 6hDvxRING]
47. American Heart Association. Stress management. Dallas, TX: AHA URL: http://www.heart.org/HEARTORG/HealthyLiving/StressManagement/Stress-Management_UCM_001082_SubHomePage.jsp [accessed 2016-05-03] [WebCite Cache ID 6hDw3RBhc]
48. National Institute of Mental Health. Fact sheet on stress. Bethesda, MD: NIMH URL: <http://www.nimh.nih.gov/health/publications/stress/index.shtml> [accessed 2016-05-03] [WebCite Cache ID 6hDwJfLZT]
49. Centers for Disease Control and Prevention. Coping with stress. Atlanta, GA: CDC; 2015. URL: http://www.cdc.gov/violenceprevention/pub/coping_with_stress_tips.html [accessed 2016-05-03] [WebCite Cache ID 6hDwNuDD5]
50. Joachims T. Text categorization with support vector machines: learning with many relevant features. 1998 Presented at: 10th European Conference on Machine Learning; Apr 21-24, 1998; Chemnitz, Germany p. 137-142.

51. McCallum A. Bow: a toolkit for statistical language modeling, text retrieval, classification and clustering. 1998 Sep 12. URL: <http://www.cs.cmu.edu/~mccallum/bow> [accessed 2016-05-03] [WebCite Cache ID 6hDwRJvFM]
52. Joachims T. Making large-scale SVM learning practical. In: Schölkopf B, Burges CJC, Smola AJ, editors. *Advances in Kernel Methods: Support Vector Learning*. Cambridge, MA: MIT Press; 1999:169-184.
53. van Rijsbergen CJ. *Information Retrieval*. Second edition. Newton, MA: Butterworth-Heinemann; 1979.
54. Yang Y. An evaluation of statistical approaches to text categorization. *Inf Retrieval J* 1999;1:69-90.
55. Manning C, Schütze H. *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press; 1999.
56. Agresti A. *An Introduction to Categorical Data Analysis*. Hoboken, NJ: John Wiley & Sons; 2007.
57. Smith A, Brenner J. *Twitter use 2012*. Washington, DC: Pew Internet & American Life Project; 2012 May 31. URL: http://www.pewinternet.org/files/old-media/Files/Reports/2012/PIP_Twitter_Use_2012.pdf [accessed 2017-05-22] [WebCite Cache ID 6qefUNewr]
58. Duggan J, Ellison N, Lampe C, Lenhart A, Madden M. *Social media update 2014*. Washington, DC: Pew Research Center; 2015 Jan 09. URL: <http://www.pewinternet.org/2015/01/09/social-media-update-2014/> [accessed 2016-05-03] [WebCite Cache ID 6hDy9qpB3]
59. McEnery T. *Swearing in English: Bad Language, Purity and Power From 1586 to the Present*. London, UK: Routledge; 2004.
60. Patel R, Chang T, Greysen SR, Chopra V. Social media use in chronic disease: a systematic review and novel taxonomy. *Am J Med* 2015 Dec;128(12):1335-1350. [doi: [10.1016/j.amjmed.2015.06.015](https://doi.org/10.1016/j.amjmed.2015.06.015)] [Medline: [26159633](https://pubmed.ncbi.nlm.nih.gov/26159633/)]
61. Huynh Q, Craig W, Janssen I, Pickett W. Exposure to public natural space as a protective factor for emotional well-being among young people in Canada. *BMC Public Health* 2013 Apr 29;13:407 [FREE Full text] [doi: [10.1186/1471-2458-13-407](https://doi.org/10.1186/1471-2458-13-407)] [Medline: [23627738](https://pubmed.ncbi.nlm.nih.gov/23627738/)]

Abbreviations

- API:** application programming interface
CDC: Centers for Disease Control and Prevention
PPV: positive predictive value
SVM: support vector machines

Edited by G Eysenbach; submitted 05.05.16; peer-reviewed by Z Zhang, J Guidry, S Kiritchenko, S Mohammad; comments to author 21.08.16; revised version received 09.11.16; accepted 05.04.17; published 13.06.17.

Please cite as:

Doan S, Ritchart A, Perry N, Chaparro JD, Conway M

How Do You #relax When You're #stressed? A Content Analysis and Infodemiology Study of Stress-Related Tweets

JMIR Public Health Surveill 2017;3(2):e35

URL: <http://publichealth.jmir.org/2017/2/e35/>

doi: [10.2196/publichealth.5939](https://doi.org/10.2196/publichealth.5939)

PMID: [28611016](https://pubmed.ncbi.nlm.nih.gov/28611016/)

©Son Doan, Amanda Ritchart, Nicholas Perry, Juan D Chaparro, Mike Conway. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 13.06.2017. 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

Health Care Professionals' Evidence-Based Medicine Internet Searches Closely Mimic the Known Seasonal Variation of Lyme Borreliosis: A Register-Based Study

Samuli Pesälä^{1,2}, MD; Mikko J Virtanen², MSc; Jussi Sane², PhD; Jukkapekka Jousimaa³, MD, PhD; Outi Lyytikäinen², MD, PhD; Satu Murtopuro², DMD; Pekka Mustonen³, MD, PhD; Minna Kaila⁴, MD, PhD; Otto Helve^{2,5}, MD, PhD

¹University of Helsinki, Helsinki, Finland

²National Institute for Health and Welfare, Helsinki, Finland

³Duodecim Medical Publications Ltd, Helsinki, Finland

⁴Public Health Medicine, Helsinki University Hospital, University of Helsinki, Helsinki, Finland

⁵Children's Hospital, Helsinki University Hospital, University of Helsinki, Helsinki, Finland

Corresponding Author:

Samuli Pesälä, MD

University of Helsinki

Tukholmankatu 8B

00029 Helsinki, Finland

Helsinki, 00029 Helsinki, Finl

Finland

Phone: 358 400205955

Fax: 358 295248468

Email: samuli.pesala@helsinki.fi

Abstract

Background: Both health care professionals and nonprofessionals seek medical information on the Internet. Using Web-based search engine searches to detect epidemic diseases has, however, been problematic. Physician's databases (PD) is a chargeable evidence-based medicine (EBM) portal on the Internet for health care professionals and is available throughout the entire health care system in Finland. Lyme borreliosis (LB), a well-defined disease model, shows temporal and regional variation in Finland. Little data exist on health care professionals' searches from Internet-based EBM databases in public health surveillance.

Objective: The aim of this study was to assess whether health care professionals' use of Internet EBM databases could describe seasonal increases of the disease and supplement routine public health surveillance.

Methods: Two registers, PD and the register of primary health care diagnoses (Avohilmo), were used to compare health care professionals' Internet searches on LB from EBM databases and national register-based LB diagnoses in order to evaluate annual and regional variations of LB in the whole country and in three selected high-incidence LB regions in Finland during 2011-2015.

Results: Both registers, PD and Avohilmo, show visually similar patterns in annual and regional variation of LB in Finland and in the three high-incidence LB regions during 2011-2015.

Conclusions: Health care professionals' Internet searches from EBM databases coincide with national register diagnoses of LB. PD searches showed a clear seasonal variation. In addition, notable regional differences were present in both registers. However, physicians' Internet medical searches should be considered as a supplementary source of information for disease surveillance.

(*JMIR Public Health Surveill* 2017;3(2):e19) doi:[10.2196/publichealth.6764](https://doi.org/10.2196/publichealth.6764)

KEYWORDS

search engine; evidence-based medicine; information systems; public health surveillance; Lyme borreliosis

Introduction

Traditionally, many syndromic surveillance systems have collected epidemiological data from health care professionals' clinical encounters when predicting several disease epidemics, such as influenza [1]. This worldwide severe disease showing seasonal and geographical variation can be predicted by using Internet search trends [1-3]. Internet-based surveillance systems have allowed good congruence with traditional (data submitted by public health authorities) surveillance approaches for monitoring emerging infectious diseases of public health concern [1,3]. To improve early detection of influenza, Google search queries were used to track influenza-like disease in population [1,2]. Because certain searches from Google correlated highly with medical visits related to influenza-like symptoms, influenza activity could be estimated geographically [2]. However, when reassessed, the reliability of the surveillance tool was shown to be problematic and to contain substantial flaws, especially in regard to assessing the correct timing and location [4]. Therefore, it has been commented that the use of near-real time electronic health data and computational methods should be incorporated [4]. The models such as Internet-based influenza monitoring have not, however, included the characterization of the populations performing the searches, also comprising nonprofessionals. Internet search engine queries and the data from social media can be combined to detect infectious diseases as well [5]. Generally, for framing the methods on health-related Internet information and epidemiological data, two terms are used: infodemiology and infoveillance. Infodemiology can be determined as a discipline inside public health informatics that studies information in an electronic medium or in a population aiming to inform public health and public policy [6]. When infodemiology data are used for surveillance purposes, the term is called infoveillance [6].

Health care professionals' use of CD-ROM-based medical searches as a tool for early detection of epidemics has been studied earlier in Finland during 1995 using the National Infectious Diseases Register (NIDR) as a reference register [7]. Microbiological laboratories notify diagnostic findings electronically to NIDR. Medical professionals use electronic sources to provide appropriate answers to clinical questions, such as Internet sites and search engines [8]. For doctors and nurses, the main reasons to seek Internet-based evidence are patient care and continuing professional development [9]. Even for those clinical hospital personnel who are concerned about the quality of the data, the most popular electronic source for information seeking was Google [10], as Google may be a portal to the MEDLINE or Pubmed database [10], one of the most popular and commonly used evidence-based medicine (EBM) electronic databases [11]. Thus, the Google search data can be expected to include searches by both health care professionals and nonprofessionals, making the user base extremely heterogeneous and sensitive to, for example, media trends.

We aimed to assess whether health care professionals' use of Internet-based EBM databases was comparable to the use of the register of public primary health care diagnoses (Avohilmo) as a part of routine public health surveillance. Avohilmo serves as an electronic database for actual primary health care

notifications collected from public sector health care units. Avohilmo data are used for example in health care decision making, planning, and research. When a patient visits the public primary health care unit in Finland, a physician makes a note of the diagnosis in the electronic patient record. From there the diagnosis will be transferred to the Avohilmo database maintained by the research and development institute, the National Institute for Health and Welfare (NIHW).

The Finnish Medical Society Duodecim is a scientific society [12] that publishes medical information and contributes to the continuous professional development of doctors in Finland. Duodecim Medical Publications Ltd, owned by the Finnish Medical Society Duodecim, carries out publication of medical information. Duodecim Medical Publications Ltd produces and maintains Internet-based, chargeable Physician's databases (PD) consisting of, for example, practically orientated point-of-care medical guidelines especially for physicians serving in outpatient treatment and hospital outpatient clinics. The database is available throughout the entire health care system of Finland (over 20,000 working age physicians in Finland in 2014 [13]) and the users of database are health care professionals working in Finland. Every keyword that health care professionals search is included in a log file.

We chose Lyme borreliosis (LB) as a model to evaluate the usability of health care professionals' queries for surveillance purposes in a dedicated EBM portal. LB is a spirochetal infectious disease caused by *Borrelia burgdorferi sensu lato* and is transmitted via ticks [14]. Seasonal variation is an important feature of LB [15,16], since annual climate changes affect tick activity that transmit the *B. burgdorferi* pathogen to humans between spring and autumn [16] when people go outdoors or on holiday. The incidence of LB has increased significantly in Europe [17,18]. Significant variation exists in both the temporal and the geographical distribution of LB in Finland [19] (personal communication by E Sajanti, MJ Virtanen, J Hytönen, J Sane, October 10, 2016). The differential diagnostics of LB presenting with erythema migrans (EM) are scarce [14], making register-based surveillance possible. The Avohilmo data on LB (mainly EM, ie, a typical tick-bite rash in LB) are registered by an International Classification of Diseases (10th Revision; ICD-10) disease classification code "A69.2" and are available since 2011. According to Avohilmo and NIDR registers, the incidence of LB in Finland has increased steadily (personal communication by E Sajanti, MJ Virtanen, J Hytönen, J Sane, October 10, 2016). Of note, in primary health care patients with EM, a clinical picture is sufficient for diagnosis and further laboratory testing is not recommended. Therefore, the cases reported to Avohilmo are not likely reported in NIDR. Our hypothesis was that the timeliness of health care professionals' queries coincides with Avohilmo findings of LB. The chosen registers, PD and Avohilmo, represent unique databases, which do not allow matching the searches and diagnoses to one another.

Methods

We carried out a descriptive register study to compare two registers, PD and the register of primary health care outpatient

diagnoses (Avohilmo), to research health care professionals' Internet-based queries of LB and public primary health care outpatient diagnoses on LB. To study the use of electronic databases in the context of a well-defined disease, we retrospectively collected logs of PD searches for LB by using the keywords: "borre*" or "lyme*" or "migrans*." These searches were further defined by the number of searches (years 2011-2015) in the whole country and all 21 health care districts in Finland annually and monthly. Three high-incidence LB regions (Helsinki and Uusimaa, Southwest Finland, and Kymenlaakso), all located in Southern Finland [19], were selected to be studied further. We chose blood pressure and diabetes to serve as comparison words to LB search word to distinguish actual increase in single search parameter from annual increase of general searches. The aim of this study was to compare PD searches to public primary health care outpatient diagnoses (Avohilmo) (1) annually in the whole country during 2011-2015, (2) monthly in the whole country during 2014-2015, and also (3) annually in three selected high-incidence LB regions during 2011-2015. An ethical approval for this study was granted by NIHW.

Results

We found visually similar patterns in annual and regional variation in LB searches and primary health care outpatient diagnoses. Figure 1 shows the annual variation of the whole

country in 2011-2015. The PD searches of the whole country in 2011-2015 start mostly in April and reach the maximum (searches peak at 25,463 and primary health care outpatient diagnoses at 835 in 2015) in July-August, and descend to the minimum in February (searches lowest at 1618 in 2011 and primary health care outpatient diagnoses at 15 in 2012). In June-September 2014, a plateau stage occurs, a pattern that is also present in Avohilmo. Comparison words, blood pressure and diabetes, to LB search word show no temporal variation (data not shown) as LB does.

The monthly variation in searches in the whole country in 2014-2015 is seen in Figure 2 and diagnoses in Figure 3. In 2014, the PD and Avohilmo data show a rapid ascent in April, both peak in July, and then start a decline forming a double-peaked pattern in July-September. The findings in PD and Avohilmo data reach the minimum in February and March 2015, respectively. In 2015, both searches and primary health care outpatient diagnoses show a rapid up-and-down pattern starting in April, peaking in August, and then declining fast.

In both PD and Avohilmo data, three hospital regions stand out as high-incidence areas of LB (Helsinki and Uusimaa, Southwest Finland, and Kymenlaakso). These regions show similar patterns between in both PD and Avohilmo data during 2011-2015 (data not shown). These regions also follow the same annual variation pattern as shown in Figure 1.

Figure 1. Physician's database searches for Lyme borreliosis (solid line) and Avohilmo diagnoses for Lyme borreliosis (dashed line) in the whole country during 2011-2015.

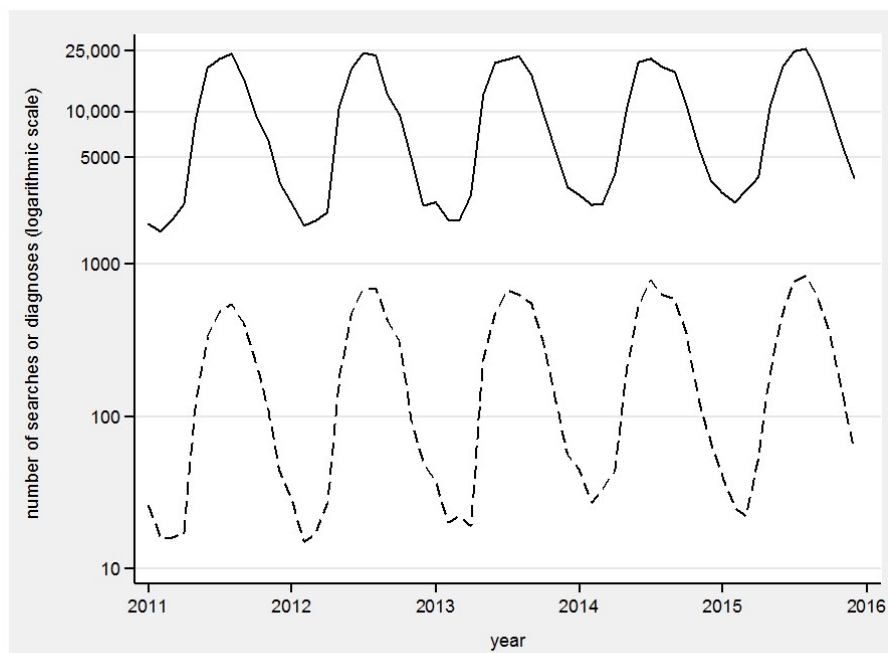


Figure 2. Physician's database searches for Lyme borreliosis in the whole country during 2014-2015.

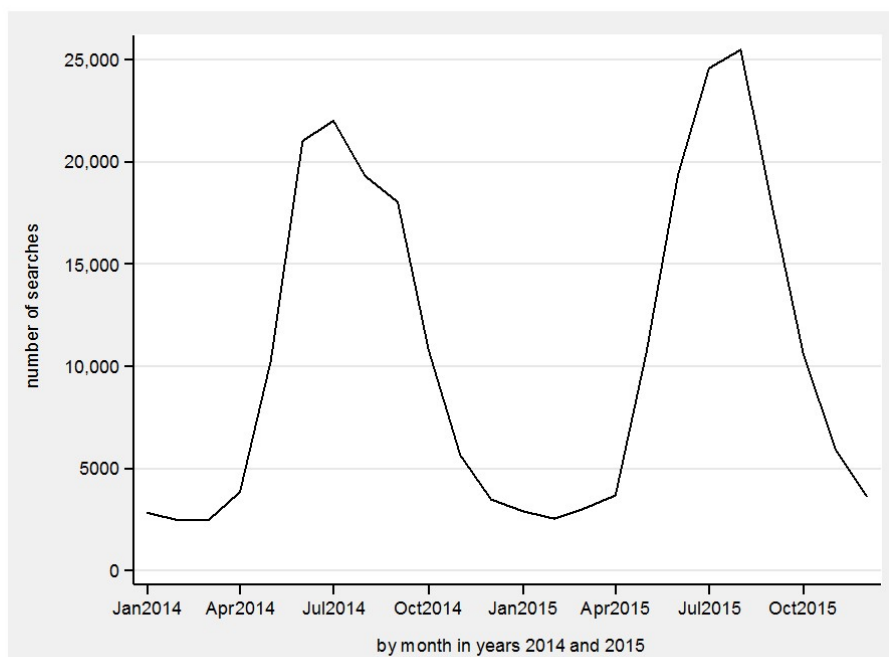
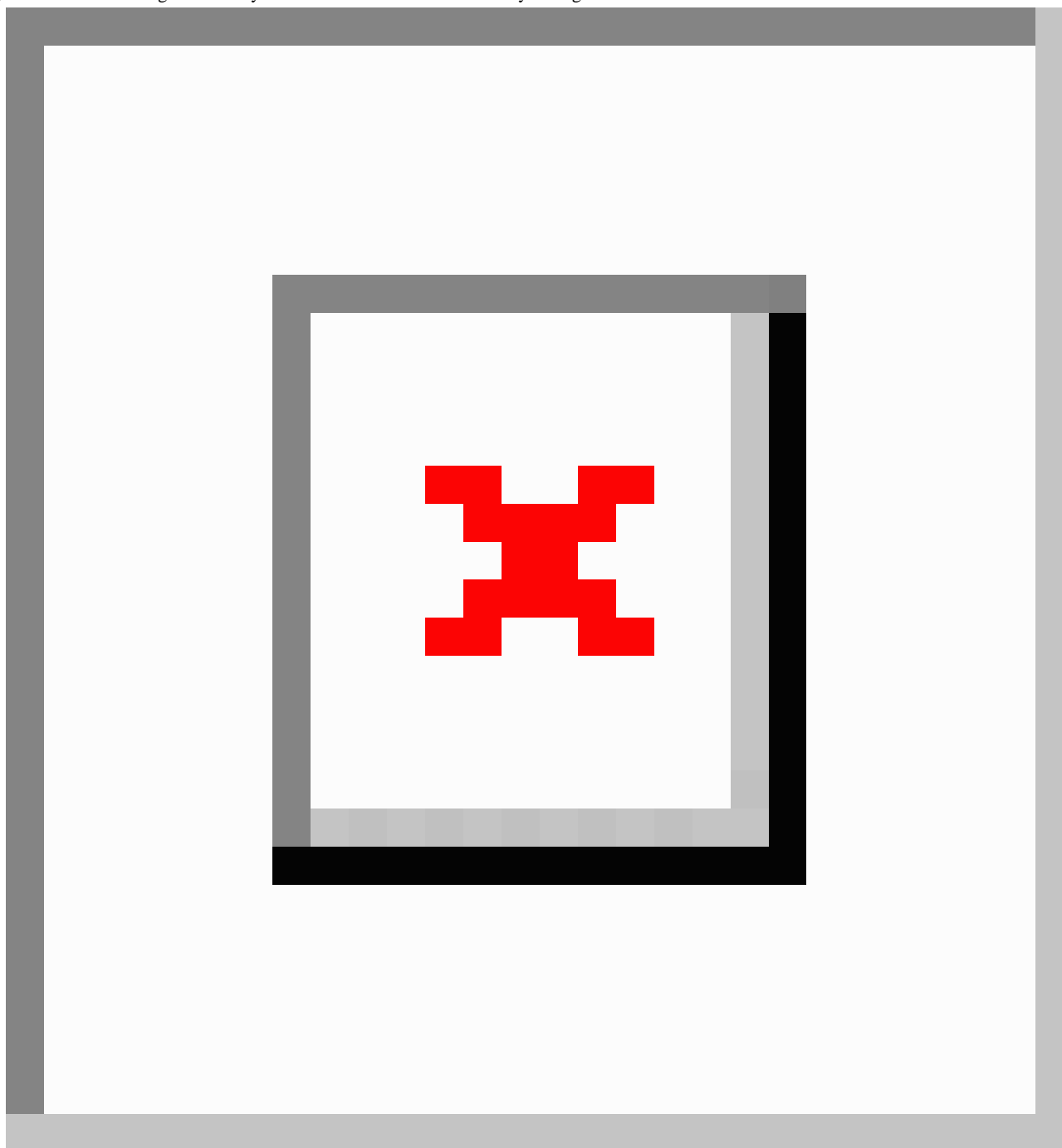


Figure 3. Avohilmo diagnoses for Lyme borreliosis in the whole country during 2014-2015.



Discussion

Our study showed visually similar patterns between PD searches and primary health care outpatient LB diagnoses both in the whole country and in three healthcare districts with high incidence of LB during 2011-2015. In PD searches, the seasonal variation was clear and regional differences appeared which were consistent with LB diagnoses registered in Avohilmo. A double-peaked pattern in July-September 2014 (Figures 2 and 3) may possibly be related to media trends.

The use of Internet-based EBM sources, such as the Cochrane Library, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and MEDLINE or Pubmed, has increased

over the years among health care professionals [20]. Log files from Internet portals such as Google could be utilized to help in detecting and predicting epidemiological patterns [1,4,5]. Although we speculate that the double-peaked pattern, formed by health care professionals' searches, is related to media trends, we hypothesize that nonprofessionals' queries are even more affected. Therefore, we believe professional queries to mimic more closely actual disease trends. The Google flu-related search trends that included searches by health care professionals and nonprofessionals appeared to fail to predict the timing of emerging infectious diseases [4]. Search logs from Internet EBM databases for health care professionals, however, are here shown to accurately mirror LB diagnoses in a nationwide discharge register.

Our study includes some limitations that should be considered. LB is often diagnosed days after the initial tick bite. It is not uncommon for a patient to seek physician's attention after coming home from a holiday in a region with a high prevalence of LB. Therefore, the weakness in our study is associated with the fact that diagnoses and searches are not necessarily stated and performed at the same geographical location of contraction of disease. Another limitation in our study is that PD searches and Avohilmo diagnoses represent unique entries in these registers, not unique users or patients. Therefore, they cannot be linked directly to one another. Thus, the conclusions on the number of diagnoses should be carefully drawn when

considering a measure of LB incidence. Our study showed strengths in representativeness (health care professionals) and timeliness (real-time Internet database).

To our knowledge, this is the first study to demonstrate an association between health care professionals' Internet-based searches and a nationwide primary care discharge register on a specific diagnosis. We state that PD searches closely mimic the known seasonal variation of LB. This finding could be used for means to strengthen the surveillance of seasonal increases of the disease. Future research should focus on the validation of the method and the applicability of the method to other specific pathogens.

Authors' Contributions

MJV, OL, SM, and OH designed the study concept. SP, JS, JJ, and MK planned the study concept. SP, JS, OL, and OH performed the literature research. PM gathered and supplied the study data. SP carried out the data analysis. SP, MJV, and JS carried out the data interpretation. JS and OH involved in the final approval and critical revision. MJV, JJ, OL, PM, SM, and MK involved in the critical revision. SP drafted the manuscript. SP and OH are the guarantors.

Conflicts of Interest

Dr Minna Kaila reports various trustee positions in the Medical Society Duodecim since the late 1990s. Dr Jukkapekka Jousimaa reports personal fees from Duodecim Medical Publications Ltd during the conduct of this study. Dr Otto Helve reports various trustee positions in the Medical Society Duodecim and Kustannus Oy Duodecim since 2009. Other authors have no competing interests.

References

1. 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]
2. 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]
3. Milinovich GJ, Williams GM, Clements AC, Hu W. Internet-based surveillance systems for monitoring emerging infectious diseases. *Lancet Infect Dis* 2014 Feb;14(2):160-168. [doi: 10.1016/S1473-3099(13)70244-5] [Medline: 24290841]
4. Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLoS Comput Biol* 2013;9(10):e1003256 [FREE Full text] [doi: 10.1371/journal.pcbi.1003256] [Medline: 24146603]
5. Woo H, Cho Y, Shim E, Lee J, Lee C, Kim SH. Estimating influenza outbreaks using both search engine query data and social media data in South Korea. *J Med Internet Res* 2016;18(7):e177 [FREE Full text] [doi: 10.2196/jmir.4955] [Medline: 27377323]
6. 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]
7. Jormanainen V, Jousimaa J, Kunnamo I, Ruutu P. Physicians' database searches as a tool for early detection of epidemics. *Emerg Infect Dis* 2001;7(3):474-476 [FREE Full text] [doi: 10.3201/eid0703.010324] [Medline: 11384534]
8. Dwairy M, Dowell AC, Stahl J. The application of foraging theory to the information searching behaviour of general practitioners. *BMC Fam Pract* 2011;12:90 [FREE Full text] [doi: 10.1186/1471-2296-12-90] [Medline: 21861880]
9. Younger P. Internet-based information-seeking behaviour amongst doctors and nurses: a short review of the literature. *Health Info Libr J* 2010 Mar;27(1):2-10 [FREE Full text] [doi: 10.1111/j.1471-1842.2010.00883.x] [Medline: 20402799]
10. Hider PN, Griffin G, Walker M, Coughlan E. The information-seeking behavior of clinical staff in a large health care organization. *J Med Libr Assoc* 2009 Jan;97(1):47-50 [FREE Full text] [doi: 10.3163/1536-5050.97.1.009] [Medline: 19159006]
11. Davies K. UK doctors awareness and use of specified electronic evidence-based medicine resources. *Inform Health Soc Care* 2011 Jan;36(1):1-19. [doi: 10.3109/17538157.2010.534211] [Medline: 21142434]
12. Duodecim. Duodecim Medical Publications Ltd URL: <http://www.duodecim.fi/english> [accessed 2016-10-07] [WebCite Cache ID 615Pr3evA]
13. laakariliitto.fi. The Finnish Medical Association: Physicians 2016 URL: http://www.laakariliitto.fi/site/assets/files/1011/116_taskutil_06_en_160524net.pdf [WebCite Cache ID 615PCE40g]

14. Shapiro ED. Borrelia burgdorferi (Lyme disease). *Pediatr Rev* 2014 Dec;35(12):500-509. [doi: [10.1542/pir.35-12-500](https://doi.org/10.1542/pir.35-12-500)] [Medline: [25452659](https://pubmed.ncbi.nlm.nih.gov/25452659/)]
15. Nelson CA, Saha S, Kugeler KJ, Delorey MJ, Shankar MB, Hinckley AF, et al. Incidence of Clinician-Diagnosed Lyme Disease, United States, 2005-2010. *Emerg Infect Dis* 2015 Sep;21(9):1625-1631 [FREE Full text] [doi: [10.3201/eid2109.150417](https://doi.org/10.3201/eid2109.150417)] [Medline: [26291194](https://pubmed.ncbi.nlm.nih.gov/26291194/)]
16. Euro.who. Lyme borreliosis in Europe: influences of climate and climate change, epidemiology, ecology and adaptation measures URL: http://www.euro.who.int/_data/assets/pdf_file/0006/96819/E89522.pdf [WebCite Cache ID 6mG7NPP0Q]
17. Bennet L, Halling A, Berglund J. Increased incidence of Lyme borreliosis in southern Sweden following mild winters and during warm, humid summers. *Eur J Clin Microbiol Infect Dis* 2006 Jul;25(7):426-432. [doi: [10.1007/s10096-006-0167-2](https://doi.org/10.1007/s10096-006-0167-2)] [Medline: [16810531](https://pubmed.ncbi.nlm.nih.gov/16810531/)]
18. Wilking H, Stark K. Trends in surveillance data of human Lyme borreliosis from six federal states in eastern Germany, 2009-2012. *Ticks Tick Borne Dis* 2014 Apr;5(3):219-224. [doi: [10.1016/j.ttbdis.2013.10.010](https://doi.org/10.1016/j.ttbdis.2013.10.010)] [Medline: [24656810](https://pubmed.ncbi.nlm.nih.gov/24656810/)]
19. Julkari.fi. Infectious diseases in Finland 2013: National Institute for Health and Welfare, Report 18/2014 URL: http://www.julkari.fi/bitstream/handle/10024/125566/URN_ISBN_978-952-302-194-5.pdf?sequence=1 [WebCite Cache ID 6mG6nXp9b]
20. Weng Y, Kuo KN, Yang C, Lo H, Shih Y, Chen C, et al. Increasing utilization of Internet-based resources following efforts to promote evidence-based medicine: a national study in Taiwan. *BMC Med Inform Decis Mak* 2013 Jan 07;13:4 [FREE Full text] [doi: [10.1186/1472-6947-13-4](https://doi.org/10.1186/1472-6947-13-4)] [Medline: [23289500](https://pubmed.ncbi.nlm.nih.gov/23289500/)]

Abbreviations

Avohilmo: the register of public primary health care diagnoses

EBM: evidence-based medicine

EM: erythema migrans

LB: Lyme borreliosis

NIDR: National Infectious Diseases Register

NIHW: National Institute for Health and Welfare

PD: Physician's databases

Edited by G Eysenbach; submitted 13.10.16; peer-reviewed by A Benetoli, AC Ianos; comments to author 10.11.16; revised version received 08.12.16; accepted 02.03.17; published 11.04.17.

Please cite as:

Pesälä S, Virtanen MJ, Sane J, Jousimaa J, Lyytikäinen O, Murtopuro S, Mustonen P, Kaila M, Helve O

Health Care Professionals' Evidence-Based Medicine Internet Searches Closely Mimic the Known Seasonal Variation of Lyme Borreliosis: A Register-Based Study

JMIR Public Health Surveill 2017;3(2):e19

URL: <http://publichealth.jmir.org/2017/2/e19/>

doi: [10.2196/publichealth.6764](https://doi.org/10.2196/publichealth.6764)

PMID: [28400357](https://pubmed.ncbi.nlm.nih.gov/28400357/)

©Samuli Pesälä, Mikko J Virtanen, Jussi Sane, Jukkapekka Jousimaa, Outi Lyytikäinen, Satu Murtopuro, Pekka Mustonen, Minna Kaila, Otto Helve. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 11.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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

Public Response to Scientific Misconduct: Assessing Changes in Public Sentiment Toward the Stimulus-Triggered Acquisition of Pluripotency (STAP) Cell Case via Twitter

Alberto Gayle¹, MSDIS; Motomu Shimaoka¹, MD, PhD

Mie University Graduate School of Medicine, Tsu, Japan

Corresponding Author:

Motomu Shimaoka, MD, PhD

Mie University Graduate School of Medicine

2-174 Edobashi

Tsu, 514-8507

Japan

Phone: 81 592315036

Fax: 81 592315209

Email: shimaoka@doc.medic.mie-u.ac.jp

Abstract

Background: In this age of social media, any news—good or bad—has the potential to spread in unpredictable ways. Changes in public sentiment have the potential to either drive or limit investment in publicly funded activities, such as scientific research. As a result, understanding the ways in which reported cases of scientific misconduct shape public sentiment is becoming increasingly essential—for researchers and institutions, as well as for policy makers and funders. In this study, we thus set out to assess and define the patterns according to which public sentiment may change in response to reported cases of scientific misconduct. This study focuses on the public response to the events involved in a recent case of major scientific misconduct that occurred in 2014 in Japan—stimulus-triggered acquisition of pluripotency (STAP) cell case.

Objectives: The aims of this study were to determine (1) the patterns according to which public sentiment changes in response to scientific misconduct; (2) whether such measures vary significantly, coincident with major timeline events; and (3) whether the changes observed mirror the response patterns reported in the literature with respect to other classes of events, such as entertainment news and disaster reports.

Methods: The recent STAP cell scandal is used as a test case. Changes in the volume and polarity of discussion were assessed using a sampling of case-related Twitter data, published between January 28, 2014 and March 15, 2015. Rapidminer was used for text processing and the popular bag-of-words algorithm, SentiWordNet, was used in Rapidminer to calculate sentiment for each sample Tweet. Relative volume and sentiment was then assessed overall, month-to-month, and with respect to individual entities.

Results: Despite the ostensibly negative subject, average sentiment over the observed period tended to be neutral (−0.04); however, a notable downward trend ($y = -0.01x + 0.09$; $R^2 = .45$) was observed month-to-month. Notably polarized tweets accounted for less than one-third of sampled discussion: 17.49% (1656/9467) negative and 12.59% positive (1192/9467). Significant polarization was found in only 4 out of the 15 months covered, with significant variation month-to-month ($P < .001$). Significant increases in polarization tended to coincide with increased discussion volume surrounding major events ($P < .001$).

Conclusions: These results suggest that public opinion toward scientific research may be subject to the same sensationalist dynamics driving public opinion in other, consumer-oriented topics. The patterns in public response observed here, with respect to the STAP cell case, were found to be consistent with those observed in the literature with respect to other classes of news-worthy events on Twitter. Discussion was found to become strongly polarized only during times of increased public attention, and such increases tended to be driven primarily by negative reporting and reactionary commentary.

(*JMIR Public Health Surveill* 2017;3(2):e21) doi:[10.2196/publichealth.5980](https://doi.org/10.2196/publichealth.5980)

KEYWORDS

scientific misconduct; retraction of publication as a topic; mass media; social media; public opinion; public policy; data mining; publication; stem cells; Japan

Introduction

Background

With the rise of social network services (SNS), all news events, no matter how large or small, have become subject to intense public scrutiny and debate [1,2]. Of course, this has gone on in some form or another since the advent of civilization. However, the democratization, reach, and consequence of public scrutiny has never before been realized to the degree seen today [3-5]. Indeed, although assessment of public opinion has traditionally been the domain of pollsters and social scientists, social media analytics are increasingly being seen as a reliable alternative [6]. With a large and increasingly diverse demographic base, Twitter has been shown to be reasonably representative in terms of demographics [7,8] and public sentiment [9], especially with respect to breaking news [10-12].

Recent investigations into communication on Twitter have uncovered common, generalizable patterns in the way sentiment changes in response to the emergence of notable events—namely, that increases in public attention are coincident with increases in negative sentiment [13]. Such patterns follow known dynamics associated with media sensationalism [14,15] and have been observed across a broad spectrum of mass media topics, including entertainment, sports, business, politics, and natural disasters [16]. Sensationalism has also been found to be a problem in the reporting of medical science [17]. This is of particular concern given the profound and lasting impact on the direction of public policy that sensationalist reporting might have [18]. And although studies have examined the role of the traditional news media in shaping public opinion as it relates to medical science and policy [19], no studies to date have explored whether such dynamics would apply to the presumably expert-driven communications on Twitter.

One area of particular interest is scientific misconduct, particularly in the areas of academic and medical science. Scientific misconduct concerns more than just a given researcher or institution; damage to public perception of, and goodwill toward scientific research itself is a driving concern [20]. Most academic research institutions derive the bulk of their research budgets from public spending, and so a loss of reputation can have a direct and far-ranging impact. Academic institutions invest heavily in anticipation of future academic trends and research demands [21]; consequently, unanticipated changes in public policy or funding may result in large, unrecoupable capital expenditures and lost opportunity [22]. Understanding the specific dynamics governing public response to reports of scientific misconduct on social media is therefore invaluable.

Stimulus-Triggered Acquisition of Pluripotency Cell Case

Here, we assess and define the patterns according to which public sentiment may change in response to reports of academic scientific misconduct on Twitter. This study focuses on public response to a recent and widely covered case of scientific misconduct—the stimulus-triggered acquisition of pluripotency (STAP) cell case that occurred in Japan in 2014 [23]. The STAP cell case is used in this study, as this represents the major

scientific misconduct in the era of SNS and was well mentioned by the mainstream media such as Nature, Science, New York Times, Cable News Network (CNN), and British Broadcasting Corporation (BBC). In this case, media reports focused on Ms Obokata, an upcoming biochemist, who attracted as much attention for her achievements, as for her gender and youth. Attention also focused on notable coauthors, Dr Y Sasai, Dr T Wakayama, and Dr C Vacanti, as well as the sponsoring institution, Riken.

Here, we have demonstrated that Twitter response to the STAP case tended to generally stay neutral, but specifically skew negative as discussion polarity and volume increased. Our results are consistent with those observed in studies covering other topics of interest [24]. These findings suggest that changes in public sentiment toward any major event—cases of scientific misconduct included—might be as much a function of the attention received as it is a function of the theme or merits of any specific case.

Methods

Data Collection

For the purpose of this analysis, an event timeline was constructed based on primary reports and press-releases [25,26]. Where such data were lacking, secondary sources were vetted and compiled [23,27-29]. All data obtained were used and reported according to the terms and privacy policies of the respective data sources. Where such no such policy existed, data were assumed to be public domain.

Twitter data (“Tweets”) were obtained directly from Apple’s now defunct subsidiary, Topsy [30]. According to Twitter, “a Tweet...is a message of 140 characters or less that is public by default” [31]. For the purposes of this study, the data obtained was handled according to established ethical precedent regarding public domain social media content, that is, consent of the authors was neither required nor obtained [32]. Our search covered a full 14 months, beginning with the day of Riken’s press release concerning STAP (January 29, 2014) and ending March 31, 2015. Prior research has demonstrated that datasets derived using a limited set of focused keywords are suitable for the purpose of analyzing public sentiment regarding a specific issue or brand [33]. In order to minimize collection of irrelevant Tweets, the following minimum set of search criteria were chosen: obokata OR 小保方 OR #obokata OR #小保方. Other, potentially relevant terms, such as “STAP” and “Riken” were excluded at this stage after having been found to return results that were likely to be either redundant or irrelevant.

An automated platform for downloading and processing Tweets was developed using RStudio [34]. Using this platform, we obtained a sample set consisting of all English-language Tweets covering the relevant timeframe polled at a rate of $n=100$ Tweets every 3 h. All content as well as associated metadata (eg, author name, date, and number of retweets) were downloaded. This initial, raw dataset contained $N=12,925$ Tweets in total. Full-text was retained for qualitative analysis.

Data Processing

Initial processing of data was conducted programmatically within RStudio. This included formatting and transformation of the downloaded data into tables. Downloaded tables were saved and processed using Excel 2013. This included the programmatic identification and removal of all Tweets containing non-English characters or text. In addition, all Tweets

were manually checked to remove irrelevant or spam posts, resulting in a final dataset of $n=9467$ Tweets total.

For the purpose of sentiment classification, RapidMiner version 6.3, Enterprise Edition (Rapid-I GmbH, Dortmund, Germany) was used. The following preprocessing steps were followed (Table 1). These steps are common across the many domains in which text-mining is applied [35], with minimal variation.

Table 1. Text processing steps.

Steps	Description
Tokenize	Parse every tweet into separate, single-element tokens (ie, words or word-parts)
Transform cases	Makes all text lower case to facilitate data processing
Filter tokens by length	Removes tokens consisting of less than 2 characters
Filter English stop words	Removes common, low-information particles (eg, “the”) and punctuation marks
Filter tokens by content	Removed hashtags and other message-irrelevant tokens such as “http”
Stemming (WordNet)	Algorithm for identifying and groups tokens as lemmas, to facilitate processing
Generate n-grams	Generates list of all two-, three-, or four-word token combinations (ie, phrases)
Word vector creation	Generate metric indicative of the measuring the important of each word in a tweet
Pruning	Remove tokens that appear in less than 1% or more than 80% of documents

This processing generated weighted word vectors, representing the weighted distribution of each processed token or n-gram within a given Tweet. Word vector statistics were calculated using the term frequency-inverse document frequency (TF-IDF) weighting scheme. TF-IDF emphasizes the importance of key but not uncommon terms [36,37] and has been demonstrated to improve the performance of text-mining tasks [38]. TF-IDF is calculated as follows: $TF\text{-}IDF = tf * \log(N/df)$, where tf is the frequency of a term within a given document, df is the frequency across all documents, and N is the number of documents total [39].

To evaluate sentiment for each Tweet, the SentiWordNet 3.0 extension was used within Rapidminer. SentiWordNet is a well-established sentiment analysis protocol and has been cited by almost 1000 (988) journal publications as of the date of this writing, according to Google Scholar search. SentiWordNet assigns three sentiment scores (“positive,” “negative,” and “objective”) to each word, based on a generalized classification system developed by the authors using a combination of manual and automated sentiment scoring algorithms [40]. SentiWordNet’s “bag-of-words” methodology has been demonstrated to be reliable for document-level sentiment analysis, with aggregate-level performance roughly on par with more sophisticated methods, including human coding [41].

For this analysis, nouns were omitted from sentiment calculation. Recent studies have demonstrated that, for automated sentiment analyses, nouns are not likely to provide additional, reliable information [42]. And in topics with terminology that is either uncommon or uncommonly applied, this is even more the case—especially when using a general purpose lexicon such as SentiWordNet [43]. All terms were, however, retained for topic-level analysis. The sentiment of each Tweet was then calculated by aggregating the scores of all relevant word tokens, as determined using SentiWordNet.

Scores were thus assigned for each Tweet, ranging from -1 to $+1$, based on the estimated degree of negative or positive sentiment. These scores are reported in unstandardized form. For the purpose of statistical analysis and visualization, scores were then standardized, to produce a distribution with mean of zero ($\bar{x}=0$) and standard deviation of one ($\sigma_x=1$). All Tweets with standardized scores less than -1 were labeled “negative,”

whereas those with standardized scores greater than $+1$ were labeled “positive;” Tweets with standardized scores less than $+1$ but greater than -1 were labeled “neutral.”

A support vector machine (SVM) analysis was then used to identify the terms and phrases that were most commonly associated with each respective sentiment label. SVM is a computational method that derives a classification scheme based on the degree to which the various input cases (ie, word vectors) predict a given binary class (eg, positive or negative sentiment or “mentions Sasai” or null) [44]. All input terms (and term combinations, ie, “n-grams”) can thus be assessed in terms of “importance” with respect to a given label [45]. Conceptually, this is similar to a logistic regression; however, the computation is far more computationally intensive [46,47]. In addition, Tweets mentioning Ms Obokata, Dr Sasai, and the Riken institution were labeled accordingly; associated terms or phrases were also extracted via SVM.

Data Analysis and Visualization

Once the Twitter data were processed as described, the data were exported to Microsoft Excel for further processing using the Pivot Table function. Sentiment as well as sampled Tweet volume were aggregated and indices were calculated for all relevant sub- and cross-tables. These tables were then used to generate visualizations either directly in Microsoft Excel or using ggplot2 and ggtern in RStudio. In cases where a given table or visualization suggested a time-trend or association with respect to aggregate sentiment or Tweet volume, statistical significance was assessed using chi-squared and Tukey’s post hoc (1-way analysis of variance, ANOVA) tests. A GLM model (with Bonferroni correction) was used to test month-to-month

mean difference versus previous months; 2-way ANOVA was used to compare metrics for individual entities. SPSS Statistics version 23 (IBM Corp) was used for all statistical tests.

Results

Analysis of Sentiment

Over the 15-month period covered, overall sentiment was found to be -0.037 on average, with a notable downward trend (Figure 1). Over this same period, sampled Tweets averaged 631.1 Tweets per month, with a maximum of 2349 Tweets (April

2014; 372.2% index) and a minimum of 75 Tweets (November 2014; 11.9% index). One-way ANOVA with Tukey's honest significant difference (HSD) confirmed significant month-to-month variation with respect to sentiment (Table 2). Overall, discussion tended to be mostly neutral (69.92%), with positive and negative discussion being far less prevalent (12.59% and 17.49%, respectively) overall. Discussion month-to-month tended to be mostly objective; however, when polarized, discussion tended negative (Figure 2). Chi-squared tests confirmed observed differences to be significant in this respect as well (Table 3).

Table 2. Tukey's post hoc test for significance (1-way analysis of variance, ANOVA). Italicized values indicate significance $P < .05$.

Year	Month	2014												2015		
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
2014	Jan	0.000	<i>-0.039</i>	<i>-0.044</i>	<i>-0.105</i>	<i>-0.135</i>	<i>-0.060</i>	<i>-0.134</i>	<i>-0.078</i>	0.015	<i>-0.036</i>	<i>-0.106</i>	<i>-0.125</i>	<i>-0.160</i>	<i>-0.348</i>	<i>-0.260</i>
	Feb	<i>0.039</i>	0.000	<i>-0.005</i>	<i>0.067</i>	<i>-0.096</i>	<i>-0.021</i>	<i>-0.095</i>	<i>-0.039</i>	0.053	0.002	<i>-0.067</i>	<i>0.087</i>	<i>-0.121</i>	<i>-0.309</i>	<i>-0.222</i>
	Mar	<i>0.044</i>	0.005	0.000	<i>-0.062</i>	<i>-0.091</i>	<i>-0.016</i>	<i>-0.090</i>	<i>-0.034</i>	<i>0.059</i>	0.008	<i>-0.062</i>	<i>-0.081</i>	<i>-0.116</i>	<i>-0.304</i>	<i>-0.217</i>
	Apr	<i>0.105</i>	<i>0.067</i>	<i>0.062</i>	0.000	<i>-0.030</i>	<i>0.045</i>	<i>-0.029</i>	0.028	<i>0.120</i>	<i>0.069</i>	<i>-0.001</i>	<i>-0.020</i>	<i>-0.054</i>	<i>-0.242</i>	<i>-0.155</i>
	May	<i>0.135</i>	<i>0.096</i>	<i>0.091</i>	<i>0.030</i>	0.000	<i>0.075</i>	0.001	<i>0.057</i>	<i>0.150</i>	<i>0.099</i>	0.029	0.010	<i>-0.025</i>	<i>-0.213</i>	<i>-0.125</i>
	Jun	<i>0.060</i>	0.021	0.016	<i>-0.045</i>	<i>-0.075</i>	0.000	<i>-0.074</i>	<i>-0.018</i>	<i>0.075</i>	0.024	<i>-0.046</i>	<i>-0.065</i>	<i>-0.100</i>	<i>-0.288</i>	<i>-0.200</i>
	Jul	<i>0.134</i>	<i>0.095</i>	<i>0.090</i>	<i>0.029</i>	<i>-0.001</i>	<i>0.074</i>	0.000	<i>0.056</i>	<i>0.149</i>	<i>0.098</i>	0.028	0.009	<i>-0.026</i>	<i>-0.214</i>	<i>-0.126</i>
	Aug	<i>0.078</i>	0.039	0.034	<i>-0.028</i>	<i>-0.057</i>	0.018	<i>-0.056</i>	0.000	<i>0.093</i>	0.041	<i>-0.028</i>	<i>-0.047</i>	<i>-0.082</i>	<i>-0.270</i>	<i>-0.183</i>
	Sep	<i>-0.015</i>	<i>-0.053</i>	<i>-0.059</i>	<i>-0.120</i>	<i>-0.150</i>	<i>-0.075</i>	<i>-0.149</i>	<i>-0.093</i>	0.000	<i>-0.051</i>	<i>-0.121</i>	<i>-0.140</i>	<i>-0.175</i>	<i>-0.363</i>	<i>-0.275</i>
	Oct	<i>0.036</i>	<i>-0.002</i>	<i>-0.008</i>	<i>-0.069</i>	<i>-0.099</i>	<i>-0.024</i>	<i>-0.098</i>	<i>-0.041</i>	0.051	0.000	<i>-0.070</i>	<i>-0.089</i>	<i>-0.123</i>	<i>-0.312</i>	<i>-0.224</i>
	Nov	<i>0.106</i>	0.067	0.062	0.001	<i>-0.029</i>	0.046	<i>-0.028</i>	0.028	<i>0.121</i>	0.070	0.000	<i>-0.019</i>	<i>-0.054</i>	<i>-0.242</i>	<i>-0.154</i>
	Dec	<i>0.125</i>	<i>0.087</i>	<i>0.081</i>	0.020	<i>-0.010</i>	<i>0.065</i>	<i>-0.009</i>	<i>0.047</i>	<i>0.140</i>	<i>0.089</i>	0.019	0.000	<i>-0.035</i>	<i>-0.223</i>	<i>-0.135</i>
2015	Jan	<i>0.160</i>	<i>0.121</i>	<i>0.116</i>	<i>0.054</i>	0.025	<i>0.100</i>	0.026	<i>0.082</i>	<i>0.175</i>	<i>0.123</i>	0.054	0.035	0.000	<i>-0.188</i>	<i>-0.101</i>
	Feb	<i>0.348</i>	<i>0.309</i>	<i>0.304</i>	<i>0.242</i>	<i>0.213</i>	<i>0.288</i>	<i>0.214</i>	<i>0.270</i>	<i>0.363</i>	<i>0.312</i>	<i>0.242</i>	<i>0.223</i>	<i>0.188</i>	0.000	<i>0.087</i>
	Mar	<i>0.260</i>	<i>0.222</i>	<i>0.217</i>	<i>0.155</i>	<i>0.125</i>	<i>0.200</i>	<i>0.126</i>	<i>0.183</i>	<i>0.275</i>	<i>0.224</i>	<i>0.154</i>	<i>0.135</i>	<i>0.101</i>	<i>-0.087</i>	0.000

Table 3. Tukey’s post hoc test for homogenous subsets (1-way analysis of variance, ANOVA).

Year	Month	N	Subset for alpha=.05									
			1	2	3	4	5	6	7	8	9	
2015	February	1034	-.2648									
	March	187		-.1774								
	January	209			-.0768							
2014	May	558			-.0521	-.0521						
	July	680			-.0510	-.0510						
	December	1092			-.0422	-.0422	-.0422					
	November	75				-.0230	-.0230	-.0230				
	April	2349				-.0224	-.0224	-.0224				
	August	395					.0052	.0052	.0052			
	June	887						.0230	.0230			
	March	691							.0391	.0391		
	February	424							.0443	.0443		
	October	136							.0467	.0467		
	January	630								.0829	.0829	
	September	120									.0978	
	Significance			<i>P</i> >.99	<i>P</i> >.99	<i>P</i> =.52	<i>P</i> =.77	<i>P</i> =.07	<i>P</i> =.09	<i>P</i> =.21	<i>P</i> =.14	<i>P</i> >.99

Both sets of trends were assessed against the timeline of events to determine whether sentiment and volume varied according to actual, real-world events. Findings are reported in terms of average sentiment, volume index (ratio of monthly or average

volume), and positivity index (ratio of positive to negative volume). [Multimedia Appendix 1](#) presents a summary timeline. A brief analysis is as follows.

Figure 1. Volume and average sentiment over time. Sentiment score calculated using unweighted aggregate sentiment scores found in the SentiWordNet database, for each valid token in each Tweet. For this analysis, verb, adjectives, and adverbs were considered valid for the purpose of sentiment scoring. Volume is based on number of Tweets retrieved per sampling interval. Sentiment increasingly negative over time; one key exception corresponds with the tragedy surrounding Dr Sasai (August to October 2014). Volume is driven by major events.

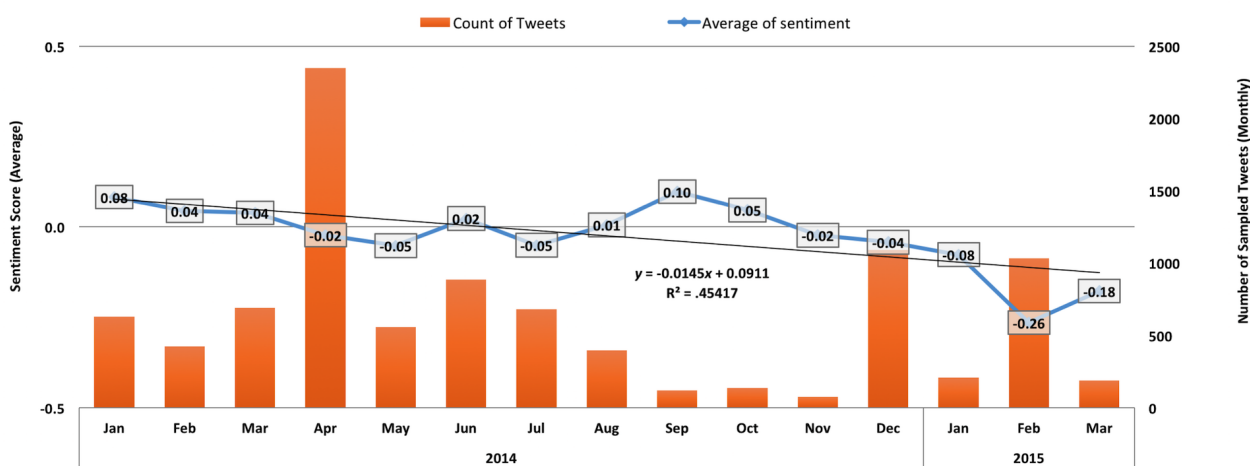
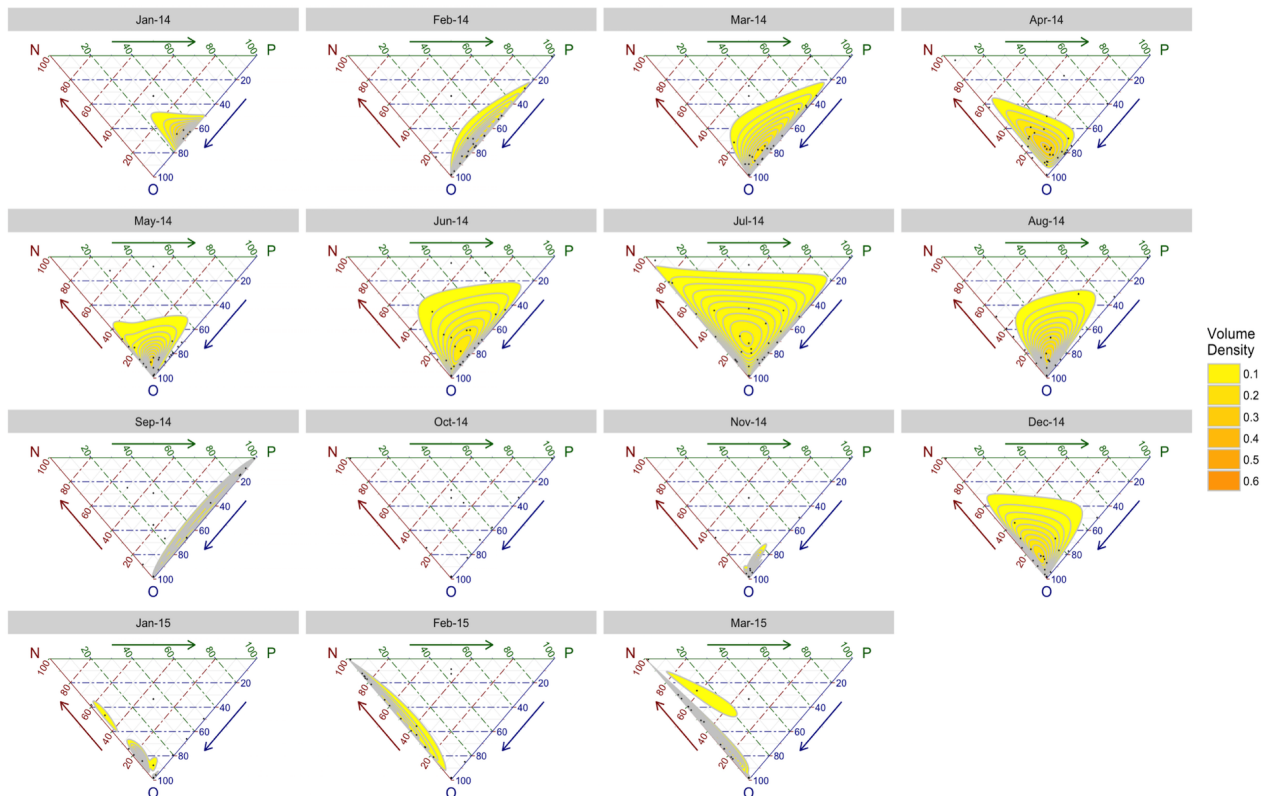


Figure 2. Month-to-month ternary sentiment or volume density chart. Density plot calculated based on the proportion of negative (N: top left), positive (P: top right), and objective or nonpolarized (O: bottom center) discussion volume (represented by the unlabeled data points). Volume density is calculated via isometric log ratio transformation.



Timeline Analysis and Analysis of Key Events

January 2014: STAP Announced to Much Fanfare; Obokata Hailed

On January 29, 2014, a letter [48] and research article [49] describing a “unique and unexpected cellular reprogramming phenomenon, called stimulus-triggered acquisition of pluripotency (STAP)” were published in Nature (ISSN: 0028-0836) by a team from the Riken Center for Developmental Biology in Japan and collaborators. These papers claimed that “stressors reprogram mammalian somatic cells, resulting in the generation of pluripotent cells”—a claim which challenged the very tenets of cell biology. Beginning one day before publication, the lead author, Ms Haruko Obokata, was hailed as a rising star in a series of media events and press releases [50].

For the 3 days comprising the month of January 2014 (January 29-31), average sentiment was found to be second highest among all months covered (0.083; $P < .001$, versus grand mean). In terms of volume, January indexed at par (99.8% index; 630/631.1); however, positive index was exceptionally high (3583.3%; 215/6). Discussion hails the “breakthrough” discovery; initial skepticism is present but muted.

February 2014: Doubts Begin to Surface; Riken Launches Investigation

Average sentiment decreased significantly, but remained overall positive (0.044; $P = .004$). In terms of volume, February indexed below par (67.2% index; 424/631.1); positivity also remained

high (888.9%; 80/9). Discussion continues to focus on the merits and implications of STAP; however, strong doubts begin to take hold.

- Initial concerns about possible figure manipulation first voiced on PubPeer (February 4) [51].
- STAP coauthor, Dr Charles Vacanti, posts images claimed to be human STAP cells (February 5) [52].
- Riken subsequently launches an investigation into “alleged irregularities” for both papers [53].

March 2014: Calls for Retraction Mount; Obokata Increasingly Implicated

Average sentiment decreased slightly, but not significantly (0.039; $P > .99$). However, volume began to increase (109.5% index; 691/631.1); discussion remained positive, albeit to a lesser degree (387.1%; 120/31). Allegations of possible misconduct increasingly take center stage; some antiscepticism still remains.

- “Essential technical tips for STAP...” published by Ms Obokata, Dr Sasai, and Dr Niwa on March 3 [54].
- STAP coauthor, Dr Wakayama, breaks from others and proposes retraction of both papers [55].
- In an interim report, the Riken investigation team finds inappropriate handling of data [56].

April 2014: Obokata Declared Guilty Amid Shake-Up at Riken; Public Interest Spikes

Discussion volume reached its peak (372.2% index; 2349/631.1); however, discussion took on a more negative tone. Average

sentiment decreased significantly (-0.022 ; $P<.001$), with a positivity now at 88.7% index (276/311). Riken's finding of misconduct, problems with the investigation team, and Ms Obokata's claims of innocence dominate; discussion centers on the drama between actors and institution, not the science.

- Riken finds Ms Obokata guilty of “two instances of research misconduct” in the STAP work [57].
- Ms Obokata holds a press conference in order to rebut Riken's conclusions [58].
- Nature publishes a strongly worded editorial on science policy in Japan, citing STAP case [59].

May 2014: Retraction Appears Inevitable; Obokata, However, Still Defiant

Average sentiment continues the negative trend (-0.052 ; $P=.003$). A substantial drop in volume was also observed (88.4% index; 558/631.1), accompanied by a sharp drop in positivity as well (57%; 29/51). Fueled by reports of prior, scathing rejections by Science and Cell, discussion now focuses on the nature of the experiments and results; Riken's initial rejection of Ms Obokata's appeal also a hot topic.

- Ms Obokata, under pressure to retract both papers, agrees to retract only the letter (May 28).
- Other senior authors continue to negotiate with Ms Obokata regarding the remaining article [60].

June 2014: STAP Papers Retracted; End Now in Sight, Calls for Punitive Action Emerge

Driven by a sharp increase in positivity (245%; 137/56) and overall discussion (140.5% index; 887/631.1), average sentiment became slightly positive (0.023; $P<.001$). Despite calls for punitive action against the senior investigators, STAP retraction is discussed in neutral terms. Decision to suspend disciplinary process against Ms Obokata and allow her to join STAP reproduction efforts is discussed favorably.

- STAP coauthor, Dr Wakayama, presents genetic evidence refuting the existence of STAP cells [61].
- Ms Obokata and coauthors finally agree to retract both papers published in Nature [62].
- Riken reform committee recommends restructuring of Center for Developmental Biology [63].

July 2014: STAP Discussion Takes on a Lighter—and Somewhat Derisive—Tone

Average sentiment was once again negative (-0.051 ; $P<.001$), with a positivity index of 67.8% (97/143) against volume only slightly above par (107.7% index; 680/631.1). Discussion focuses on the fallout of the STAP retraction.

- On July 3, the two Nature papers reporting the STAP cells are retracted [64,65].
- Ms Obokata sustains injuries while being pursued by television reporters [66].
- Concurrent with Riken's investigation of the STAP case, Waseda University begins investigation of alleged plagiarism in Ms Obokata's doctoral dissertation [67]; Ms Obokata retains PhD, for now.

August 2014: Sasai's Suicide Draws Muted Response

Discussion volume was subpar (62.6% index; 395/631.1); despite a positivity index of 115% (45/39), average sentiment was nevertheless mixed (0.005; $P<.001$). Discussion focuses on the apparent suicide of Dr Sasai; commentary is reserved, though some take the opportunity to critique the academic culture in Japan.

- On August 5, Dr Sasai, a STAP coauthor, found dead at the Riken center due to apparent suicide.
- Dr Sasai leaves behind a note addressed to Ms Obokata, urging her to verify existence of STAP [68].
- A STAP coauthor, Dr Niwa, announces his lab's failure to replicate STAP results (August 27) [69].

September 2014: Discussion More Sympathetic, Pensive as STAP Retrial Continues

Driven by a large increase in positivity (1575%; 63/4), average sentiment improved considerably (0.098; $P<.001$). Nevertheless, discussion volume was remarkably low (19% index; 120/631.1). No single topic stands out; discussion instead touches on themes ranging from Ms Obokata's mentoring to the role of the media.

- Vacanti et al release new STAP protocol; addition of adenosine triphosphate (ATP) now asserted to be key (September 3) [70].
- Dr Endo publishes a report suggesting that STAP cells may have been embryonic stem cells (September 21) [71].

October 2014: Waseda University's Threat to Revoke PhD Draws Mixed Response

In October 2014, average sentiment decreased significantly, but remained overall positive (0.047; $P=.23$). Despite low overall discussion volume (21.5% index; 136/631.1), positivity was comparably high (1267%; 38/3). Discussion volume, however, remained overall low. Discussion mostly focuses on Waseda University's demand that Ms Obokata correct her dissertation [72]; some see this as an “opportunity.”

November 2014: Attention Focused Elsewhere as STAP Retrial Draws to a Close

In November 2014, discussion volume fell to an all-time low (11.9% index; 75/631.1). Despite above average positivity (150%; 3/2), average sentiment continued to trend negative (-0.023 ; $P=.06$). Discussion centers on speculation regarding the STAP replication efforts. No other coherent themes emerge.

December 2014: Riken Unexpectedly Halts STAP Retrial; Obokata Resigns

In December 2014, average sentiment decreased slightly, but not to a significant degree (-0.042 ; $P>.99$). However, a large increase in discussion volume (173% index; 1092/631.1) precipitated a sharp drop in positivity (54.2%; 71/131). Ms Obokata's failure to replicate STAP, along with her abrupt resignation attracts the most discussion; “disgraced” is the now most commonly used term to describe Ms Obokata.

- Riken halts STAP verification experiments, announcing them to have failed [73].
- Ms Obokata resigns from her position at Riken [74].

January 2015: Obokata Accused of Stealing Materials; Criminal Charges Threatened

In January 2015, the downward trend in average sentiment continued—albeit not to a significant degree (-0.08 ; $P=.10$). Discussion volume (33.1% index; 209/631.1) and positivity were both subpar (23%; 6/26). Discussion focuses on allegations of criminal wrongdoing.

February 2015: Riken Openly Discusses Punitive Measures; Guardian Piece Attracts Mixed Response

In February 2015, a large uptick in discussion (163.8% index; 1034/631.1) drove a precipitous decline in average sentiment (-0.26 ; $P<.001$) and positivity (1.0%; 8/766). Discussion focuses on the following:

- Riken’s announcement of penalties related to the STAP research and publication.
- Riken’s public announcement of plans to pursue criminal charges against Ms Obokata [75].
- The Guardian piece, “What pushes scientists to lie,” [76] is highly cited during this period.

March 2015: Riken Announces Intent Not to Sue Obokata

In March 2015, average sentiment improves slightly but remains extremely negative (-0.18 ; $P<.001$). Positivity remains low (5%; 4/78) amid low overall discussion volume (29.6% index; 187/631.1). Despite a few supportive messages, discussion overwhelmingly highlights Ms Obokata’s continuing woes.

- Riken’s decision not to sue Ms Obokata [77] garners mixed attention.
- However, reported demands that she return publication-related expenses [78] draws derision.

Sentiment Surrounding Various, Individual Stakeholders

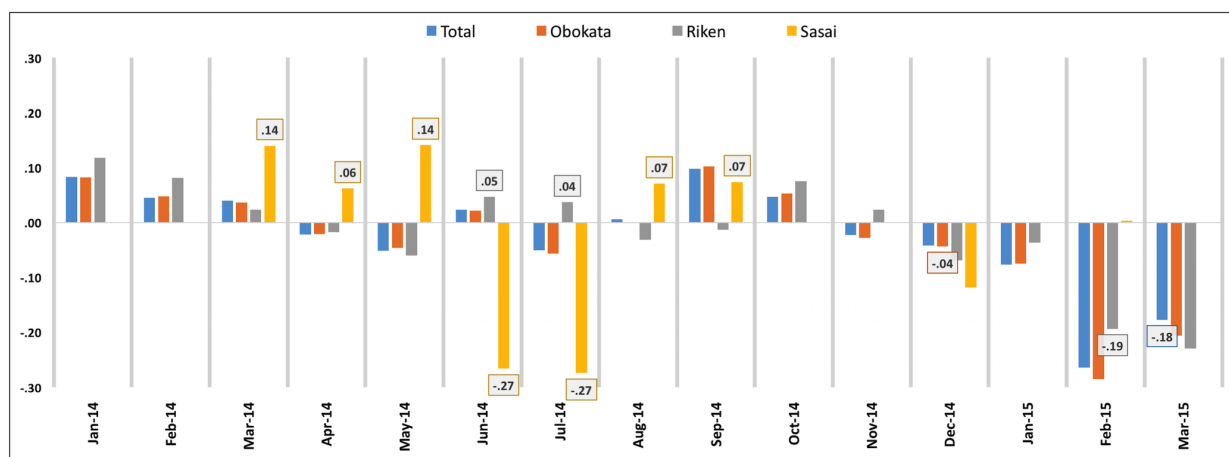
Significant differences were found with respect to the sentiment surrounding various parties. Tweet data were aggregated

according to whether Ms “Obokata,” Dr “Sasai,” or (inclusive) the “Riken” institute were mentioned. Overall, sentiment surrounding Ms Obokata and the Riken institute was found to be consistent with broader trends (-0.04 and -0.03 , respectively; no significant difference). However, sentiment toward Dr Sasai was found to be significantly more positive overall (0.03; $P<.04$), with visible trends coinciding with relevant events; but these results were not statistically significant (Figure 2).

Dr Sasai initially received minimum attention. However, once allegations of misconduct began to emerge, Dr Sasai’s continued, public support of Ms Obokata becomes increasingly associated with discussion that was significantly more favorable (0.14, 0.06, and 0.14 respectively, in March 2014, April 2014, and May 2014; $P<.001$, $P=.05$, $P=.02$ respectively). Indeed, discussion mentioning Dr Sasai remained positive even as sentiment toward Ms Obokata (-0.02 and -0.05 respectively, in April 2014 and May 2014) and Riken (-0.02 and -0.06 respectively, in April 2014 and May 2014) became decidedly negative. Incident to and just before retraction of the STAP papers, discussion increasingly turned to the role of Ms Obokata’s mentors and institutional culpability. Accordingly, sentiment associated with mentions of Dr Sasai fell precipitously during this period (-0.27 in both June 2014 and July 2014; $P=.09$ and $P<.001$, respectively). By contrast, sentiment associated with Riken was positive during the same timeframe (0.05 and 0.04 respectively, in June 2014 and July 2014; $P=.01$ and $P=.62$, respectively; Figure 3).

These trends, however, reversed in August and September of that year. Coincident with and following the tragedy surrounding Dr Sasai, the sentiment associated with mentions of Dr Sasai became positive (0.07 and 0.07 respectively, in August 2014 and September 2014; $P=.04$ for both), whereas sentiment associated with Riken followed the opposite trend (-0.03 and -0.01 respectively, in August 2014 and September 2014; $P<.001$ and $P=.03$, respectively).

Figure 3. Sentiment comparison for key actors. Month-to-month sentiment for key figures and entities corresponds with associated timeline events. Month-to-month sentiment scores were independently aggregated for Tweets mentioning Ms Obokata, Dr Sasai, or Riken. Data labels shown where mean differences are significant versus total.



Content Analysis

A simplified “grounded theory” approach [79] was used to examine the generalizable themes that were most commonly expressed within this dataset. For this purpose, entities specific to this case were omitted. Each Tweet was manually categorized according to major topic or polarity-driving theme. Coding was based on an analysis of the terms and concepts within a given Tweet that correspond with the degree and direction of polarity assigned by the SentiWordNet algorithm. Categories were added, merged, and/or eliminated progressively, as Tweets were reviewed. Higher-level codes were manually assigned to similarity clusters in an iterative fashion, to uncover overarching themes. Tweets were allowed to be assigned to multiple coding classes. The primary themes thus found to be driving polarity were the following: (1) “unfocused negativity—outrage or mockery,” (2) “cynicism toward academic science,” (3) “defense of academic science,” and (4) “miscellaneous conspiracy theories.”

Discussion

Principal Findings

Only a few studies have covered the publicity of science using SNS [80,81]; and thus, very little is currently known about how public sentiment may change in response to reported cases of scientific misconduct. A study conducted on behalf of the UK House of Lords found that, although public interest in science is high, “a culture of...secrecy...invites suspicion” [82]. However, whether such suspicions are based on rational or justifiable criticism is still a matter of controversy [83]. In addition, little is currently known as to whether and how news related to scientific misconduct is relevant to or impacts public opinion on a broad scale.

This study found that STAP-related discussion volume varied significantly month-to-month, coincident with new events. Furthermore, we found that month-to-month sentiment was generally neutral or of mixed composition, tending to skew negative when polarized. This is consistent with previous findings concerning the characteristics of public sentiment as expressed on Twitter [84]. In this study, it was noted that such increases in polarity generally corresponded with large increases in overall volume, that is, major events attracted more attention, most of which tended negative. It is interesting to note that such increases in volume and corresponding negativity were generally short-lived; positivity, on the other hand, tended to be much less prone to such fluctuation.

In addition, this analysis found that sentiment surrounding various stakeholders differed significantly with respect to specific events. Of particular note is the sentiment surrounding Dr Sasai, the researcher whose tragic fate was found to correspond with an increase in positive sentiment. The relationship between the death of a key stakeholder in a public crisis and subsequent improvement of the public mood—from criticism to sympathy—has been covered in the Japanese literature in the 1980s [85] and also in a recent Time article reporting Dr Sasai’s death [86]. And in this study, an analysis of public sentiment, expressed via social media, was able to detect and observe this effect.

The results presented here provide an important case study for understanding the impact of scientific misconduct on public sentiment. The coverage received by the STAP cell case can be attributed to many factors, but the instrumentality of social media cannot be ignored. Although this manuscript was undergoing review, a related study was published that provided rudimentary analysis of the print and social media coverage of the STAP cell case in Japan [87]. This study confirmed that print coverage tended to lag behind social media coverage—a finding which supports other reports suggesting that news initiated and driven by public interest may indeed influence public opinion [9,88], or at least mainstream reporting. Here, we have presented a more robust perspective by including all major timeline events in the STAP cell case and by employing a well-established, objective method for sentiment analysis. This research is the first to establish the patterns according to which public opinion evolves in response to reports of scientific misconduct in the popular press.

The question still remains, however, whether sentiment expressed on Twitter regarding future cases of misconduct will accurately reflect overall public sentiment. Prior research has demonstrated that the Twitter medium most effectively influences or reflects public response with respect to high-volume events or crises [89]. And some research has even suggested superiority to traditional polling [6] in some respects. For this study, the case covered was considered to be one of the largest and most impactful cases of scientific misconduct in recent memory [27]. Moreover, this case is the first and only major scientific misconduct case to occur in an era where social media coverage and documentation is so ubiquitous. Indeed, much of the initial concerns and evidence regarding STAP originated in social media [90-93]. And recent research and commentary on evolving mass media trends suggests that social media is increasingly becoming both an initiator and driver of public attention and news cycles [88]. Consequently, we expect the impact of future cases of major scientific misconduct to be generalizable, using social media metrics in the fashion demonstrated here [94,95].

Limitations

The text- and sentiment-analysis procedures employed in this study were robust and well-validated. The reported metrics are limited by the analytical processes that were used to derive them from the text. In this case, SentiWordNet was used to obtain sentiment scores. Reported sentiment scores are thus limited by the accuracy and precision of the SentiWordNet database with respective material covered. In addition, volume estimations were based on and limited by the distribution characteristics of the sample obtained from the data provider. Furthermore, the reported metrics are estimations, as is the case with all sampling-based analytical approaches. That having been said, the analytical and data retrieval methods used are well established and have been verified to be sufficiently robust for such analyses [92,94]. Future studies would benefit from larger sample sizes and more precise sentiment estimation methods; however, based on previous studies, improvements are expected to be marginal.

Conclusions

This study represents the first objective analysis of public response to a major case of scientific misconduct. This study observes and tracks changes in public sentiment over a 15-month sequence of events associated with the STAP cell case, which was one of the most publicized cases of major scientific misconduct in recent memory. Here, we demonstrated that public response to this particular case tended to be generally neutral or of mixed composition, particularly during times of lower public attention. This was observed in the large majority of months covered in this study. Also observed was that sentiment tended to skew negative as discussion polarity and volume increased. These findings are generally consistent with those observed in the literature with respect to major events across a wide range of topics, including entertainment, sports, business, politics, and even natural disasters. These findings support the notion that changes in public sentiment toward any major event—cases of scientific misconduct included—might

be as much a function of the attention received as it is a function of the theme or merits of any specific case. As the saying goes, “no news is good news”—and this study demonstrates this quite clearly. Once the STAP story becomes tainted by allegations of misconduct, increases in public attention—driven mostly by the public relations (PR) efforts of the respective actors—consistently corresponded with increases in overall negativity. The only event that broke this trend was one of the only events not staged for publicity—the apparent suicide of a key stakeholder. Here, we observed a clear and significant positive shift in overall sentiment; however, this was also accompanied by a notable subsequent decrease in volume. Overall, these results strongly suggest that, in cases of research misconduct, public opinion—and by extension, public policy—is likely to be more influenced by negative-leaning news and reporting. Academic researchers, policy makers, and those with associated interests are advised to carefully consider the implications.

Conflicts of Interest

None declared.

Multimedia Appendix 1

A summary timeline.

[[JPG File, 800KB - publichealth_v3i2e21_app1.jpg](#)]

References

1. Becker H, Naaman M, Gravano L. Columbia University Libraries. 2011. Beyond Trending Topics: Real-World Event Identification on Twitter URL: <https://academiccommons.columbia.edu/catalog/ac:135415> [accessed 2017-03-22] [WebCite Cache ID 6p9liDYJr]
2. Phuvipadawat S, Murata T. Breaking news detection and tracking in Twitter. 2010 Presented at: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology; 2010; New York. [doi: [10.1109/WI-IAT.2010.205](https://doi.org/10.1109/WI-IAT.2010.205)]
3. Kent ML. Using social media dialogically: public relations role in reviving democracy. *Public Relat Rev* 2013 Nov;39(4):337-345. [doi: [10.1016/j.pubrev.2013.07.024](https://doi.org/10.1016/j.pubrev.2013.07.024)]
4. Auger GA. Fostering democracy through social media: evaluating diametrically opposed nonprofit advocacy organizations' use of Facebook, Twitter, and YouTube. *Public Relat Rev* 2013 Nov;39(4):369-376. [doi: [10.1016/j.pubrev.2013.07.013](https://doi.org/10.1016/j.pubrev.2013.07.013)]
5. Berthon PR, Pitt LF, Plangger K, Shapiro D. Marketing meets web 2.0, social media, and creative consumers: Implications for international marketing strategy. *Bus Horiz* 2012 May;55(3):261-271. [doi: [10.1016/j.bushor.2012.01.007](https://doi.org/10.1016/j.bushor.2012.01.007)]
6. Murphy J, Link MW, Childs JH, Tesfaye CL, Dean E, Stern M, et al. Social media in public opinion research: executive summary of the Aapor task force on emerging technologies in public opinion research. *Public Opin Q* 2014 Nov 25;78(4):788-794. [doi: [10.1093/poq/nfu053](https://doi.org/10.1093/poq/nfu053)]
7. Culotta A, Kumar N, Cutler J. Illinois Institute of Technology. 2015. Predicting the demographics of Twitter users from website traffic data URL: <http://www.cs.iit.edu/~culotta/pubs/culotta15predicting.pdf> [accessed 2016-05-09] [WebCite Cache ID 6hMos6s9O]
8. Mislove A, Lehmann S, Ahn Y, Onnela J, Rosenquist J. Understanding the demographics of Twitter users. 2011 Presented at: Proceedings of the International AAAI Conference on Web and Social Media; 2011; Menlo Park URL: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2816>
9. O'Connor B, Balasubramanyan R, Routledge B, Smith N. From tweets to polls: linking text sentiment to public opinion time series. 2010 Presented at: the International AAAI Conference on Weblogs and Social Media; 11(2); 2010; Washington DC p. 122-129 URL: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1536>
10. Vis F. Twitter as a reporting tool for breaking news. *Digit Journal* 2013 Feb;1(1):27-47. [doi: [10.1080/21670811.2012.741316](https://doi.org/10.1080/21670811.2012.741316)]
11. Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? 2013 Presented at: WWW '10 Proceedings of the 19th international conference on World wide web; April 26 - 30, 2010; Raleigh, NC URL: <http://dl.acm.org/citation.cfm?id=1772751>

12. Petrovic S, Osborne M, McCreddie R, Macdonald C, Ounis I. Can Twitter replace newswire for breaking news? In: International AAAI Conference on Web and Social Media, North America, Jun. 2013 Presented at: Seventh International AAAI Conference on Weblogs and Social Media; July 8– 10, 2013; Cambridge URL: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6066>
13. Killough K. University of Windsor. 2008. Changing the narrative: how media sensationalism affects public policy and how activists can use it to their advantage URL: <http://scholar.uwindsor.ca/etd/4628/> [accessed 2017-04-05] [WebCite Cache ID 6pUTuC3kt]
14. Ahern KR, Sosyura D. Rumor has it: sensationalism in financial media. *Rev Financ Stud* 2015 Jan 24;28(7):2050-2093. [doi: [10.1093/rfs/hhv006](https://doi.org/10.1093/rfs/hhv006)]
15. Johnston CD, Bartels BL. Sensationalism and sobriety differential media exposure and attitudes toward American courts. *Public Opin Q* 2010 Feb 10;74(2):260-285. [doi: [10.1093/poq/nfp096](https://doi.org/10.1093/poq/nfp096)]
16. Molek-Kozakowska K. Towards a Pragma-linguistic framework for the study of sensationalism in news headlines. *Discourse & Communication* 2013;7(2):173-197.
17. Ransohoff DF, Ransohoff RM. Sensationalism in the media: when scientists and journalists may be complicit collaborators. *Eff Clin Pract* 2001;4(4):185-188. [Medline: [11525108](https://pubmed.ncbi.nlm.nih.gov/11525108/)]
18. Lilford RJ, Braunholtz D. For debate: the statistical basis of public policy: a paradigm shift is overdue. *BMJ* 1996 Sep 07;313(7057):603-607. [doi: [10.1136/bmj.313.7057.603](https://doi.org/10.1136/bmj.313.7057.603)]
19. Otten AL. The influence of the mass media on health policy. *Health Affairs* 1992 Nov 01;11(4):111-118. [doi: [10.1377/hlthaff.11.4.111](https://doi.org/10.1377/hlthaff.11.4.111)]
20. Miller S. Public understanding of science at the crossroads. *Public Understand Sci* 2001 Jan 26;10(1):115-120. [doi: [10.1088/0963-6625/10/1/308](https://doi.org/10.1088/0963-6625/10/1/308)]
21. Stern AM, Casadevall A, Steen RG, Fang FC. Financial costs and personal consequences of research misconduct resulting in retracted publications. *eLife Internet* 2014;3:e02956. [doi: [10.7554/eLife.02956](https://doi.org/10.7554/eLife.02956)]
22. Michalek AM, Hutson AD, Wicher CP, Trump DL. The costs and underappreciated consequences of research misconduct: a case study. *PLoS Med* 2010;7(8):e1000318 [FREE Full text] [doi: [10.1371/journal.pmed.1000318](https://doi.org/10.1371/journal.pmed.1000318)] [Medline: [20808955](https://pubmed.ncbi.nlm.nih.gov/20808955/)]
23. Suda M. Netsuzo no kagakusha : Sutappu saibo jiken (Scientific Misconduct: the STAP Scandal). Tokyo, Japan: Bungeishunju Ltd; 2015.
24. Thelwall M, Buckley K, Paltoglou G. Sentiment in Twitter events. *J Am Soc Inf Sci* 2010 Dec 06;62(2):406-418. [doi: [10.1002/asi.21462](https://doi.org/10.1002/asi.21462)]
25. Riken. 2014. Riken Press Releases URL: <http://www.riken.jp/en/pr/press/2014/> [accessed 2017-03-22] [WebCite Cache ID 6p9JzaG0v]
26. Riken. 2015. Riken press releases URL: <http://www.riken.jp/en/pr/press/2015/> [accessed 2016-04-28] [WebCite Cache ID 6hMpkQhGp]
27. Cyranoski D. Collateral damage: how a case of misconduct brought a leading Japanese biology institute to its knees. *Nature* 2015 Apr 30;520(7549):600-603. [doi: [10.1038/520600a](https://doi.org/10.1038/520600a)] [Medline: [25925457](https://pubmed.ncbi.nlm.nih.gov/25925457/)]
28. Thomson H. New Scientist Internet. 2014 Jul 2. Acid-bath stem cell papers are finally retracted URL: <https://www.newscientist.com/article/dn25831-acid-bath-stem-cell-papers-are-finally-retracted/> [accessed 2016-05-09] [WebCite Cache ID 6hMpuID4s]
29. Lancaster C. The acid test for biological science: STAP cells, trust, and replication. *Sci Eng Ethics* 2016 Feb;22(1):147-167. [doi: [10.1007/s11948-015-9628-2](https://doi.org/10.1007/s11948-015-9628-2)] [Medline: [25649070](https://pubmed.ncbi.nlm.nih.gov/25649070/)]
30. Topsy. Twitter search, monitoring, & analytics URL: <https://web.archive.org/web/20150331010859/http://topsy.com/> [accessed 2016-05-09] [WebCite Cache ID 6hMq4Mame]
31. Twitter. Privacy Policy URL: <https://twitter.com/privacy?lang=en> [accessed 2016-05-16] [WebCite Cache ID 6hXKnUAGV]
32. Wilkinson D, Thelwall M. Researching personal information on the public web: methods and ethics. *Soc Sci Comput Rev* 2010 Aug 17;29(4):387-401. [doi: [10.1177/0894439310378979](https://doi.org/10.1177/0894439310378979)]
33. Russell Neuman W, Guggenheim L, Mo Jang S, Bae SY. The dynamics of public attention: agenda-setting theory meets big data. *J Commun* 2014 Mar 26;64(2):193-214. [doi: [10.1111/jcom.12088](https://doi.org/10.1111/jcom.12088)]
34. RStudio. URL: <https://www.rstudio.com/> [accessed 2016-05-09] [WebCite Cache ID 6hMqEqPH3]
35. Liew WT, Adhitya A, Srinivasan R. Sustainability trends in the process industries: a text mining-based analysis. *Comput Ind* 2014 Apr;65(3):393-400. [doi: [10.1016/j.compind.2014.01.004](https://doi.org/10.1016/j.compind.2014.01.004)]
36. Guru D, Suhil M. A novel term_class relevance measure for text categorization. *Procedia Computer Science* 2015;45:13-22. [doi: [10.1016/j.procs.2015.03.074](https://doi.org/10.1016/j.procs.2015.03.074)]
37. Abilhoa WD, de Castro LN. A keyword extraction method from Twitter messages represented as graphs. *Appl Math Comput* 2014 Aug;240:308-325. [doi: [10.1016/j.amc.2014.04.090](https://doi.org/10.1016/j.amc.2014.04.090)]
38. Gebre B, Zampieri M, Wittenburg P, Heskes T. Proceedings of the Eighth Workshop on Innovative Use of NLP for Building Educational Applications. 2013. Improving native language identification with TF-IDF weighting URL: <http://pubman.mpdl.mpg.de/pubman/faces/viewItemOverviewPage.jsp?itemId=escidoc:1740046:12> [accessed 2015-11-10] [WebCite Cache ID 6k1LUqkWy]

39. Srivastava AN, Sahami M. Text Mining: Classification, Clustering, and Applications. London: Chapman & Hall/CRC; 2009.
40. Baccianella S, Esuli A, Sebastiani F. NMIS. 2010. SentiWordNet 3: an enhanced lexical resource for sentiment analysis and opinion mining URL: <http://nmis.isti.cnr.it/sebastiani/Publications/LREC10.pdf> [accessed 2016-05-09] [WebCite Cache ID 6hMqNghxE]
41. Annett M, Kondrak G. A comparison of sentiment analysis techniques: polarizing movie blogs. In: Advances in Artificial Intelligence. New York: Springer; 2008:25-35.
42. Bhadane C, Dalal H, Doshi H. Sentiment analysis: measuring opinions. Procedia Computer Science 2015;45:808-814. [doi: [10.1016/j.procs.2015.03.159](https://doi.org/10.1016/j.procs.2015.03.159)]
43. Thelwall M, Buckley K. Topic-based sentiment analysis for the social web: the role of mood and issue-related words. J Am Soc Inf Sci Tec 2013 Jun 19;64(8):1608-1617. [doi: [10.1002/asi.22872](https://doi.org/10.1002/asi.22872)]
44. Ben-Hur A, Weston J. A user's guide to support vector machines. Methods Mol Biol 2010;609:223-239. [doi: [10.1007/978-1-60327-241-4_13](https://doi.org/10.1007/978-1-60327-241-4_13)] [Medline: [20221922](https://pubmed.ncbi.nlm.nih.gov/20221922/)]
45. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. Mach Learn 2002;46(1-3):389-422. [doi: [10.1023/A:1012487302797](https://doi.org/10.1023/A:1012487302797)]
46. Salazar D, Vélez J, Salazar J. Comparison between SVM and logistic regression: which one is better to discriminate? Rev Colomb Estad 2012;35(2):223-237 [FREE Full text]
47. Martin-Barragan B, Lillo R, Romo J. Interpretable support vector machines for functional data. Eur J Oper Res 2014 Jan;232(1):146-155. [doi: [10.1016/j.ejor.2012.08.017](https://doi.org/10.1016/j.ejor.2012.08.017)]
48. Obokata H, Sasai Y, Niwa H, Kadota M, Andrabi M, Takata N, et al. Bidirectional developmental potential in reprogrammed cells with acquired pluripotency. Nature 2014 Jan 30;505(7485):676-680. [doi: [10.1038/nature12969](https://doi.org/10.1038/nature12969)] [Medline: [24476891](https://pubmed.ncbi.nlm.nih.gov/24476891/)]
49. Obokata H, Wakayama T, Sasai Y, Kojima K, Vacanti MP, Niwa H, et al. Stimulus-triggered fate conversion of somatic cells into pluripotency. Nature 2014 Jan 30;505(7485):641-647. [doi: [10.1038/nature12968](https://doi.org/10.1038/nature12968)] [Medline: [24476887](https://pubmed.ncbi.nlm.nih.gov/24476887/)]
50. Riken. 2014 Jan 29. Stress turns ordinary cells pluripotent URL: http://www.riken.jp/~media/riken/pr/topics/2014/20140702_1/140702_1_1-en.pdf [accessed 2016-05-09] [WebCite Cache ID 6hMqR8jQj]
51. PubPeer. 2014 Jan 29. Stimulus-triggered fate conversion of somatic cells into pluripotency URL: <https://pubpeer.com/publications/24476887> [accessed 2016-05-09] [WebCite Cache ID 6hMqUgVgg]
52. Thomson H. New Scientist. 2014 Feb 5. Extraordinary stem cell method tested in human tissue URL: <https://www.newscientist.com/article/dn25004-extraordinary-stem-cell-method-tested-in-human-tissue/> [accessed 2016-05-13] [WebCite Cache ID 6hSripS05]
53. Cyranoski D. Nature Internet. 2014 Feb 17. Acid-bath stem-cell study under investigation URL: <http://www.nature.com/doifinder/10.1038/nature.2014.14738> [accessed 2017-03-22] [WebCite Cache ID 6p9Kkctdi]
54. Niwa H, Obokata H, Sasai Y, Niwa H. Essential technical tips for STAP cell conversion culture from somatic cells. Protocol Exchange 2014 Mar 5. [doi: [10.1038/protex.2014.008](https://doi.org/10.1038/protex.2014.008)]
55. Martin A, Naik G. Wall Str J. 2014 Mar 10. Japanese institute weighs retracting stem-cell studies URL: <http://www.wsj.com/news/articles/SB10001424052702304020104579430600241800532> [accessed 2016-05-09] [WebCite Cache ID 6hMqZMHdz]
56. Riken. 2014 Mar 14. Interim report on the investigation of the Obokata URL: http://www.riken.jp/en/pr/press/2014/20140314_1/ [accessed 2016-05-05] [WebCite Cache ID 6hMqdxjDc]
57. Riken. 2014 Apr 1. Report on STAP cell research paper investigation URL: http://www.riken.jp/en/pr/press/2014/20140401_2/ [accessed 2016-05-05] [WebCite Cache ID 6hMqguqWb]
58. Cyranoski D. Biologist defiant over stem-cell method. Nature 2014 Apr 17;508(7496):299. [doi: [10.1038/508299a](https://doi.org/10.1038/508299a)] [Medline: [24740046](https://pubmed.ncbi.nlm.nih.gov/24740046/)]
59. Editorial. Agency for change. Nature 2014 May 1;509(7498):8. [doi: [10.1038/509008a](https://doi.org/10.1038/509008a)] [Medline: [24784180](https://pubmed.ncbi.nlm.nih.gov/24784180/)]
60. Japan Times. 2014 May 28. Embattled STAP cell scientist Obokata to retract research paper URL: <http://www.japantimes.co.jp/news/2014/05/28/national/science-health/embattled-stap-cell-scientist-obokata-retract-research-paper> [accessed 2016-05-05] [WebCite Cache ID 6hMqIQKSq]
61. Cyranoski D. Nature. 2014 Jun 17. Gene tests suggest acid-bath stem cells never existed URL: <http://www.nature.com/doifinder/10.1038/nature.2014.15425> [accessed 2017-03-22] [WebCite Cache ID 6p9KvK5Q3]
62. Normile D. Science. 2014 Jun 4. Key researcher agrees to retract both disputed stem cell papers internet URL: <http://www.sciencemag.org/news/2014/06/key-researcher-agrees-retract-both-disputed-stem-cell-papers> [accessed 2016-05-09] [WebCite Cache ID 6hMqr32gq]
63. Normile D. Science. 2014 Jun 12. Japanese stem cell debacle could bring down center Internet URL: <http://www.sciencemag.org/news/2014/06/japanese-stem-cell-debacle-could-bring-down-center> [accessed 2016-05-09] [WebCite Cache ID 6hMqv5r2U]
64. Obokata H, Wakayama T, Sasai Y, Kojima K, Vacanti MP, Niwa H, et al. Retraction: stimulus-triggered fate conversion of somatic cells into pluripotency. Nature 2014 Jul 3;511(7507):112. [doi: [10.1038/nature13598](https://doi.org/10.1038/nature13598)] [Medline: [24990753](https://pubmed.ncbi.nlm.nih.gov/24990753/)]
65. Obokata H, Sasai Y, Niwa H, Kadota M, Andrabi M, Takata N, et al. Retraction: bidirectional developmental potential in reprogrammed cells with acquired pluripotency. Nature 2014 Jul 3;511(7507):112. [doi: [10.1038/nature13599](https://doi.org/10.1038/nature13599)] [Medline: [24990752](https://pubmed.ncbi.nlm.nih.gov/24990752/)]

66. Japan Times. 2014 Jul 24. Obokata injured while fleeing pack of NHK reporters, lawyer says URL: <http://www.japantimes.co.jp/news/2014/07/24/national/obokata-injured-fleeing-pack-nhk-reporters-lawyer-says/> [accessed 2016-05-05] [WebCite Cache ID 6hMqyjuHv]
67. Retraction Watch. 2014 Apr 8. Waseda University checking dissertations for plagiarism in wake of STAP stem cell misconduct finding Internet URL: <http://retractionwatch.com/2014/04/08/waseda-university-checking-dissertations-for-plagiarism-in-wake-of-stap-stem-cell-misconduct-finding/> [accessed 2016-05-09] [WebCite Cache ID 6hMr3Pth]
68. Cyranoski D. Nature. 2014 Aug 13. Stem-cell pioneer blamed media “bashing” in suicide note URL: <http://www.nature.com/news/stem-cell-pioneer-blamed-media-bashing-in-suicide-note-1.15715> [accessed 2017-03-22] [WebCite Cache ID 6p9LOpHab]
69. Science. 2014 Aug 27. In Japan, official effort to replicate STAP stem cells comes up empty URL: <http://www.sciencemag.org/news/2014/08/japan-official-effort-replicate-stap-stem-cells-comes-empty> [accessed 2016-05-13] [WebCite Cache ID 6hSuZGbBV]
70. Vacanti CA, Kojima K. Web.archive. 2014 Sep 14. Revised STAP cell protocol URL: https://web.archive.org/web/20141219103753/https://research.bwhanesthesia.org/site_assets/51520d191eea6679ce000001/cterm/Revised_STAP_protocol-28bcd7e61d02a23624eb590717e241fe.pdf [accessed 2016-05-13] [WebCite Cache ID 6hSvDmaPk]
71. Endo TA. Quality control method for RNA-seq using single nucleotide polymorphism allele frequency. Genes Cells 2014 Nov;19(11):821-829 [FREE Full text] [doi: [10.1111/gtc.12178](https://doi.org/10.1111/gtc.12178)] [Medline: [25243705](https://pubmed.ncbi.nlm.nih.gov/25243705/)]
72. Martin A. WSJ. 2014 Oct 7. Japanese University may strip stem-cell scientist of doctorate URL: <http://www.wsj.com/articles/japanese-university-may-strip-stem-cell-scientist-of-doctorate-1412683955> [accessed 2016-05-09] [WebCite Cache ID 6hMr7R0eA]
73. Lies E. Reuters. 2014 Dec 18. Disgraced Japan researcher fails to replicate 'game changing' stem cell results URL: <http://www.reuters.com/article/us-japan-science-scandal-idUSKBN0JXOB820141219> [accessed 2016-05-09] [WebCite Cache ID 6hMrE1b2a]
74. Gallagher J. BBC News Internet. 2014 Dec 19. Stem cell scandal scientist Haruko Obokata resigns URL: <http://www.bbc.com/news/health-30534674> [accessed 2016-05-09] [WebCite Cache ID 6hMrLeaAI]
75. Normile D. Science. 2015 Feb 10. RIKEN announces penalties related to stem cell fiasco URL: <http://www.sciencemag.org/news/2015/02/riken-announces-penalties-related-stem-cell-fiasco> [accessed 2016-05-09] [WebCite Cache ID 6hMrR9zff]
76. Rasko J, Power C. The Guardian. 2015 Feb 18. What pushes scientists to lie? the disturbing but familiar story of Haruko Obokata URL: <https://www.theguardian.com/science/2015/feb/18/haruko-obokata-stap-cells-controversy-scientists-lie> [accessed 2016-04-25] [WebCite Cache ID 6hMrVJobf]
77. Japan Times. 2015 Mar 17. Riken won't sue Obokata over STAP cell scandal URL: <http://www.japantimes.co.jp/news/2015/03/17/national/crime-legal/riken-wont-sue-obokata-over-stap-cell-scandal/> [accessed 2016-05-05] [WebCite Cache ID 6hMskaqFM]
78. Nose T. Web.archive. 2015 Mar 19. Riken to get Obokata to return portion of research funds over stem cell scandal URL: https://web.archive.org/web/20151213214710/http://ajw.asahi.com/article/behind_news/social_affairs/AJ201503190077 [accessed 2016-05-09] [WebCite Cache ID 6hMsorkrW]
79. Strauss A, Corbin JM. Grounded Theory in Practice. London: SAGE; 1997.
80. Letierce J, Passant A, Breslin J, Decker S. Understanding how Twitter is used to spread scientific messages. 2010 Presented at: Proceedings of the WebSci10: Extending the Frontiers of Society On-Line; 2010; Raleigh, NC URL: <http://journal.webscience.org/314/>
81. Shuai X, Pepe A, Bollen J. How the scientific community reacts to newly submitted preprints: article downloads, Twitter mentions, and citations. PLoS One 2012;7(11):e47523 [FREE Full text] [doi: [10.1371/journal.pone.0047523](https://doi.org/10.1371/journal.pone.0047523)] [Medline: [23133597](https://pubmed.ncbi.nlm.nih.gov/23133597/)]
82. Parliament.uk. 2000 Feb 23. Select committee on science and technology third report URL: <http://www.publications.parliament.uk/pa/ld199900/ldselect/ldsctech/38/3802.htm> [WebCite Cache ID 6hXMpQ6gu]
83. Miller S. Public understanding of science at the crossroads. Public Understand Sci 2001 Jan 26;10(1):115-120. [doi: [10.1088/0963-6625/10/1/308](https://doi.org/10.1088/0963-6625/10/1/308)]
84. Wilkinson D, Thelwall M. Trending Twitter topics in english: an international comparison. J Am Soc Inf Sci Tec 2012 Jul 02;63(8):1631-1646. [doi: [10.1002/asi.22713](https://doi.org/10.1002/asi.22713)]
85. Tsunoda F. Isshi taizai o shasu : Rikugun Daijin Anami Korechika [a biography of an imperial army general Korechika Anami]. Tokyo, Japan: Shinchosha; 1980.
86. Spitzer K. Time. 2014 Aug 8. Science scandal triggers suicide, soul-searching in Japan URL: <http://time.com/3091584/japan-yoshiki-sasai-stem-cells-suicide-haruko-obokata/> [accessed 2016-05-05] [WebCite Cache ID 6hMtGURWm]
87. Sugawara Y, Tanimoto T, Miyagawa S, Murakami M, Tsuya A, Tanaka A, et al. Scientific misconduct and social media: role of twitter in the stimulus triggered acquisition of pluripotency cells scandal. J Med Internet Res 2017 Feb 28;19(2):e57 [FREE Full text] [doi: [10.2196/jmir.6706](https://doi.org/10.2196/jmir.6706)] [Medline: [28246071](https://pubmed.ncbi.nlm.nih.gov/28246071/)]

88. Chung S, Shih C, Huang C. Can collective sentiment expressed on twitter predict political elections? 2011 Presented at: Twenty-Fifth AAAI Conf Artif Intell Internet; 2011; San Francisco URL: <http://www.aaai.org/ocs/index.php/AAAI/AAAI11/paper/view/3549> [doi: [10.1016/j.ultras.2016.09.020](https://doi.org/10.1016/j.ultras.2016.09.020)]
89. Bollen J, Pepe A, Mao H. Arxiv. 2009 Nov 8. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena URL: <http://arxiv.org/abs/0911.1583> [accessed 2017-03-22] [WebCite Cache ID 6p9XvP8FR]
90. PubPeer. 2014 Feb 5. Stimulus-triggered fate conversion of somatic cells into pluripotency URL: <https://pubpeer.com/publications/8B755710BADFE6FB0A848A44B70F7D#fb6119> [accessed 2016-05-09] [WebCite Cache ID 6hMqUgVgg]
91. Stapcell. Haruko Obokata, STAP stem cells URL: <http://stapcell.blogspot.com/> [accessed 2016-05-09] [WebCite Cache ID 6hMsyOFuB]
92. Retraction Watch. 2014 Sep 10. "Truly extraordinary," "simply not credible," "suspiciously sharp:" a STAP stem cell peer review report revealed URL: <http://retractionwatch.com/2014/09/10/truly-extraordinary-simply-not-credible-suspiciously-sharp-a-stap-stem-cell-peer-review-report-revealed/> [accessed 2016-05-09] [WebCite Cache ID 6hMt4KD7I]
93. Knoepfler lab. The Niche. 2014 Jan 29. Review of Obokata stress reprogramming Nature papers Internet URL: <http://www.ipscell.com/2014/01/review-of-obokata-stress-reprogramming-nature-papers/> [accessed 2016-05-09] [WebCite Cache ID 6hMtAaBr1]
94. Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? 2010 Presented at: Proceedings of the 19th international conference on World wide web; 2010; New York URL: <http://dl.acm.org/citation.cfm?id=1772751>
95. Schultz F, Utz S, Göritz A. Is the medium the message? perceptions of and reactions to crisis communication via twitter, blogs and traditional media. Public Relat Rev 2011 Mar;37(1):20-27. [doi: [10.1016/j.pubrev.2010.12.001](https://doi.org/10.1016/j.pubrev.2010.12.001)]

Abbreviations

ANOVA: analysis of variance
ATP: adenosine triphosphate
BBC: British Broadcasting Corporation
CNN: Cable News Network
HSD: honest significant difference
PR: public relations
SNS: social network services
STAP: stimulus-triggered acquisition of pluripotency
SVM: support vector machine
TF-IDF: term frequency-inverse document frequency

Edited by G Eysenbach; submitted 17.05.16; peer-reviewed by D Broniatowski, D Fanelli; comments to author 21.08.16; revised version received 11.11.16; accepted 07.03.17; published 20.04.17.

Please cite as:

Gayle A, Shimaoka M

Public Response to Scientific Misconduct: Assessing Changes in Public Sentiment Toward the Stimulus-Triggered Acquisition of Pluripotency (STAP) Cell Case via Twitter

JMIR Public Health Surveill 2017;3(2):e21

URL: <http://publichealth.jmir.org/2017/2/e21/>

doi: [10.2196/publichealth.5980](https://doi.org/10.2196/publichealth.5980)

PMID: [28428163](https://pubmed.ncbi.nlm.nih.gov/28428163/)

©Alberto Gayle, Motomu Shimaoka. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 20.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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

Zika in Twitter: Temporal Variations of Locations, Actors, and Concepts

Anthony Stefanidis¹, PhD; Emily Vraga², PhD; Georgios Lamprianidis¹, PhD; Jacek Radzikowski¹, MS Comp Sc; Paul L Delamater¹, PhD; Kathryn H Jacobsen³, MPH, PhD; Dieter Pfoser¹, PhD; Arie Croitoru¹, PhD; Andrew Crooks⁴, PhD

¹Center for Geospatial Intelligence, Department of Geography and Geoinformation Science, George Mason University, Fairfax, VA, United States

²Department of Communication, George Mason University, Fairfax, VA, United States

³Department of Global and Community Health, George Mason University, Fairfax, VA, United States

⁴Center for Geospatial Intelligence, Department of Computational and Data Sciences, George Mason University, Fairfax, VA, United States

Corresponding Author:

Anthony Stefanidis, PhD
Center for Geospatial Intelligence
Department of Geography and Geoinformation Science
George Mason University
4400 University Drive, MS 6C3
Fairfax, VA, 22030
United States
Phone: 1 7039931212
Fax: 1 7039939299
Email: demisn@gmail.com

Abstract

Background: The recent Zika outbreak witnessed the disease evolving from a regional health concern to a global epidemic. During this process, different communities across the globe became involved in Twitter, discussing the disease and key issues associated with it. This paper presents a study of this discussion in Twitter, at the nexus of location, actors, and concepts.

Objective: Our objective in this study was to demonstrate the significance of 3 types of events: location related, actor related, and concept related, for understanding how a public health emergency of international concern plays out in social media, and Twitter in particular. Accordingly, the study contributes to research efforts toward gaining insights on the mechanisms that drive participation, contributions, and interaction in this social media platform during a disease outbreak.

Methods: We collected 6,249,626 tweets referring to the Zika outbreak over a period of 12 weeks early in the outbreak (December 2015 through March 2016). We analyzed this data corpus in terms of its geographical footprint, the actors participating in the discourse, and emerging concepts associated with the issue. Data were visualized and evaluated with spatiotemporal and network analysis tools to capture the evolution of interest on the topic and to reveal connections between locations, actors, and concepts in the form of interaction networks.

Results: The spatiotemporal analysis of Twitter contributions reflects the spread of interest in Zika from its original hotspot in South America to North America and then across the globe. The Centers for Disease Control and World Health Organization had a prominent presence in social media discussions. Tweets about pregnancy and abortion increased as more information about this emerging infectious disease was presented to the public and public figures became involved in this.

Conclusions: The results of this study show the utility of analyzing temporal variations in the analytic triad of locations, actors, and concepts. This contributes to advancing our understanding of social media discourse during a public health emergency of international concern.

(*JMIR Public Health Surveill* 2017;3(2):e22) doi:[10.2196/publichealth.6925](https://doi.org/10.2196/publichealth.6925)

KEYWORDS

Zika virus; social media; Twitter messaging; geographic information systems

Introduction

The emergence of social media has presented the general public with a novel avenue to disseminate information, exchange views, and network. Given that over 2.3 billion people worldwide are currently active social media users [1], social media play a significant role in communicating news and opinions. When it comes to health communication, in particular [2], social media have been studied to support a broad spectrum of activities: predicting disease outbreaks by monitoring Twitter references to certain terms [3], devising effective communication campaigns [4,5], supporting behavior change interventions [6,7], and tracking the general public's views on a variety of issues such as vaccination policies [8]. However, the tools for discerning patterns in these social media discussions pertaining to health are still in their formative stages [8-11]. In this paper, we present a study of the recent discourse in Twitter regarding the Zika outbreak to demonstrate the significance of three types of events: (1) geographical events capturing the evolution of the narrative over time and across locations, (2) social media presence events capturing the impact of and interactions over time among key actors, and (3) concept events that capture the emergence and evolution of key concepts that frame this narrative. Combined, these 3 types of events are capturing the evolution and provide valuable insight of this public discourse process.

Studies of social media content for public health issues tend to address one or more of 3 dimensions of these contributions [8-13]:

1. Their geographical dimension, studying the locations of the participating communities,
2. Their social dimension, studying the acting participants (ie actors) in this exchange, and
3. Their linguistic dimension, studying trending patterns of concepts and associations among terms in the social media discourse.

To advance this emerging field of study, we need to improve our understanding of the structure and evolution mechanisms along these 3 dimensions. Our objective in this study was to demonstrate the significance of 3 types of events: location related, actor related, and concept related for understanding how a public health emergency of international concern plays out in social media, and Twitter in particular. To better articulate this approach to studying health discourse in social media, we use the discussion in Twitter related to the Zika virus outbreak of late 2015 and early 2016 as a test case. We present sample results of the analysis of its content along the above-mentioned 3 dimensions. More specifically, we show that:

- The spatiotemporal analysis of data reveals locational events, capturing the progression of the epidemic (and of the discussion about it) from a localized South and Central American issue to a global one
- The analysis of the Twitter announcements and presence of two key public health organizations, namely, the U.S. Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO), reveals events related to the social media presence of major actors; thereby

capturing their gravity relative to the overall structure of the social network of participants over time

- The analysis of the temporal variation of references to certain Zika-related health issues (such as pregnancy and microcephaly) captures concept events, as terms emerge and connect over time, to frame the social media narrative that emerges through this discourse.

We argue that, jointly, events along these 3 dimensions capture the emergence, convergence, and evolution of such health narratives, and support the systematic study of health-related discourse in social media.

Methods

Design and Data Collection

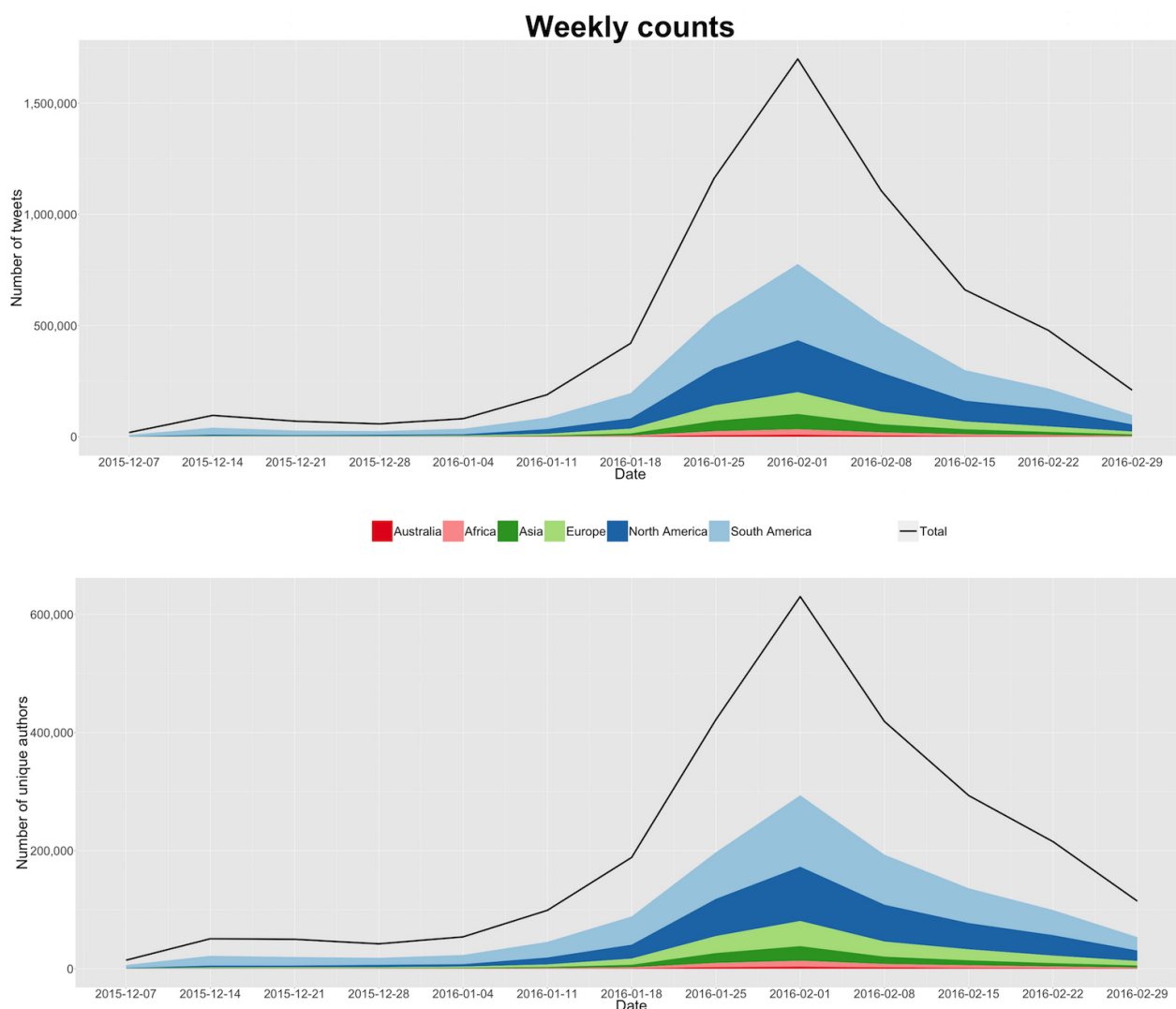
Beginning on December 11, 2015, we accessed the Twitter application program interface (API) to collect tweets referring to the “Zika virus” outbreak. We used the GeoSocial Gauge system prototype [14] to collect tweets mentioning the word “Zika” (and all variants of this string, such as #Zika, #ZikaVirus, etc) and the variant “zikkv” that is also used at times to refer to the Zika virus. The GeoSocial Gauge system harvests data from Twitter's streaming API, and retrieves the tweet content as well as accompanying metadata, including information such as user name, timestamp, and location. Earlier studies have shown the suitability of this sample of data for capturing the geographical footprint and the key actors associated with the social media narrative (eg, [15]).

Summary Data Characteristics

A total of 6,249,626 tweets about Zika were collected during the 12 weeks from December 11, 2015, through March 4, 2016. From among them, 48.16% (3,010,091), had a geographical location associated with them in their metadata, indicating where these tweets were posted from. Among these geolocated tweets, 1.12% were geolocated using coordinates provided directly by the user's device (eg, tweets posted using a handheld device with GPS location turned on), 2.14% were geolocated from Twitter using user IP information, and the remaining were geolocated through metadata content in the form of toponyms (ie, place names) from the users profile description. Such toponyms were then resolved into the finest available geographical granularity, using the GeoNames geographical database.

Figure 1 shows the temporal variation of Twitter traffic related to the search terms over our study period, with the data binned in weeks. Traffic grew progressively until it reached its peak during the first week of February, 2016, with a total of 1,698,883 tweets during the first 7 days of the months (882,727 of them geolocated). This corresponds to an average of 126,104 tweets per day during that week. In comparison, during the study period overall there were on average of 74,400 tweets daily. After that peak week, traffic started gradually dropping, approaching prepeak volume patterns. Contributions from South America were prominent throughout the study period. At the peak of Zika traffic (on February 3, 2016), we can identify in our data corpus that 630,142 unique Twitter users who were participating in this discussion, with 293,785 of them being geolocated.

Figure 1. Top: The number of tweets per week. The continuous black line shows the total volume of Twitter traffic mentioning our search terms during our study period. The portion of this traffic that is geolocated is color coded according to the continent of origin. Bottom: The number of unique users contributing to this discussion in the same period, color coded by their continent of origin.



Social Media Interaction as a Narrative Structure

Since its beginnings almost 50 years ago, the formal study of narrative has crossed many disciplines (such as sociolinguistics, anthropology, and psychology) and applications (eg, medicine, journalism, law); yet it still lacks a clear-cut definition [16]. Broadly defined, the concept of “narrative” is commonly used to refer to storytelling, typically as an orderly sequence of events [17]. Earlier studies addressed narrative in the context of spoken [18] and written [19] expression. A seminal early classification of narrative stories identified 5 distinct types of narratives: (1) official stories from authorities; (2) invented or adapted stories in which the narrative is refined as the public adopts and adapts stories; (3) firsthand stories in which individuals recount their experiences; (4) secondhand stories that are repetition of firsthand stories that have been heard or read; and (5) culturally common stories that reflect common perceptions emerging from the social environment rather than particular individuals [20].

Social media interaction, viewed as narrative, moves beyond individual articulations of past experiences toward encompassing collaboration and greater complexity. The nature of

contributions and interactions within social media platforms makes them align most closely with the fifth type of story, culturally common ones [20]. However, all 4 other types of narratives are also present in social media content, as users and authorities alike contribute content, views, and opinions. This makes social media narrative, particularly convoluted, and necessitates novel frameworks for analysis and understanding of them.

In a manner comparable with the triadic nature of Meade and Emch’s [21] triangle of human ecology, health narratives in social media have been shown to comprise 3 fundamental and interacting components: locations, actors, and concepts [8]. These 3 components are interconnected, forming a health narrative triangle as shown in Figure 2. Concepts that emerge out of a particular narrative may better engage a specific set of people to participate in that discourse. On the other hand, specific participants may affect with which concepts are introduced and discussed when an overall narrative is formed. As these 3 components encapsulate the overall narrative and represent its foundational elements, here we are focusing on studying its temporal variation through the detection of events

along each component, in the form of temporal variations of its defining parameters.

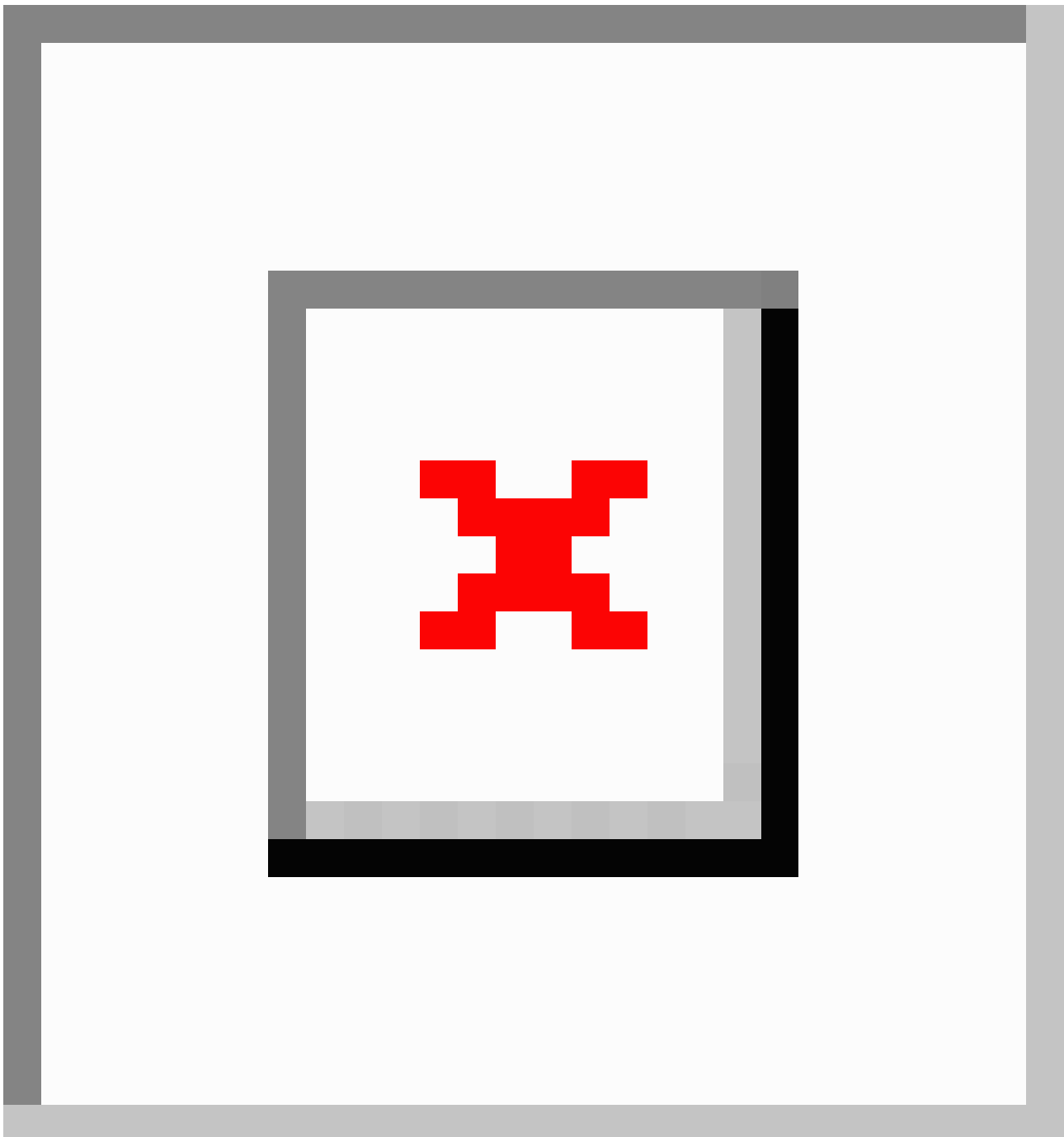
Locations correspond to “where” the narrative is taking place, representing the geographical footprint of the participating communities. Events along this dimension would primarily correspond to changes in the footprint over time, and they can be detected through the application of various spatiotemporal statistical analysis techniques [22].

Actors correspond to “who” participates in and shapes the narrative, reflecting the social dimension of the discourse. They are the contributors to the narrative, with their level of influence dependent upon their overall presence in social media, including their contributions, and the impact of these contributions on the

participating community. Events along this dimension can be detected through the analysis of an actor’s contributions and the responses they elicit (ie, in the form of retweets, replies, or mentions), these events can be detected through the analysis of network or trend metrics [23].

Concepts correspond to “what” is part of the narrative, in the form of keywords that are used by the participants in the context of this discourse. They represent the structure of the linguistic dimension of the discourse and encapsulate the associations among terms that are established through crowd interaction. Events along this dimension reveal the emergence of certain terms that expand or redefine the scope of the narrative, and they can be detected through term-trend analysis techniques [24].

Figure 2. The health narrative triangle.



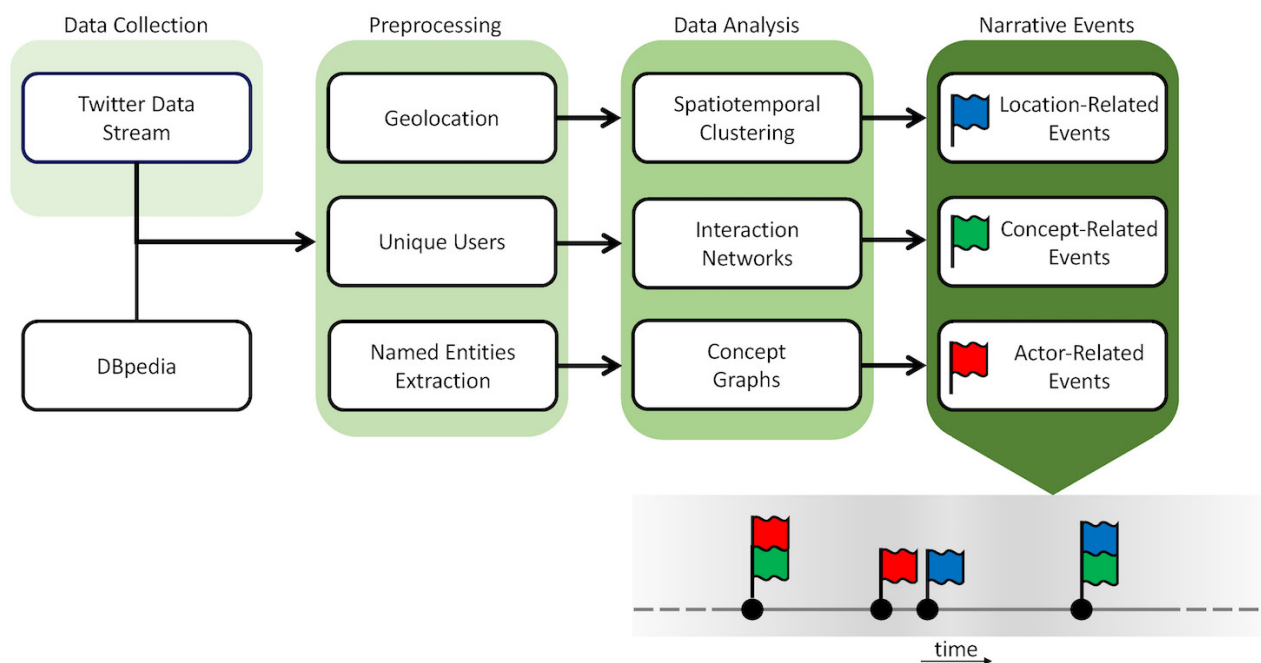
Overview of Data Processing Steps

Our overall approach to process the data comprised a number of automated processes that extract content from streaming Twitter data and process that information to identify corresponding events as shown in Figure 3. The input comprises the Twitter data harvested using Twitter’s API. These data are preprocessed to derive preliminary actor, location, and concept information. Location and actor (ie, user Twitter account) information are derived directly from tweet metadata. Concept information is derived through the identification of named entities. A named entity can be broadly defined in the context of this work as a concept that has a specific definition within an ontology (ie, Zika, or WHO). In our study, the reference ontology is dbpedia [25,26]. These entities are extracted from the content of the harvested tweets and are classified within the corresponding ontology (ie, Zika as disease, and WHO as an organization).

In the next step, we carry out data analysis across these 3 dimensions of the data. Regarding locations, we perform

spatiotemporal clustering to identify regions that exhibit higher levels of Twitter traffic clusters compared with the rest of the world. These clusters indicate new communities joining this global discussion. Actor and concept information can be aggregated as corresponding networks. Regarding actors (users), we construct interaction networks of links among them, aggregating interactions in the form of retweets, mentions, and replies within our data corpus. Regarding concepts (named entities), the corresponding graphs capture the co-occurrence of these concepts (named entities) within single tweets. Actor related and concept-related events are derived as changes in the corresponding frequency counts over time. These changes can be monitored either in terms of their absolute values (eg, exceeding thresholds) or the variations of these values (eg, corresponding gradients). When combined, these 3 events serve as the constructs of event storyline summaries of the Twitter narrative (represented in the lower right part of Figure 3) regarding the monitored event (the spread of Zika in this case). These processes are illustrated in more detail under the Results section below.

Figure 3. An overview of the Twitter narrative analysis approach, starting with data collection and proceeding with preprocessing and data analysis to identify narrative events, which can be used to build an event storyline.



Results

Location-Related Events: Spatiotemporal Patterns

Figure 4 shows the spatiotemporal evolution of our Twitter data corpus. Each map presents the spatial distribution of geolocated tweets (as colored dots) during the corresponding week. The top left panel shows the data during the first week, and time progresses from left to right and from top to bottom towards the 12th week (bottom right). Together, these 12 panels capture the progressive expansion of the Twitter discussion about Zika from a localized (primarily in the Americas top-left panel) to a global phenomenon during this time period.

Spatial clusters of tweets in a particular week indicate a high level of local participation in this discussion in Twitter. Clusters reveal the regions most engaged in this discourse, and they serve as proxy of local interest for the issue. These clusters therefore are essential constructs of the corresponding health narrative. Discernible changes in these clusters correspond to narrative events, as they reflect instances where participation spread to different communities, driven either by new reported cases or by media coverage.

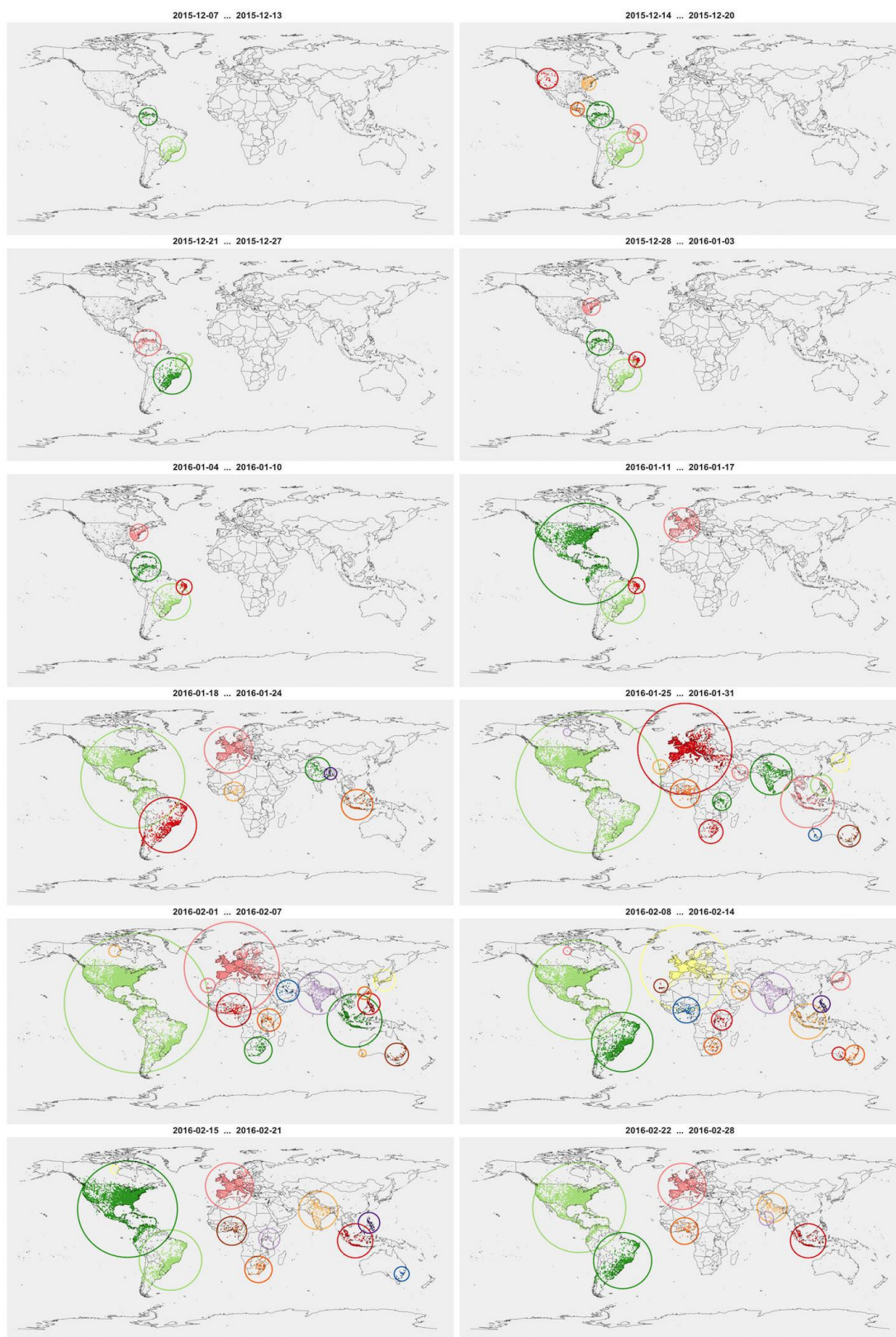
Several tools automatically detect such clusters in spatial and spatiotemporal data [27-29], primarily focusing on local density and proximity analysis. For our analysis, we used the Density Based Spatial Clustering of Application with Noise (DBSCAN) [30]. DBSCAN is a well-established clustering algorithm that

builds on density as a measure for defining and detecting clusters. The algorithm aims to find the maximal set of points that meet a certain density property based on 2 user-defined parameters: a neighborhood radius ϵ and a minimum number of points d . These parameters can be determined through domain knowledge or through a heuristic estimation algorithm (such as the “Automatic Epsilon Calculation” method [31]). DBSCAN offers several distinct advantages, including the ability to distinguish noise in the data, accommodate arbitrary cluster shapes, and perform clustering without previous knowledge or assumptions on the number of clusters. The results of the application of DBSCAN in our datasets are shown in Figure 4. Detected clusters are depicted as colored dots (different colors for different clusters). Additionally, each cluster is delineated by its circular outline, to better signify the corresponding region.

The temporal emergence of these clusters reveals the progressive participation of different communities in this discourse about

Zika as it became a global rather than local concern. In early December 2015, there were only two distinct clusters, in Brazil (light green) and the Venezuela/Colombia region (darker green). This is a close match with the locations of reported cases that week [32]. During the following week in mid-December, 2015, new cases reported in Haiti and Puerto Rico, and an intensification of cases occurred in Guatemala and southern Mexico. This corresponds with increased Twitter posts from these areas, with new Twitter clusters in Central America. Additional expansion of global interest continued in the coming weeks, with Europe becoming more involved in Zika social media by week 6 (third row, right panel) and Asia and Africa joining in by week 7 and further intensifying in weeks 9 and 10). In the context of health narrative analysis, events along the location dimension correspond to the distinct instances where additional locations join the Zika discourse, with the virus (and concerns about it) spreading from its original hotbed of South and Central America to communities across the globe.

Figure 4. Spatiotemporal participation patterns and identifiable clusters over our 12-week study. The top left panel shows the data during the first week, and time progresses from left to right and from top to bottom toward the 12th week (bottom right).



Actor-Related Events: Social Media Presence of Health Organizations

To showcase events related to the social dimension of the narrative, we examined the involvement of the CDC and WHO in the Zika narrative. CDC and WHO were selected for our analysis because they were the top two organizations among the identified named entities in our dataset (ahead eg, of Reuters, Google, Monsanto, and the US Congress, which were some of the other top-ranked entities). Furthermore, and arguably not uncorrelated to this fact, these two organizations were clearly the top information disseminators or communicators for health information worldwide, rendering their analysis even more interesting.

Figure 5 shows the levels of presence in our data corpus of WHO (top) and CDC (bottom). For each organization, the top graph shows activity and the bottom graph depicts impact. All four charts show time along the horizontal axis (with December 2015 at the extreme left and March 2016 at the extreme right). The vertical axis captures metrics of social media presence for these organizations, as follows. The upper part of each chart is a social media activity subchart, expressing how active these organizations were during our study period, and comprises (1) a blue line, showing the daily variations of the number of contributions (original tweets) made by these organizations over time; (2) an orange line, showing how often these organizations retweeted the posts of another user; and (3) a purple line, showing how often these organizations replied to another user. The bottom part of each chart is a social media impact subchart, expressing how well this activity resonated within the Twitter community, and comprises (1) a red line, showing how often these organizations were mentioned in tweets by other users in our data corpus (ie, in the context of Zika); (2) a green line, showing the number of times that a message originally posted by these organizations was retweeted by the general public in our data corpus; and (3) a cyan line, showing how often other users replied directly to these organizations' posts.

Overall, the level of reply activity is substantially lower than retweeting or mentions. This suggests that Twitter users view these organizations' accounts as information sources rather than portals for active interaction with them. The vertical axis of both social media activity subcharts spans the range 0 - 50. The maximum number of Zika tweets per day for WHO was 52, while for CDC it was 49. The vertical axes for the other colored charts span the range 0 - 6000 for WHO and 0 - 1500 for CDC. Together, social media activity and impact express the overall social media presence of these organization's accounts.

CDC had higher levels of activity, publishing more original tweets (241 vs 152, respectively, shown in the blue chart), replying to more tweets (93 vs 18, purple lines), and retweeting more (45 vs 9, orange lines) than WHO. In contrast, impact shows a reversal of the activity pattern, with WHO receiving more retweets (22,798 vs 10,935, green lines), more mentions (22,468 vs 10,789, red lines) and more replies (185 vs 170, cyan lines) than CDC. This reflects the global scope of WHO, compared with the primarily US-focused CDC.

The contrast between impact and activity expresses the "amplification potential" of these organizations. Numerically,

an expression of this potential can be the "amplification factor," which can be defined as the ratio between the sum of the 3 impact metrics over the sum of the 3 activity metrics. In our data corpus, the WHO had an amplification factor of 254 compared with only 58 for the CDC. In this particular case, the amplification factor is a direct reflection of followership in Twitter. For this particular disease, the WHO amplification factor is 4.4 times that of CDC's, which is exactly the same as the ratio of their corresponding followership (3.32 million vs 755,000, respectively at the time of writing).

The footprints of these organizations in the full Twitter community can be illustrated with "interaction networks" that capture the aggregate connections among these nodes. In these retweet-mention-and-reply networks, two nodes are connected each time one retweets, mentions, or replies to the other one. Figure 6 (top) shows the placement of WHO and CDC in the broader interaction network when the top 3000 connections are visualized. The size of the nodes corresponds to node activity. In this figure, two distinct identifiable clusters correspond to political leaders and are geographically oriented. One cluster is formed around Venezuelan political figures such as President Nicolas Maduro (@NicolasMaduro) and Tareck El Aissami (@TareckPSUV), whereas the second is formed around Honduran Minister of Communication Hilda Hernandez (@HildaHernandezA).

Figure 6 (bottom) shows the subsets of the full interaction network related to WHO (left) and CDC (right). WHO is primarily connected through its links through news organizations, such the Spanish newspaper El Pais (@el_pais), the Spanish language branches of BBC and CNN (@bbcunited and @cnnee), and some Venezuelan news websites such as La Patilla and El Nacional (@la_patilla and @ElNacionalWeb). In contrast, the CDC presence is shared by a number of official accounts, including the primary CDC Twitter account (@CDCgov) and those of CDC Director Frieden (@drfriedencdc) and CDC Travel (@cdctravel). Contrasting this with the above presented observations regarding the amplification factor and followership of the WHO and the CDC suggests that engaging news agencies in a health organization's social media network footprint tends to bear fruit in terms of message penetration and public outreach. This is consistent with earlier work [8] that demonstrated that news outlets played a substantial role (through news stories) in disseminating health-related information through social media than traditional health organizations' announcements.

In the context of our analysis, key actor-related narrative events (such as major announcements by official agencies) correspond to discernible instances of growth in organizational impact. When Dr Margaret Chan, Director-General of WHO, announced on January 28, 2016, that the Zika virus is "spreading explosively" [33], and again on February 1, 2016, when WHO issued a statement calling Zika a "Public Health Emergency of International Concern" [34], tweets interacting with WHO spiked. Similar peaks were observed for CDC, after it issued an official health advisory on January 15, 2016 [35], and when it updated its guidelines for screening pregnant women on February 9, 2016 [36].

Figure 5. Temporal variations of the social media presence of WHO (top) and CDC (bottom) in the Zika-related discussion in Twitter during our study period. The horizontal shows time, and the vertical numbers of tweets or retweets or mentions as appropriate.

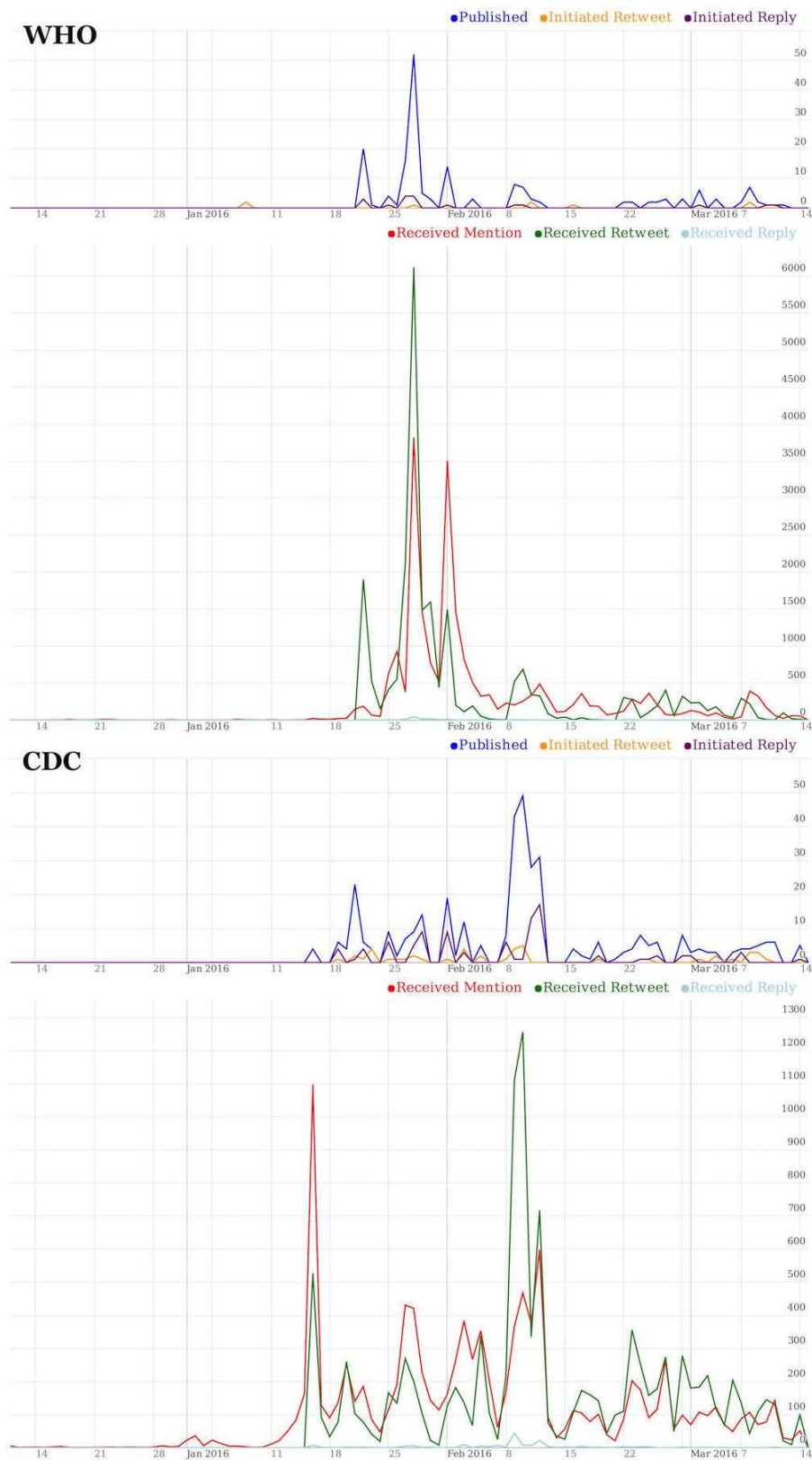
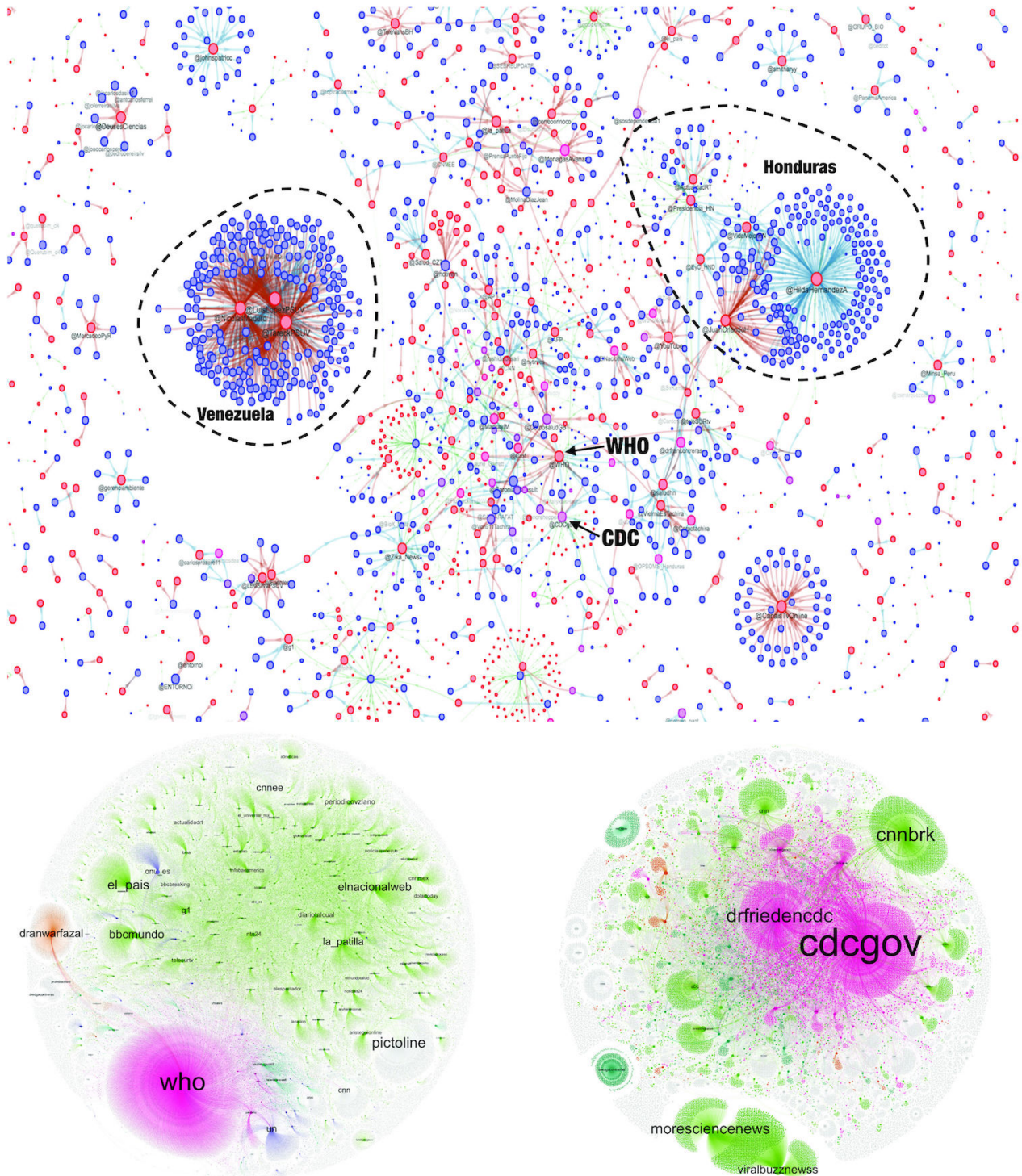


Figure 6. Top: The presence of WHO and CDC in an aggregate interaction network. In this particular figure, we visualized only the top 3,000 connections. Red nodes are originators of information (ie, original authors), blue are consumers (eg, retweeting, mentioning, or replying to these original posts). Bottom: Subsets of the full retweet network pertaining to the WHO (left) and CDC (right), and clusters identified within them. Magenta clusters are centered upon health entities, green upon news organizations, orange upon political entities.



Concept-Related Events: Pregnancy and Abortion

In our Zika analysis, key concept events correspond to the discernible instances of notable changes in references to various terms in the context of the social media, Zika-related interaction. As we mentioned above, concepts are extracted as named entities corresponding to dbpedia ontology entries. In order to do so, we used the online TextRazor natural language processing

(NLP)-based API to parse the text and extract entities. These entities are classified according to the dbpedia ontology.

Figure 7 shows the trending of references to pregnancy and abortion in our data corpus. These two terms were chosen due to their strong emergence during our study period. More specifically, while both terms were rarely encountered during the first weeks of our study, references to pregnancy grew

dramatically after January 15, 2016, when the CDC issued its above-referenced interim travel guidance for pregnant women because of growing evidence for the risk of microcephaly in infants born to infected mothers. After that date, the term pregnancy became a key concept associated with the virus, reaching a peak frequency on February 5, 2016, when the CDC offered Zika testing to all pregnant women with potential virus exposure [37]. The term abortion reached a peak on February 6, 2016. Although abortion mentions followed a path overall comparable to pregnancy mentions, they also had an additional off-cycle peak on February 17, 2016, in response to comments by Pope Francis regarding contraception and abortion in the context of Zika [38].

Figure 8 illustrates the evolution of pregnancy and abortion as key components of the Zika narrative by showing their concept graphs for 2 distinct 4-day periods: January 1 to 4, 2016 (middle) and February 14 to 18, 2016 (bottom). The concept graph captures the terms most frequently encountered in conjunction with the word abortion in our Zika data corpus. The size of the

node reflects how frequently a particular term was encountered overall in the data corpus, and the width of the line joining two nodes reflects how frequently these nodes were encountered in the same tweet (expressing a level of mental association of these two terms). Red links are directly related to abortion (radiating outward from that term in the graphs), whereas gray links are connections among these other terms. Abortion was infrequently mentioned in January, 2016. The terms associated with it include ones that match our health narrative triangle: green terms are geographical entities (countries like Brazil and Colombia), red terms are concepts (such as pregnancy and microcephaly) and blue are actors (such as WHO). Mentions of abortion were much more frequent in February, 2016. Although the geographical references remained similar, the concepts had expanded to include terms such as fetus and birth control, and the pool of actors increased to include the Pope and the Catholic Church as key players in the discourse on abortion in the context of Zika. Thus, our triad captures the evolving perceptions of Zika by the general public and the actors that shape this evolution.

Figure 7. References to pregnancy (blue) and abortion (red) in our data corpus. Horizontal axis: time. Vertical axis: number of mentions.

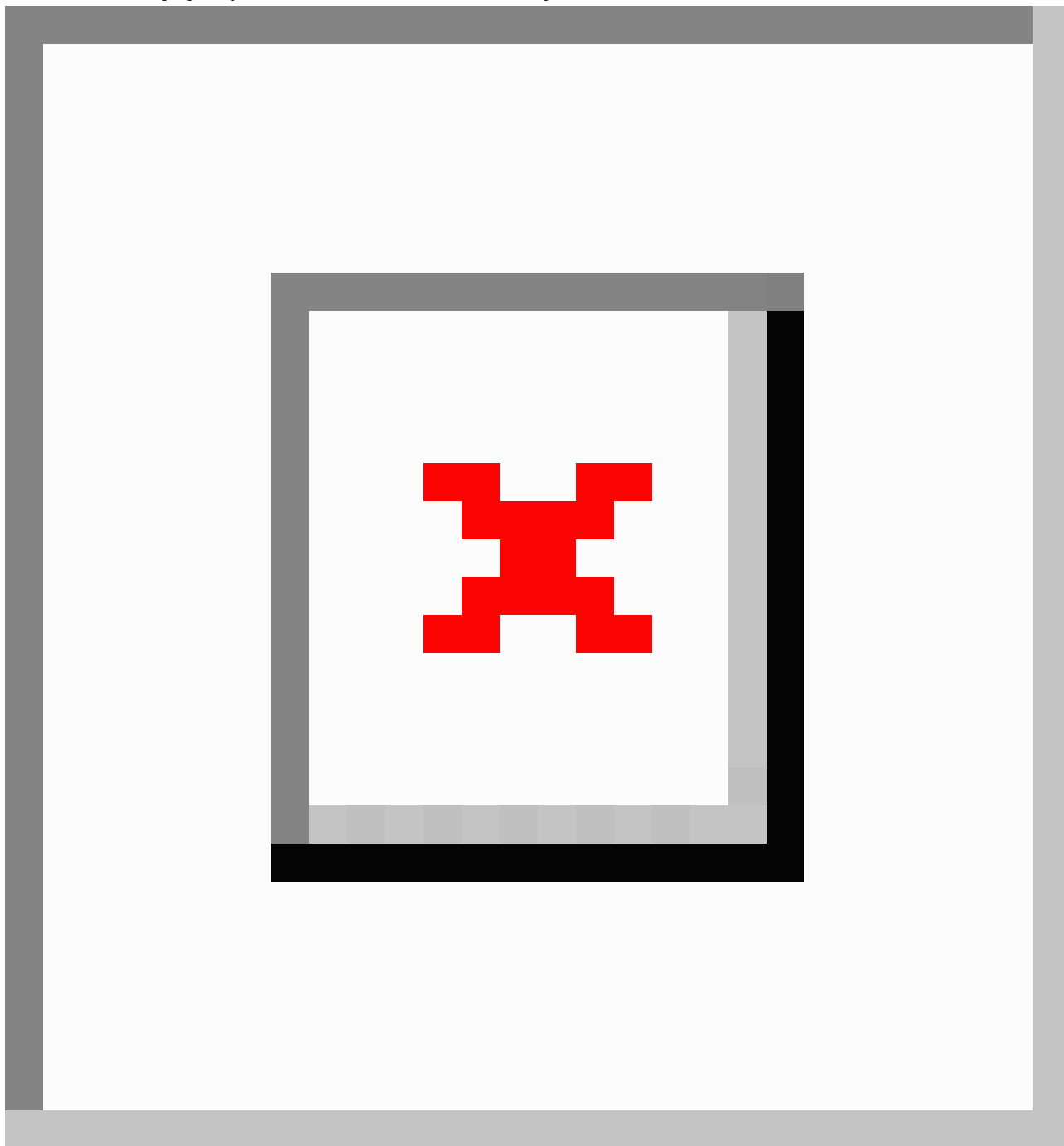
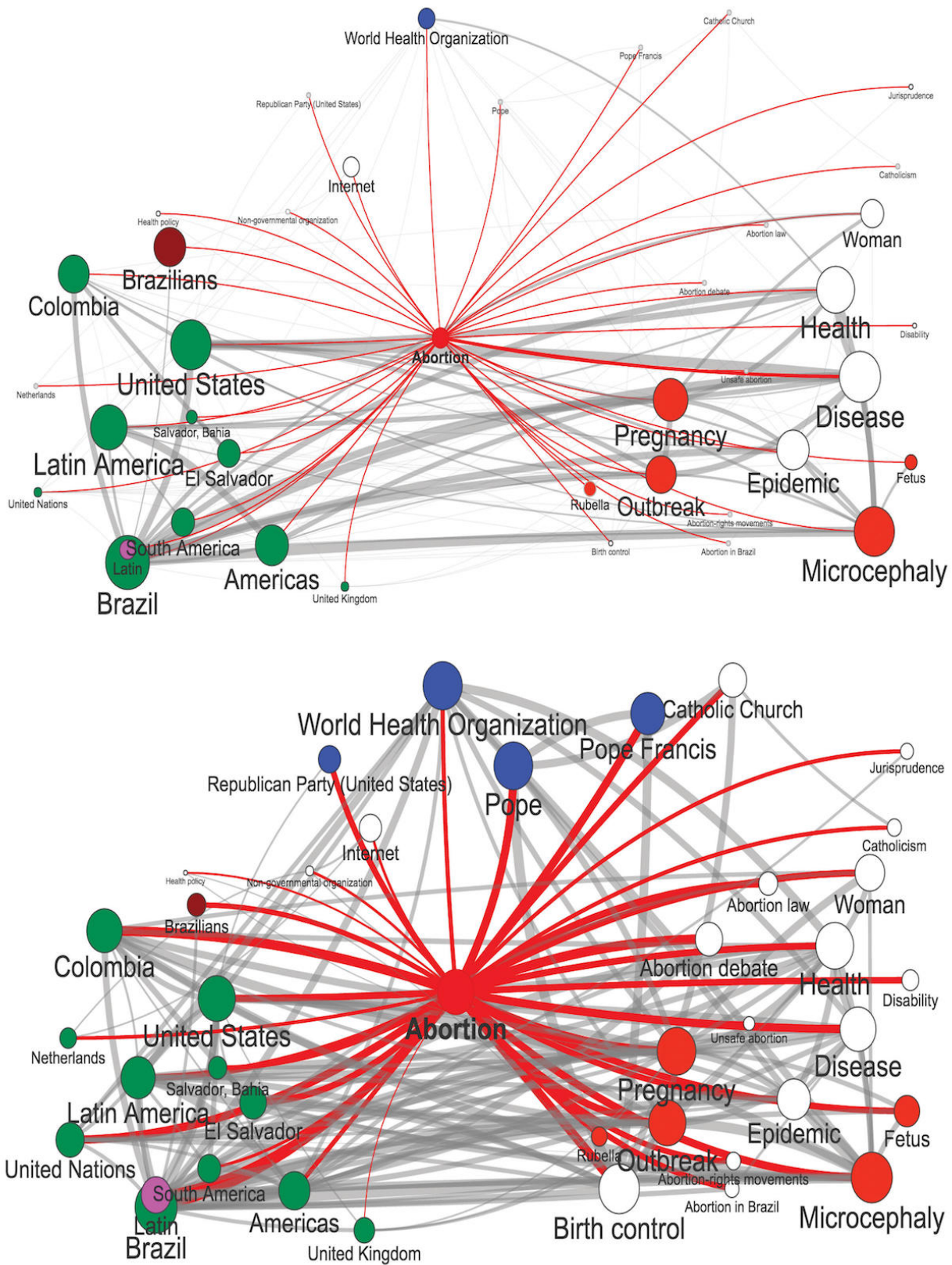


Figure 8. Top: Concept graph for abortion mentions over the 5-day period of 1/1-1/5. Bottom: Concept graph for abortion mentions over the 5-day period of 2/14-2/18.



Narrative Storyline Communication

The information captured through the processes outlined above can be viewed as a summarization of the discussion in Twitter

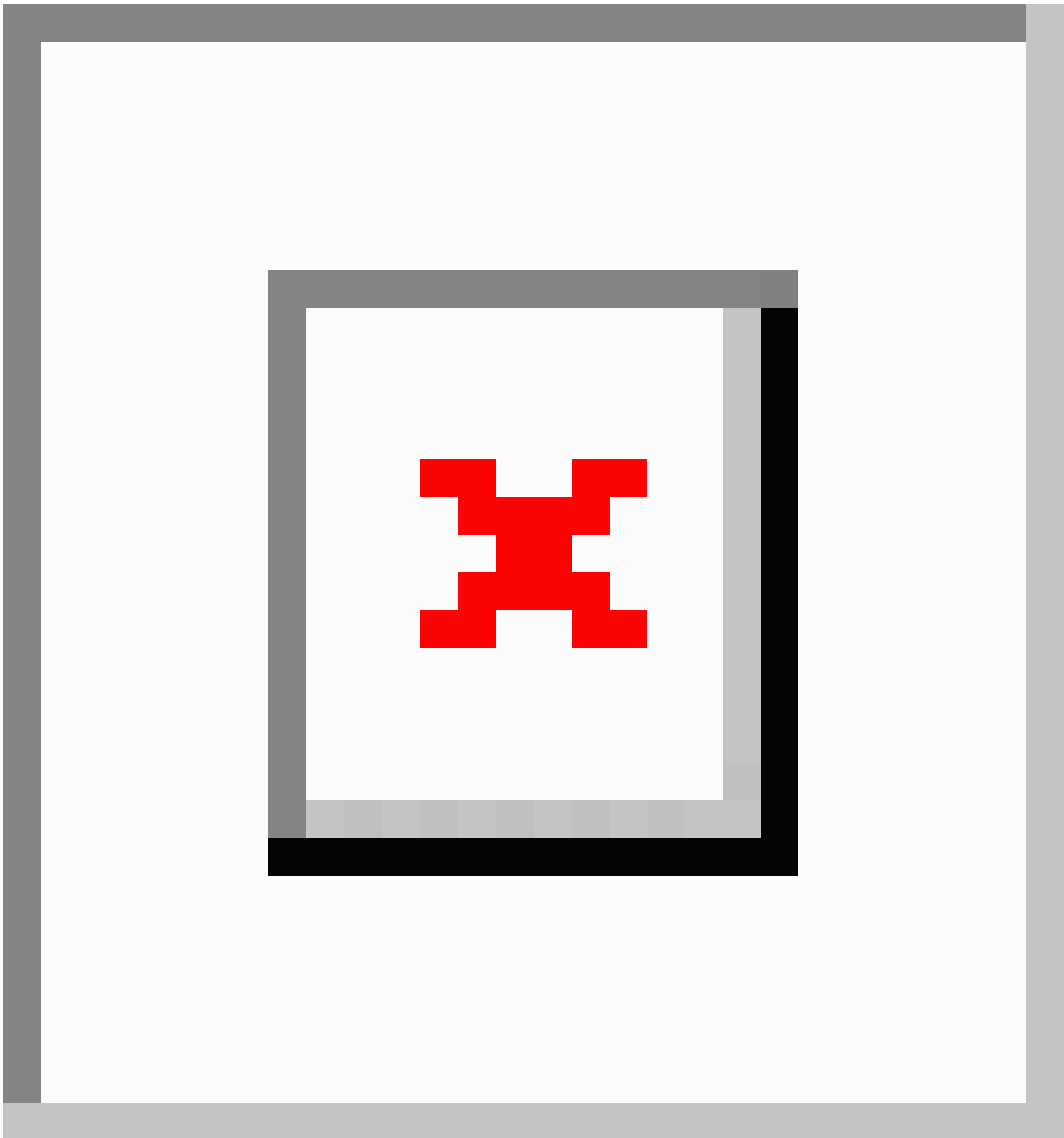
regarding the Zika crisis, revealing essential narrative components of the process (eg, in the context of computational narrative models [39-42]). These narrative components are aligned to event storylines as they were addressed in structuralist

approaches to narrative formalization [43]. As such, they can inform narrative-driven visualizations [44] of the interaction in social media.

In order to demonstrate the potential use of such events for narrative communication, in Figure 9 we present a sample visualization of such summaries, as they were captured through our analysis. We show storylines for representative subsets from each of the 3 dimensions of the narrative, namely 5 locations (blue), actors (red), and concepts (green) over a period of 6

weeks. In this visualization, horizontal black lines indicate continuous activity regarding the corresponding entry, lack of them indicates that a particular topic was below a threshold, and colored disks indicate activity events. The size of the disk is proportional to the magnitude of the event relative to the overall activity. In this manner, we are able to capture both global events (eg, the spikes associated with February 3, 2016, when traffic reached a maximum) as well as localized events, for example, the various spikes before or after that date.

Figure 9. Visualizing a narrative storyline across locations (blue), actors (red), and concepts (green).



Discussion

Principal Findings

This study analyzed how the public engaged with the Zika outbreak of 2015-2016 via social media, from its initial outbreak in South America to its global spread. Our analysis showed how the variations in locations, actors, and concepts involved in Twitter traffic reflect how the health narrative changed as Zika went from being seen as a relatively benign, regional health concern to a global epidemic.

Our findings show how location plays a central role delineating various discrete conversations about Zika on Twitter. The unfolding discussion closely matched maps of the spread of the disease across the globe and expanded news coverage of the emerging infectious disease event. Geographical cohesion is also observed in the clusters of relationships created by retweeting behaviors, with two of the largest clusters emerging from government and political groups within countries dealing with the Zika spread.

At the same time, some of the actors engaged in this conversation crossed geographical boundaries and engaged new actors in the global discussion. Although initially both CDC

and WHO largely existed on the periphery of the conversation, we showed how they were able to drive substantial traffic on Twitter when making official announcements.

The combination of geographical clusters and key actors both contribute and respond to the concepts that defined the Zika outbreak. Discussions about pregnancy and abortion as they related to Zika corresponded with public releases from health officials. The Catholic Church also entered into this narrative framework following Pope Francis's remarks about contraception in response to Zika, in part because of the largely Catholic populations in highly affected areas of the Americas.

The study showed how different populations engaged in the discourse and how actors shape and reshape the discussion and the concepts that the public associates with this public health emergency. Advancing our capabilities to study these temporal variations across these 3 dimensions brings us one step closer to deciphering the complex mechanisms through which the public participates in the discourse. Further work should therefore explore how health campaigns can better use the power of social media for information dissemination campaigns that better align with the dynamic changes of the public health emergency concerns.

Acknowledgments

This research was funded through a Multidisciplinary Research Award by the Office of the Provost of George Mason University. Publication of this article was funded by the George Mason University Libraries Open Access Publishing Fund.

Authors' Contributions

JR and GL led the acquisition and preliminary analysis of the data. All the authors contributed substantially to the design of the study and the analysis and interpretation of the data. AS, EV, and KHJ led the write-up of the manuscript, and all authors contributed to its preparation and approved the final version.

Conflicts of Interest

None declared.

References

1. Wearesocial. Digital in 2016 URL: <http://wearesocial.com/uk/special-reports/digital-in-2016> [accessed 2017-02-08] [WebCite Cache ID 6o70AcQNI]
2. Chou W, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. *J Med Internet Res* 2009 Nov 27;11(4):e48 [FREE Full text] [doi: [10.2196/jmir.1249](https://doi.org/10.2196/jmir.1249)] [Medline: [19945947](https://pubmed.ncbi.nlm.nih.gov/19945947/)]
3. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of tweets during the 2009 H1N1 outbreak. *PLoS One* 2010 Nov 29;5(11):e14118 [FREE Full text] [doi: [10.1371/journal.pone.0014118](https://doi.org/10.1371/journal.pone.0014118)] [Medline: [21124761](https://pubmed.ncbi.nlm.nih.gov/21124761/)]
4. 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/)]
5. 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/)]
6. Laranjo L, Arguel A, Neves A. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. *JAMIA* 2015;22(1):243-256. [doi: [10.1136/amiajnl-2014-002841](https://doi.org/10.1136/amiajnl-2014-002841)]
7. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelandotte C. Are health behavior change interventions that use online social networks effective? A systematic review. *J Med Internet Res* 2014 Feb 14;16(2):e40 [FREE Full text] [doi: [10.2196/jmir.2952](https://doi.org/10.2196/jmir.2952)] [Medline: [24550083](https://pubmed.ncbi.nlm.nih.gov/24550083/)]
8. Radzikowski J, Stefanidis A, Jacobsen K, Croitoru A, Crooks A, Delamater PL. The measles vaccination narrative in Twitter: a quantitative analysis. *JMIR Public Health Surveill* 2016;2(1):e1 [FREE Full text] [doi: [10.2196/publichealth.5059](https://doi.org/10.2196/publichealth.5059)] [Medline: [27227144](https://pubmed.ncbi.nlm.nih.gov/27227144/)]

9. Paul M, Dredze M. You are what you tweet: analyzing Twitter for public health. : Association for the Advancement of Artificial Intelligence; 2011 Presented at: Proceedings of the 5th International AAAI Conference on Weblogs and Social Media; 2011; Barcelona p. 265-272.
10. 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 23;15(4):e85 [FREE Full text] [doi: [10.2196/jmir.1933](https://doi.org/10.2196/jmir.1933)] [Medline: [23615206](https://pubmed.ncbi.nlm.nih.gov/23615206/)]
11. Lazard A, Scheinfeld E, Bernhardt J, Wilcox GB, 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/)]
12. Khoury M, Ioannidis JP. Big data meets public health. *Science* 2014;346(6213):1054. [doi: [10.1126/science.aaa2709](https://doi.org/10.1126/science.aaa2709)]
13. Harris JK, Choucair B, Maier RC, Jolani N, Bernhardt JM. Are public health organizations tweeting to the choir? Understanding local health department Twitter followership. *J Med Internet Res* 2014 Feb 26;16(2):e31 [FREE Full text] [doi: [10.2196/jmir.2972](https://doi.org/10.2196/jmir.2972)] [Medline: [24571914](https://pubmed.ncbi.nlm.nih.gov/24571914/)]
14. Croitoru A, Crooks A, Radzikowski J, Stefanidis A. GeoSocial gauge: a system prototype for knowledge discovery from geosocial media. *Int J Geogr Inf Sci* 2013;27(12):2483-2508. [doi: [10.1080/13658816.2013.825724](https://doi.org/10.1080/13658816.2013.825724)]
15. Valkanas G, Katakis I, Gunopulos D, Stefanidis A. Mining twitter data with resource constraints. 2014 Presented at: IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT); 2014; Warsaw p. 157-164.
16. Riessman C, Quinney L. Narrative in social work a critical review. *Qual Soc Work* 2005;4(4):391-412. [doi: [10.1177/1473325005058643](https://doi.org/10.1177/1473325005058643)]
17. Labov W. Narrative pre-construction. *Narrat Inq* 2006;16(1):37-45.
18. Labov W, Waletzky J. Narrative analysis: oral versions of personal experience. *JNLH* 1997;7:3-38.
19. Graesser AC, Singer M, Trabasso T. Constructing inferences during narrative text comprehension. *Psychol Rev* 1994 Jul;101(3):371-395. [Medline: [7938337](https://pubmed.ncbi.nlm.nih.gov/7938337/)]
20. Schank R. Tell me a story: a new look at real and artificial memory. New York: Charles Scribner's Sons; 1990.
21. Meade M, Emch M. Medical Geography, Third Edition. Surrey, UK: The Guilford Press; 2010.
22. Cressie N, Wikle C. Statistics for Spatio-Temporal Data. New York, NY: Wiley; 2011.
23. Kwak H, Lee C, Park H. What is Twitter, a Social network or a news media? 2010 Presented at: 19th International Conference on World Wide Web; 2010; Raleigh, NC p. 591-600.
24. Lee K, Palsetia D, Narayanan R. Twitter trending topic classification. 2011 Presented at: 11th International Conference on Data Mining Workshops; 2011; Vancouver, BC p. 251-258.
25. Auer S, Bizer C, Kobilarov G, Lehmann J, Cyganiak R. DBpedia: A Nucleus for a Web of Open Data. In: *The Semantic Web*. New York, NY: Springer; 2007:722-735.
26. mappings.dbpedia. Ontology Classes URL: <http://mappings.dbpedia.org/server/ontology/classes/> [accessed 2017-02-22] [WebCite Cache ID 6oSSUXapV]
27. Berkhin P. A survey of clustering data mining techniques. In: Nicholas C, Kogan J, Teboulle M, editors. *Grouping Multidimensional Data: Recent Advances in Clustering*. Berlinpringer. New York, NY: Springer; 2006:25-71.
28. Kisilevich S, Mansmann F, Nanni M. Spatio-temporal clustering. In: Maimon O, Rokach L, editors. *Data Mining and Knowledge Discovery Handbook*. New York, NY: Springer; 2009:855-874.
29. Amini A, Wah TY, Saboohi H. On density-based data streams clustering algorithms: a survey. *J Comput Sci Technol* 2014;29(1):116-141. [doi: [10.1007/s11390-014-1416-y](https://doi.org/10.1007/s11390-014-1416-y)]
30. Ester M, Kriegel H, Sander J. A density-based algorithm for discovering clusters in large spatial databases with noise. New York, NY: AAAI; 1996 Presented at: 2nd International Conference on Knowledge DiscoveryData Mining; 1996; Portland, OR p. 226-231.
31. Gorawski M, Malczok R. AEC algorithm: a heuristic approach to calculating density-based clustering eps parameter. In: *Advances in Information Systems*. New York, NY: Springer; 2006:90-99.
32. HealthMap. 2016 Zika outbreak URL: <http://www.healthmap.org/zika/> [accessed 2017-02-08] [WebCite Cache ID 6o70WRPzo]
33. CNN. Zika virus 'spreading explosively,' WHO leader says URL: <http://www.cnn.com/2016/01/28/health/zika-virus-global-response/> [WebCite Cache ID 6l5Qoty3O]
34. WHO. WHO director-general summarizes the outcome of the emergency committee regarding clusters of Microcephaly and Guillain-Barré Syndrome URL: <http://www.who.int/mediacentre/news/statements/2016/emergency-committee-zika-microcephaly/en/> [WebCite Cache ID 6l7Z4mprJ]
35. CDC. Recognizing, managing, and reporting Zika virus infections in travelers returning from Central America, South America, the Caribbean, and Mexico URL: <http://emergency.cdc.gov/han/han00385.asp> [WebCite Cache ID 6l7ZGbb02]
36. CDC. Questions and answers for healthcare providers caring for pregnant women and women of reproductive age with possible Zika virus exposure URL: <http://www.cdc.gov/zika/hc-providers/qa-pregnant-women.html> [accessed 2016-10-09] [WebCite Cache ID 6l7ZKhHGw]

37. CDC. Update: Interim guidelines for health care providers caring for pregnant women and women of reproductive age with possible Zika virus exposure ? United States, 2016 URL: <http://www.cdc.gov/mmwr/volumes/65/wr/mm6505e2.htm> [WebCite Cache ID 617ZOoNqu]
38. Washington Post. Pope Francis suggests contraception could be permissible in Zika fight URL: <https://www.washingtonpost.com/news/acts-of-faith/wp/2016/02/17/mexico-confirms-zika-virus-cases-in-pregnant-women-as-pope-francis-exits-the-country/> [accessed 2016-10-09] [WebCite Cache ID 617ZTVxFh]
39. Finlayson MA. Corpus Annotation in Service of Intelligent Narrative Technologies. In: Intelligent Narrative Technologies IV. Menlo Park, CA: AAAI Press; 2011 Presented at: AIIDE Workshop; 2011; Las Vegas, NV p. 17-20.
40. Finlayson M, Richards W, Winston PH. AI Magazine. 2010. Computational models of narrative: Review of a workshop URL: <https://users.cs.fiu.edu/~markaf/doc/o6.finlayson.2010.aimag.31.97.pdf> [accessed 2017-04-16] [WebCite Cache ID 6pmYau5fB]
41. Wolff A, Mulholland P, Collins T. Storyspace: a story-driven approach for creating museum narratives. New York, NY: ACM; 2012 Presented at: 23rd ACM conference on Hypertext and social media; 2012, June 25-28; Milwaukee, WI p. 89-98. [doi: [10.1145/2309996.2310012](https://doi.org/10.1145/2309996.2310012)]
42. Niehaus J, Romero V, Koelle D, Palmon N, Bracken B, Pfautz J, et al. drops.dagstuhl.de.: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik; 2014. A Flexible Framework for the Creation of Narrative-Centered Tools URL: <http://drops.dagstuhl.de/opus/volltexte/2014/4651/pdf/18.pdf> [accessed 2017-04-16] [WebCite Cache ID 6pmYyGVjf]
43. Polkinghorne D. Narrative knowing and the human sciences. Albany: State University of New York Press; 1988.
44. Wood J, Beecham R, Dykes J. Moving beyond sequential design: reflections on a rich multi-channel approach to data visualization. IEEE Trans Vis Comput Graph 2014 Dec;20(12):2171-2180. [doi: [10.1109/TVCG.2014.2346323](https://doi.org/10.1109/TVCG.2014.2346323)] [Medline: [26356931](https://pubmed.ncbi.nlm.nih.gov/26356931/)]

Abbreviations

API: application program interface

CDC: Centers for Disease Control

DBSCAN: Density Based Spatial Clustering of Application with Noise

WHO: World Health Organization

Edited by G Eysenbach; submitted 31.10.16; peer-reviewed by MH Tsou, J Sutton, A MacKinlay; comments to author 28.01.17; revised version received 01.03.17; accepted 05.03.17; published 20.04.17.

Please cite as:

Stefanidis A, Vraga E, Lamprianidis G, Radzikowski J, Delamater PL, Jacobsen KH, Pfoser D, Croitoru A, Crooks A

Zika in Twitter: Temporal Variations of Locations, Actors, and Concepts

JMIR Public Health Surveill 2017;3(2):e22

URL: <http://publichealth.jmir.org/2017/2/e22/>

doi: [10.2196/publichealth.6925](https://doi.org/10.2196/publichealth.6925)

PMID: [28428164](https://pubmed.ncbi.nlm.nih.gov/28428164/)

©Anthony Stefanidis, Emily Vraga, Georgios Lamprianidis, Jacek Radzikowski, Paul L Delamater, Kathryn H Jacobsen, Dieter Pfoser, Arie Croitoru, Andrew Crooks. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 20.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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

What Are People Tweeting About Zika? An Exploratory Study Concerning Its Symptoms, Treatment, Transmission, and Prevention

Michele Miller¹, MS; Tanvi Banerjee^{2,3}, PhD; Roopteja Muppalla^{2,3}, BTech; William Romine¹, PhD; Amit Sheth^{2,3}, PhD

¹Department of Biological Sciences, Wright State University, Dayton, OH, United States

²Department of Computer Science and Engineering, Wright State University, Dayton, OH, United States

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

Corresponding Author:

Michele Miller, MS

Department of Biological Sciences

Wright State University

3640 Colonel Glenn Hwy

Dayton, OH,

United States

Phone: 1 937 775 3333

Fax: 1 937 775 3320

Email: millerme91@gmail.com

Abstract

Background: In order to harness what people are tweeting about Zika, there needs to be a computational framework that leverages machine learning techniques to recognize relevant Zika tweets and, further, categorize these into disease-specific categories to address specific societal concerns related to the prevention, transmission, symptoms, and treatment of Zika virus.

Objective: The purpose of this study was to determine the relevancy of the tweets and what people were tweeting about the 4 disease characteristics of Zika: symptoms, transmission, prevention, and treatment.

Methods: A combination of natural language processing and machine learning techniques was used to determine what people were tweeting about Zika. Specifically, a two-stage classifier system was built to find relevant tweets about Zika, and then the tweets were categorized into 4 disease categories. Tweets in each disease category were then examined using latent Dirichlet allocation (LDA) to determine the 5 main tweet topics for each disease characteristic.

Results: Over 4 months, 1,234,605 tweets were collected. The number of tweets by males and females was similar (28.47% [351,453/1,234,605] and 23.02% [284,207/1,234,605], respectively). The classifier performed well on the training and test data for relevancy (F1 score=0.87 and 0.99, respectively) and disease characteristics (F1 score=0.79 and 0.90, respectively). Five topics for each category were found and discussed, with a focus on the symptoms category.

Conclusions: We demonstrate how categories of discussion on Twitter about an epidemic can be discovered so that public health officials can understand specific societal concerns within the disease-specific categories. Our two-stage classifier was able to identify relevant tweets to enable more specific analysis, including the specific aspects of Zika that were being discussed as well as misinformation being expressed. Future studies can capture sentiments and opinions on epidemic outbreaks like Zika virus in real time, which will likely inform efforts to educate the public at large.

(*JMIR Public Health Surveill* 2017;3(2):e38) doi:[10.2196/publichealth.7157](https://doi.org/10.2196/publichealth.7157)

KEYWORDS

viruses; epidemiology; social media; machine learning

Introduction

Background

The 2014 and 2015 Ebola outbreak caused fear and misinformation to spread wildly across the globe. It was shown that the spread of misinformation led to deaths due to improper practice of appropriate preventative measures [1].

Experts at the Center for Disease Control (CDC) and the World Health Organization (WHO) admit that they mishandled the response for Ebola by not responding to the threat sooner [2]. One year after the Ebola outbreak ended, the Zika outbreak started and also caused fear and misinformation to spread. In the recent years, citizen sensing has picked up greatly with the rise of mobile device popularity and social media sites such as Facebook and Twitter. The idea with citizen sensing is that citizens play the role of sensors in the environment [3], providing information regarding health care issues such as disease outbreaks like Ebola and Zika [4].

Big social data eliminate the time lag caused by traditional survey-based methods, allowing for studying public opinions on issues while addressing privacy concerns of users through group-level analyses of public behavior with respect to specific issues in real time. In particular, public opinion mining has facilitated exploration of public views on important social issues such as gender-based violence [5] and health-related beliefs [6-7].

With respect to Zika, Twitter served as a source of misinformation. To counter, the CDC responded with correct information, either by tweeting general statements about Zika or by responding to questions and comments directed at them. For example, 1 user tweeted, "Apparently Florida is immune to the Zika virus," whereas the CDC had tweeted about Zika in Florida several times, including this tweet: "Updated: CDC travel and testing recommendations for Miami-Dade county b/c of continued local #Zika transmission."

Zika

Many people do not even realize they are sick from Zika, let alone the need to go to the hospital; and death due to Zika is extremely rare [8]. The Zika virus usually causes only mild symptoms such as headache, rash, fever, conjunctivitis, and joint pain, which can last from a few days to a week after being infected [8]. Guillain-Barré syndrome and microcephaly have been linked to Zika and as this is the first outbreak of Zika associated with these defects, management is still an important challenge [9]. There are 3 main ways by which one can contract Zika: (1) being bitten by an infected Aedes mosquito, (2) through sexual contact, and (3) from mother to fetus [8]. There is currently no medicine or vaccine to treat the Zika virus; however, there are several methods of prevention [8].

Related Works

A study by Oyeyemi et al [10] concerning misinformation about Ebola on Twitter found that 44.0% (248/564) of the tweets about Ebola were retweeted at least once, with 38.3% (95/248) of those tweets being scientifically accurate, whereas 58.9%

(146/248) were inaccurate. Furthermore, most of the tweets containing misinformation were never corrected. Another study about Ebola by Tran and Lee [4] found that the first reported incident of the doctor with Ebola had more impact and received more attention than any other incident, showing that people pay more attention and react more strongly to a new issue.

Majumder et al attempted to estimate the basic R_0 and R_{obs} for Zika using HealthMap and Google Trends [11]. R_0 is known as the basic reproduction number and is the number of expected new infections per first infected individual in a disease-free population. R_{obs} is the observed number of secondary cases per infected individual. Their results indicate that the ranges for R_{obs} were comparable between the traditional method and the novel method. However, traditional methods had higher R_0 estimates than the HealthMap and Google Trend data. This indicates that digital surveillance methods can estimate transmission parameters in real time in the absence of traditional methods.

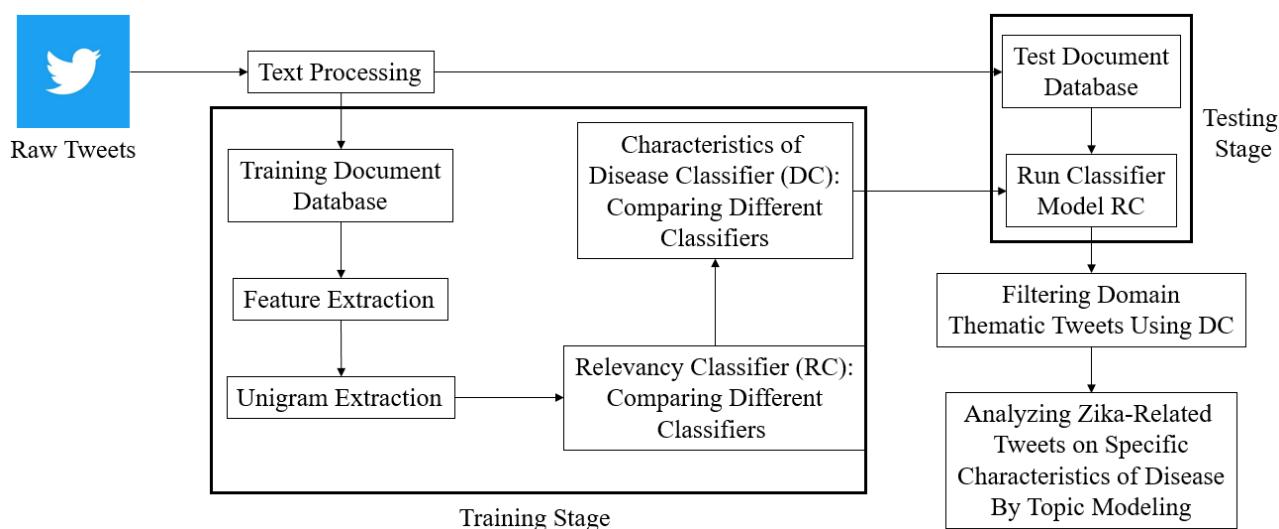
Another study collected tweets on Zika for 3 months [12]. They found that citizens were more concerned with the long-term issues than the short-term issues such as fever and rash. Using hierarchical clustering and word co-occurrence analysis, they found underlying themes related to immediate effects such as the spread of Zika. Long-term effects had themes such as pregnancy. One issue with this paper was that they never employed experts to check the relevance of the tweets with respect to these topics, which is a common problem in mining social media data.

A study by Glowacki et al [13] collected tweets during an hour-long live CDC twitter chat. They only included words used in more than 4 messages to do a topic analysis and found that the 10-topic solution best explained the themes. Some of the themes were virology of Zika, spread, consequences for infants and pregnant women, sexual transmission, and symptoms. This was a curated study where only tweets to and from the CDC were explored, whereas the aim of our larger study was to determine what the general public was discussing about Zika.

A study by Fu et al [14] analyzed tweets from May 1, 2015 to April 2, 2016 and found 5 themes using topic modeling: (1) government, private and public sector, and general public response to the outbreak; (2) transmission routes; (3) societal impacts of the outbreak; (4) case reports; and (5) pregnancy and microcephaly. This study did not check for noise within the social media data. Moreover, the computational analysis was limited to 3 days of data, which may not reflect the themes in the larger dataset.

In many of these studies, the need for checking the performance of the system as well as a post hoc error analysis on checking for the generalizability of their method is overlooked. We address this in our study by employing machine learning techniques on an annotated data set, as well as a post hoc error analysis on a test dataset, to ensure the generalizability of our system.

Figure 1. Block diagram of the pragmatic function-oriented content retrieval using a hierarchical supervised classification technique, followed by deeper analysis for characteristics of disease content.



In this study, an exploratory analysis focused on finding important subcategories of discussion topics from Zika-related tweets was performed. Specifically, we addressed 4 key characteristics of Zika: symptoms, transmission, treatment, and prevention. Using the system described in Figure 1, the following research questions were addressed:

R1. Dataset Distribution Analysis: What proportion of male and female users tweeted about Zika, what were the polarities of the tweets by male and female users, and what were the proportions of tweets that discussed topics related to the different disease characteristics—symptoms, transmission, treatment, and prevention?

R2. Classification Performance Analysis: What was the agreement among annotators' labels that were used as the ground truth in this study, what was the classification performance to detect the tweets relevant to Zika, and how well were the classifiers able to distinguish between tweets on the different disease characteristics?

R3. Topical Analysis: What were the main discussion topics in each of these categories, and what were the most persistent concerns or misconceptions regarding the Zika virus?

Methods

In this exploratory study, a combination of natural language processing and machine learning techniques was used to determine what information about Zika symptoms, transmission, prevention, and treatment people were discussing on Twitter. Specifically, a 2-stage classifier system was built for finding relevant tweets on Zika and then categorizing these into 4 disease categories: symptoms, transmission, prevention, and treatment (Figure 1).

Dataset Distribution (Addressing R1)

Data Collection

Tweets were collected between February 24, 2016 and April 27, 2016 for a total of 1,234,605 tweets using Twitris 2.0 [15]. During this time frame, a lot of people were tweeting about

their concern about hosting the Olympics, new information about Zika was being found weekly, and it was right after Zika was linked to microcephaly and Guillain-Barré syndrome. We used a streaming application program interface (API) from the Twitris system [15] to collect the tweets, which means we only had access to a small percent of the tweets. We initially started the search using only the keyword “Zika” but quickly realized that the search was capturing a large number of tweets unrelated to Zika virus. We then created a semantic concept called Zika that utilized 2 terms “Zika” and “Zika virus,” which improved the quality of tweets for the data collection. This may be due to the fact that the Twitter streaming API allows collection of around 1% of the total tweets streaming at a given time [16]. Finally, the keyword “treatment” was added to the Zika concept as there were hardly any tweets about treatment, which was not surprising because there is currently no treatment for Zika. Adding the keyword “treatment” allowed us to check for social media responses to the significant drug and vaccine research being implemented during the time of tweet collection. The other disease-related category titles (prevention, transmission, and symptoms) did not need to be included as keywords as we observed that more than enough tweets were being collected for those categories.

Labeling Process and Data Annotation

Three microbiology and immunology experts annotated 1467 random tweets as being relevant or nonrelevant. Tweets were considered relevant if it contained information about Zika and the focus of the tweet was on Zika. For example, “Millions of GM mosquitoes to fight Zika virus in Caymans” was annotated as relevant as the tweet is about using genetically modified (GM) mosquitoes to fight Zika, whereas “#MoreTrustedThanHillary going to Brazil during Zika virus season” was annotated as nonrelevant as the focus of the tweet is on making fun of Hillary Clinton and is not about Zika. The relevant tweets were then further categorized as pertaining to the topic of (1) symptoms, (2) treatment, (3) transmission, or (4) prevention by the same 3 experts. Tweets were categorized as “symptoms” if they pertained to any of the symptoms associated with Zika as seen in this tweet: “WHO sees scientific

consensus on Zika virus as cause for disorders.” Tweets were categorized as “treatment” if they mentioned the fact that there is no treatment, research related to treatments, or included information about fake treatments. Here is an example of a treatment tweet: “Zika virus cloned in step toward vaccine.” Tweets were categorized as “transmission” if they mentioned modes of transmission, mosquitoes, or the Olympics. Here is an example of a transmission tweet: “Zika virus strain responsible for the outbreaks in Brazil has been detected in Africa.” Finally, tweets were categorized as “prevention” if they discussed ways to prevent the spread of Zika, or funding to fight Zika. Here is an example of a prevention tweet: “Senate Nears Deal for at Least \$1.1 Billion to Fight Zika Virus.” These 4 categories were used because they are characteristics of disease used in many medical journals and by the CDC and WHO. Fleiss kappa [17] was used to quantify the interrater reliability of our expert annotators.

Preprocessing

Before analysis, the data were preprocessed to remove the URL, screen handles (@username), retweet indicators, and non-ascii characters. Data were further normalized by removing capital letters, numbers, punctuations, and whitespaces from the tweets. Terms were filtered out to remove single characters like “d,” “e,” which do not convey any meaning about the topics in the corpus, and top words like “and,” “so,” etc were removed for the classification stage. Each tweet was represented as a feature vector of the words present in the tweet using unigrams.

Classification Performance (Addressing R2)

Supervised classification techniques including the decision tree (J48), multinomial Naive Bayes (MNB), Bayesian networks (Bayes Net), sequential minimal optimization (SMO) using support vector machine (SVM), Adaboost, as well as bagging or bootstrapping (Bagging) techniques were implemented on the Zika dataset for (1) classifying whether a tweet was relevant or nonrelevant, and (2) if relevant, further categorizing the tweets into the disease characteristics. Supervised techniques rely on labeled data, in this case tweets that are manually labeled as relevant to Zika virus, as well as the category it belongs to: Zika symptoms, Zika treatment, Zika transmission, and Zika prevention. They “learn” the nature of the tweets in the different groups and subgroups.

The performance of each classifier was assessed using the tenfold cross-validation, which is a commonly used method for the evaluation of classification algorithms that diminishes the bias in the estimation of classifier performance [18]. This approach uses the entire dataset for both training and testing, and is especially useful when the manually labeled dataset is relatively small. The study reports the average of the precision, recall, F-scores, and area under the curve (AUC) as measures of classification performance.

Topical Analysis (Addressing R3)

Studies such as Hong and Davison [19] have shown the utility of using traditional topic modeling methods like latent Dirichlet allocation (LDA) for grouping of themes occurring in short text documents. The basic idea in LDA is that 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 [20]. In this study, we use topic modeling for finding the underlying topics in each of the 4 disease characteristics to facilitate more detailed qualitative exploration of the types of discussions that occur within each disease characteristic.

Perplexity is a common measure to evaluate the topic models generated by LDA [21]. We use this measure to evaluate the topic modeling results by testing out different numbers of topic models from 2 to 20 for all 4 disease categories—symptoms, transmission, prevention, and treatment—using the well-established 10-fold cross-validation technique to ensure repeatability as well as generalizability.

Results

Dataset Distribution (Addressing R1)

Overall, 41.88% (517,070/1,234,605) of tweets contained a retweet and 84.60% (1,044,489/1,234,605) contained a URL. Tweets by gender were found by twitter usernames using the genderize API [22]. According to genderize, 28.47% (351,453/1,234,605) of the tweets were by males, 23.02% (284,207/1,234,605) by females, and 48.51% (598,945/1,234,605) were by unknown gender. The polarity of the individual tweets was found using the sentiment package in R [23] (Figure 2). The polarity of the tweets between males and females was similar. We further found a class imbalance in the categories (Figure 3). As there is no treatment for Zika, not many people tweeted about it. Transmission and prevention tweets were most frequent, indicating that they were the most discussed topics concerning Zika.

Classification Performance (Addressing R2)

In the first stage of the categorization process for the ground truth tweets, tweets were first classified as being relevant or not relevant to Zika. Tweets that were relevant were then categorized as being about symptoms, treatment, transmission, or prevention. To train the classifiers and evaluate their performance, 1467 tweets were manually labeled. Figure 4 provides the distribution of the relevant tweets in the 4 categories. As seen from Figure 3, the distribution of the labeled gold standard dataset was similar to the distribution of the large data corpus, except for a larger portion of tweets related to treatment.

Figure 2. Polarity and proportion of tweets divided in the gender categories.

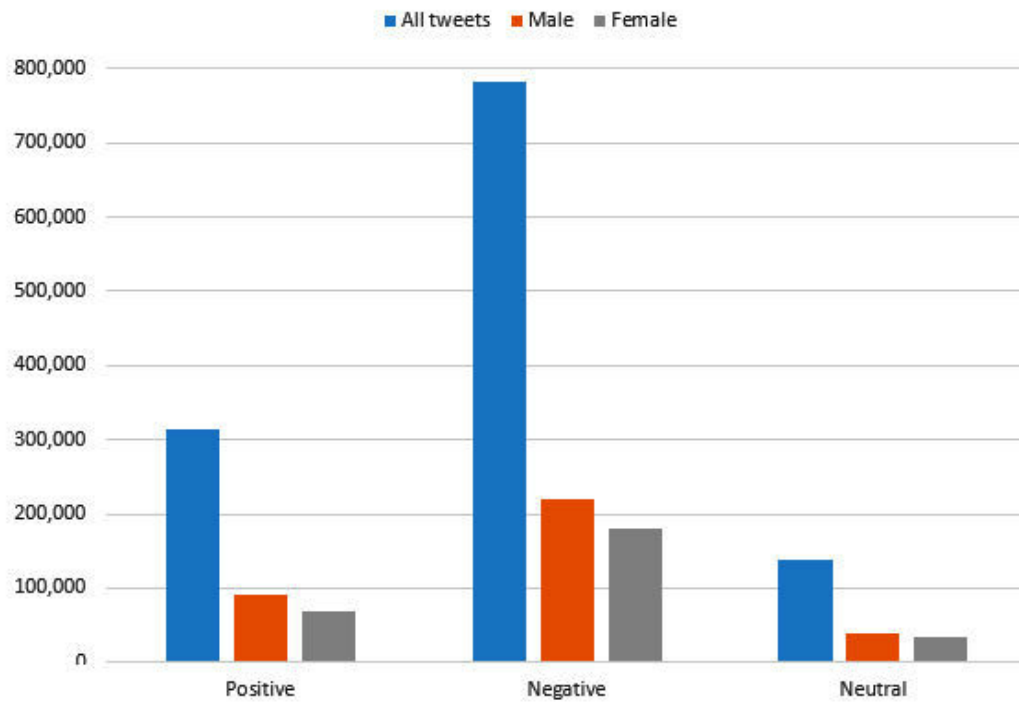


Figure 3. Number of tweets in each disease category after classifying all tweets (1.2 million tweets) using the best classification model multinomial Naive Bayes (discussed in the Classification and Performance Using 10-fold Cross-Validation section).

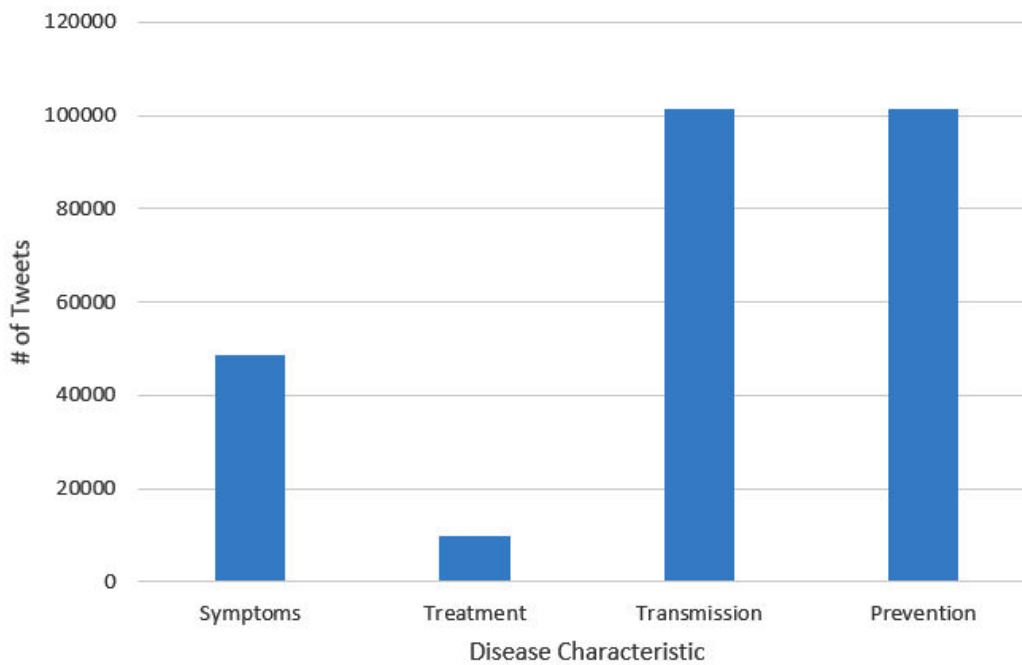
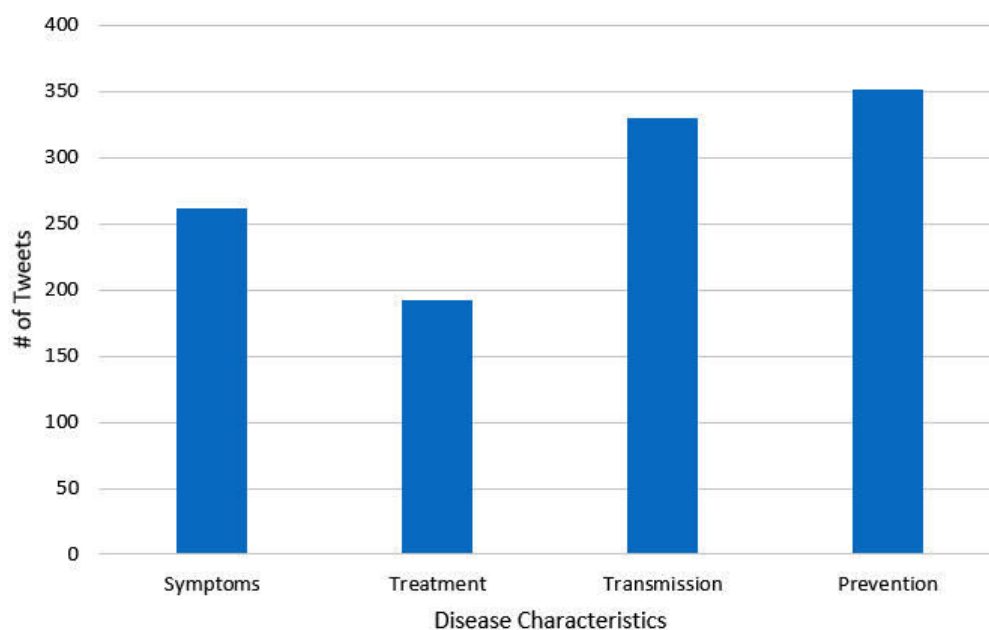


Figure 4. Number of tweets from the labeled dataset for each of the 4 categories of disease characteristics.**Table 1.** Different classifier performances for detecting relevant tweets using decision tree (J48), multinomial Naive Bayes (MNB), Bayesian networks (Bayes Net), sequential minimal optimization (SMO) using support vector machine (SVM), and bagging or bootstrapping (Bagging) techniques.

Classifier	TP ^a	FP ^b	Precision	Recall	F1 score	AUC ^c
J48	0.821	0.390	0.812	0.821	0.815	0.784
MNB (bayes)	0.880	0.368	0.881	0.880	0.868	0.943
Bayes Net	0.832	0.479	0.821	0.832	0.812	0.837
SMO	0.895	0.252	0.892	0.895	0.892	0.822
Bagging	0.857	0.411	0.852	0.857	0.843	0.877

^aTP: true positive.

^bFP: false positive.

^cAUC: area under the curve.

Interrater Reliability

Fleiss kappa values for relevant or not was .71. Fleiss kappa values for symptoms, treatment, transmission, and prevention were .93, .62, .92, and .87, respectively. This indicates substantial to almost perfect agreement among the raters [24]. Given substantial interrater reliability, a model needed to be built based on the gold standard dataset.

Classification and Performance Using Tenfold Cross-Validation

Table 1 gives the performance of different classifiers on the 1467 preprocessed Twitter data to find the relevancy of the tweet toward Zika. Unigram features were extracted from the texts using the Weka toolbox [25]. For this dataset, the classifiers performed well, with AUC values ranging from 0.78 to 0.94. MNB outperformed other classifiers based on the

F-measure (0.86) and AUC (0.94) (Table 1). MNB classifiers perform better for data sets that have a large variance in document length (in this case, the length of the tweets) by incorporating the evidence of each appearing word into its model [26].

The class imbalance was affecting the classifier performance. Although the AUC value was high (0.94), the classifier predicted a tweet was relevant more often than not relevant as 77.44% (1136/1467) of the tweets belonged to the relevant category.

Table 2 gives the performance of different classifiers on 1135 preprocessed Twitter data to find the categorical classification (symptoms, treatment, transmission, and prevention) of the tweets. Again, the classifiers performed well with AUC values ranging from 0.83 to 0.94. With this dataset, MNB outperforms other classifiers again.

Table 2. Different classifier performances for detecting the 4 disease categories within the relevant tweets using decision tree (J48), multinomial Naive Bayes (MNB), Bayesian networks (Bayes Net), sequential minimal optimization (SMO) using support vector machine (SVM), as well as bagging or bootstrapping (Bagging) techniques.

Classifier	TP ^a	FP ^b	Precision	Recall	F1 score	AUC ^c
J48	0.694	0.122	0.702	0.694	0.695	0.838
MNB	0.784	0.084	0.787	0.784	0.785	0.940
Bayes Net	0.697	0.121	0.729	0.697	0.702	0.885
SMO (SVM)	0.775	0.088	0.780	0.775	0.777	0.877
Bagging	0.727	0.112	0.741	0.727	0.730	0.901

^aTP: true positive.

^bFP: false positive.

^cAUC: area under the curve.

Table 3. Precision, recall, and F-measure for each of the 4 disease characteristics.

Category	Symptoms	Treatment	Transmission	Prevention	Average
precision	0.98	0.97	0.86	0.94	0.94
Recall	0.81	0.97	0.88	0.83	0.87
F1 score	0.89	0.97	0.87	0.88	0.90

On the basis of the above results, the 2-stage classifier system was found to have a high precision and recall performance for categorizing the tweets into relevant and not relevant, and further classifying the relevant tweets into the 4 disease categories. Once the performance of the model based on the gold standard dataset was confirmed to have high precision and recall, the model needed to be tested on a new set of tweets.

Classification and Performance Based on Error Analysis Using Hold-Out Dataset

As a post hoc analysis of generalizability, 530 new tweets (also known as hold-out data) that were not included in the gold standard data set were analyzed using the 2-stage classifier model. High precision and recall values were obtained for the relevance classifier, with Precision =0.99 and Recall =0.99. Hence, the F-measure was also 0.99 (harmonic mean of precision and recall). This high performance of the classifier indicates that the gold standard dataset was a good representation of the distribution of the tweets in the large data corpus.

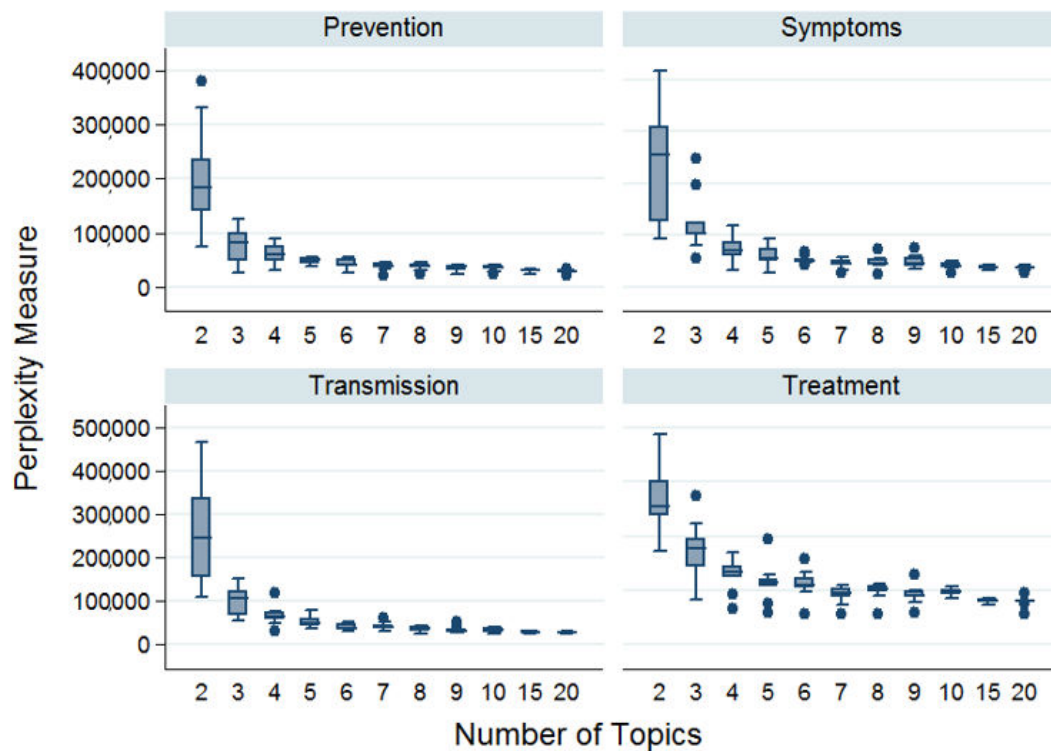
Even though the classes were unbalanced, high precision and recall values were still obtained for disease characteristics (Table

3) in the second classification stage. An overall high F-measure of 0.9 was obtained. This further indicated that the gold standard dataset was a good representation of the tweets, as well as the disease categories in the larger corpus.

The error analysis indicates that the classifiers performed well with the unseen test data and were generalizable enough to work with the large dataset. The dataset was further examined with a focus on the insights provided within each disease category. More specifically, the topics discussed on Twitter in the symptoms category were examined to discover the latent semantic topics discussed therein. The symptoms category was chosen because the researchers felt it was the topic of most concern due to all the defects associated with Zika.

Topical Analysis (Addressing R3)

From Figure 5, we observe that the perplexity values decrease rapidly until about 5, and then level off after 5 for all the 4 categories, indicating that increasing the number of topics after 5 does not significantly improve the performance of the LDA models (the lower the perplexity value the better). Therefore, the number of topics was restricted to 5 while discussing the topics for each category.

Figure 5. Prevention, symptoms, transmission, and treatment perplexity measure plots.

The results of LDA are discussed for each of the 4 disease characteristics in Tables 4-5. Topic modeling results are shared here [27] for the research community to examine the outcome of using topic modeling, as well as the overlap among the topics generated. First, the results for the 3 categories, that is, prevention, transmission, and treatment, will be discussed. Then, a more detailed analysis of the topic modeling results for the “symptoms” category will be discussed along with the misinformation tweeted by users within that domain.

Table 4 provides the topics for the 3 categories: (1) prevention, (2) transmission, and (3) treatment, along with representative tweets within each topic.

Prevention: Within the prevention topics, topic #1 was *need to control and prevent spread*, topic #2 was *the need for money to combat mosquitoes and research treatments*, topic #3 was *ways to actually prevent spread*, topic #4 was *introducing a bill to get funds*, and topic #5 was *research* (Table 4).

Transmission: In transmission, there was a strong overlap in topics #1 (vector, ie, mosquitoes for Zika) and #4 (disease spread) that highlight the overlap between spread by mosquitoes and the spread of disease in general. Another topic (#2) was *sexual spread*, which is another mode of transmission besides through mosquitoes. The next topic (#3) was *infants*, who are most affected by this epidemic due to the risk of microcephaly. The final topic (#5) was *sports*.

Treatment: There was a slight overlap between topics #1 (lack of treatment) and #3 (vaccine development) primarily due to

the large co-occurrence of the word “vaccine” in both these topics. *Blood testing* (#4) was another major topic as some people got infected with Zika after receiving a blood transfusion. As no treatment exists, a lot of research is focused on developing a drug for Zika, which is why *test development* (#5) was the final topic.

Symptoms: In the topic model results for symptoms, topics #1 (zika effects), #2 (brain defects), and #4 (zika scarier than thought) were well separated, whereas topics #3 (confirmation of defects) and #5 (initial reports) overlap significantly (Figure 6 and Table 5). The topics are described in Table 5. Topics #3 and #5 overlap for symptoms because a lot of the initial reports for different locations were about new cases of microcephaly in that location as seen in this example tweet: “Colombia Reports First Cases of Microcephaly Linked to Zika Virus.” Topic #3 more strongly addressed the defects that were confirmed, whereas topic #5 focused on where reports came from.

In this section, the topic modeling results generate insightful results that allow researchers to understand the citizens’ concerns, as well as the spread of misinformation. According to the theory of LDA, each topic represents certain common properties that reflect the pattern in the tweets. Finding out the exact meanings of the topics requires additional information and domain knowledge. We see that for each of the disease characteristics, the discovered topics can be interpreted straightforwardly through the lens of domain-specific knowledge about Zika.

Table 4. Prevention, transmission, and treatment topic modeling results.

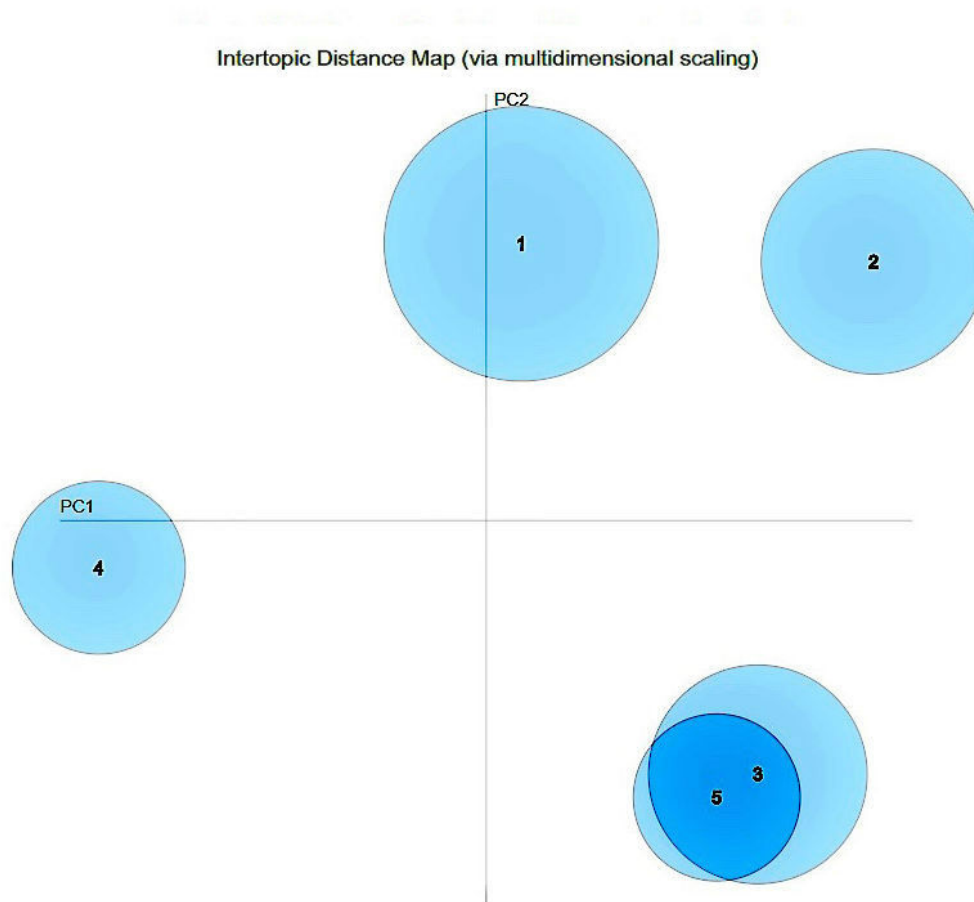
Disease characteristic	Topic	Sample tweets for each topic
Prevention	(#1) Control	RT ^a @DrFriedenCDC: A2. The best way to prevent #Zika & other diseases spread by mosquitoes is to protect yourself from mosquito bites. #Reut
	(#2) Money need	#healthy Congress has not yet acted on Obama's \$2 billion in emergency funding for Zika, submitted in February
	(#3) Prevention	RT @bmj_latest: Couples at risk from exposure to Zika virus should consider delaying pregnancy, says @CDCgov
	(#4) Bill	https://t.co/Ke12LOdypf Senate Approves \$1.1 Billion In Funding To Fight The Zika Virus #NYCnowApp
	(#5) Research	Florida is among those at greatest risk for Zika. @FLGovScott's sweeping abortion bill blocks scientists' access to conduct research
Transmission	(#1) Vectors (mosquitoes)	This map shows the Northeast is at risk for Zika mosquitoes this summer
	(#2) Sexual	@user1 First Sexually Transmitted Case Of Zika Virus In U.S. Confirmed
	(#3) Infants	CDC ^b reports 157 cases of U.S. pregnant women infected with Zika virus.
	(#4) Spread	Zika strain from Americas outbreak spreads in Africa for first time: WHO ^c (Update)
	(#5) Sports	MLB ^d moves games from Puerto Rico due to Zika concerns....uh..what about the Olympics?? Can't be good.
Treatment	(#1) Lack of treatment	RT @DrFriedenCDC: Much is still unknown about #Zika and there is no current medicine for treatment or vaccine to prevent the virus.
	(#2) Zika test	Rapid Zika Test Is Introduced by Researchers The test, done with a piece of paper that changes color if the virus...
	(#3) Vaccine development	Researchers discover structure of Zika virus, a key discovery in development of antiviral treatments and vaccines
	(#4) Blood test	Experimental blood test for Zika screening approved
	(#5) Test development	New mouse model leads way for #Zika drug, vaccine tests

^aRT: ReTweet.^bCDC: Center for Disease Control.^cWHO: World Health Organization.^dMLB: Major League Baseball.**Table 5.** Symptoms topic modeling results.

Topic	Words	Tweets
(#1) Zika effects	infect, babies, mosquito, cause, microcephaly, symptom, pregnancy	RT ^a @USATODAYhealth: Zika affects babies even in later stages of pregnancy. Microcephaly seen in babies from moms infected in 6th month
(#2) Brain defects	brain, link, studies, microcephaly, baby, disorder, cause, damage, infect, fetal	Zika Virus May Cause Microcephaly by Hijacking Human Immune Molecule: Fetal brain model provides first clues on how Z...
(#3) Confirmed defects	defect, cause, birth, confirm, health, severe, link, official	Enough conspiracy theories; nature is nasty enough: U.S. health officials confirm Zika cause of severe birth defects
(#4) Scariest than thought	scariest, than, thought, us, official, health, CDC ^b , warn, learn, first	#breakingnews Zika Virus "Scariest Than We First Thought," Warn US Health Officials
(#5) Initial reports	first, report, death, case, puerto, confirm, rico, cause, colombia, defect	Colombia Reports First Cases of Microcephaly Linked to Zika Virus—Sun Jan 09 15:13:20 EST

^aRT: ReTweet.^bCDC: Center for Disease Control.

Figure 6. A 2-dimensional principal components plot of topics discussed pertaining to Zika symptoms.



Discussion

The number of tweets and polarity of tweets were similar between male and female Twitter users with the majority of tweets being negative. There was a similar class imbalance in the random sample of labeled tweets and total corpus showing that the gold standard labeled by experts was an accurate representation of the total corpus. The 2-stage classifier performed well for both levels (relevancy and disease characteristics). Finally, the most persistent topics concerning the disease characteristics were uncovered using topic models.

Sentiment Analysis and Word Polarity

Although a majority of the tweets were categorized as having negative polarity, the percentage of positive tweets was higher than expected. Some examples of tweets that were classified as positive are as follows: “Case report: assoc btw #Zika/teratogenicity strengthened & evidence shows impact on fetus may take time to manifest,” “RT : At recent int’l meeting about , experts exchanged insights, identified knowledge gaps, and agreed on a plan,” and “91,387 Cases of Zika Confirmed in Brazil This Year: Brazil has confirmed 91,387 cases of...” Words such as “strengthened,” “agreed,” and “confirmed” may be why some tweets were being classified as positive. Sentiment analysis is complex as most sentiment analysis tools just use the word “polarities.” However, contextual information needs to be incorporated for topic-specific sentiment analysis in this

domain. We are currently looking into this issue but leave detailed analysis and discussion to a future study.

Classification Analysis

One of the interesting findings of our analysis was the fact that the MNB classifier outperformed the other more popular classifiers in text analytics: random forest (J48) and SVM. According to 1 study [28], this has to do with the class imbalance issue in our dataset, for both the first (relevancy) and second (disease category) stage of our classifier. This also highlights the possible orthogonality of the features used in our study: the unigrams. Specifically, in this dataset, measuring the likelihood of the features in a given class independently outperforms other complex models such as J48 and SVM. This possibly also relates to the fact that the data are less noisy as they have been evaluated by expert annotators. Naive Bayes is one of the simplest classification models available to us, but it is nonetheless among the most effective for this dataset. This result is non intuitive but not surprising when we consider that using text for classification is relatively imprecise compared with other types of data. In datasets with large amounts of error, simpler models are less likely to overfit the data. Hence we recommend that future research on text analytics begin with Naive Bayes and then proceed to using more complex models to see if these actually improve classification accuracy.

Annotation Observations

One major issue when annotating tweets was what to do about news tweets like this one: “Your Wednesday Briefing: Bernie

Sanders, Hillary Clinton, Zika Virus: Here's what you need to know to start..." The issue was that this does give relevant information about Zika in that it tells what news stations were discussing and what else was going on at the same time as the Zika outbreak. However, the tweet itself does not give any information about Zika symptoms, treatment, transmission, and prevention. The expert annotators indicated that they decided to code these tweets as relevant because they were about Zika, but we did not include them in the disease characteristics annotations as they do not have any information about the disease characteristics that we outlined *a priori* as our domain of inquiry. This said, they did decide to include information about Zika and sporting events because these could be sources of transmission from athletes and fans not taking proper precautions. We recognize, however, that the sporting context may not have been viewed as important had the Olympics not occurred during the same time as our data collection. Any qualitative deductive coding scheme is underlain by specific assumptions and theoretical constraints that can be highly context-specific, and we feel that it is important for research using citizen sensing to incorporate experts who are able to delineate scientifically accepted contextual boundaries for inquiry.

Topic Modeling

The perplexity plot (Figure 5) indicates that while we could use a larger number of topics for very small improvements, using a number of topics greater than 5 quickly becomes a case of diminishing returns, especially if we choose to use a parsimonious model to represent our data. Moreover, as we wanted to conduct an exploratory analysis of the topics for this study, the results rationalize our choice of 5 topics.

The emergent topics in prevention (need for control and prevent spread, need for money, ways to prevent spread, bill to get funds, and research) were not surprising considering that there has been much discussion about how to prevent Zika, the need for funding to prevent Zika, and the research required to find a cure for Zika as it is an emerging disease. There is also a need to better understand Zika virus, the disease it causes, and ways to combat it [29]. Looking at the tweets for topics #1 and #4 in transmission, both highlight the concerns and risks associated with Zika spread, which is most likely why they both overlap. Sports was most likely a topic because the tweets were collected during baseball season and just before the Olympics, and many athletes were concerned about getting infected with Zika while competing in the 2016 Olympics in Rio de Janeiro.

In symptoms, topics #1, #3, and #5 were closely related in that they addressed the defects caused by Zika, but nonetheless point to slightly different concepts (Table 5). For example, microcephaly is not the only defect; there is also Guillain-Barré, which would be topic #1. Topic #2 focuses on microcephaly because that is perhaps the most persistent concern related to Zika. Such discussion is seen in these tweets: "RT @USATODAYhealth: Zika affects babies even in later stages of pregnancy. Microcephaly seen in babies from moms infected in 6th month" and "Zika Virus May Cause Microcephaly by Hijacking Human Immune Molecule: Fetal brain model provides first clues on how Z..." Topic #3 contains tweets that occurred

when the defects were confirmed to be caused by Zika and not something else: "Zika linked to fetal brain damage: Finnish study: infectious Zika virus from fetal tissue in cell culture. The virus," and "Enough conspiracy theories; nature is nasty enough: U.S. health officials confirm Zika cause of severe birth defects."

Topic #4 for symptoms was primarily generated through discussion of a British Broadcasting Corporation article [30] on how more birth defects have been linked to Zika and that the virus was expected to travel further than initially thought, leading to experts admitting that Zika is scarier than was first thought. The statement of Zika being "scarier than we first thought" by the CDC was a big topic on Twitter: "CDC says zika virus scarier than thought as US prepares for outbreak: On Monday, the U.S. Centers for Disease..." This also affected the US political environment: "#2016elections U.S. Officials Warn Zika Scarier Than Initially Thought: By Timothy Gardner and Jeff Mason WA..." This discussion led to additional tweets about the danger of Zika virus: "The Edge: Zika Is Now Even More Terrifying," "Zika virus 'shrinks brains' in tests," and "#Zika Survivor Says 'I Could Feel My Skin Shrinking' CBS Boston." These tweets demonstrate how a statement by the CDC can be spread and how users can tweak the wording of these CDC statements to generate more concern than is warranted by the actual impact of the disease. Finally, topic #5 includes tweets about initial reports of Zika outbreaks and deaths.

Within symptoms, several tweets in topic #1 were calling Zika a hoax, "Zika HOAX exposed by South American doctors: Brain deformations caused by larvicide chemical," "The Zika Virus is a hoax! It is like calling the common cold an epidemic. It's what they put in the drinking water," and "CDC likely fabricating link between Zika virus and microcephaly cases." However, the CDC has stated multiple times that Zika and microcephaly are definitely linked: "CDC: Zika definitely causes severe birth defects" and "Here's a #Zika basic: Zika infection during pregnancy can cause some severe birth defects." Some of the people saying Zika is a hoax are misunderstanding this quote from the CDC: "People usually don't get sick enough to go to the hospital, and they very rarely die of Zika. For this reason, many people might not realize they have been infected. Once a person has been infected, he or she is likely protected from future infections." This statement is true for the majority of healthy adults. However, for infants it can cause microcephaly and in some cases Guillain-Barré syndrome in healthy adults: "Symptoms of Guillain-Barré syndrome include weaknesses in arms & legs. GBS is linked w/ #Zika." There also have already been multiple deaths due to Zika as was detailed in topic #5. The CDC has also been directly answering questions about Zika on Twitter. One user tweeted at the CDC, "Why is of particular concern to women who are pregnant or considering becoming pregnant?," to which the CDC responded, "Zika infection in pregnancy can cause microcephaly and other severe brain defects. ." This shows that while some misinformation is still getting tweeted, the CDC is working to get the correct information out there. This is useful because it shows that the CDC could potentially target specific user groups directly through our classifier and topic modeling approach, and respond to users within a topic group with a similar response that can

allow correct information to get transmitted to a larger population with less effort.

All of the topics under the different disease characteristics fit the characteristic. For example, control, money need, prevention, bill, and research were all major topics of prevention discussions. This indicates that the classification model accurately labeled tweets. It also indicates that tweets about major topics were collected and accurately reflected in our topic model. Also, whereas all 4 disease characteristics are important, symptoms was discussed in detail because the researchers felt it included the most important information for public health officials to know especially once the misconceptions and misinformation, such as Zika being a hoax, were found. Categorizing the symptoms into the different topics using topic modeling also allowed us to get deeper into the themes within the symptoms category. This can allow a more targeted interaction with agencies like CDC and specific users to provide interventions against the spread of misinformation. If we are able to make the persistent misconceptions that people have about Zika clear, then public health agencies can inform accordingly.

Limitations

Although we feel that our methods and findings are trustworthy and robust, we would like to point out some limitations we face in our dataset, and the use of social media.

Language Constraint

We have restricted our study to English-language tweets, which certainly limits the strength of our study. This is more critical to address given that South American countries were initially affected by Zika. This also restricts our analysis of measuring disease outbreak, which is why we refrained from doing so in our study. Future studies could address this limitation through analysis of tweets written in Spanish or Portuguese.

Keyword Constraint

As described in the Data Collection section, we used the keywords *Zika*, *Zika virus*, *Zika treatment*, and *Zika virus treatment* in our study. Hence, we can expect that this search would overlook tweets that referred to the disease in a different name or talked about the disease without using the word Zika. The keywords *Zika treatment* and *Zika virus treatment* were added because there were few tweets about the treatment of Zika. This was not surprising as there is currently no treatment for Zika. By including those 2 keywords, researchers could download relevant tweets containing those keywords. This was done because we still felt that treatment needed to be included because there was ongoing drug and vaccine research being implemented during the time of tweet collection. From a preliminary manual data analysis, we observed that the other category titles did not need to be included as keywords because more than enough tweets were being collected for those categories. One interesting observation here is that although the

keyword “treatment” was part of the crawling process, the treatment subcategory was still the smallest class in the distribution of the dataset (see Figures 3 and 4).

Gender and Polarity Constraint

Only 51.49% (635,660/1,234,605) of the tweets were labeled by the gender API using the profile name (Figure 2). Similarly, 11.14% (137,536/1,234,605) of the tweets were not labeled on their polarity. Given the concerns about infant microcephaly and sexual transmission, gender is an important factor to consider when contextualizing discussions around Zika. Gender needs to be addressed moving forward with this study by creating a customized gender recognition tool using machine learning specifically for Twitter data.

Conclusion

The proportion of tweets between male and female Twitter users was similar by number of tweets in general and by polarity. The majority of tweets were negative but there were more positive tweets than expected, which may be due to the use of positive words such as “strengthened,” “agreed,” and “confirmed.” There was a class imbalance in the ground truth and overall tweets; however, the imbalance was similar between the two, showing that the tweets used in the ground truth were a good representation of the tweets overall. There were hardly any tweets about treatment, which was not surprising because there is no treatment for Zika. The classification performance was very high for relevancy ($F=0.86$) and disease characteristics ($F=0.94$) for the ground truth ($F=0.99$) and for the overall tweets ($F=0.90$). The 5 topics for prevention were control, money need, prevention, bill, and research. The 5 topics for transmission were vectors (mosquitoes), sexual transmission, infants, spread, and sports. The 5 topics for treatment were lack of treatment, Zika test, vaccine development, blood test, and test development. Finally, the 5 topics for symptoms were Zika effects, brain defects, confirmed defects, scarier than thought, and initial reports.

This is one of the first studies to report successful creation of an automated content classification tool to analyze Zika-related tweets, specifically in the area of epidemiology. Through citizen sensing, such a system will help advance the field’s technological and methodological capabilities to harness social media sources for disease surveillance research.

Future Work

Future studies should include creation and evaluation of an automated technique to detect misinformation using tweets to allow for well-targeted, timely interventions. Such a platform will generate data on emerging temporal trends for more timely interventions and policy responses to misinformation on Zika. We would encourage such studies to leverage multiple information sources including blogs, news articles, as well as social media.

Acknowledgments

We would like to thank Dr Megan Rúa for her help with interpreting the topic models, Scott Holdgreve, Ryan, Becker, and Dr. Amber Todd for their help with annotating, and Sampath Gogineni for his help with some of the initial analysis. Romine and Miller would like to acknowledge support from the Department of Education I3 project U411C140081 and the Institute of Educational Sciences (IES) award R305A150364. Twitris is funded through the National Science Foundation award IIP 1542911. Banerjee and Muppalla are funded through NIH 1K01LM012439-01. Sheth is funded through NIH grant 1 R01 HD087132-01.

Conflicts of Interest

None declared.

References

1. Nanlong M. AllAfrica. 2016. Nigeriabola - two die after drinking salt water in Jos URL: <http://allafrica.com/stories/201408111640.html> [accessed 2016-12-12] [WebCite Cache ID 6mhjzekwJ]
2. Centers for Disease Control and Prevention (CDC). CDC. 2016. Transcript for CDC telebriefing: Zika summit press conference 2016 URL: <https://www.cdc.gov/media/releases/2016/t0404-zika-summit.html> [accessed 2017-06-06] [WebCite Cache ID 6r1Oj46i9]
3. Berg N. Greenbiz. 2013. How citizens have become sensors URL: <https://www.greenbiz.com/news/2013/03/20/how-citizens-have-become-sensors> [WebCite Cache ID 6mhnJvcJ8]
4. Tran T, Lee K. Understanding citizen reactions and Ebola-related information propagation on social media. 2016 Presented at: International Conference on Advances in Social Networks Analysis and Mining; August 18, 2016; San Francisco.
5. Purohit H, Banerjee T, Hampton A, Shalin V, Bhandutia N, Sheth A. Arxiv. 2016. Gender-based violence in 140 characters or fewer: a #BigData case study of Twitter URL: <https://arxiv.org/abs/1503.02086> [accessed 2017-06-10] [WebCite Cache ID 6r7WfwX7K]
6. Paul M, Dredze M. You are what you tweet: analyzing twitter for public health. 2011 Presented at: ICWSM; July 2011; Barcelona.
7. Bhattacharya S, Tran H, Srinivasan P. Discovering health beliefs in twitter. 2012 Presented at: AAAI- Fall Symposium on Information Retrieval Knowledge Discovery in Biomedical Text; 2012; Washington, DC.
8. Center for Disease Control (CDC). CDC. Zika virus URL: <http://www.cdc.gov/zika/> [accessed 2016-12-12] [WebCite Cache ID 6mhnTZk4b]
9. Wong SS, Poon RW, Wong SC. Zika virus infection-the next wave after dengue? J Formos Med Assoc 2016 Apr;115(4):226-242 [FREE Full text] [doi: [10.1016/j.jfma.2016.02.002](https://doi.org/10.1016/j.jfma.2016.02.002)] [Medline: [26965962](https://pubmed.ncbi.nlm.nih.gov/26965962/)]
10. Oyeyemi S, Gabarron E, Wynn R. Ebola, Twitter, and misinformation: a dangerous combination? BMJ 2014;349:g6178. [doi: [10.1136/bmj.g6178](https://doi.org/10.1136/bmj.g6178)]
11. Majumder MS, Santillana M, Mekar SR, McGinnis DP, Khan K, Brownstein JS. Utilizing nontraditional data sources for near real-time estimation of transmission dynamics during the 2015-2016 Colombian Zika virus disease outbreak. JMIR Public Health Surveill 2016 Jun 01;2(1):e30 [FREE Full text] [doi: [10.2196/publichealth.5814](https://doi.org/10.2196/publichealth.5814)] [Medline: [27251981](https://pubmed.ncbi.nlm.nih.gov/27251981/)]
12. Khatua A, Khatua A. Immediate and long-term effects of 2016 Zika outbreak: a twitter-based study. 2016 Presented at: IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom); 2016; Munich p. 1-6. [doi: [10.1109/HealthCom.2016.7749496](https://doi.org/10.1109/HealthCom.2016.7749496)]
13. Glowacki EM, Lazard AJ, Wilcox GB, 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/)]
14. Fu K, Liang H, Saroha N, Tse ZT, Ip P, Fung IC. How people react to Zika virus outbreaks on twitter? a computational content analysis. Am J Infect Control 2016 Dec 01;44(12):1700-1702. [doi: [10.1016/j.ajic.2016.04.253](https://doi.org/10.1016/j.ajic.2016.04.253)] [Medline: [27566874](https://pubmed.ncbi.nlm.nih.gov/27566874/)]
15. Jadhav A, Purohit H, Kapanipathi P, Ananthram P, Ranabahu A, Nguyen V, et al. Twitris 2.0: semantically empowered system for understanding perceptions from social data. 2010 Presented at: Semantic Web Application Challenge at ISWC; 2010; Shanghai, China.
16. Twitter. Twitter streaming API information URL: <https://dev.twitter.com/streaming/overview/processing> [accessed 2017-02-22] [WebCite Cache ID 6oT3cftAk]
17. McHugh ML. 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/)]
18. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: Data mining, inference, and prediction. New York: Springer; 2009.
19. Hong L, Davison B. Empirical study of topic modeling in twitter. 2010 Presented at: Proceedings of the first workshop on social media analytics; 2010; Washington, DC p. 80-88.
20. Blei D, Ng A, Jordan M. Latent dirichlet allocation. J Mach Learn Res 2003;3:1022.
21. Chemudugunta C, Smyth P, Steyvers M. NIPS. 2006. Modeling general and specific aspects of documents with a probabilistic topic model URL: <https://papers.nips.cc/paper/>

- [2994-modeling-general-and-specific-aspects-of-documents-with-a-probabilistic-topic-model.pdf](#) [accessed 2017-06-10] [WebCite Cache ID 6r7XaxpYZ]
22. Genderize. Genderize API URL: <https://genderize.io/> [accessed 2017-06-06] [WebCite Cache ID 6r1WPw9BO]
 23. R-project. Sentiment URL: <https://cran.r-project.org/src/contrib/Archive/sentiment/> [WebCite Cache ID 6mhnirAFy]
 24. Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):174.
 25. Frank E, Hall M, Witten I. The WEKA workbench. In: *Data mining: Practical machine learning tools and techniques*. Burlington: Morgan Kaufmann; 2016.
 26. McCallum A, Nigam K. A comparison of event models for naive bayes text classification. 1998 Presented at: AAAI-98 workshop on learning for text categorization; 1998; Madison p. 41-48.
 27. Bitbucket. Zika study URL: <https://roopteja.bitbucket.io/zikastudy/> [accessed 2017-03-07] [WebCite Cache ID 6onCPDSVm]
 28. Frank E, Bouckaert R. Naive bayes for text classification with unbalanced classes. 2006 Presented at: PKDD; 2006; Berlin.
 29. National Institute of Allergy Infectious Disease (NIAID). NIAID.NIH. Zika virus URL: <https://www.niaid.nih.gov/diseases-conditions/zika-virus> [accessed 2016-12-12] [WebCite Cache ID 6mhnqsP3D]
 30. British Broadcasting Corporation (BBC). BBC. Zika virus scarier than thought? says US URL: <http://www.bbc.com/news/world-us-canada-36020165> [accessed 2016-12-12] [WebCite Cache ID 6mhp7CZwI]

Abbreviations

- API:** application program interface
AUC: area under the curve
CDC: Center for Disease Control
DC: disease classifier
GBS: Guillain-Barré syndrome
J48: decision tree
LDA: latent Dirichlet allocation
MNB: Multinomial Naive Bayes
RC: relevancy classifier
SMO: sequential minimal optimization
SVM: support vector machine
WHO: World Health Organization

Edited by G Eysenbach; submitted 13.12.16; peer-reviewed by D Mowery, J Bian, MA Mayer; comments to author 11.02.17; revised version received 08.03.17; accepted 14.04.17; published 19.06.17.

Please cite as:

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;3(2):e38

URL: <http://publichealth.jmir.org/2017/2/e38/>

doi: [10.2196/publichealth.7157](https://doi.org/10.2196/publichealth.7157)

PMID: [28630032](https://pubmed.ncbi.nlm.nih.gov/28630032/)

©Michele Miller, Tanvi Banerjee, Roopteja Muppalla, William Romine, Amit Sheth. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 19.06.2017. 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

Measuring Audience Engagement for Public Health Twitter Chats: Insights From #LiveFitNOLA

Kristina M Rabbarison^{1*}, DrPH, MS; Merriah A Croston¹, MPH; Naomi K Englar^{2*}, BA; Connie L Bish^{1*}, MPH, PhD; Shelbi M Flynn^{3*}, BA; Carolyn C Johnson^{2*}, PhD, FAAHB

¹Centers for Disease Control and Prevention, Division of Population Health, Atlanta, GA, United States

²Tulane University School of Public Health and Tropical Medicine, Tulane Prevention Research Center, New Orleans, LA, United States

³City of New Orleans Health Department, New Orleans, LA, United States

*these authors contributed equally

Corresponding Author:

Kristina M Rabbarison, DrPH, MS
Centers for Disease Control and Prevention
Division of Population Health
4770 Buford Hwy NE MS F78
Atlanta, GA, 30341
United States
Phone: 1 770 488 5810
Email: krabbarison@cdc.gov

Abstract

Background: Little empirical evidence exists on the effectiveness of using Twitter as a two-way communication tool for public health practice, such as Twitter chats.

Objective: We analyzed whether Twitter chats facilitate engagement in two-way communications between public health entities and their audience. We also describe how to measure two-way communications, incoming and outgoing mentions, between users in a protocol using free and publicly available tools (Symplur, OpenRefine, and Gephi).

Methods: We used a mixed-methods approach, social network analysis, and content analysis. The study population comprised individuals and organizations participating or who were mentioned in the first #LiveFitNOLA chat, during a 75-min period on March 5, 2015, from 12:00 PM to 1:15 PM Central Time. We assessed audience engagement in two-way communications with two metrics: engagement ratio and return on engagement (ROE).

Results: The #LiveFitNOLA chat had 744 tweets and 66 participants with an average of 11 tweets per participant. The resulting network had 134 network members and 474 engagements. The engagement ratios and ROEs for the #LiveFitNOLA organizers were 1:1, 40% (13/32) (@TulanePRC) and 2:1, -40% (-25/63) (@FitNOLA). Content analysis showed information sharing (63.9%, 314/491) and health information (27.9%, 137/491) as the most salient theme and sub-theme, respectively.

Conclusions: Our findings suggest Twitter chats facilitate audience engagement in two-way communications between public health entities and their audience. The #LiveFitNOLA organizers' engagement ratios and ROEs indicated a moderate level of engagement with their audience. The practical significance of the engagement ratio and ROE depends on the audience, context, scope, scale, and goal of a Twitter chat or other organized hashtag-based communications on Twitter.

(*JMIR Public Health Surveill* 2017;3(2):e34) doi:[10.2196/publichealth.7181](https://doi.org/10.2196/publichealth.7181)

KEYWORDS

social media; Twitter; Twitter chat; public health, communication; content analysis; social network analysis

Introduction

Approximately a quarter (23%) of American adult (age 18+) Internet users and a fifth (19%) of the entire American adult population use Twitter [1]. Twitter is a free social channel where registered individuals or organizations (ie, users) can share

140-character messages called tweets. There are 302 million active, monthly users and 500 million tweets posted per day [2]. Users are identified with usernames preceded by the "@" symbol. A retweet occurs when a user repeats another user's tweet; sometimes, it is designated with "RT." A hashtag is any

word or phrase preceded by the pound symbol (#); it enables users to organize tweets by topics.

Public health practitioners use Twitter for disease surveillance [3-7], information dissemination [8-11], emergency response [12-16], and community building [17-21]. Evidence shows public health entities need to better harness Twitter's potential as a two-way communication tool to establish relationships and increase audience engagement to improve the reach of their health promotion activities [22-24]. In response to this need, many public health practitioners now lead or participate in Twitter chats as part of their efforts to engage audiences in two-way communications. Currently, little empirical evidence exists on the effectiveness of using Twitter as a two-way communication tool for public health practice, and this warrants further research [24-30].

Commercial Twitter engagement measuring tools exist; many include metrics such as the number of "likes" and the click-through rate in their engagement metrics. However, they do not focus on the incoming and outgoing communications that occur when one user mentions another user in a tweet. In addition, they require a purchasing fee or a recurring membership fee.

Our primary aim was to analyze whether Twitter chats facilitate engagement in two-way communications between public health entities and their audience. We also describe how to measure this two-way communication, incoming and outgoing mentions, between users in a protocol using free and publicly available tools (Symplur [31], OpenRefine [32], and Gephi [33]). We chose the tools used in the protocol because they are user-friendly and do not require advanced analytical skills. The protocol's target audiences are public health practitioners, such as health communication specialists and social media managers who are comfortable with point and click applications.

Methods

Study Design

We used a case study of one Twitter chat (#LiveFitNOLA) organized by Tulane Prevention Research Center (@TulanePRC), a university-based research and education center, and the City of New Orleans Health Department's Fit NOLA Initiative (@FitNOLA). Twitter chats are interactive, organized, and curated communications on Twitter. They focus on a specific topic and take place at predesignated times.

With a mixed-method approach, social network analysis (SNA), and content analysis, we analyzed (1) whether Twitter chats facilitated audience engagement among public health entities and their audience and (2) whether functional themes were present in the engagements between Twitter chat participants.

Definition of Engagement

Twitter defines engagement as the number of times a user interacted with a tweet, including clicks, retweets, replies, follows, likes, links, cards, hashtags, embedded media, username, profile photo, or tweet expansion [34]. Similarly, studies on health-related Twitter engagement define engagement as the number of user-mentions (retweets and replies), favorites,

clicks, or detail expansions [28,29,35]. For the scope of our study, we define engagement in two-way communications as the number of incoming and outgoing mentions between users. A mention can be a direct mention, a retweet, or a reply. Assessing the number of incoming and outgoing mentions between public health entities and their audience during a Twitter chat is a first step in addressing the evidence gap on the effective use of Twitter as an audience engagement tool in two-way communications for public health entities.

#LiveFitNOLA

@TulanePRC and @FitNOLA organize, host, and curate a monthly Twitter chat using the hashtag #LiveFitNOLA. The chat is focused on health and wellness in New Orleans. It is used to inform and engage participants in open discussions about the New Orleans' culture of health. The target audiences for the #LiveFitNOLA chat are organizations and individuals in New Orleans interested in health and wellness. As hosts, @TulanePRC and @FitNOLA coordinated the chat plans and preparations, created visual promotions and tweets, and invited Twitter followers and community partners to participate.

The first #LiveFitNOLA Twitter chat occurred on March 5, 2015, between 12:00 PM and 1:00 PM Central Time (CT). *Health & Fitness Magazine* (@HealthFitMag), a Louisiana-based and operated magazine, participated as a guest host and determined the chat topic; approved all questions in advance; and invited their Twitter followers and community partners to participate. In addition, the Centers for Disease Control and Prevention's National Center for Chronic Disease Prevention and Health Promotion (@CDCChronic) played a supporting role by participating as a subject matter expert. @CDCChronic provided evidence-based information on healthy living but did not intend to initiate or seek discussions with chat participants. The chat organizers selected the chat's date and time (lunch hour on a weekday in New Orleans) after consulting other national organizations that host recurring, public health-related Twitter chats. Posting tweets during lunch hour toward the end of the week is a recommended best practice for social media scheduling, including Twitter chats [36-38].

Data and Study Population

We collected the #LiveFitNOLA Twitter chat transcript from the Symplur Healthcare Hashtag Project [31]. Symplur tracks and archives tweets associated with registered health care-related hashtags. Symplur has a free and publicly available interface, where anyone can register and search for health care-related hashtags. To access hashtag-specific Twitter transcripts, Symplur can be queried for defined time frames. Using OpenRefine, we created a directed, relational dataset with source and target data points from the #LiveFitNOLA transcript by isolating the @username from tweet content and removing all other data elements. OpenRefine is compatible with both Windows and Mac Operating Systems, and is a publicly available data-cleaning tool [32].

For this study, we collected the #LiveFitNOLA chat transcript during a 75-min period on March 5, 2015, from 12:00 PM to 1:15 PM CT. This time frame included the hour-long #LiveFitNOLA chat and 15 min of residual conversations. Thus,

our study population was composed of the individuals and organizations that participated and were mentioned in the #LiveFitNOLA chat during the previously defined time frame. We identified the chat participants' geographic location based on self-disclosed location listed in user profiles. We categorized the #LiveFitNOLA chat participants' locations as New Orleans, other US locations, and not available.

All information included in our analysis is secondary data obtained from publicly available Twitter data. The CDC Human Research Protection Office (HRPO) designated the research activity conducted in this case study as exempt—HRPO Exemption Determination for Protocol #6803 “#LiveFitNOLA: A dissemination and translation case study.” The Tulane Institutional Review Board determined this research activity does not constitute human subjects research—ID #15-857090U “#LiveFitNOLA: Monthly Twitter Chats – A Community Engagement Project.”

Social Network Analysis (SNA)

We conducted an SNA to measure and visualize engagement in two-way communications between #LiveFitNOLA chat organizers and other chat participants. The SNA allowed us to create an engagement ratio of outgoing to incoming mentions among Twitter chat participants, and a return on engagement measure (ROE). The engagement ratio and ROE provide a simple way for public health practitioners to measure audience engagement on Twitter.

An SNA is the study of relationships between a connection's source and its target within a specifically defined and bounded network [39,40]. We defined relationships among the #LiveFitNOLA participants as the engagement in incoming and outgoing communications occurring when one participant mentions another participant. Specifically, a connection occurred when a participant mentioned another user by including their @username in a tweet, whether by quoting a tweet, retweeting, or directly addressing another user. We conducted the SNA with Gephi, a free, publicly available, and interactive network visualization and exploration software. Gephi is compatible with both Windows and Mac Operating Systems [33].

In our study, SNA focused on simple network components and metrics, including node, edge, degree, in degree, out degree, and a number of communities. A node (ie, a network member) is an individual or an organization in a network; it is represented by a circle in a network map [39,40]. An edge (ie, a connection) is the relational tie between a source node and a target node; a line or an arrow between a source node and a target node represents an edge [39,40]. An arrow represents unidirectional communication between the two nodes, and points from the source node to the target node [39,40]. A line represents bidirectional communications between the source node and the target node [39,40]. The #LiveFitNOLA network was composed of chat participants and any other users mentioned during the chat. An edge occurs when a source participant mentioned a

network member [39,40]. The total number of edges, independent of direction, is called degree [39,40]. “In degree” is the total number of incoming edges or incoming mentions for a #LiveFitNOLA network member [39,40]. For example, the @TulanePRC in degree is the total number of times other #LiveFitNOLA participants mentioned @TulanePRC in their tweets. “Out degree” is the total number of outgoing edges or outgoing mentions for a #LiveFitNOLA network member [39,40]. For example, the @TulanePRC out degree is the total number of times @TulanePRC mentioned other #LiveFitNOLA network members in their tweets. We did not use other SNA metrics such as network density and centrality measures, which are difficult to interpret because of the practical scope of this case study. Instead, we developed two metrics—an engagement ratio and an ROE measure—to assess two-way communications, incoming and outgoing mentions, between public health entities, and their audience on Twitter.

The engagement ratio compares a network member's total number of outgoing mentions (out degree) with their total number of incoming mentions (Figure 1, where, @username out degree is the number of outgoing mentions for a network member and @username in degree is the number of incoming mentions for a network member).

The ROE is based on the concept of return on investment (ROI), an investment performance measure. Like ROI does for monetary investments, the ROE measures the engagement gain or loss generated relative to the amount of engagement invested. In other words, the engagement gain or loss calculated as incoming mentions related to the number of outgoing mentions invested (Figure 2).

To estimate the level of incoming and outgoing communications between public health entities and their audience, we focused our analysis on the #LiveFitNOLA Twitter chat organizers (@TulanePRC and @FitNOLA), the chat guest host (@HealthFitMag), and the supporting organization (@CDCChronic). We identified these usernames on the network map (Figure 3). Not all other participants were identified.

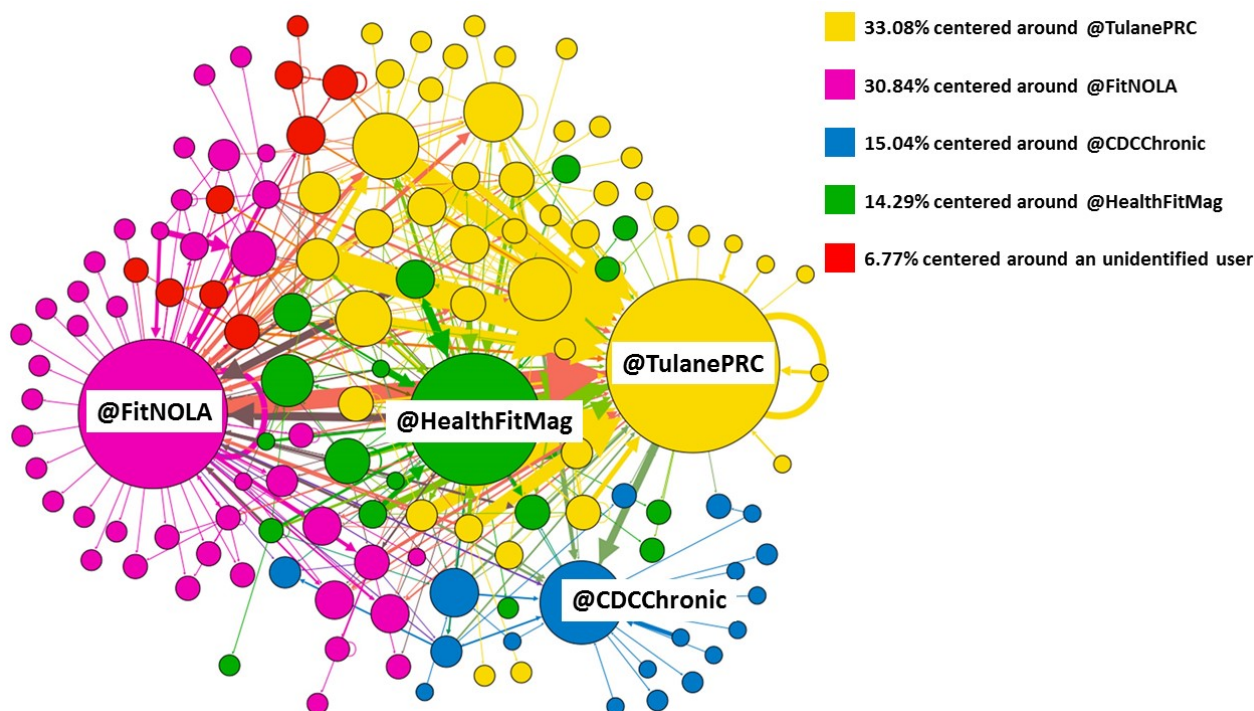
The step-by-step protocol for creating the SNA dataset is described in Multimedia Appendix 1. This protocol includes the Twitter chat transcript acquisition from Symplur, the transcript transformation into a relational dataset with source and target data points (or nodes) using OpenRefine, and network map visualization and measures with Gephi. An annotated and editable R code is in Multimedia Appendix 2, and can be used to collect hashtag-based tweet transcripts from Symplur. Public health professionals able to use point and click apps and with basic quantitative skills likely have the appropriate skills to conduct this analysis for organizations' Twitter chats. The edits required for the provided R code in Multimedia Appendix 2 do not require R program language knowledge; instead edits are changes to the usernames of interests and transcript pages.

Figure 1. The engagement ratio equation.

$$\text{@username engagement ratio} = \frac{\text{@username out degree}}{\text{@username in degree}}$$

Figure 2. The return on engagement equation.

$$\text{@username ROE (\%)} = \frac{\text{@username incoming mentions} - \text{@username outgoing mentions}}{\text{@username outgoing mentions}} \times 100$$

Figure 3. #LiveFitNOLA Twitter chat, March 5, 2015 from 12:00 PM to 1:15 PM Central Time, network map—Circles represent network members, which include #LiveFitNOLA chat participants and other Twitter users mentioned during the chat. Colors represent communities within the network; #LiveFitNOLA Twitter chat hosts are identified; other network members are not. Lines and arrows represent communications between network members. Lines are for two-way communications and arrows are for one-way communications.

Content Analysis

The content analysis assessed the primary function of tweets posted during the #LiveFitNOLA Twitter chat. We adapted a preexisting coding scheme originally developed by Lovejoy and Saxton to characterize tweets posted by public health entities [41]. They identified and described three primary themes: (1) information sharing, (2) community building, and (3) action- or activism-related [41]. Xu et al revised their coding scheme (herein: Information-Community-Action [I-C-A] framework) to characterize tweets that included health-related hashtags [41,42]. We applied a variant of Xu et al's version [42] of the I-C-A framework to the #LiveFitNOLA Twitter chat transcript. First, we piloted their version of the I-C-A framework using a

random sample of 50 original tweets posted during the chat. Based on the pilot, we altered the I-C-A framework themes, sub-themes, and definitions to include 3 themes and 4 subthemes: information sharing (health information, health opinion, health experience, and asking); positive affect or interpersonal closeness; action or activism or advocacy.

We applied the adjusted (I-P-A) coding scheme to a dataset of original tweets, which excluded all retweets to avoid biasing the distribution of themes and sub-themes. Of the 744 tweets in the full #LiveFitNOLA Twitter chat transcript, 491 (66%) were original tweets and 253 (34%) were retweets. Two analysts independently coded all original tweets for the most applicable theme and subtheme. They collaboratively reconciled coding

disagreements resulting in percentages of agreement of 89% for themes and 74% for subthemes. We calculated inter-coder reliability for theme and subtheme using Cohen Kappa and Scott Pi. For theme, Cohen Kappa and Scott Pi were both .78. For subtheme, Cohen Kappa and Scott Pi were both .66.

Results

Data and Study Population

The #LiveFitNOLA Twitter chat had 744 tweets and 66 participants, with an average of 11 tweets per participant, during the 75-min study period. More than half of the #LiveFitNOLA chat participants were individual users (n=35, 53%). Twenty-six participants (39%) were organizational users, and 5 (8%) were uncategorized users. As expected, the majority (n=39, 60%) of the #LiveFitNOLA chat participants were from New Orleans. Fifteen (23%) were from other US locations, and 12 (18%) were from unidentified locations (Table 1).

Social Network Analysis

The resulting #LiveFitNOLA network had 135 network members and 474 edges connections.

The network members included all 66 #LiveFitNOLA participants and 69 other Twitter users mentioned during the chat. Independent of direction, the overall number of

connections (ie, outgoing and incoming mentions) per network member ranged from 1 to 101, with an average of 8 (SD=16) connections per network member. On average, there were 4 (SD=9) outgoing mentions per network number, and 4 (SD=7) incoming mentions per network member (Table 2).

The network had 5 distinct communities. More than half of the network's connections formed 2 distinct communities, centered on the 2 #LiveFitNOLA chat organizers, @TulanePRC (33.3%) and @FitNOLA (31.8%) (Figure 3). Among all #LiveFitNOLA chat participants, @HealthFitMag (84) and @TulanePRC (82) tweeted the most. @TulanePRC had the most incoming mentions (in degree=45), and @FitNOLA had the most outgoing mentions (out degree=63; Table 3). The engagement ratios and ROEs for the #LiveFitNOLA organizers were 1:1 and 40% (13/32) (@TulanePRC) and 2:1 and -40% (-25/63) (@FitNOLA).

Content Analysis

Table 4 shows the percentage of original tweets stratified by theme and sub-theme. Information sharing (63.9%, 314/291) and health information (27.9%, 137/491) were the most common theme and sub-theme, respectively. The second most common theme was positive affect or interpersonal closeness (31.7%, 156/491). The minority of original tweets was categorized as action or activism or advocacy (3.8%, 19/491) and unable to determine (0.4%, 2/491).

Table 1. March 5, 2015 (12:00 PM to 1:15 PM Central Time) #LiveFitNOLA Twitter chat participant characteristics.

Participant characteristics	Total number of participants, N=66 Total number of tweets, N=744
Average number of tweets per participant, mean (SD)	11 (17)
Type of participant, n (%)	
Individual users	35 (53)
Organizational users	26 (39)
Uncategorized users	5 (8)
Participant locations, n (%)	
New Orleans	39 (60)
Other US locations	15 (23)
Unavailable locations	12 (18)

Table 2. March 5, 2015 (12:00 PM to 1:15 PM Central Time) #LiveFitNOLA Twitter chat network characteristics.

Network characteristics	Total number of network members ^a , N=135
	Total number of connections ^b , N=474
Degree ^c , mean (SD)	8 (16)
Out degree ^d , mean (SD)	4 (9)
In degree ^e , mean (SD)	4 (7)

^aNetwork members include 66 #LiveFitNOLA Twitter chat participants and 69 Twitter users mentioned during the chat.

^bNumber of communications: The number of mentions during the #LiveFitNOLA Twitter chat ie, the total number of times a Twitter user mentions another user by including their @username in a tweet, whether by quoting a tweet, retweeting (repeating another user's tweet), or directly addressing another user.

^cDegree: Number of undirected communications between a source and a target Twitter user within the network regardless of the direction of the communication.

^dOut degree: Number of outgoing communications a source Twitter user sent to other Twitter users within the network.

^eIn degree: Number of incoming communications a target Twitter user received from other Twitter users within the network.

Table 3. March 5, 2015 (12:00 PM to 1:15 PM Central Time) #LiveFitNOLA Twitter chat engagement ratios and return on engagement.

Usernames	Number of tweets	Outgoing ^a	Incoming ^b	Out:In ^c	ROE ^d
@TulanePRC	82	32	45	1:1	40%
@FitNOLA	70	63	38	2:1	-40%
@HealthFitMag	84	40	33	1:1	-18%
@CDCChronic	24	8	19	1:2	138%

^aOutgoing: Number of outgoing communications a source Twitter user sent to other Twitter users within the network.

^bIncoming: Number of incoming communications a target Twitter user received from other Twitter users within the network.

^cEngagement ratio: Ratio of outgoing to incoming communications a Twitter user of interest has. The engagement ratio is rounded to the next integer.

^dROE: The engagement gain or loss generated relative to the amount of engagement invested.

Table 4. March 5, 2015 (12:00 PM to 1:15 PM Central Time) #LiveFitNOLA Twitter chat Information-Community-Action (I-C-A) Framework: Coding categories for communication themes, tweet exemplars, and percentage.

Theme	Subtheme	Definitions and examples	n (%)
Information sharing	Health information	Disseminating research findings, tip or advice, tools or resources, health news, and general information about health-related events <i>A2: Social support influences all aspects of health. Proper support has been shown to health increase physical activity. #LiveFitNOLA</i>	137 (27.9)
	Health opinion	Expressing original (not merely affirming another tweet or statement) opinions on health-related issues (unbacked by information or data included in the tweet) <i>@TulanePRC A3. I also think we do not treasure our health or bodies and so we do not treat them as gifts! #LiveFitNOLA</i>	90 (18.3)
	Health experience	Sharing personal or family or friend or experience with health problems or regarding health-related topics <i>@CDCChronic yep! I don't go to a Gym. I do workouts on YouTube #Live-FitNOLA</i>	73 (14.8)
	Asking	Asking questions about health or health-related issues <i>@Tastedat is all about food so wheres your favorite place to eat out health? #LiveFitNOLA</i>	14 (2.8)
Positive affect or Interpersonal closeness		Showing positive affect such as appreciation, greeting, agreement, affirmation, and congratulation; showing interpersonal closeness <i>@frenchktnola @tulaneprc so awesome! Thank you for joining! #LiveFit-NOLA</i>	156 (31.7)
Action or activism or advocacy		Raising awareness, promoting health-related causes, and prompting receivers to take actions such as signing petition, making donation, sharing information, and participating in events <i>A3. Find out if you are at risk for diabetes and ways to prevent it bit.ly/1EORNDQ (2/2) #LiveFitNOLA</i>	19 (3.8)
Unable to determine		Lacks sufficient contextual information to determine the category or unrelated to health <i>@TulanePRC #LiveFitNOLA</i>	2 (0.4)

Discussion

Principal Findings

The trends in our findings indicate public health entities can use Twitter chats as a two-way communication audience engagement tool with their audience. To our knowledge, this is the first study that empirically examined Twitter audience engagement based on the incoming and outgoing communications between public health entities and their audience. To respond for the need to further assess the effectiveness of public health entities in two-way communications on Twitter, we developed metrics and a protocol to assess the incoming and outgoing communications between public health entities and their audience during a Twitter chat.

The #LiveFitNOLA chat organizers targeted New Orleans residents to engage in open discussions about the culture of health in the New Orleans area. A hashtag trends on Twitter when it is algorithmically determined to be one of the most popular hashtags or topics at a particular time [43]. Trends are hashtags or topics that are popular for a specific time and specific location. They are not the popular hashtags or topics that have been popular for a while or occur regularly [44]. Although trends are not representative of the public, they are representative of active Twitter users who are at a specific

location during a specific time. A trending hashtag is a proxy measure of reach. For this particular chat, @TulanePRC and @FitNOLA reached members of the intended audience because #LiveFitNOLA was the sixth trending hashtag in New Orleans on March 5th, 2015 [45]. This was reflected in our findings, the majority of the #LiveFitNOLA chat participants were from New Orleans.

@TulanePRC was the main driver of this chat. Its engagement ratio revealed it received one incoming mention from a chat participant for each outgoing mention to another network member. Furthermore, @TulanePRC had a 40% ROE, which means it generated a 40% gain of incoming engagement related to the outgoing engagement it invested. @FitNOLA experienced a 40% engagement loss related to its invested engagement. By reading the chat transcript, we learned @FitNOLA retweeted most of @TulanePRC's original tweets, which doubled its outgoing mentions compared with @TulanePRC's outgoing mentions. @FitNOLA's engagement level might be improved in the future if both #LiveFitNOLA organizers meet and coordinate to send original tweets rather than retweeting each other.

The guest host, @HealthFitMag, largely played a supporting role. It repeated most of the information shared by both #LiveFitNOLA organizers and did not provide original content. In future chats, the chat organizers should encourage guest hosts

to share original content and actively respond to the chat participants. Interestingly, while playing a supporting role, @CDCChronic had the biggest return on its engagement at 138% and a 1:2 engagement ratio. @CDCChronic received twice as many incoming mentions as its outgoing mentions. Popular and influential Twitter accounts like @CDCChronic attract incoming communications from other users participating in hashtag and time-bound Twitter activities. As popularity begets attention, this might introduce false positives or false negatives to the engagement ratio and ROE. In this instance, #LiveFitNOLA chat participants sought to engage directly in two-way communications with @CDCChronic. @CDCChronic did not respond to the participants because its role was to disseminate evidence-based information on healthy living during the chat, not initiate or seek discussions like the Twitter chat hosts (@TulanePRC and @FitNOLA). The negative or positive popularity of certain users participating in hashtag-based Twitter interactions might influence how others will react to them; and in some cases, this is something beyond the control of the said user. For public health entities seeking to increase audience engagement in two-way communications, we recommend they address the roles of each account at the beginning of the chat. We also recommend taking into consideration the popularity or influence of certain accounts when interpreting the result of the engagement ratio and ROE.

Tweet sentiments might also influence the engagement ratio and ROE. For the scope of this case study, we did not conduct any sentiment analyses on tweet contents. We do recommend such analyses whenever possible to complement the engagement ratio and ROE results. An example of the importance of sentiment analyses would be the disclosure of a negative action, which would likely get a high ROE but might have highly negative reputational impact on the user. Mathematically, the ROE would be high but its impact on relationship building could be detrimental. Stylistic differences might influence the engagement ratio and ROE, and introduce false positives or false negatives because bombastic or offensive tweets might incite strong negative or positive reactions from other users. To this end, when interpreting the findings of the engagement ratio and ROE, we recommend skimming through the transcript to see if a user with a surprisingly low or surprisingly high engagement ratio and ROE have a very different tweeting style to other users.

The depth and breadth of incoming and outgoing communications between users might influence the interpretation of the engagement ratio. Depth represents how many times the same source mentions a specific target user (incoming) and how many times a specific source user mentions the same target user (outgoing). Breadth measures mention volume; how many different users mention a specific target user (incoming), and how many different users mention a specific source user (outgoing). For example, of @TulanePRC's 32 outgoing communications, 31 were sent to distinct different users and one was to itself, which indicates a broad-scale engagement with other users. In contrast, if a particular user had mentioned the same user 31 times, a single-source engagement would occur. If a single-source and a broad-scale engagement have the same engagement ratio and ROE, the

engagement quality depends on the context. If the goal was to engage with more people such as in a Twitter chat, then a broad-scale engagement is better. If the goal was to engage in a one-on-one conversation, as might be the case for customer service interactions on Twitter, then a single-source engagement would be better.

Bots might interfere with the activities around a particular hashtag, and affect the engagement ratio and ROE with false positives. Tweets sent by bots need to be removed from the transcript before analyzing the incoming and outgoing engagement between users around a particular hashtag. The #LiveFitNOLA chat described here did not include any bot. However, the #LiveFitNOLA hashtag was hijacked by a bot in later periods of its use. We recommended the chat organizers to remove the bot's activities from any analyses they might conduct within the period of the bot's activities. In conclusion, user popularity, tweet sentiment and stylistic difference, and bots might influence the engagement ratio and ROE. Their practical significance depends on the audience, context, scope, scale, and goal of a Twitter chat or other organized hashtag-based communications on Twitter.

More advanced analyses could be conducted on networks formed on Twitter around hashtags. For example, in our case study, granular community-based analysis could be used to assess individual communities, cross-community, and within-community interactions. We did not conduct such analyses and did not include them in our protocol because they are beyond the practical purpose of our study. In addition, they require advanced knowledge of SNA beyond what is needed to obtain the engagement ratio and ROE. As previously mentioned, our intended audience are public health professionals able to use point and click apps and with basic quantitative skills.

In line with previous studies, results of the content analysis revealed the majority of tweets shared health information or showed positive affect or interpersonal closeness. The distribution of themes, however, deviated from prior research using the I-C-A framework [41,42]. Specifically, we identified more positive affect or interpersonal closeness tweets and fewer action or activism or advocacy tweets than other research that used the I-C-A framework [41,42]. These results highlight the potential of Twitter chats, which are typically prearranged, time-specific, moderated, and topically focused to enable participants cut through the millions of Tweets posted per day and establish directed connections. This might account for the increase in percentage of tweets that show positive affect or interpersonal closeness in our analysis, compared with other research using the I-C-A framework to analyze Twitter datasets that were not specific to a single Twitter chat.

Limitations

A limitation for this case study is @FitNOLA and @HealthFitMag retweeted most of @TulanePRC's tweets. Hosts retweeting hosts create an echo chamber and do not add new content, which might hinder host engagements with other users. Based on our findings, the hosts' audience engagement might be improved by ensuring Twitter chats are organized and led by two or more organizations that do not retweet each other and individually tweet original contents. Each host should share

carefully planned and curated original contents and refrain from sharing (retweeting) the same information as other hosts.

Another limitation is our use of self-disclosed geographical locations in user profiles. This is not an ideal way to identify geographic origins of tweets. We chose this approach because using the Global Positioning System (GPS) tags in tweets or Internet Protocol (IP) addresses would require more advanced analytical skills than the ones needed for the protocol we used in our study. In addition, users might not always enable their GPS-based location in their tweets, and IP addresses might not be accurate location indicators.

Conclusions

Further studies are needed to establish quantifiable parameters on what is a low, medium, or high engagement ratio and ROE such as through longitudinal analyses of a recurring Twitter chat. In addition, such a study could provide more nuanced information about a public health entity's Twitter audience for their hashtag-based communications such as their constituents' health topic of interests or the type of guest host who will engage more with their audience.

The protocol described in [Multimedia Appendix 1](#) on how to visualize and measure engagement levels for hashtag-based communications on Twitter is not limited to Twitter chats. It can be used for any time-bound hashtag-based Twitter communications where transcripts are available, such as conference hashtags (eg, #APHA15 for the 2015 American Public Health Association Annual Meeting), disease or condition hashtags (eg, #Ebola and #Zika), or public health initiatives or campaigns (eg, #CultureofHealth, a Robert Wood Johnson Foundation initiative to “enable all to live longer, healthier lives”).

Based on the lessons learned from this case study, to build further evidence on the effective use of Twitter as an audience engagement in two-way communication tool for public health entities our future research includes: (1) conducting longitudinal ROE assessments with recurring Twitter chats and (2) determining audience engagement effectiveness based on the hashtag-based communication's purposes by examining larger public health campaigns.

Acknowledgments

We thank Jennifer L Kline, Lourdes M Martinez, April S Smith, and Toye Williams for their programmatic, scientific, and technical support. The authors do not have any financial disclosure to report. The Prevention Research Center at Tulane University School of Public Health and Tropical Medicine is supported by the Centers for Disease Control and Prevention, Atlanta, Georgia, through the Health Promotion and Disease Prevention Research Centers Program under cooperative agreement #U48DP005050.

Conflicts of Interest

None declared.

Multimedia Appendix 1

How to visualize and measure engagement level for hashtag-based Twitter conversations.

[[PDF File \(Adobe PDF File\), 9MB - publichealth_v3i2e34_app1.pdf](#)]

Multimedia Appendix 2

R code to collect the #LiveFitNOLA chat transcript from Symplur, on March 5th 2015 from 1:00 PM to 2:15 PM ET, as an example (www.tinyurl.com/LiveFitNOLAMarch52015).

[[PDF File \(Adobe PDF File\), 275KB - publichealth_v3i2e34_app2.pdf](#)]

References

1. Duggan M, Ellison N, Lampe C, Lenhart A, Madden M. Pew Research Center. 2014. Social media update URL: http://www.pewinternet.org/files/2015/01/PI_SocialMediaUpdate20144.pdf [accessed 2016-12-16] [[WebCite Cache ID 6mnup29wW](#)]
2. Twitter. About - Twitter URL: <https://about.twitter.com/company> [accessed 2016-12-16] [[WebCite Cache ID 6mnv1Oter](#)]
3. Aslam AA, Tsou M, Spitzberg BH, An L, Gawron JM, Gupta DK, et al. The reliability of tweets as a supplementary method of seasonal influenza surveillance. *J Med Internet Res* 2014;16(11):e250 [[FREE Full text](#)] [doi: [10.2196/jmir.3532](https://doi.org/10.2196/jmir.3532)] [Medline: [25406040](https://pubmed.ncbi.nlm.nih.gov/25406040/)]
4. 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/)]
5. Nagar R, Yuan Q, Freifeld CC, Santillana M, Nojima A, Chunara R, et al. A case study of the New York City 2012-2013 influenza season with daily geocoded twitter data from temporal and spatiotemporal perspectives. *J Med Internet Res* 2014;16(10):e236 [[FREE Full text](#)] [doi: [10.2196/jmir.3416](https://doi.org/10.2196/jmir.3416)] [Medline: [25331122](https://pubmed.ncbi.nlm.nih.gov/25331122/)]

6. Santos JC, Matos S. Analysing twitter and web queries for flu trend prediction. *Theor Biol Med Model* 2014 May 07;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/)]
7. Harris JK, Mansour R, Choucair B, Olson J, Nissen C, Bhatt J, Centers for Disease Control and Prevention. Health department use of social media to identify foodborne illness - Chicago, Illinois, 2013-2014. *MMWR Morb Mortal Wkly Rep* 2014 Aug 15;63(32):681-685 [FREE Full text] [Medline: [25121710](https://pubmed.ncbi.nlm.nih.gov/25121710/)]
8. Harris JK, Mueller NL, Snider D, Haire-Joshu D. Local health department use of twitter to disseminate diabetes information, United States. *Prev Chronic Dis* 2013;10:E70 [FREE Full text] [doi: [10.5888/pcd10.120215](https://doi.org/10.5888/pcd10.120215)] [Medline: [23639765](https://pubmed.ncbi.nlm.nih.gov/23639765/)]
9. Scanfeld D, Scanfeld V, Larson EL. Dissemination of health information through social networks: twitter and antibiotics. *Am J Infect Control* 2010 Apr;38(3):182-188 [FREE Full text] [doi: [10.1016/j.ajic.2009.11.004](https://doi.org/10.1016/j.ajic.2009.11.004)] [Medline: [20347636](https://pubmed.ncbi.nlm.nih.gov/20347636/)]
10. Step MM, Bracken CC, Trapl ES, Flocke SA. User and content characteristics of public tweets referencing little cigars. *Am J Health Behav* 2016 Jan;40(1):38-47. [doi: [10.5993/AJHB.40.1.5](https://doi.org/10.5993/AJHB.40.1.5)] [Medline: [26685812](https://pubmed.ncbi.nlm.nih.gov/26685812/)]
11. Capurro D, Cole K, Echavarría MI, Joe J, Neogi T, Turner AM. The use of social networking sites for public health practice and research: a systematic review. *J Med Internet Res* 2014;16(3):e79 [FREE Full text] [doi: [10.2196/jmir.2679](https://doi.org/10.2196/jmir.2679)] [Medline: [24642014](https://pubmed.ncbi.nlm.nih.gov/24642014/)]
12. Duke JC, Hansen H, Kim AE, Curry L, Allen J. The use of social media by state tobacco control programs to promote smoking cessation: a cross-sectional study. *J Med Internet Res* 2014;16(7):e169 [FREE Full text] [doi: [10.2196/jmir.3430](https://doi.org/10.2196/jmir.3430)] [Medline: [25014311](https://pubmed.ncbi.nlm.nih.gov/25014311/)]
13. Black DR, Dietz JE, Stirratt AA, Coster DC. Do social media have a place in public health emergency response? *J Emerg Manag* 2015;13(3):217-226. [doi: [10.5055/jem.2015.0235](https://doi.org/10.5055/jem.2015.0235)] [Medline: [26150365](https://pubmed.ncbi.nlm.nih.gov/26150365/)]
14. Towers S, Afzal S, Bernal G, Bliss N, Brown S, Espinoza B, et al. Mass media and the contagion of fear: the case of Ebola in America. *PLoS One* 2015;10(6):e0129179 [FREE Full text] [doi: [10.1371/journal.pone.0129179](https://doi.org/10.1371/journal.pone.0129179)] [Medline: [26067433](https://pubmed.ncbi.nlm.nih.gov/26067433/)]
15. Odlum M, Yoon S. What can we learn about the Ebola outbreak from tweets? *Am J Infect Control* 2015 Jun;43(6):563-571. [doi: [10.1016/j.ajic.2015.02.023](https://doi.org/10.1016/j.ajic.2015.02.023)] [Medline: [26042846](https://pubmed.ncbi.nlm.nih.gov/26042846/)]
16. Gurman TA, Ellenberger N. Reaching the global community during disasters: findings from a content analysis of the organizational use of twitter after the 2010 Haiti earthquake. *J Health Commun* 2015;20(6):687-696. [doi: [10.1080/10810730.2015.1018566](https://doi.org/10.1080/10810730.2015.1018566)] [Medline: [25928401](https://pubmed.ncbi.nlm.nih.gov/25928401/)]
17. Buzzelli MM, Morgan P, Muschek AG, Macgregor-Skinner G. Information and communication technology: connecting the public and first responders during disasters. *J Emerg Manag* 2014;12(6):441-447. [doi: [10.5055/jem.2014.0207](https://doi.org/10.5055/jem.2014.0207)] [Medline: [25530560](https://pubmed.ncbi.nlm.nih.gov/25530560/)]
18. Awad NI, Cocchio C. Use of twitter at a major national pharmacy conference. *Am J Health Syst Pharm* 2015 Jan 1;72(1):65-69. [doi: [10.2146/ajhp140175](https://doi.org/10.2146/ajhp140175)] [Medline: [25511841](https://pubmed.ncbi.nlm.nih.gov/25511841/)]
19. Lu X, Brelsford C. Network structure and community evolution on twitter: human behavior change in response to the 2011 Japanese earthquake and tsunami. *Sci Rep* 2014 Oct 27;4:6773 [FREE Full text] [doi: [10.1038/srep06773](https://doi.org/10.1038/srep06773)] [Medline: [25346468](https://pubmed.ncbi.nlm.nih.gov/25346468/)]
20. Anderson G, Gleeson S, Rissel C, Wen LM, Bedford K. Twitter tweets and twaddle: twittering at AHPA. National Health Promotion Conference. *Health Promot J Austr* 2014 Aug;25(2):143-146. [doi: [10.1071/HE13112](https://doi.org/10.1071/HE13112)] [Medline: [25200470](https://pubmed.ncbi.nlm.nih.gov/25200470/)]
21. Chung JE. A smoking cessation campaign on twitter: understanding the use of twitter and identifying major players in a health campaign. *J Health Commun* 2016 May;21(5):517-526. [doi: [10.1080/10810730.2015.1103332](https://doi.org/10.1080/10810730.2015.1103332)] [Medline: [27115179](https://pubmed.ncbi.nlm.nih.gov/27115179/)]
22. Wilkinson SE, Basto MY, Perovic G, Lawrentschuk N, Murphy DG. The social media revolution is changing the conference experience: analytics and trends from eight international meetings. *BJU Int* 2015 May;115(5):839-846 [FREE Full text] [doi: [10.1111/bju.12910](https://doi.org/10.1111/bju.12910)] [Medline: [25130687](https://pubmed.ncbi.nlm.nih.gov/25130687/)]
23. Harris JK, Choucair B, Maier RC, Jolani N, Bernhardt JM. Are public health organizations tweeting to the choir? Understanding local health department twitter followership. *J Med Internet Res* 2014 Feb 26;16(2):e31 [FREE Full text] [doi: [10.2196/jmir.2972](https://doi.org/10.2196/jmir.2972)] [Medline: [24571914](https://pubmed.ncbi.nlm.nih.gov/24571914/)]
24. Duke JC, Hansen H, Kim AE, Curry L, Allen J. The use of social media by state tobacco control programs to promote smoking cessation: a cross-sectional study. *J Med Internet Res* 2014 Jul 10;16(7):e169 [FREE Full text] [doi: [10.2196/jmir.3430](https://doi.org/10.2196/jmir.3430)] [Medline: [25014311](https://pubmed.ncbi.nlm.nih.gov/25014311/)]
25. Maher CA, Lewis LK, Ferrar K, Marshall S, De BI, Vandelanotte C. Are health behavior change interventions that use online social networks effective? a systematic review. *J Med Internet Res* 2014 Feb;16(2):e40 [FREE Full text] [doi: [10.2196/jmir.2952](https://doi.org/10.2196/jmir.2952)] [Medline: [24550083](https://pubmed.ncbi.nlm.nih.gov/24550083/)]
26. Neiger BL, Thackeray R, Burton SH, Giraud-Carrier CG, Fagen MC. Evaluating social media's capacity to develop engaged audiences in health promotion settings: use of twitter metrics as a case study. *Health Promot Pract* 2013 Mar;14(2):157-162. [doi: [10.1177/1524839912469378](https://doi.org/10.1177/1524839912469378)] [Medline: [23271716](https://pubmed.ncbi.nlm.nih.gov/23271716/)]
27. Neiger BL, Thackeray R, Burton SH, Thackeray CR, Reese JH. Use of twitter among local health departments: an analysis of information sharing, engagement, and action. *J Med Internet Res* 2013 Aug;15(8):e177 [FREE Full text] [doi: [10.2196/jmir.2775](https://doi.org/10.2196/jmir.2775)] [Medline: [23958635](https://pubmed.ncbi.nlm.nih.gov/23958635/)]
28. Park H, Reber BH, Chon M. Tweeting as health communication: Health Organizations' use of twitter for health promotion and public engagement. *J Health Commun* 2016 Dec;21(2):188-198. [doi: [10.1080/10810730.2015.1058435](https://doi.org/10.1080/10810730.2015.1058435)] [Medline: [26716546](https://pubmed.ncbi.nlm.nih.gov/26716546/)]

29. Welch V, Petkovic J, Pardo Pardo J, Rader T, Tugwell P. Interactive social media interventions to promote health equity: an overview of reviews. *Health Promot Chronic Dis Prev Can* 2016 Apr;36(4):63-75. [doi: [10.24095/hpcdp.36.4.01](https://doi.org/10.24095/hpcdp.36.4.01)]
30. Thackeray R, Neiger BL, Burton SH, Thackeray CR. Analysis of the purpose of state health departments' tweets: information sharing, engagement, and action. *J Med Internet Res* 2013;15(11):e255 [FREE Full text] [doi: [10.2196/jmir.3002](https://doi.org/10.2196/jmir.3002)] [Medline: [24217361](https://pubmed.ncbi.nlm.nih.gov/24217361/)]
31. Symplur. Healthcare hashtag project URL: <https://www.symplur.com/healthcare-hashtags/> [accessed 2016-12-16] [WebCite Cache ID 6mntw3Lzn]
32. OpenRefine. URL: <http://openrefine.org/download.html> [accessed 2016-12-16] [WebCite Cache ID 6mnvLBW4Z]
33. Gephi. URL: <https://gephi.org/users/download/> [accessed 2016-12-16] [WebCite Cache ID 6mnvYG241]
34. Twitter. Tweet activity dashboard URL: <https://support.twitter.com/articles/20171990> [accessed 2017-02-11] [WebCite Cache ID 6oCpov1Ex]
35. Bhattacharya S, Srinivasan P, Polgreen P. Engagement with health agencies on twitter. *PLoS One* 2014;9(11):e112235 [FREE Full text] [doi: [10.1371/journal.pone.0112235](https://doi.org/10.1371/journal.pone.0112235)] [Medline: [25379727](https://pubmed.ncbi.nlm.nih.gov/25379727/)]
36. Englar N, Mitchell W, Cabredo E, O'Malley K. Hashtags for health: leveraging partnerships to expand community engagement through local, collaborative twitter chats. 2015 Nov 02 Presented at: APHA annual meeting & expo; October 31 - November 4, 2015; Chicago, IL URL: <https://apha.confex.com/apha/143am/webprogram/Paper332362.html>
37. Bennett S. Adweek. 2015. What are the best times to post on #Facebook, #Twitter and #Instagram? [INFOGRAPHIC] URL: <http://www.adweek.com/socialtimes/best-time-to-post-social-media/504222> [accessed 2016-12-16] [WebCite Cache ID 6mnu8bogN]
38. Maxwell L. Marketmymarket. 2015. Getting tweets out at the best times on twitter URL: <http://www.marketmymarket.com/getting-tweets-out-at-the-best-times-on-twitter/> [accessed 2016-12-16] [WebCite Cache ID 6mnuRI610]
39. Valente T. *Social Networks and Health: Models, Methods, and Applications*. Oxford, United Kingdom: Oxford University Press; Mar 25, 2010.
40. Borgatti S, Everett M, Johnson J. *Analyzing social networks*. Thousand Oaks, CA: Sage Publications; May 29, 2013.
41. Lovejoy K, Saxton GD. Information, community, and action: How nonprofit organizations use social Media. *J Comput Mediat Commun* 2012 Apr 10;17(3):337-353. [doi: [10.1111/j.1083-6101.2012.01576.x](https://doi.org/10.1111/j.1083-6101.2012.01576.x)]
42. Xu WW, Chiu I, Chen Y, Mukherjee T. Twitter hashtags for health: applying network and content analyses to understand the health knowledge sharing in a Twitter-based community of practice. *Qual Quant* 2014 Jul 8;49(4):1361-1380. [doi: [10.1007/s11135-014-0051-6](https://doi.org/10.1007/s11135-014-0051-6)]
43. Twitter. The twitter glossary URL: <https://support.twitter.com/articles/166337?lang=en> [accessed 2017-03-07] [WebCite Cache ID 6omwzVpfx]
44. Twitter. FAQs about trends on twitter URL: <https://support.twitter.com/articles/101125> [accessed 2017-03-07] [WebCite Cache ID 6omyVnkVY]
45. Trendogate. Trends in New Orleans/United States: trends for 2015-03-05 URL: <https://trendogate.com/placebydate/2458833/2015-03-05> [accessed 2017-03-07] [WebCite Cache ID 6omwf1K7W]

Abbreviations

- ROE:** return of engagement
ROI: return on investment
SNA: social network analysis

Edited by T Sanchez; submitted 16.12.16; peer-reviewed by B Spitzberg, D Broniatowski; comments to author 11.01.17; revised version received 18.02.17; accepted 28.03.17; published 08.06.17.

Please cite as:

Rabarison KM, Croston MA, Englar NK, Bish CL, Flynn SM, Johnson CC
Measuring Audience Engagement for Public Health Twitter Chats: Insights From #LiveFitNOLA
JMIR Public Health Surveill 2017;3(2):e34
URL: <http://publichealth.jmir.org/2017/2/e34/>
doi: [10.2196/publichealth.7181](https://doi.org/10.2196/publichealth.7181)
PMID: [28596149](https://pubmed.ncbi.nlm.nih.gov/28596149/)

©Kristina M Rabarison, Merriah A Croston, Naomi K Englar, Connie L Bish, Shelbi M Flynn, Carolyn C Johnson. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 08.06.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public*

Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

The Use of Intervention Mapping to Develop a Tailored Web-Based Intervention, Condom-HIM

Joyal Miranda^{1*}, PhD; José Côté^{2*}, PhD

¹Faculty of Community Services, Daphne Cockwell School of Nursing, Ryerson University, Toronto, ON, Canada

²Faculté des sciences infirmières, Université de Montréal, Montreal, ON, Canada

* all authors contributed equally

Corresponding Author:

Joyal Miranda, PhD

Faculty of Community Services

Daphne Cockwell School of Nursing

Ryerson University

350 Victoria Street

Toronto, ON, M5B 2K3

Canada

Phone: 1 416 979 5000 ext 7850

Fax: 1 416 979 5332

Email: joyal.miranda@ryerson.ca

Abstract

Background: Many HIV (human immunodeficiency virus) prevention interventions are currently being implemented and evaluated, with little information published on their development. A framework highlighting the method of development of an intervention can be used by others wanting to replicate interventions or develop similar interventions to suit other contexts and settings. It provides researchers with a comprehensive development process of the intervention.

Objective: The objective of this paper was to describe how a systematic approach, intervention mapping, was used to develop a tailored Web-based intervention to increase condom use among HIV-positive men who have sex with men.

Methods: The intervention was developed in consultation with a multidisciplinary team composed of academic researchers, community members, Web designers, and the target population. Intervention mapping involved a systematic process of 6 steps: (1) needs assessment; (2) identification of proximal intervention objectives; (3) selection of theory-based intervention methods and practical strategies; (4) development of intervention components and materials; (5) adoption, implementation, and maintenance; and (6) evaluation planning.

Results: The application of intervention mapping resulted in the development of a tailored Web-based intervention for HIV-positive men who have sex with men, called Condom-HIM.

Conclusions: Using intervention mapping as a systematic process to develop interventions is a feasible approach that specifically integrates the use of theory and empirical findings. Outlining the process used to develop a particular intervention provides clarification on the conceptual use of experimental interventions in addition to potentially identifying reasons for intervention failures.

(*JMIR Public Health Surveill* 2017;3(2):e20) doi:[10.2196/publichealth.7052](https://doi.org/10.2196/publichealth.7052)

KEYWORDS

Internet; condoms; HIV seropositivity; intention; self-efficacy; sexual behavior

Introduction

Recent studies have found that individuals who use the Internet to meet sexual partners are at an elevated risk of contracting human immunodeficiency virus (HIV) infection and other sexually transmitted diseases (STDs) [1-3]. Individuals who

seek partners via the Internet tend to engage in riskier sexual behavior (ie, a seropositive man having insertive condomless anal sex with a partner who is seronegative or whose serostatus is unknown) [4-6]. Several researchers studying the population of men who have sex with men (MSM) have identified a significant association between online partner seeking and an

increased incidence of HIV transmission [2,5] and have speculated that using the Internet to seek sexual partners may be fueling HIV risk behaviors within this group [7].

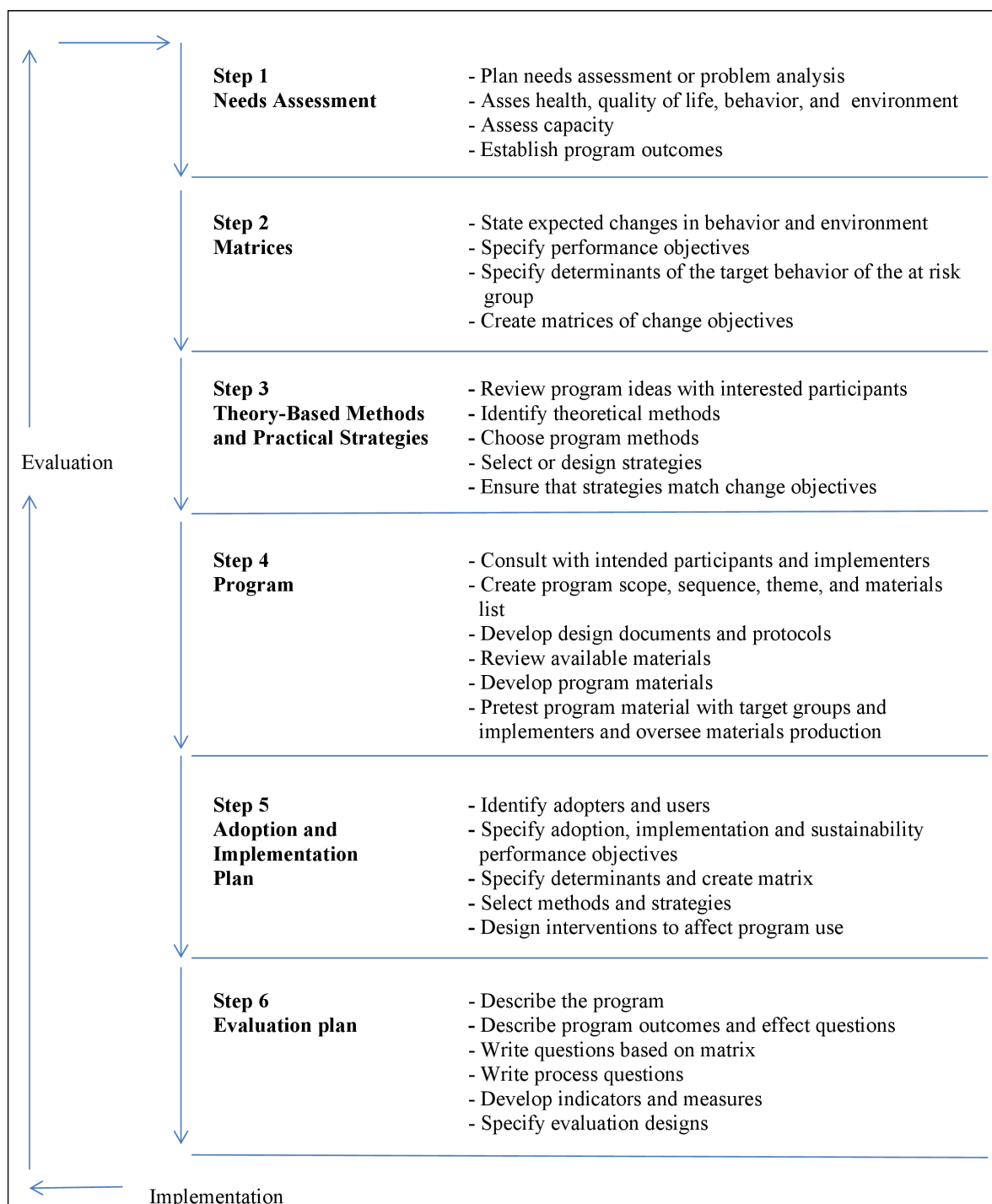
Using the Internet as a method of meeting sexual partners came to the attention of public health officials after a syphilis outbreak in San Francisco in 1999 was associated with increasing transmission among MSM, who were HIV-positive and also met their partners online [8]. Since then, other studies have also examined the transmission of HIV and other STDs among MSM who met online [7,9,10]. In particular, a meta-analysis involving Internet-using MSM found that men who seek sexual partners via the Internet, especially HIV-positive men, were more likely than men who did not seek sexual partners on the Internet to engage in unprotected anal intercourse [7]. For many, high-risk sexual behaviors continued after HIV diagnosis. The reductions seen in high-risk behaviors were short-lived, and the individuals tended to revert to engaging in HIV transmission high-risk behaviors within 1 year of diagnosis [11]. From a public health perspective, efforts to prevent HIV-positive, Internet-using MSM from transmitting HIV have been focused on the sexual high-risk behavior of having unprotected insertive and receptive anal sex with HIV-negative partners and those of unknown serostatus. As MSM were early adopters of the Internet, and given the increase in high-risk sexual behaviors, strategies using the same method of socializing should be considered when interventions are developed to prevent HIV transmission among this particular target population [12].

To date, only a small number of studies have been reported on the efficacy of Internet-based interventions targeting HIV-positive MSM in influencing risky behaviors such as inconsistent condom use [13,14]. Moreover, many of these have neglected to address the process by which the intervention was developed. Systematically developing an intervention based on empirical evidence and theory has been found to substantially improve the chances of the intervention's success, as well as of

identifying possible causes for its failures [15]. Lack of information on the systematic development and content of an intervention poses limits not only to adequate intervention evaluation, but also to any meta-analytic reviews of the intervention's effectiveness that may be conducted in the future [16]. The purpose of this paper is to describe the systematic process followed to develop a tailored Web-based intervention, Condom-HIM. In this case, intervention mapping was used to design an intervention aimed at increasing condom use among HIV-positive MSM who do not consistently use condoms with their partners who are either HIV-negative or of unknown serostatus.

Methods

Intervention mapping, the systematic process applied to develop the tailored Web-based intervention, is composed of the following 6 steps (see Figure 1) that build on each other: (1) conducting a needs assessment; (2) creating a detailed map of intervention objectives and the behavioral and environmental determinants of the problem that the intervention would address; (3) selecting theory-based methods and practical strategies to modify the behavioral and environmental determinants; (4) producing the intervention's components and materials; (5) planning for adoption, implementation, and sustainability; and (6) creating evaluation plans and instruments [17]. The intervention mapping process allows the intervention's developers to make effective decisions at each step in its development based on empirical evidence, theory, and information collected from the target population. In this paper, the first 4 steps of the intervention mapping process and the particular tasks carried out within them are described in detail. Steps 5 and 6, which involve intervention implementation, adoption, monitoring, and evaluation, are only briefly highlighted because the paper's focus is on the development of the intervention.

Figure 1. The intervention mapping process.

Results

Development Team

A multidisciplinary research team was first established in order to develop the intervention using the various steps of the

intervention mapping process. Individuals were selected to represent various fields of expertise that were required for the intervention's design, and such individuals included the following:

- a behavioral researcher with expertise in adopting, maintaining, and changing health-related behaviors such as condom use, as well as in Web-based intervention design and delivery;
- two sexologist researchers with a focus on psychosocial and sociocultural factors associated with at-risk or preventive behaviors, particularly in relation to STDs within the MSM population;
- two nursing researchers with expertise in the field of HIV and AIDS, as well as delivery and development of Web-based interventions;
- one computer and software research engineer;
- three sexual health community-based clinicians with access to the target population;
- individuals representing the target population;
- a Web design team to support the visual development of the Web-based intervention.

Intervention Mapping Step 1: Needs Assessment

The first intervention mapping step of conducting a needs assessment helps to identify the health problem and its related behavioral and environmental determinants for the at-risk population [18]. A call was put out by one of the local sexual health community centers for HIV-positive MSM willing to participate in interviews asking about their needs in relation to HIV and AIDS prevention issues. The principal investigator of the study conducted a total of 10 interviews with HIV-positive MSM from the local sexual health community center as well as 2 interviews with HIV counselors from the center who also fit into the target population in order to identify the current needs of MSM. It was identified by the target population that there was indeed a need to have more innovative and interactive support through the use of the Internet in relation to condom use, while still allowing for anonymity of their HIV status. The current methods to date in relation to increasing condom use required interventions that were face-to-face, thereby requiring individuals to identify their HIV status if wanting to participate. In addition to the needs assessment, the multidisciplinary team also met to discuss the prevalent issues within the community and review the empirical literature. They found that, of the groups within Canada considered most at risk for contracting HIV, MSM accounted for the greatest proportion (54.3%) of new HIV infections in 2014 [19]. One main source of transmission identified was unsafe sex among people who know they are infected with HIV [20]. Various studies targeting the sexual risk behaviors in MSM living with HIV have found that 10%-60% do not consistently practice safer sex behaviors [20]. In addition, MSM recently infected with HIV reported that they had a median of 20 sexual partners within the previous year and that they continued to repeatedly engage in high-risk sexual behaviors, specifically unprotected anal intercourse [11]. Of particular interest is the fact that 34% of MSM with recently diagnosed HIV did not change their risk behaviors following diagnosis, and 20% actually increased their risk behaviors after diagnosis [11]. Such potential results of engaging in unprotected anal sex results in putting oneself at risk of contracting secondary infections (eg, syphilis, gonorrhea, and herpesvirus), which in turn have the potential to accelerate the HIV disease [20,21].

Evidence has indicated that 35%-45% of MSM access the Internet to seek sexual partners [4]. Furthermore, this group of MSM was more likely to report unprotected anal intercourse with nonconcordant casual partners they met online [3]. Halkitis and Parsons [22] found that 84% of MSM reported engaging in unprotected anal intercourse in the past 3 months and that 43% of the men reported recent unprotected anal intercourse with a partner of unknown serostatus. These results indicated the potential for widespread transmission of HIV to uninfected men by the partners they met on the Internet.

On the basis of the needs assessment and review of the literature, a Web-based intervention was considered to be potentially useful in targeting HIV-positive MSM seeking partners online. Although most studies targeted HIV-negative or untested individuals, both the National Institutes of Health and the Centers for Disease Control and Prevention have proclaimed the need for behavioral interventions directed toward individuals living with HIV [23]. Prevention interventions are needed to help HIV-positive MSM adopt and maintain safer sex behaviors such as condom use. Without such interventions, the growing number of MSM who use the Internet to seek sexual partners will continue to form a potential source of HIV transmission to HIV-negative MSM or those of unknown status [24]. Preventing new infections has represented the only long-term, sustainable way to curb the spread of HIV and AIDS [25]. Consequently, the research team reviewing the empirical evidence in relation to decreasing high-risk behaviors with HIV-positive MSM who use the Internet to seek sexual partners identified unprotected anal intercourse as the relevant risk behavior within this population.

Intervention Mapping Step 2

Matrices of Change Objectives

The second step of intervention mapping provides the foundation for the intervention's design by specifying who and what will change as a result of the intervention. A set of matrices was generated by the research team in order to identify the intervention's performance objectives, which is what the at-risk group members must do to accomplish the health-related behavior of HIV-positive MSM to consistently and correctly use condoms for anal sex with their partners who were either HIV-negative or of unknown serostatus.

Condom use was chosen as the behavioral outcome over other forms of HIV prevention such as serosorting and viral load status. Serosorting is defined as the practice of having sex with partners of concordant HIV status. This form of HIV prevention has been practiced by many HIV-positive MSM. Research has shown, however, that serosorting is unlikely to be beneficial in many populations of MSM and can actually be expected to lead to increased risk of HIV transmission [26]. In a systematic review to assess the association between serosorting and HIV conducted among MSM, it was found that compared with condom use serosorting was associated with a higher risk of HIV (odds ratio, OR, 1.80, 95% CI 1.21-2.70) [27].

Results of studies regarding viral loads have shown that, since the introduction of antiretroviral therapy, MSM have perceived that having an undetectable HIV load reduces the risk of HIV

transmission [28-31]. For instance, Lampe [28] found that, among MSM receiving antiretroviral therapy, those who reported undetectable viral load had higher rates of condomless sex with a partner of different serostatus than those who did not. In addition, a recent study by Van Den Boom et al [29] found that, among MSM, between 20% and 57% of reported practices of engaging in unprotected anal sex were related to having an undetectable viral load. Also of interest was that, among HIV-discordant partners, undetectable viral loads were considered before engaging in unprotected anal sex with sex buddies (40% of the time) and with casual partners (57% of the time). The problem with this particular strategy of prevention is that, even if the viral load is undetectable, unknown risk parameters, viral load variability, and the possibility of drug-resistant strains of HIV still makes this strategy an unreliable and inconsistent way to prevent HIV transmission [32].

Performance Objectives

In order to specifically target the behavioral change among MSM, the research team next identified performance objectives that would clarify the exact behavior performance expected of an individual affected by the intervention [33]. In this case, they asked themselves what the target population exposed to the intervention must do in order to engage in consistent condom use. The team's review of various meta-analyses on the efficacy of condom use interventions indicated that behavioral interventions rather than just educational type interventions have been shown to promote the greatest changes in condom use [34-36]. Therefore, the following performance objectives for the intervention were selected: (1) to plan condom use when having anal intercourse, (2) to negotiate with a partner the use of a condom during anal intercourse, and (3) to choose to avoid anal intercourse without a condom.

Determinants of Condom Use

Once the target behavior and performance objectives had been specified, the next step was to identify determinants, that is, the factors found to be associated with the performance of the behavior [37]. The research team undertook a review of available evidence about determinants contributing to condom use. Two researchers from our multidisciplinary team had previously conducted a study examining the determinants of condom use among HIV-positive men having anal sex with HIV-negative men or men of unknown HIV status [38]. Their results indicated that intention (OR 3.13, 95% CI 1.25-7.81) and self-efficacy (OR 3.62, 95% CI 1.40-9.37) were the main predictors of condom use. Self-efficacy was found to interact with intention, thereby moderating the relationship between intention and behavior (OR 20.96, 95% CI 2.90-151.51) [38]. Their findings were similar to those of a meta-analysis of studies on condom use, which indicated that self-efficacy was a strong predictor of intention to use condoms and of condom use, especially among HIV-positive MSM [39]. Additional studies identified intention as a significant predictor of unprotected anal sex among HIV-positive MSM, indicating that HIV-positive MSM usually act on their stated intentions [38,40,41]. Finally, in a meta-analysis examining sexual risk behaviors of people living with HIV, psychological factors such as a low intention to engage in safer sex, lack of confidence in one's ability to engage in safer sex practices, belief that one has little control over condom use, and lack of communication were also reported as predictors of risky sex among MSM [20]. On the basis of the evidence reviewed, the research team concluded that the intervention should target individuals' intention and self-efficacy in order to increase condom use among HIV-positive MSM. [Table 1](#) provides an example of the matrices with one of the performance objectives, which is to plan condom use when having anal intercourse, aligned with the change objectives of intention and self-efficacy (see [Table 1](#)).

Table 1. Intervention matrices.

Performance objectives (POs) for the Individual	Change objectives and personal determinants Intention	Self-efficacy
PO 1: To plan condom use when having anal intercourse	<p>Formulate an action plan to use condoms every time during sexual intercourse:</p> <ol style="list-style-type: none"> Buy or obtain condoms; carry a condom or keep condoms nearby; Discuss or negotiate the use of condoms before sexual intercourse; communicate intentions to use a condom; convince or persuade partner to use a condom; Make an agreement to use condoms or not have sex; use condoms correctly; maintain using condoms for every act of sexual intercourse. <p>Formulate a coping strategies plan for potential barriers or difficulties in planning to use condoms when having anal intercourse:</p> <p>Step 1: Recognize and list situations or barriers that may lead you to have unprotected anal intercourse (ie, positive emotional feelings for the partner, wishing to please the partner, not having condoms nearby at the time of sex, becoming “caught up in the passion of the moment,” and negative connotations about condom use);</p> <p>Step 2: Appraise the situation or barriers for potential problems;</p> <p>Step 3: Generate coping alternatives such as active coping strategies—active use of cognitive self-guidance, recall of both AIDS fears and safety benefits, and experience in safer sex.</p>	<p>Express confidence in their ability to use condoms during sexual intercourse</p> <ol style="list-style-type: none"> Confidence in their ability to buy or obtain condoms: Identifies where he can buy condoms, also after shop closing time Identifies appropriate condoms to buy and use Chooses the right condom size and brand and types of lube Confidence in their ability to always have a condom with them: Describes where he puts away his condoms, to have them available always Describes how he plans to carry a condom always Confidence in their action plan to use a condom: Describes a plan how he takes a condom out at the right moment Demonstrates how to use condoms correctly Ability to identify situations where they feel they may lack confidence in their ability to use condoms: Able to identify and manage situations where they may lack confidence in their ability to use condoms and devise a plan for each situation identified (eg, influence of alcohol and/or drugs, being sexually excited or turned on, experiencing pressure to engage in unprotected anal sex from partner).

Intervention Mapping Step 3: Theory-Based Methods and Practical Strategies

The third step in intervention mapping consists of determining which theories and theoretically based methods would be most effective in achieving the intervention’s performance objectives and then deciding which practical strategies would best operationalize those theoretical methods [18].

A review of the HIV prevention literature demonstrated that interventions informed by a theory were more successful than those that were not [42]. A systematic review indicated that extensive use of theory in an intervention’s design and evaluation was associated with high effect size estimates ($P=.049$) [43]. Theory informs the development of interventions by identifying theoretical constructs to be targeted and by explaining the mechanisms underlying specific behavioral change techniques [43], which increases the likelihood of their positive impact on behavioral change.

When examining various behavioral theories that addressed self-efficacy and intention, 2 particular theories emerged as relevant: the Theory of Planned Behavior and Social Cognitive Theory. Interventions based on these theories were found to be successful in increasing individuals’ self-efficacy and intention to use condoms [44,45]. The Theory of Planned Behavior proposes that intention plays a key role in the prediction of behavior [46]. According to a meta-analysis, interventions based on this theory tended to have moderate to large effects on behavior ($d=0.36-0.66$) [43,46]. Similarly, in a meta-analytical study examining the theory in the context of condom use,

intention and planned behavior construct accounted for 12.4% of the variance in condom use behavior [47].

Social Cognitive Theory, which is one of the most widely used models in studies of sexual transmission risk behaviors, identifies self-efficacy as a key determinant of behavior [48]. Specifically, the theory proposes that effective health behaviors are more likely to be adopted when individuals believe that they can implement them [49]. This proposition was supported by a meta-analysis that showed a positive correlation between self-efficacy and condom use [39].

Upon identifying the relevant theories to be used in the development of the intervention, possible theoretical methods and practical strategies to use in the intervention were then identified in addition to considerations to the theoretical parameters; the conditions under which the methods will work.

The 3 approaches that guided this process were to review the literature, engage in discussions with the clinicians in our multidisciplinary team, and consult with the target population. The empirical literature was reviewed to examine the types of theoretical methods and strategies included in interventions found to be effective in increasing condom use among the target population. One component of effective interventions was the provision of facts and information on HIV and STDs, which was expected to change cognitive factors such as attitudes and beliefs. Another targeted technical skills, such as the correct use of condoms, and interpersonal skills, such as those related to negotiating safer sex and assertiveness. The most common methods used to increase self-efficacy and intention involved modeling, role-playing, guided practice, building skills for

resistance to social pressure and shifting focus, and enhancing positive social supports. These theoretical methods were derived from both the Social Cognitive Theory and the Theory of Planned Behavior [33,50,51].

Meetings were also held with the local sexual health community clinicians on our team to determine from their perspective which methods and strategies would work in a Web-based intervention. They were asked about methods and strategies currently used in their clinics that they found to be effective to promote condom use within the HIV-positive MSM population. They highlighted the following methods as being of possible use in a Web-based intervention: knowledge acquisition, skills training, problem solving, and persuasive communication.

A focus group of 10 HIV-positive MSM was also conducted to explore their perspectives on acceptable methods and strategies for a Web-based intervention. After they were given information on the intended intervention's objectives and the theoretical methods and strategies under consideration, they identified the method of using role models and skill building as methods that would interest potential participants. Strategies in implementing

the methods included modeling scenarios, interactive activities, skills training, and problem solving. When discussing the possible methods, the parameters of the methods were also discussed with the focus group. The parameters are the conditions under which the methods will work. For example, when discussing the method of "modeling," some of the parameters include the learner identifying with the model, the model demonstrating feasible skills, model receives reinforcement, and learner perceives a coping model. Discussion with the focus group included the types of individuals whom the target population would identify with. The focus group identified that "peers" would best suit the modeling method. Throughout the intervention, videos using HIV-positive MSM were used. Also, throughout the intervention the models demonstrated their development of an action plan to use condoms and negotiating skills in relation to condom use with their partners. When demonstrating their negotiating skills, various scenarios were demonstrated in order to provide various coping strategies to the possible responses regarding the use of condoms by their partners (see Table 2).

Table 2. Selected examples of determinants, theoretical methods, and practical strategies.

Determinants	Theoretical method	Strategy
Intention	Mastery of skill building, modeling resistance, training refusal skills	Real video of individuals speaking about their action plans to always have a condom with them Scenario of two individuals negotiating condom use and the coping strategies used when partner refuses to use a condom
Self-efficacy	Active learning, modeling, guided practice	Character video of individuals negotiating condom use Interactive activity where individuals present responses to common phrases used in negotiating condom use

Focus group participants also noted the importance of having the information tailored to their individual needs and offering the interventions in a peer-to-peer format.

Coincidentally, tailoring has been recognized as an important component associated with the development and evaluation of Web-based interventions [52]. Tailoring has been defined as a process of creating individualized communications by gathering and assessing personal data related to a given health outcome to determine what information or strategies would be most appropriate to include to meet an individual's unique needs [53]. Compared with generic information, tailored feedback messages are more effective because they are more likely to be heard, read, or remembered and viewed as being personally relevant. They foster an active approach to the learning process, engage individuals in building their intention and self-efficacy, and alter their health risk behaviors [54-58]. Web-based interventions offer the possibility of using a constructed computer algorithm to tailor information supplied to participants based on their needs [52]. Delivering the tailored feedback through this Web-based method provides the ability to toggle between modalities, which enhances the user's experience and understanding of the material [54]. After reviewing the evidence gathered through the 3 previously discussed approaches, the multidisciplinary team confirmed that the Internet would be a viable mode to deliver the intervention.

Intervention Mapping Step 4

Intervention Components and Materials

Step 4 of intervention mapping entails determining the scope and sequence of the intervention's components and producing the materials for them [18]. Data from the focus group's perspectives regarding acceptable activities, the kind of visuals to be used in the intervention, and the type of language that would most resonate with our target population were all important factors that the team took into account when developing materials for the Web-based intervention. For example, with regard to making the visual aspect of the intervention attractive to the target population, the focus group participants stressed that it should contain vibrant colors, be interactive, and engage users with pictures of real individuals as well as animated characters. The focus group conceptualized the intervention as representing a journey in which participants would be able to gather information, engage in interactive activities, and receive tailored intervention messages based on their specific needs. They also highlighted the importance of having a peer who would guide users through the intervention and indicated that the language used in the intervention should remain neutral, in order to refrain from offending anyone. Thus, any crude or vulgar terms were to be avoided. Meetings with the Web design team took place to incorporate these aspects

from the focus group's feedback into the design of the intervention (see [Figures 2](#) and [3](#) for example).

Focus group participants also pointed to the need to include discussions of coping strategies as part of the intervention. They raised issues with previous interventions that gave information and strategies for condom use but neglected to discuss coping strategies to persuade partners to use a condom. A review of the literature and consultation with the health behavior expert on the multidisciplinary team identified a related gap between intention and behavior. Although intentions are considered the best predictors of behavior within the Theory of Planned Behavior, Sheeran's [59] meta-analysis has shown that intentions alone are not sufficient to predict behavioral change as there are large amounts of behavioral variance that are unexplained. This unexplained phenomenon, which has been termed *the intention to behavior gap* [60], occurs mainly among individuals who have intentions but who fail to act on them [60].

Bridging the gap between intention and behavior could be achieved by means of self-regulation skills, which are crucial for the uptake and maintenance of intended behavioral changes. Planning has been suggested as a self-regulatory skill that mediates between intention and behavior [60-62]. Two types

of self-regulatory planning are identified in the literature. Action planning, that is, plans regarding the when, where, and how of implementing the intended behavior, links the behavioral responses to situational cues, thereby facilitating initiation of the intended behavior [60]. Coping planning concerns anticipating the difficulties or barriers that may impede one's behavioral intentions. It involves making a detailed plan about how to pursue a behavior in the face of obstacles and represents a mental link between anticipated risk situations (ie, situations that endanger the performance of the intended behavior) and suitable coping responses [60,62,63]. Thus, based on the focus group's recommendations, the review of the literature, and discussions with the team's health behavior expert, a component was added to the intervention that consisted of action planning and coping planning. As a result, the final intervention is composed of 3 sessions, each lasting approximately 60 minutes, and participants have a total of 2 weeks to complete all 3 sessions. The first session will focus on planning condom use when having anal intercourse; session 2 will focus on negotiating the use of a condom with a partner; and the final session will focus on choosing not to have sexual intercourse without a condom. Each of the 3 sessions is composed of various activities in order to support the individual in accomplishing the session objectives (see [Table 3](#)).

Table 3. Intervention components.

Session	Objectives	Activities
Session one: 60 minutes	To plan condom use when having anal intercourse	<ul style="list-style-type: none"> Formulate an action plan to use condoms every time during sexual intercourse Express confidence in their ability to use condoms during sexual intercourse Confidence in their ability to buy or obtain condoms Confidence in their ability to always have a condom with them Confidence in their action plan to use a condom Confidence in their ability to use a condom when either not expected by the partner or there is pressure from partner to not use a condom Ability to identify situations where they feel they may lack confidence in their ability to use condoms adequately (eg, influence of alcohol and/or drugs).
Session two: 60 minutes	Negotiate with partner the use of a condom during sexual intercourse	<ul style="list-style-type: none"> Formulate an action plan to communicate your intentions of using a condom Express confidence in their ability to negotiate or discuss with their partners the use of a condom during anal intercourse Express confidence in their ability to convince partner to use a condom.
Session three: 60 minutes	Choosing not to have sexual intercourse without a condom	<ul style="list-style-type: none"> Formulate an action plan to refuse to have sex without a condom Express confidence in their ability to refuse having sexual intercourse without a condom Express confidence in their ability to refrain from unprotected sexual intercourse even if decision results in loss of their partner.

Figure 2. Peer-to-peer tailored video message of a peer's experiences with consistent condom use represents the method of using role models.

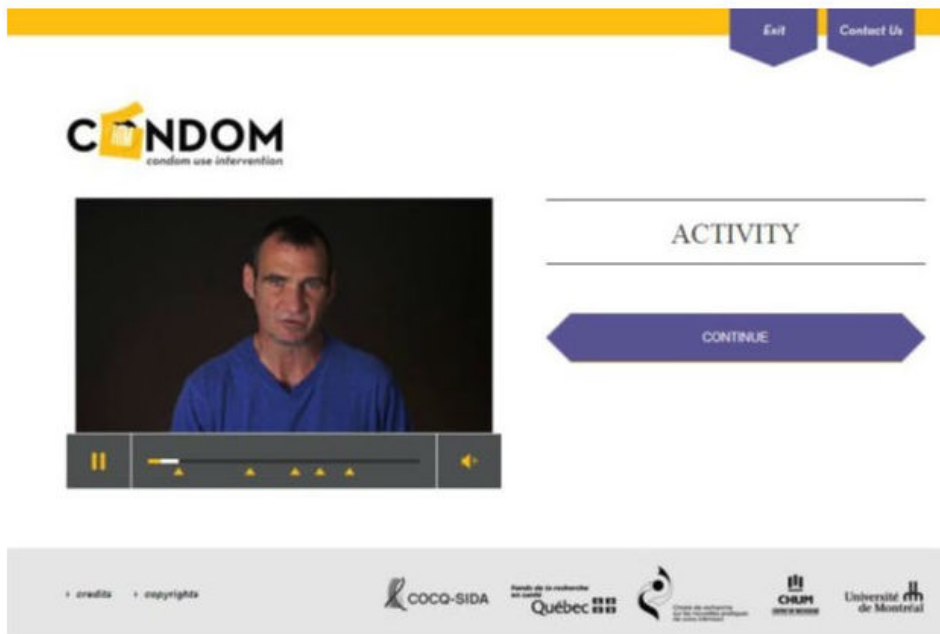
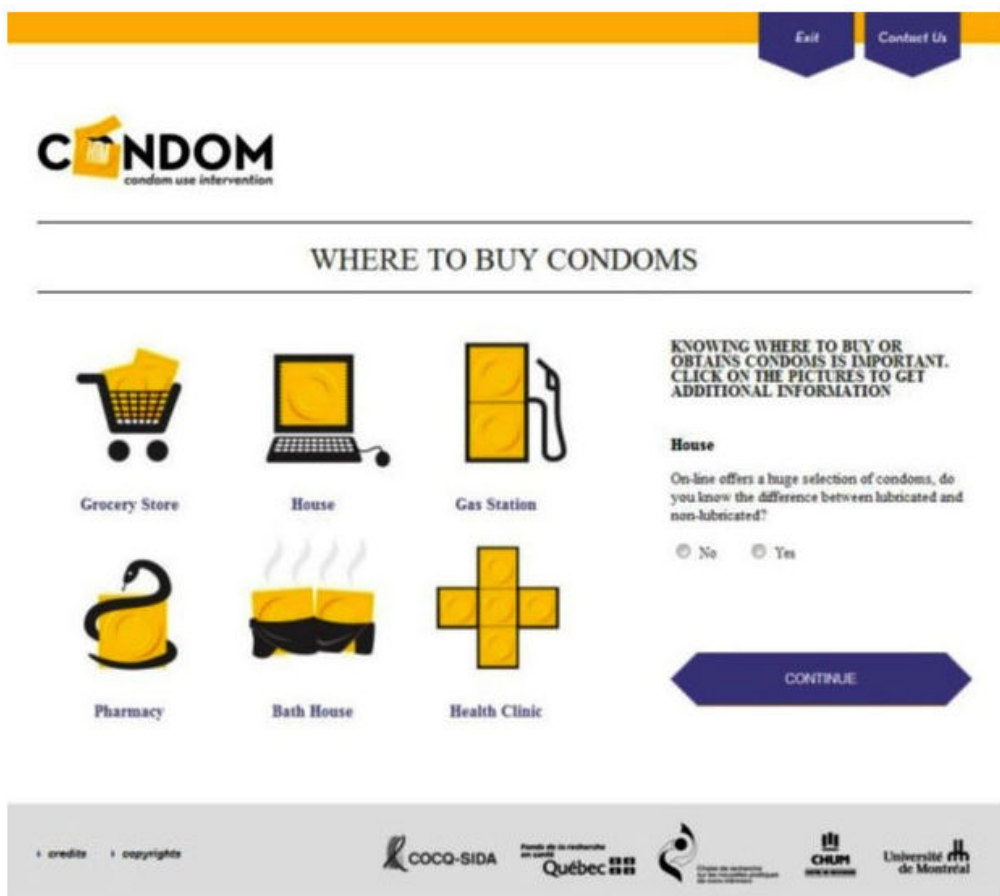


Figure 3. Interactive activity component of the Web-based intervention that facilitates knowledge acquisition.



Tailoring the Intervention's Components

Once the team selected what kind of components to include in the intervention, deciding on the sequence through which the intervention would be delivered, that is, tailoring, was the next step. The first task in generating an algorithm to tailor the intervention was to determine the process by which the intervention's messages would be tailored to the participants. Within the literature, message tailoring has been defined as a means of creating communications in which information about individuals is used to determine what specific content they will receive [64]. Hawkins et al [64] identified 3 distinct strategies through which tailoring goals can be achieved: personalization, feedback, and content matching. Discussions were held with the multidisciplinary team in order to debate the most appropriate method for tailoring our intervention.

A decision was made to tailor the intervention's messages by using a 3-stage approach that incorporated each of the 3 strategies identified by Hawkins et al [64]. The first stage of this process involved assessing participants' intention and self-efficacy to use condoms. The total score on the included measures would be used to assist in categorizing participants into 4 possible profiles (high intention and high self-efficacy, low intention and low self-efficacy, high intention and low self-efficacy, and low intention and high self-efficacy).

The second stage was based on participants' individual responses to each of the 12 items on the self-efficacy questionnaire and the 3 items on the intention questionnaire. Participants would be categorized into either a high or a low self-efficacy and intention category according to their response to each item on the questionnaires. The rationale for including this stage was that individuals who had been categorized as either low or high in intention and self-efficacy based on the total sum scores may not mimic the same categorization for each item of the questionnaires. For example, individuals who were categorized as having a high self-efficacy overall could possibly score themselves as having a low self-efficacy in a specific item. This second stage in the approach to tailoring the messages, therefore, would act as a safeguard in addressing specific item responses to the questionnaires that would not be apparent in the overall total sum scoring used in stage 1 of the tailoring process.

All 3 strategies identified by Hawkins et al [64] informed the third stage in our approach toward tailoring the intervention. In it, participants would be given a personalized message intended to validate whether or not they had been categorized correctly according to their responses to the questionnaires. Their response would determine the content of the message they would receive next. [Multimedia Appendix 1](#) presents an excerpt from the intervention algorithm of the 3-stage intervention message tailoring procedure.

After the algorithm for the tailored Web-based intervention messages was completed, a meeting with the Web design team and computer programmer was initiated in order to develop the front end and back end of the intervention. The front end concerned the visual and graphic nature of the Web-based intervention, while the back end concerned the programming component.

Intervention Mapping Step 5: Adoption, Implementation, and Maintenance

Although step 5 of intervention mapping is focused on the intervention's adoption, implementation, and maintenance, which must be considered to ensure that the intervention is delivered at acceptable levels of completeness and fidelity, considerations for program implementation actually begin as early as the needs assessment. Throughout the development process the research team did keep in mind the implementation of the intervention.

In addition to considerations for implementation, step 5 in intervention mapping also requires that a detailed plan be drawn up that would influence the behavior of individuals or groups who will make decisions about adopting and using the intervention. Before our intervention goes live on the Internet, both our Web design team and our computer programmer will initiate trial runs with it to test its functioning and implementation over the Internet. The maintenance of the Web-based intervention is a collaborative responsibility of the principal investigator and computer programmer. The computer-tailored Web-based intervention will be housed on the research center's servers. All data collection and intervention maintenance will be conducted by the principal investigator of the study and the computer programmer, both of whom have access to the intervention.

Intervention Mapping Step 6: Evaluation Planning

Step 6 of intervention mapping involves designing a plan to evaluate either the efficacy or the effectiveness of the intervention. This plan should be taken into consideration from the beginning of the needs assessment for the intervention and developed along with the rest of the intervention map. A pilot randomized controlled trial is planned to examine the feasibility, acceptability, and preliminary efficacy of this tailored Web-based intervention [65]. Subsequently, the research team will meet to discuss plans to make any adjustments needed in the intervention based on the results of the pilot study. After any adjustments are made, a full-scale, randomized controlled trial is planned in order to evaluate the effectiveness of the Web-based intervention in increasing condom use among HIV-positive MSM.

Discussion

Principal Findings

Reporting on the process used to design a particular intervention can serve as a valuable guide for the development of other new interventions and the refinement or revision of existing ones. It also provides a framework and methodology upon which others can build, as well as increases the transparency of the development process and enhances the interpretation of the particular intervention's effects. In this study, intervention mapping provided the research team with a systematic approach toward developing a theory- and evidence-based health prevention intervention. It afforded guidelines and tools that helped select the theoretical foundations for the intervention and apply the theory in the actual materials and activities produced for the intervention. In addition, the intervention

mapping approach allowed for the inclusion of individuals from the target population and clinicians from the field in the development process. The focus group from the target population assisted in validating the methods and strategies developed for the Web-based intervention and was also useful in articulating the visual aspect of the intervention.

The innovations of this particular intervention are its implementation over the Internet and the tailoring process used to assemble its Web-based messages. Given the association between HIV or STD transmission and the high level of Internet use by MSM, the Internet represents a logical method to deliver an intervention to this population. A Web-based intervention allows for the delivery of low-cost prevention messages to a greater number of people, in addition to connecting with individuals who are currently not being reached through more traditional methods [10]. Using the Internet to deliver prevention interventions is still in its infancy stage. To date, very few interventions are delivered completely over the Internet. The majority of them rely on traditional face-to-face interactions and paper-pencil assessments to recruit participants and collect data. This particular Web-based intervention relies solely on the Internet to recruit participants, collect data, and deliver the intervention. Providing a fully Web-based intervention increases efficiency as it involves low costs and has the potential to reach a larger population.

The 3-stage process used to tailor the Web-based messages for this intervention represents another advance in computer tailoring, which has become an increasingly common strategy to alter health risk behaviors [55]. Our tailoring process involves using an electronic algorithm developed by the researchers. Yet, to date, no other study has implemented a participant validation strategy within the tailoring algorithm.

Using intervention mapping to develop a tailored Web-based intervention turned out to be a lengthy and time-consuming process in our case. The main challenges encountered related to the iterative process and working within a large multidisciplinary team. Each development decision made was discussed with members of the research team. This process led to multiple revisions of the intervention. Yet, although time consuming, this process resulted in the production of a concrete intervention that will aim to increase condom use in HIV-positive MSM.

Conclusions

This paper describes the development process and key components of the Condom-HIM Web-based intervention using intervention mapping. Although a tedious process, the systematic process contributed to the development of a tailored, Web-based theory-driven intervention that has the potential to reach a larger population with minimal human resource costs associated with its implementation and evaluation.

Acknowledgments

The authors would like to thank Fonds de la Recherche en Santé du Québec for funding granted toward the development of the Condom-HIM Web-based application.

Authors' Contributions

JM conceived the intervention and drafted the manuscript. JC participated in the design of the intervention and provided feedback on the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Tailored messaging.

[[PDF File \(Adobe PDF File\), 34KB - publichealth_v3i2e20_app1.pdf](#)]

References

1. Landovitz RJ, Tseng C, Weissman M, Haymer M, Mendenhall B, Rogers K, et al. Epidemiology, sexual risk behavior, and HIV prevention practices of men who have sex with men using GRINDR in Los Angeles, California. *J Urban Health* 2013 Aug;90(4):729-739 [FREE Full text] [doi: [10.1007/s11524-012-9766-7](https://doi.org/10.1007/s11524-012-9766-7)] [Medline: [22983721](https://pubmed.ncbi.nlm.nih.gov/22983721/)]
2. Young SD, Szekeres G, Coates T. The relationship between online social networking and sexual risk behaviors among men who have sex with men (MSM). *PLoS One* 2013;8(5):e62271 [FREE Full text] [doi: [10.1371/journal.pone.0062271](https://doi.org/10.1371/journal.pone.0062271)] [Medline: [23658716](https://pubmed.ncbi.nlm.nih.gov/23658716/)]
3. Bolding G, Davis M, Hart G, Sherr L, Elford J. Where young MSM meet their first sexual partner: the role of the Internet. *AIDS Behav* 2007 Jul;11(4):522-526. [doi: [10.1007/s10461-007-9224-9](https://doi.org/10.1007/s10461-007-9224-9)] [Medline: [17347876](https://pubmed.ncbi.nlm.nih.gov/17347876/)]
4. Liao A, Millett G, Marks G. Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men. *Sex Transm Dis* 2006 Sep;33(9):576-584. [doi: [10.1097/01.olq.0000204710.35332.c5](https://doi.org/10.1097/01.olq.0000204710.35332.c5)] [Medline: [16540884](https://pubmed.ncbi.nlm.nih.gov/16540884/)]

5. Cruess DG, Burnham KE, Finitis DJ, Cherry C, Grebler T, Goshe BM, et al. Online partner seeking and sexual risk among HIV+ gay and bisexual men: a dialectical perspective. *Arch Sex Behav* 2016 Sep 26:- Epub ahead of print. [doi: [10.1007/s10508-016-0833-0](https://doi.org/10.1007/s10508-016-0833-0)] [Medline: [27671781](https://pubmed.ncbi.nlm.nih.gov/27671781/)]
6. Garofalo R, Herrick A, Mustanski BS, Donenberg GR. Tip of the Iceberg: young men who have sex with men, the internet, and HIV risk. *Am J Public Health* 2007 Jun;97(6):1113-1117. [doi: [10.2105/AJPH.2005.075630](https://doi.org/10.2105/AJPH.2005.075630)] [Medline: [17463378](https://pubmed.ncbi.nlm.nih.gov/17463378/)]
7. Horvath KJ, Rosser B R Simon, Remafedi G. Sexual risk taking among young internet-using men who have sex with men. *Am J Public Health* 2008 Jun;98(6):1059-1067. [doi: [10.2105/AJPH.2007.111070](https://doi.org/10.2105/AJPH.2007.111070)] [Medline: [18445804](https://pubmed.ncbi.nlm.nih.gov/18445804/)]
8. Bernstein KT, Stephens SC, Strona FV, Kohn RP, Philip SS. Epidemiologic characteristics of an ongoing syphilis epidemic among men who have sex with men, San Francisco. *Sex Transm Dis* 2013 Jan;40(1):11-17. [doi: [10.1097/OLQ.0b013e31827763ea](https://doi.org/10.1097/OLQ.0b013e31827763ea)] [Medline: [23254114](https://pubmed.ncbi.nlm.nih.gov/23254114/)]
9. Eaton LA, Maksut JL, Gamarel KE, Siembida EJ, Driffin DD, Baldwin R. Online sex partner meeting venues as a risk factor for testing HIV positive among a community-based sample of black men who have sex with men. *Sex Transm Dis* 2016 Jun;43(6):360-364. [doi: [10.1097/OLQ.0000000000000454](https://doi.org/10.1097/OLQ.0000000000000454)] [Medline: [27200520](https://pubmed.ncbi.nlm.nih.gov/27200520/)]
10. Chan PA, Towey C, Poceta J, Rose J, Bertrand T, Kantor R, et al. Online hookup sites for meeting sexual partners among men who have sex with men in Rhode Island, 2013: a call for public health action. *Public Health Rep* 2016;131(2):264-271 [FREE Full text] [doi: [10.1177/003335491613100210](https://doi.org/10.1177/003335491613100210)] [Medline: [26957661](https://pubmed.ncbi.nlm.nih.gov/26957661/)]
11. Eaton LA, Kalichman SC. Changes in transmission risk behaviors across stages of HIV disease among people living with HIV. *J Assoc Nurses AIDS Care* 2009;20(1):39-49 [FREE Full text] [doi: [10.1016/j.jana.2008.10.005](https://doi.org/10.1016/j.jana.2008.10.005)] [Medline: [19118770](https://pubmed.ncbi.nlm.nih.gov/19118770/)]
12. Rosser BR, Miner MH, Bockting WO, Ross MW, Konstan J, Gurak L, et al. HIV risk and the internet: results of the men's internet sex (MINTS) study. *AIDS Behav* 2009 Aug;13(4):746-756 [FREE Full text] [doi: [10.1007/s10461-008-9399-8](https://doi.org/10.1007/s10461-008-9399-8)] [Medline: [18512143](https://pubmed.ncbi.nlm.nih.gov/18512143/)]
13. Bull SS, Lloyd L, Rietmeijer C, McFarlane M. Recruitment and retention of an online sample for an HIV prevention intervention targeting men who have sex with men: the Smart Sex Quest Project. *AIDS Care* 2004 Nov;16(8):931-943. [doi: [10.1080/09540120412331292507](https://doi.org/10.1080/09540120412331292507)] [Medline: [15511725](https://pubmed.ncbi.nlm.nih.gov/15511725/)]
14. Davidovich U, de Wit John B F, Stroebe W. Behavioral and cognitive barriers to safer sex between men in steady relationships: implications for prevention strategies. *AIDS Educ Prev* 2004 Aug;16(4):304-314. [doi: [10.1521/aeap.16.4.304.40398](https://doi.org/10.1521/aeap.16.4.304.40398)] [Medline: [15342333](https://pubmed.ncbi.nlm.nih.gov/15342333/)]
15. Brug J, Oenema A, Ferreira I. Theory, evidence and intervention mapping to improve behavior nutrition and physical activity interventions. *Int J Behav Nutr Phys Act* 2005 Apr 4;2(1):2 [FREE Full text] [doi: [10.1186/1479-5868-2-2](https://doi.org/10.1186/1479-5868-2-2)] [Medline: [15807898](https://pubmed.ncbi.nlm.nih.gov/15807898/)]
16. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. *Health Psychol* 2008 May;27(3):379-387. [doi: [10.1037/0278-6133.27.3.379](https://doi.org/10.1037/0278-6133.27.3.379)] [Medline: [18624603](https://pubmed.ncbi.nlm.nih.gov/18624603/)]
17. Kok G, Mesters I. Getting inside the black box of health promotion programmes using intervention mapping. *Chronic Illn* 2011 Sep;7(3):176-180. [doi: [10.1177/1742395311403013](https://doi.org/10.1177/1742395311403013)] [Medline: [21900338](https://pubmed.ncbi.nlm.nih.gov/21900338/)]
18. Bartholomew K, Parcel G, Kok G, Gottlieb N, Fernandez M. *Planning Health Promotion Programs: An Intervention Mapping Approach*, 3rd Edition. San Francisco, CA: Jossey-Bass A Wiley Imprint; 2011.
19. Public Health Agency of Canada. Ottawa: Public Health Agency of Canada; 2015. HIV/AIDS in Canada: surveillance report to December 2014 URL: <https://www.canada.ca/content/dam/canada/health-canada/migration/healthy-canadians/publications/diseases-conditions-maladies-affections/hiv-aids-surveillance-2014-vih-sida/alt/hiv-aids-surveillance-2014-vih-sida-eng.pdf> [accessed 2017-03-16] [WebCite Cache ID [6p0dB1ffz](https://www.webcitation.org/6p0dB1ffz)]
20. Crepaz N, Marks G. Towards an understanding of sexual risk behavior in people living with HIV: a review of social, psychological, and medical findings. *AIDS* 2002 Jan 25;16(2):135-149. [Medline: [11807297](https://pubmed.ncbi.nlm.nih.gov/11807297/)]
21. Barnabas RV, Wasserheit JN, Huang Y, Janes H, Morrow R, Fuchs J, NIAID HIV Vaccine Trials Network. Impact of herpes simplex virus type 2 on HIV-1 acquisition and progression in an HIV vaccine trial (the step study). *J Acquir Immune Defic Syndr* 2011 Jul 1;57(3):238-244 [FREE Full text] [doi: [10.1097/QAI.0b013e31821acab5](https://doi.org/10.1097/QAI.0b013e31821acab5)] [Medline: [21860356](https://pubmed.ncbi.nlm.nih.gov/21860356/)]
22. Halkitis PN, Parsons JT. Intentional unsafe sex (barebacking) among HIV-positive gay men who seek sexual partners on the internet. *AIDS Care* 2003;15(3):367-378. [doi: [10.1080/0954012031000105423](https://doi.org/10.1080/0954012031000105423)]
23. Gilliam PP, Straub DM. Prevention with positives: a review of published research, 1998-2008. *J Assoc Nurses AIDS Care* 2009;20(2):92-109. [doi: [10.1016/j.jana.2008.11.001](https://doi.org/10.1016/j.jana.2008.11.001)] [Medline: [19286122](https://pubmed.ncbi.nlm.nih.gov/19286122/)]
24. Crepaz N, Lyles CM, Wolitski RJ, Passin WF, Rama SM, Herbst JH, HIV/AIDS Prevention Research Synthesis (PRS) Team. Do prevention interventions reduce HIV risk behaviours among people living with HIV? a meta-analytic review of controlled trials. *AIDS* 2006 Jan 9;20(2):143-157. [doi: [10.1097/01.aids.0000196166.48518.a0](https://doi.org/10.1097/01.aids.0000196166.48518.a0)] [Medline: [16511407](https://pubmed.ncbi.nlm.nih.gov/16511407/)]
25. The U.S. President's Emergency Plan for AIDS Relief (PEPFAR). 2009 Dec. PEPFAR and prevention, care, and treatment URL: https://www.pepfar.gov/about/strategy/prevention_care_treatment/133293.htm [accessed 2017-03-16] [WebCite Cache ID [6p0fXlg22](https://www.webcitation.org/6p0fXlg22)]
26. Wilson DP, Regan DG, Heymer K, Jin F, Prestage GP, Grulich AE. Serosorting may increase the risk of HIV acquisition among men who have sex with men. *Sex Transm Dis* 2010 Jan;37(1):13-17. [doi: [10.1097/OLQ.0b013e3181b35549](https://doi.org/10.1097/OLQ.0b013e3181b35549)] [Medline: [20118674](https://pubmed.ncbi.nlm.nih.gov/20118674/)]

27. Kennedy CE, Bernard LJ, Muessig KE, Konda KA, Akl EA, Lo Y, et al. Serosorting and HIV/STI infection among HIV-negative MSM and transgender people: a systematic review and meta-analysis to inform WHO guidelines. *J Sex Transm Dis* 2013;2013:583627 [FREE Full text] [doi: [10.1155/2013/583627](https://doi.org/10.1155/2013/583627)] [Medline: [26316960](https://pubmed.ncbi.nlm.nih.gov/26316960/)]
28. Lampe FC, Antiretrovirals, Sexual Transmission Risk Attitudes (ASTRA) Study Group. Sexual behaviour among people with HIV according to self-reported antiretroviral treatment and viral load status. *AIDS* 2016 Dec 17;30(11):1745-1759 [FREE Full text] [doi: [10.1097/QAD.0000000000001104](https://doi.org/10.1097/QAD.0000000000001104)] [Medline: [27045375](https://pubmed.ncbi.nlm.nih.gov/27045375/)]
29. Van Den Boom W, Stolte IG, Witlox R, Sandfort T, Prins M, Davidovich U. Undetectable viral load and the decision to engage in unprotected anal intercourse among HIV-positive MSM. *AIDS Behav* 2013 Jul;17(6):2136-2142. [doi: [10.1007/s10461-013-0453-9](https://doi.org/10.1007/s10461-013-0453-9)] [Medline: [23525838](https://pubmed.ncbi.nlm.nih.gov/23525838/)]
30. Crepaz N, Hart TA, Marks G. Highly active antiretroviral therapy and sexual risk behavior: a meta-analytic review. *JAMA* 2004 Jul 14;292(2):224-236. [doi: [10.1001/jama.292.2.224](https://doi.org/10.1001/jama.292.2.224)] [Medline: [15249572](https://pubmed.ncbi.nlm.nih.gov/15249572/)]
31. Cox J, Beauchemin J, Allard R. HIV status of sexual partners is more important than antiretroviral treatment related perceptions for risk taking by HIV positive MSM in Montreal, Canada. *Sex Transm Infect* 2004 Dec;80(6):518-523 [FREE Full text] [doi: [10.1136/sti.2004.011288](https://doi.org/10.1136/sti.2004.011288)] [Medline: [15572627](https://pubmed.ncbi.nlm.nih.gov/15572627/)]
32. Van de Ven P, Mao L, Fogarty A, Rawstorne P, Crawford J, Prestage G, et al. Undetectable viral load is associated with sexual risk taking in HIV serodiscordant gay couples in Sydney. *AIDS* 2005 Jan 28;19(2):179-184. [Medline: [15668543](https://pubmed.ncbi.nlm.nih.gov/15668543/)]
33. Bartholomew K, Parcel G, Kok G, Gottlieb N. *Planning Health Promotion Programs: Intervention Mapping, 2nd Edition*. San Francisco, CA: Jossey-Bass A Wiley Imprint; 2006.
34. Huang Z, Wang M, Fu L, Fang Y, Hao J, Tao F, et al. Intervention to increase condom use and HIV testing among men who have sex with men in China: a meta-analysis. *AIDS Res Hum Retroviruses* 2013 Mar;29(3):441-448. [doi: [10.1089/AID.2012.0151](https://doi.org/10.1089/AID.2012.0151)] [Medline: [23083341](https://pubmed.ncbi.nlm.nih.gov/23083341/)]
35. Noar SM, Black HG, Pierce LB. Efficacy of computer technology-based HIV prevention interventions: a meta-analysis. *AIDS* 2009 Jan 2;23(1):107-115. [doi: [10.1097/QAD.0b013e32831c5500](https://doi.org/10.1097/QAD.0b013e32831c5500)] [Medline: [19050392](https://pubmed.ncbi.nlm.nih.gov/19050392/)]
36. Scott-Sheldon LA, Huedo-Medina TB, Warren MR, Johnson BT, Carey MP. Efficacy of behavioral interventions to increase condom use and reduce sexually transmitted infections: a meta-analysis, 1991 to 2010. *J Acquir Immune Defic Syndr* 2011 Dec 15;58(5):489-498. [doi: [10.1097/QAI.0b013e31823554d7](https://doi.org/10.1097/QAI.0b013e31823554d7)] [Medline: [22083038](https://pubmed.ncbi.nlm.nih.gov/22083038/)]
37. Kok G, Harterink P, Vriens P, Zwart O, Hoppers HJ. The gay cruise: Developing a theory- and evidence-based Internet HIV-prevention intervention. *Sex Res Soc Policy* 2006 Jun;3(2):52-67. [doi: [10.1525/srsp.2006.3.2.52](https://doi.org/10.1525/srsp.2006.3.2.52)]
38. Schutz M, Godin G, Kok G, Vézina-Im L, Naccache H, Otis J. Determinants of condom use among HIV-positive men who have sex with men. *Int J STD AIDS* 2011 Jul;22(7):391-397. [doi: [10.1258/ijsa.2011.010205](https://doi.org/10.1258/ijsa.2011.010205)] [Medline: [21729958](https://pubmed.ncbi.nlm.nih.gov/21729958/)]
39. Casey MK, Timmermann L, Allen M, Krahn S, Turkiewicz KL. Response and self-efficacy of condom use: a meta-analysis of this important element of AIDS education and prevention. *Southern Communication Journal* 2009 Feb 02;74(1):57-78. [doi: [10.1080/10417940802335953](https://doi.org/10.1080/10417940802335953)]
40. Miner MH, Peterson JL, Welles SL, Jacoby SM, Rosser BR. How do social norms impact HIV sexual risk behavior in HIV-positive men who have sex with men?: multiple mediator effects. *J Health Psychol* 2009 Sep;14(6):761-770 [FREE Full text] [doi: [10.1177/1359105309338976](https://doi.org/10.1177/1359105309338976)] [Medline: [19687113](https://pubmed.ncbi.nlm.nih.gov/19687113/)]
41. Teng Y, Mak WW. The role of planning and self-efficacy in condom use among men who have sex with men: an application of the Health Action Process Approach model. *Health Psychol* 2011 Jan;30(1):119-128. [doi: [10.1037/a0022023](https://doi.org/10.1037/a0022023)] [Medline: [21299300](https://pubmed.ncbi.nlm.nih.gov/21299300/)]
42. Albarracín D, Gillette JC, Earl AN, Glasman LR, Durantini MR, Ho M. A test of major assumptions about behavior change: a comprehensive look at the effects of passive and active HIV-prevention interventions since the beginning of the epidemic. *Psychol Bull* 2005 Nov;131(6):856-897 [FREE Full text] [doi: [10.1037/0033-2909.131.6.856](https://doi.org/10.1037/0033-2909.131.6.856)] [Medline: [16351327](https://pubmed.ncbi.nlm.nih.gov/16351327/)]
43. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010;12(1):e4 [FREE Full text] [doi: [10.2196/jmir.1376](https://doi.org/10.2196/jmir.1376)] [Medline: [20164043](https://pubmed.ncbi.nlm.nih.gov/20164043/)]
44. Baban A, Craciun C. Changing health-risk behaviors: a review of theory and evidence-based interventions in health psychology. *J Cogn Behav Psychother* 2007;7(1):45-67.
45. Franssens D, Hoppers HJ, Kok G. Social-cognitive determinants of condom use in a cohort of young gay and bisexual men. *AIDS Care* 2009 Nov;21(11):1471-1479. [doi: [10.1080/09540120902883127](https://doi.org/10.1080/09540120902883127)] [Medline: [20024726](https://pubmed.ncbi.nlm.nih.gov/20024726/)]
46. Webb TL, Sheeran P. Does changing behavioral intentions engender behavior change? a meta-analysis of the experimental evidence. *Psychol Bull* 2006 Mar;132(2):249-268. [doi: [10.1037/0033-2909.132.2.249](https://doi.org/10.1037/0033-2909.132.2.249)] [Medline: [16536643](https://pubmed.ncbi.nlm.nih.gov/16536643/)]
47. Andrew BJ, Mullan BA, de Wit JB, Monds LA, Todd J, Kothe EJ. Does the theory of planned behaviour explain condom use behaviour among men who have sex with men? a meta-analytic review of the literature. *AIDS Behav* 2016 Dec;20(12):2834-2844. [doi: [10.1007/s10461-016-1314-0](https://doi.org/10.1007/s10461-016-1314-0)] [Medline: [26860535](https://pubmed.ncbi.nlm.nih.gov/26860535/)]
48. Safren SA, Traeger L, Skeer MR, O'Cleirigh C, Meade CS, Covahey C, et al. Testing a social-cognitive model of HIV transmission risk behaviors in HIV-infected MSM with and without depression. *Health Psychol* 2010 Mar;29(2):215-221 [FREE Full text] [doi: [10.1037/a0017859](https://doi.org/10.1037/a0017859)] [Medline: [20230095](https://pubmed.ncbi.nlm.nih.gov/20230095/)]

49. Reid AE, Aiken LS. Integration of five health behaviour models: common strengths and unique contributions to understanding condom use. *Psychol Health* 2011 Nov;26(11):1499-1520 [[FREE Full text](#)] [doi: [10.1080/08870446.2011.572259](https://doi.org/10.1080/08870446.2011.572259)] [Medline: [21678166](#)]
50. Mausbach BT, Semple SJ, Strathdee SA, Zians J, Patterson TL. Efficacy of a behavioral intervention for increasing safer sex behaviors in HIV-positive MSM methamphetamine users: results from the EDGE study. *Drug Alcohol Depend* 2007 Mar 16;87(2-3):249-257 [[FREE Full text](#)] [doi: [10.1016/j.drugalcdep.2006.08.026](https://doi.org/10.1016/j.drugalcdep.2006.08.026)] [Medline: [17182196](#)]
51. Michie S, Johnston M, Francis J, Hardeman W, Eccles M. From theory to intervention: mapping theoretically derived behavioural determinants to behaviour change techniques. *Appl Psychol* 2008;57(4):660-680. [doi: [10.1111/j.1464-0597.2008.00341](https://doi.org/10.1111/j.1464-0597.2008.00341)]
52. Noar SM, Grant Harrington N, Van Stee SK, Shemanski Aldrich R. Tailored health communication to change lifestyle behaviors. *Am J Lifestyle Med* 2010 Nov 19;5(2):112-122. [doi: [10.1177/1559827610387255](https://doi.org/10.1177/1559827610387255)]
53. Rimer BK, Kreuter MW. Advancing tailored health communication: a persuasion and message effects perspective. *J Commun* 2006 Aug;56(s1):S184-S201. [doi: [10.1111/j.1460-2466.2006.00289.x](https://doi.org/10.1111/j.1460-2466.2006.00289.x)]
54. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the web: review and analysis of key components. *Patient Educ Couns* 2009 Feb;74(2):156-173. [doi: [10.1016/j.pec.2008.08.023](https://doi.org/10.1016/j.pec.2008.08.023)] [Medline: [18947966](#)]
55. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. *Prev Med* 2010;51(3-4):214-221 [[FREE Full text](#)] [doi: [10.1016/j.ypmed.2010.06.004](https://doi.org/10.1016/j.ypmed.2010.06.004)] [Medline: [20558196](#)]
56. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive health communication applications for people with chronic disease. *Cochrane Database Syst Rev* 2004(4):CD004274. [doi: [10.1002/14651858.CD004274.pub2](https://doi.org/10.1002/14651858.CD004274.pub2)] [Medline: [15495094](#)]
57. Samoocha D, Bruinvels DJ, Elbers NA, Anema JR, van der Beek AJ. Effectiveness of web-based interventions on patient empowerment: a systematic review and meta-analysis. *J Med Internet Res* 2010;12(2):e23 [[FREE Full text](#)] [doi: [10.2196/jmir.1286](https://doi.org/10.2196/jmir.1286)] [Medline: [20581001](#)]
58. Bailey JV, Webster R, Hunter R, Griffin M, Freemantle N, Rait G, et al. The men's safer sex project: intervention development and feasibility randomised controlled trial of an interactive digital intervention to increase condom use in men. *Health Technol Assess* 2016 Dec;20(91):1-124 [[FREE Full text](#)] [doi: [10.3310/hta20910](https://doi.org/10.3310/hta20910)] [Medline: [27966409](#)]
59. Sheeran P. Intention-behavior relations: a conceptual and empirical review. *Eur J Soc Psychol* 2002 Jan;12(1):1-36. [doi: [10.1080/14792772143000003](https://doi.org/10.1080/14792772143000003)]
60. Scholz U, Schüz B, Ziegelmann JP, Lippke S, Schwarzer R. Beyond behavioural intentions: planning mediates between intentions and physical activity. *Br J Health Psychol* 2008 Sep;13(Pt 3):479-494. [doi: [10.1348/135910707X216062](https://doi.org/10.1348/135910707X216062)] [Medline: [17553212](#)]
61. Lippke S, Ziegelmann JP, Schwarzer R. Stage-specific adoption and maintenance of physical activity: testing a three-stage model. *Psychol Sport Exerc* 2005 Sep;6(5):585-603. [doi: [10.1016/j.psychsport.2004.11.002](https://doi.org/10.1016/j.psychsport.2004.11.002)]
62. Ziegelmann JP, Lippke S. Planning and strategy use in health behavior change: a life span view. *Int J Behav Med* 2007;14(1):30-39. [doi: [10.1080/10705500701316999](https://doi.org/10.1080/10705500701316999)] [Medline: [17511531](#)]
63. Sniehotta FF, Schwarzer R, Scholz U, Schüz B. Action planning and coping planning for long-term lifestyle change: theory and assessment. *Eur J Soc Psychol* 2005 Jul;35(4):565-576. [doi: [10.1002/ejsp.258](https://doi.org/10.1002/ejsp.258)] [Medline: [25855820](#)]
64. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. *Health Educ Res* 2008 Jun;23(3):454-466 [[FREE Full text](#)] [doi: [10.1093/her/cyn004](https://doi.org/10.1093/her/cyn004)] [Medline: [18349033](#)]
65. Miranda J, Côté J, Godin G, Blais M, Otis J, Guéhenec Y, et al. An internet-based intervention (condom-him) to increase condom use among hiv-positive men who have sex with men: protocol for a randomized controlled trial. *JMIR Res Protoc* 2013;2(2):e39 [[FREE Full text](#)] [doi: [10.2196/resprot.2723](https://doi.org/10.2196/resprot.2723)] [Medline: [24132072](#)]

Abbreviations

- HIV:** human immunodeficiency virus
 - MSM:** men who have sex with men
 - OR:** odds ratio
 - STD:** sexually transmitted disease
-

Edited by G Eysenbach; submitted 25.11.16; peer-reviewed by G Kok, L Scott-Sheldon, J Buttolph; comments to author 04.01.17; revised version received 07.02.17; accepted 18.02.17; published 19.04.17.

Please cite as:

Miranda J, Côté J

The Use of Intervention Mapping to Develop a Tailored Web-Based Intervention, Condom-HIM

JMIR Public Health Surveill 2017;3(2):e20

URL: <http://publichealth.jmir.org/2017/2/e20/>

doi: [10.2196/publichealth.7052](https://doi.org/10.2196/publichealth.7052)

PMID: [28428162](https://pubmed.ncbi.nlm.nih.gov/28428162/)

©Joyal Miranda, José Côté. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 19.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Viewpoint

Measuring Sexual Behavior Stigma to Inform Effective HIV Prevention and Treatment Programs for Key Populations

Shauna Stahlman¹, MPH, PhD; James R Hargreaves², PhD; Laurel Sprague^{3,4}, PhD; Anne L Stangl⁵, MPH, PhD; Stefan D Baral¹, MPH, MD

¹Center for Public Health and Human Rights, Department of Epidemiology, Johns Hopkins University, Baltimore, MD, United States

²Department of Social and Environmental Health Research, London School of Hygiene & Tropical Medicine, London, United Kingdom

³HIV Justice Network, Detroit, MI, United States

⁴Irvin D Reid Honors College, Wayne State University, Detroit, MI, United States

⁵Department of Global Health, Youth and Development, International Center for Research on Women, Washington, DC, United States

Corresponding Author:

Shauna Stahlman, MPH, PhD

Center for Public Health and Human Rights

Department of Epidemiology

Johns Hopkins University

615 N Wolfe St

Baltimore, MD, 21205

United States

Phone: 1 443 287 2370

Fax: 1 410 614 8371

Email: sstahlm1@jhu.edu

Abstract

Background: The levels of coverage of human immunodeficiency virus (HIV) treatment and prevention services needed to change the trajectory of the HIV epidemic among key populations, including gay men and other men who have sex with men (MSM) and sex workers, have consistently been shown to be limited by stigma.

Objective: The aim of this study was to propose an agenda for the goals and approaches of a sexual behavior stigma surveillance effort for key populations, with a focus on collecting surveillance data from 4 groups: (1) members of key population groups themselves (regardless of HIV status), (2) people living with HIV (PLHIV) who are also members of key populations, (3) members of nonkey populations, and (4) health workers.

Methods: We discuss strengths and weaknesses of measuring multiple different types of stigma including perceived, anticipated, experienced, perpetrated, internalized, and intersecting stigma as measured among key populations themselves, as well as attitudes or beliefs about key populations as measured among other groups.

Results: With the increasing recognition of the importance of stigma, consistent and validated stigma metrics for key populations are needed to monitor trends and guide immediate action. Evidence-based stigma interventions may ultimately be the key to overcoming the barriers to coverage and retention in life-saving antiretroviral-based HIV prevention and treatment programs for key populations.

Conclusions: Moving forward necessitates the integration of validated stigma scales in routine HIV surveillance efforts, as well as HIV epidemiologic and intervention studies focused on key populations, as a means of tracking progress toward a more efficient and impactful HIV response.

(*JMIR Public Health Surveill* 2017;3(2):e23) doi:[10.2196/publichealth.7334](https://doi.org/10.2196/publichealth.7334)

KEYWORDS

stigmatization; social stigma; HIV; male homosexuality; sex workers

Introduction

Sexual Behavior Stigma

As defined by Joint United Nations Program on human immunodeficiency virus infection and acquired immune deficiency syndrome (HIV/AIDS; UNAIDS), key populations are groups of people who are more likely to acquire or transmit HIV, and whose engagement is necessary for a successful HIV response [1]. Key populations include people living with HIV (PLHIV), men who have sex with men (MSM), transgender people, people who inject drugs, sex workers and their clients, and others. However, in this paper, we will focus primarily on MSM and sex workers and the sexual behavior-related stigma affecting these groups. Incidence and prevalence data indicate high and sustained HIV epidemics among MSM and sex workers across all income levels [2-18], as well as HIV care cascade outcomes that are far below the 90-90-90 target set by UNAIDS for 2020 [19].

Stigma toward key populations has been linked with adverse HIV-related outcomes that prevent reaching 90-90-90 testing and treatment targets. For example, among MSM, experience of stigma has been associated with reduced rates of HIV testing, increased fear and avoidance of seeking health care, and having condomless anal sex [20-26]. Interaction with health workers, who have not been properly trained in MSM clinical and cultural competency, may result in MSM avoiding HIV prevention or treatment services. In addition, reduced utilization of health and HIV services by MSM, due to enacted or perceived discrimination, may limit knowledge of the risks of unprotected anal intercourse and opportunities for access to prevention services [27,28]. These associations with HIV risk-related outcomes are likely amplified through the intersection of other types of stigma including stigma due to HIV status, race or ethnicity, or gender identity, as well as membership in multiple key population groups (eg, MSM sex workers) [29-31]. Similar patterns have been reported for female sex workers [32-34]. Consequently, the World Health Organization, UNAIDS, and the White House National HIV and AIDS Strategy recognize the important role that stigma plays in contributing to these negative outcomes and stress the need to reduce stigma affecting key populations [35-37]. Better understanding of the role that stigma plays in shaping HIV-related outcomes, tracking trends in stigma, and understanding the impact of interventions and policy changes on stigma all require accurate approaches to stigma surveillance and measurement.

Stigma is defined here as a social process that labels some people within a larger community as less valuable than others based on certain characteristics [38,39]. Several types of stigma have been characterized. Perceived stigma, also referred to as felt [40], or felt-normative [41] stigma, refers to the belief that individuals or societies treat people differently due to a stigmatized characteristic. Anticipated stigma is the fear or expectation of stigma or discrimination [42], whereas enacted or experienced stigma is the explicit experience of poor treatment on the basis of specific characteristics [43,44]. Internalized or self-stigma, which refers to the acceptance of one's lesser status within a society, manifests in low self-esteem,

self-isolation, and social withdrawal [45,46]. More recently, concern has also been raised about the severity and potential deadliness of internalized stigma in the form of internalized homophobia and violence [47]. Finally, intersecting stigma occurs when multiple forms of stigma interact within one individual or group, such as those related to being gay and living with HIV [48,49]. Stigma can result in discrimination, which has been defined as being treated differently based on an ascribed or perceived trait [50].

Stigma and discrimination affecting key populations can be caused by unequal access to social, economic, and political power that allows for separation, labeling, and stereotyping of groups [39]. In a social process, HIV is presented as preventable, with "immoral" behaviors causing the infections. Key populations are thus blamed for their own HIV infection, which can result in even further disadvantage [51]. Given the high levels of stigma affecting key populations, the consistent associations between stigma and adverse health outcomes, and the demonstrated ability to intervene on stigma through structural interventions or reduction of internalized stigma [52-56], sexual behavior stigma should be a priority to implement in surveillance efforts. Yet despite significant rhetoric from multiple HIV and health organizations highlighting the need to measure and address stigma, we found in a systematic review that most epidemiological and surveillance studies among MSM and female sex workers continue to focus on individual-level risks including numbers of partners and condomless anal sex, which have been well-established as risk factors of HIV infection [3,12,34]. Another systematic review of 549 English, French, and Spanish peer-reviewed articles that measured sex worker- or MSM-associated stigma from 2004 to 2014 identified a dearth of validated stigma metrics toward key populations in low- and middle-income settings (particularly in Africa and in the Middle East), a disproportionate lack of data on sex workers, and lack of studies utilizing mixed methods [57].

Outline for This Study

In this paper, we outline a vision for appropriate surveillance of sexual behavior stigma, defined as stigma that is anticipated, perceived, or experienced as a result of one's sexual practices [58,59]. Conceptually, sexual behavior stigma has considerable overlap with HIV-related stigma. In the early days of the HIV epidemic in high-income settings, HIV was designated as a "gay disease" and was linked to homosexuality and also to injection drug use [60]. Later, in the context of generalized HIV epidemics such as in southern and eastern sub-Saharan Africa, HIV was associated with several stigmatized behaviors including sexual promiscuity (eg, having multiple sexual partnerships) and sex work [61]. However, much current data on stigma relates explicitly to HIV infection. Although much of what we discuss in this paper has relevance for measurement of both HIV-related and sexual behavior stigma, our focus here is on the neglected area of measurement of stigma related to sexual behavior that can affect MSM and sex workers regardless of HIV status, and thereby impact access to testing, prevention, and treatment services for these groups.

The paper first outlines the goals and approaches of a sexual behavior stigma surveillance effort for which to strive. We then discuss the different population groups from whom relevant data might be collected and discuss the strengths and weaknesses of the potential data collection platforms reaching these groups. Second, we discuss the strengths and weaknesses of current sexual behavior stigma measurement methods among these different groups. We conclude by outlining an agenda for both action and research to strengthen surveillance of sexual behavior stigma.

A Vision for Sexual Behavior Stigma Surveillance

The primary objective of health surveillance is to (1) monitor trends in health issues for specific populations and (2) characterize the determinants of those most at risk for adverse health outcomes in order to inform public health action [62]. Key components of surveillance systems thus include accurate measurement or diagnostic tools for the phenomenon of interest and the application of epidemiological study designs and analysis that allow comparisons across time and place. The basic components of HIV surveillance for key populations have traditionally focused on monitoring risk practices and coverage of individual level interventions including HIV testing, condom

distribution, and treatment [34]. We suggest that this agenda should be expanded to routinely include information on sexual behavior stigma.

In Table 1, we map the major approaches for potential surveillance of sexual behavior stigma. This focuses on 4 population groups from which data could be collected: (1) members of key population groups themselves (regardless of HIV status), who may experience sexual behavior stigma, (2) PLHIV, who are also members of key populations and thus at particularly high risk of experiencing intersecting stigma, (3) members of nonkey population groups whose attitudes and behaviors toward members of key populations are the drivers of the stigma experienced by key populations in communities, and (4) health workers, whose attitudes and behaviors toward members of key populations are the drivers of the stigma experienced by key populations within health settings. In the case of the first two groups, opportunities to capture data on experiences of stigma are feasible through surveys, population-based cohorts, and through service delivery platforms. Relevant data for groups three and four can be collected through surveys and cohort studies of nonkey populations and health workers. Over the next section, we discuss current experiences from each of these approaches, reflecting on potential strengths and weaknesses of each.

Table 1. Summary of relevant populations, platforms, and areas of measurement.

Population	Data collection platforms	What can be asked
Key populations (regardless of HIV ^a status)	Surveys or cohorts using specialized sampling methodology such as RDS ^b , TLS ^c , or Internet Targeted service delivery platforms contacts	Experienced, perpetrated, perceived, anticipated, and internalized stigma as a result of sexual behavior
PLHIV ^d who are also members of key populations	Representative surveys of PLHIV ^d that are also able to collect data on stigmatized sexual practices; surveys or cohorts using specialized sampling methodology such as RDS ^b , TLS ^c , or Internet HIV ^a treatment and care programs that are also able to collect data on stigmatized sexual practices	Experienced, perpetrated, perceived, anticipated, and internalized stigma as a result of sexual behavior or HIV ^a status
Nonkey populations (regardless of HIV ^a status)	Population surveys or cohorts	Stigmatizing attitudes toward sexual practice; Perceptions or observations of stigma or discrimination;
Health workers	Health worker surveys or cohorts	Report of own discriminatory behaviors

^aHIV: human immunodeficiency syndrome.

^bRDS: respondent-driven sampling.

^cTLS: time-location sampling.

^dPLHIV: people living with HIV.

Data From Key Populations, Regardless of HIV Status

Respondent-driven, time-location, and other hidden-population sampling methods offer survey-based approaches to reaching key population groups and asking about their experiences, perceptions, or anticipations of sexual behavior stigma. These approaches have already been implemented in a variety of settings. For example, in a range of surveys, 20.1% (68/338) of MSM in Malawi reported being afraid to seek health services [63], 46.1% (149/323) reported experiencing stigma in

Swaziland [64], 76.2% (170/223) reported experiencing at least one form of discrimination in Lesotho [65], and up to 40% reported experiencing at least one discrimination event in Malawi (34.3%; 68/198), Botswana (56.9%; 66/116), and Namibia (41.5%; 88/212) [66,67]. About one-quarter (25.9%; 117/451) of MSM from a respondent-driven sampling (RDS) study in Vietnam reported high enacted homosexual stigma, 24% (108/451) reported high perceived community stigma, and 28.8% (130/451) reported high internalized homosexual stigma [68]. Also in Vietnam, 12.8% (38/297) of male sex workers

reported being hit, beaten, or sexually assaulted for engaging in sex work, and 61.1% (182/298) reported feeling afraid of being harassed or arrested by the police for engaging in sex work [69]. More than 60% of female sex workers responded with agreement or strong agreement to each item of a 10-item sex worker stigma index, which measured perceived stigma from the community and family members [70]. In China, 33.6% (242/721) of surveyed female sex workers reported that AIDS is a punishment for bad behavior, and 36.3% (262/721) agreed that people who get HIV through sex or drug use get what they deserve [71].

More recently, data have suggested a marked similarity in the prevalence of experienced, anticipated, and perceived sexual behavior stigma across settings. Recent analyses from Johns Hopkins and Emory University compared the prevalence of sexual behavior-related stigma toward MSM using data from the US and western and southern sub-Saharan Africa [72]. These data indicated that MSM in the US report similar, and in many cases, higher levels of stigma than MSM in sub-Saharan Africa because of their sexual practices, particularly from family members and health care workers. However, these comparisons are limited because the studies employed different sampling and survey methodologies across settings.

Cohort studies of key populations that measure stigma over time are rarer; although at least a few examples have been documented. The TRUST or RV368 study of MSM in Nigeria measured prevalence of stigma toward MSM before and after passage of the Same Sex Marriage Prohibition Act, demonstrating significantly higher reported occurrences of stigma after the passage of this discriminatory law [24]. Another pilot cohort of MSM in Senegal is currently assessing change in stigma over time as a result of community-based stigma mitigation interventions and MSM sensitization training for health workers [73]. Internalized homophobia has been measured over time in the Multicenter AIDS Cohort Study, an ongoing prospective study of MSM in the United States [74]. In Canada, the Maka Project Partnership includes a community-based cohort of female sex workers that reports on barriers to accessing health services [33,75]. Key population cohort studies lend further evidence for causality between experienced stigma and adverse health effects, and can be used to monitor change in stigma over time. However, these studies are often limited by time, feasibility, and cost constraints.

In addition to survey and cohort studies, another opportunity to ask questions about the experience of stigma might be through HIV testing and other service delivery platforms that reach key populations. However, this approach does not appear to have been widely implemented to date, reflected by a current lack of literature on this topic.

Data From People Living With HIV Including Those Who Also Self-Report Stigmatized Sexual Behaviors

A hugely important source of data on HIV-related stigma has been the PLHIV stigma index, a survey-based surveillance effort to understand the experiences and trends of stigma toward PLHIV. This PLHIV led and designed tool has been implemented in more than 70 countries, more than 1500 PLHIV have been trained as interviewers, and 85,000 PLHIV have been

interviewed [76]. Although focused on PLHIV overall, the results can provide useful information on HIV stigma experienced by key populations living with HIV, in that individuals are asked to disclose membership in a key group such as sex workers, gay, bisexual or other MSM, and people who inject drugs, among others. The standard survey allows for respondents to identify as being a “member” of up to as many as 9 key populations in keeping with the definitions suggested by the UNAIDS terminology guidance [77]. For example, in Belize, 23% of respondents who reported experiences of stigma for reasons other than their HIV status indicated that the stigma they faced was from their lesbian, gay, bisexual, or transgender (LGBT) identity [78]. In the Philippines, those who identified as gay or MSM reported much higher rates of exclusion from religious activities (16.2% to 7%), family activities (13.5% to 9.3%), and of verbal harassment (35.1% to 23.3%), as compared with non-MSM men. In particular, 48.6% of Filipino MSM living with HIV reported suicidal feelings because of their HIV status, a percentage more than 30% higher than the next highest gender group (women at 36.7%) [79]. In a focus group with MSM conducted as part of the Jamaican PLHIV stigma index, high levels of fear about discrimination based on HIV status were expressed, including one respondent’s characterization of disclosure as a “death sentence” [80]. In Vietnam, 6.5% of female sex workers and 2% of MSM reported physical assaults in the last 12 months, compared with 1.5% of women who were not sex workers and 0.9% of non-MSM men. Overall, in the Vietnam study, MSM reported the highest levels of social stigma and self-stigma of any group of PLHIV, followed by female sex workers and people who inject drugs [81].

Beyond violating the human rights of key populations living with HIV, the stigma experiences captured by the PLHIV stigma index can have serious health implications. In Ukraine, for example, sexual behavior stigma appears to increase hesitancy to take an HIV test. The percentage of PLHIV respondents in Ukraine indicating that they feared HIV testing because people may learn about their sexual or drug use behaviors increased from 10% to 18% between 2010 and 2013 [82]. In Philippines, MSM living with HIV were more likely to report that they avoided going to the hospital (27%) or clinic (37.8%) when they needed care than non-MSM men (23.3% and 32.6%) or women (23.3% and 33.3%) [79]. Further, in Malawi, lower percentages of LGBT PLHIV reported taking antiretroviral therapy (ART) or medications to prevent or to treat opportunistic infections than non-LGBT respondents [83].

However, a key challenge to successfully collecting these data through the PLHIV stigma index relies on reaching enough PLHIV survey respondents who are also members of key populations and on ensuring that key population respondents feel comfortable to accurately report their identifications. Apart from the PLHIV stigma index, several other smaller cross-sectional studies have collected behavioral data on MSM and sex workers, including the measurement of HIV status both self-reported and laboratory diagnosed. However, a smaller fraction of these studies have collected stigma information pertaining both to HIV status and sexual behaviors [84,85].

Data From Other Populations

Most conceptions of stigma hold within them the notion that stigma arises from a separation between those who do and do not carry a stigmatized trait: the “us” and “them,” in short hand. Stigma experienced by key populations therefore results from the actions, beliefs, or perceived beliefs about sexual behavior and key populations held in the communities in which they live. Monitoring these attitudes from other population groups is therefore another important component of sexual behavior stigma surveillance.

A number of tools for this purpose have been developed and applied. For example, one of the most commonly cited scales for measuring stigma toward MSM includes the Attitudes toward Lesbians and Gay Men Scale, which has been used in several studies since 1984 and revised multiple times and most recently in 2004 [86]. Other examples include the Reactions to Homosexuality Scale [87], and the Modern Homonegativity Scale [88]. For sex workers, scales are used less consistently across studies but include the Attitudes Toward Prostitution Scale [89] and the Attitudes Toward Prostitutes and Prostitution Scale [90]. In addition, validated indicators for HIV-related stigma have been developed for monitoring the 2016 United Nations Political Declaration on HIV and AIDS, including negative manifestations of HIV-related stigma, fear of HIV transmission, and discriminatory attitudes toward PLHIV [91].

Validated scales have been used across several settings to determine the attitude toward homosexuality and sex work among specific populations such as students or health workers, or among members of the general population [70,92-95]. An even greater number of studies have measured attitudes toward key populations, but without using validated scales. For example, in a survey of over 1000 university students in Jamaica measuring attitudes toward PLHIV, less than half reported sympathetic attitudes toward MSM (40%) or female sex workers (44%) living with HIV, although 67% and 81% reported sympathetic attitudes toward heterosexual men and nonsex worker women, respectively [96].

A more recent study estimated population-level trends in HIV stigma using data from the Demographic and Health Surveys and AIDS Indicator Surveys of 31 African countries between 2003 and 2013 [97]. These findings pointed to a decline in social distancing from PLHIV, supported by responses to two questions that have been indicated in field tests to be useful for measuring discriminating attitudes [98]. The authors also conclude an increase in anticipated stigma toward PLHIV. However, this conclusion was based on survey questions that have been suggested to be problematic based on cognitive interviewing, in that respondents likely have varying interpretations of these questions. Thus, we caution the interpretation of this trend and emphasize the importance of validating stigma-related survey instruments.

Data From Health Workers

Data from key population stigma surveys suggest that stigma experienced and perceived in health settings is widespread and acts as a barrier to accessing health services. Many of the same kinds of scales and questions as described above could be

applied to health workers as they have been in other nonkey populations. Such work is ongoing, for example, a survey of 332 staff members from health facilities and social service agencies in Jamaica and The Bahamas found that 77% believed that homosexuality is immoral, 72% believed that sex work is immoral, and 51% believed that HIV spreads due to immoral behavior [99]. A brief measurement tool to assess HIV among health care workers validated in 6 countries included questions assessing willingness to provide services to members of various key population groups if the provider had a choice. The study found that on average 13.1% and 12.4% of 1593 surveyed health care workers would prefer not to care for MSM and sex workers, respectively [100].

Another example of implementing stigma measures in a health service setting is the ancillary study of the HPTN071 or PopART trial. In this study, a cluster-randomized trial is being conducted in 21 communities in Zambia and South Africa. The ancillary study enrolled a take-all sample of both facility and community-based health workers, who are involved in the delivery of HIV testing and treatment services across all 21 communities, into a 3-year cohort study running alongside the main trial. Questions included in the cohort surveys include attitudes toward PLHIV, “women who sell sex” and “men who have sex with other men,” as well as perceptions of the way that these groups are treated by coworkers [84].

Addressing the Challenges to Strengthening Sexual Behavior Stigma Surveillance

As mentioned in the above section, there are clear opportunities to undertake surveillance of HIV-related stigma within surveys, cohort studies, and through service delivery platforms, and much work has already been undertaken. Mainstreaming sexual behavior stigma surveillance to more fully meet the vision described above will require this work to continue and to address directly the challenges and potential biases, which fall broadly into two types: measurement biases and selection biases.

Measurement Biases

Ongoing surveillance of sexual behavior stigma has the potential to be undermined through three major types of measurement bias. Confronting these is a key challenge that must be addressed. First, the validation of appropriate stigma metrics is needed. As described above, several stigma metrics for MSM and sex workers exist, and careful testing and validation is required for these.

Second, greater harmonization across time and settings in the use of these validated metrics is needed to support comparisons. Stigma takes many forms and is conceptually complex. It is perhaps therefore not surprising that a range of measures have been developed covering overlapping perspectives on stigma. However, in working toward a vision for surveillance of sexual behavior stigma, more widespread adoption of a much smaller number of approaches will be essential to facilitate comparisons over time and place. Although some details on the source or intersection of stigma may be lost by such harmonization, there is a much greater gain to be had in strengthening the capacity

to track trends. The current lack of trend data on sexual behavior stigma, as well as HIV stigma, is a direct result of this problem and must urgently be addressed going forward. In addition, there is limited consensus on the appropriate time periods of exposure for stigma, necessitating understanding the acute and chronic effects of exposure to stigma among key populations. Specifically, this includes understanding whether the effects of stigma in potentiating HIV risks among key populations are cumulative, leading to chronic elevated stress levels, or whether acute instances of stigma are more determinative in deciding the likelihood of uptake of health services. Ensuring consistency of measurement similar to other surveillance measures necessitates standardization of data collection instruments and sampling methods.

Third, the potential for social desirability and other reporting biases to bias assessments of prevalence and trends in sexual behavior stigma must be addressed. One method of evaluating the extent of social desirability is by measuring the prevalence of other sensitive topics, such as prevalence of condomless anal sex, to estimate the rate of under- or overreporting of stigma. Another method is by including social desirability scales and testing associations between the scale and variables of interest [101]. However, this remains a relatively poorly understood area. Biased reporting has proved a major barrier to the widespread adoption of sexual behavior surveillance data. On the other hand, methodological research has identified approaches that can improve reporting such as the use of trained and key population-friendly interviewers, Web-based assessments, or the use of audio computer-assisted self-interview tools [102,103].

Other potential measurement biases pertain to misclassification, as well as to the particular phrasing of questions. Exposure misclassification (eg, failure to disclose membership in a key population) can also occur in treatment program data and may be differential with respect to HIV status or other factors. Factors such as age, education level, medical history, internalized homophobia, and connectedness with the LGBT community can affect disclosure of sexual orientation to health care providers, according to a 2012 study of lesbian, gay, or bisexual (LGB) individuals in New York City, which is a city considered to be a “safe space” for gay men and other MSM [104]. In the same study, 90% of gay men disclosed their sexual orientation to a health care provider although only 60% of bisexual men disclosed [104]. Indeed, the likely underreporting of key population membership in health care or HIV treatment program settings would universally challenge surveillance efforts for MSM and sex workers.

Finally, survey questions that ask whether key population participants have ever witnessed a stigmatizing event occur may avoid social desirability bias, but they are problematic in the sense that they are dependent on a particular respondent’s network size. For example, respondents with larger social networks may be more likely to witness or hear about stigmatizing events occurring to the people within their network, as compared with those with smaller networks. Questions that ask about one’s own perpetration of stigma toward key population could be used to avoid this network size bias, but

responses to these questions would also be highly susceptible to social desirability biases.

Selection Biases

The second set of potential biases comprises selection biases that relate to the sampling approaches used to measure stigma toward key populations. Although a full discussion of such biases is beyond the scope of this paper, we note a few generic issues that would be of relevance to sexual behavior stigma. First, research remains ongoing on the representativeness of sampling strategies often used to reach hidden, key populations such as time-location or network-based sampling approaches. Novel approaches such as Web-based surveillance also hold great promise [105]. In all cases, a growing literature describes approaches to operationalizing, quality assurance checking, and analyzing data from such approaches so that the representativeness of the estimates derived from these samples is understood and the uncertainty in estimates is appropriately described. One growing area of interest is extrapolation of information from location-based samples to national-level estimation, and this would be as relevant for stigma metrics as it is for other indicators [106].

Furthermore, there has been debate regarding the appropriate sampling methodology for large surveillance efforts including the PLHIV stigma index, particularly for generating a representative sample across settings. It is presently argued that RDS, a peer-driven chain referral recruitment method, is best practice for sampling of hard-to-reach populations for HIV surveillance [107]. It has also been indicated that RDS can succeed at reaching members of key populations that are less engaged in HIV testing and less likely to be aware of living with HIV [108]. RDS is recommended in settings where time-location or venue-based sampling might overestimate measures of interest such as exposure to HIV prevention services or involvement in the gay community.

Data collected through service delivery platforms have both advantages and disadvantages. Such routine data are often cheap to collect and voluminous. However, the obvious disadvantage is that data are not collected from those who do not access services. This poses a challenge for representativeness and generalizability, but has not undermined the central importance of, for example, data on HIV infection from antenatal-clinic based surveillance or on HIV testing from voluntary counseling and testing clinics. Care and consideration must also be taken to ensure participant safety and sustained access to services, particularly in settings where sexual behaviors are criminalized, or settings with high levels of discriminatory attitudes affecting key populations. Service settings in all contexts should consider establishing plans to learn and respond to these data as a means of optimizing the implementation of their programs. We believe that service delivery platforms offer a potentially huge and important source of information on stigma, and that the utility of these data should be investigated as a matter of urgency.

Discussion

Summary

There is a sustained and often increasing burden of HIV in key populations worldwide, with uptake of HIV prevention and treatment services impeded by stigma. As a result, evidence-based stigma reduction interventions may ultimately be the key to overcoming the barriers to coverage and retention in ART and preexposure prophylaxis programs for key populations. However, to effectively measure and evaluate potential stigma mitigation or reduction interventions, consistent and validated stigma metrics for key populations should be integrated as a core component of HIV surveillance systems. To achieve this, there is a need to further validate stigma measures for key populations and attention must be given to the potential biases that might undermine surveillance aims.

Appropriate stigma surveillance measures are required to evaluate change in stigma over time and reduction of adverse HIV-related outcomes through specific mediated and modifiable pathways. Already, several studies have employed sensitization trainings for health care workers to increase competency in meeting the health care needs of key populations in many settings worldwide [52-56]. For example, in South Africa, the Anova Health Institute's *Health4Men* project aims to institutionalize competence in serving MSM in existing public clinics, which has resulted in over 160 competent sites nationally [56]. Only with validated and consistent stigma surveillance metrics, researchers and policy makers will be able to evaluate programming such as what is described here and to determine the ingredients of successful comprehensive intervention strategies.

Methods for testing the linkage of stigma surveillance data with individual-level biological outcomes among key populations would be needed to provide further evidence for adverse health effects of stigma, as well as potentially provide new perspectives to the surveillance data. There is currently much data to support the notion that stigma increases risk for adverse health outcomes, including mental health, depression, and also risk for HIV via fear or avoidance of health care seeking. However, the majority of studies rely on cross-sectional surveys and evidence would be bolstered through the use of prospective cohort data. Strong, high-quality surveillance data will ultimately be needed to change administrative policies such as the criminalization of sex work or same-sex practices that persist in many countries. However, although legal protections may reduce stigma in some instances, legislative protection alone is not sufficient to

eliminate stigma [109]. The first step to reducing the harmful effects of stigma on key populations is the appropriate surveillance and documentation of stigma phenomena.

Another key remaining issue is that there is significant work to be done to characterize intersectionality of HIV-related stigmas. Most studies have focused on internalized stigma, without reconciling the potential nonadditive effects of membership in one or more key populations, as well as HIV status, socioeconomic status, gender identity, and so on [110-113]. Moving forward, it may be useful to compare the strength of association between HIV-status stigma and treatment outcomes with the strength of association between sexual behavior stigma and treatment outcomes, among those living with HIV. Among those not living with HIV, we suggest further assessment of whether anticipated HIV stigma or the various types of sexual behavior stigma (anticipated, internalized, perceived, experienced, and so on) more strongly predict HIV risk. The results of such analyses may be used in future efforts to harmonize these measures for surveillance purposes. Although quantifying the intersectionality of these stigmas may allow for a more robust HIV response, there may be challenges to selecting the fewest number of survey items to accurately assess intersectionality while also attempting to reduce participant burden. It remains unclear whether it is important to collect the reason or attribute (eg, sexual practices, sexual or gender identities, occupation, socioeconomic status) thought to be causing the increased stigma or whether it sufficient to characterize the nonadditive burden of stigma [114].

Recommendations

Once the best practices of stigma surveillance have been established, monitoring trends over time will be important for many reasons. These include measuring the burden of stigma; guiding the planning, implementation, and evaluation of programs to prevent and mitigate stigma; detecting changes in public opinion or behavior; anticipating future trends; and providing a basis for epidemiologic investigation. Because very limited longitudinal data exist characterizing stigma affecting key populations, there are limited opportunities to assess the causality of stigma and its potential health effects, which is necessary for filling data gaps and motivating intervention resources. Thus, moving forward necessitates the integration of validated stigma scales during HIV epidemiologic, surveillance, and intervention studies focused on key populations to inform and establish comprehensive HIV and general health programming.

Acknowledgments

This work was supported by the Measurement & Surveillance of HIV Epidemics (MeSH) Consortium funded by the Bill & Melinda Gates Foundation. We thank Julian Hows, program officer at the Global Network of People Living with HIV (GNP+), for sharing information and advising on the PLHIV stigma index.

Conflicts of Interest

None declared.

References

1. UNAIDS. 2015. UNAIDS strategy 2016-2021: on the fast-track to end AIDS URL: http://www.unaids.org/sites/default/files/media_asset/20151027_UNAIDS_PCB37_15_18_EN_rev1.pdf [WebCite Cache ID 6nYpF7Cln]
2. Centers for Disease Control/Prevention (CDC). HIV infections attributed to male-to-male sexual contact - metropolitan statistical areas, United States and Puerto Rico, 2010. *MMWR Morb Mortal Wkly Rep* 2012 Nov 30;61(47):962-966 [FREE Full text] [Medline: 23190569]
3. Beyrer C, Sullivan P, Sanchez J, Baral SD, Collins C, Wirtz AL, et al. The increase in global HIV epidemics in MSM. *AIDS* 2013 Nov 13;27(17):2665-2678. [doi: 10.1097/01.aids.0000432449.30239.fe] [Medline: 23842129]
4. Feng Y, Bu K, Li M, Zhang X, Jin S, Wang L. [Meta-analysis of HIV infection incidence and risk factors among men who have sex with men in China]. *Zhonghua Liu Xing Bing Xue Za Zhi* 2015 Jul;36(7):752-758. [Medline: 26564708]
5. Zhang M, Chu Z, Wang H, Xu J, Lu C, Shang H. A rapidly increasing incidence of HIV and syphilis among men who have sex with men in a major city of China. *AIDS Res Hum Retroviruses* 2011 Nov;27(11):1139-1140. [doi: 10.1089/aid.2010.0356] [Medline: 21388334]
6. Peng Z, Yang H, Norris J, Chen X, Huan X, Yu R, et al. HIV incidence and predictors associated with retention in a cohort of men who have sex with men in Yangzhou, Jiangsu Province, China. *PLoS One* 2012;7(12):e52731 [FREE Full text] [doi: 10.1371/journal.pone.0052731] [Medline: 23285172]
7. Sanders EJ, Graham SM, Okuku HS, van der Elst EM, Muhaari A, Davies A, et al. HIV-1 infection in high risk men who have sex with men in Mombasa, Kenya. *AIDS* 2007 Nov 30;21(18):2513-2520. [doi: 10.1097/QAD.0b013e3282f2704a] [Medline: 18025888]
8. van Griensven F, Varangrat A, Wimonasate W, Tanpradech S, Kladsawad K, Chemnasiri T, et al. Trends in HIV prevalence, estimated HIV incidence, and risk behavior among men who have sex with men in Bangkok, Thailand, 2003-2007. *J Acquir Immune Defic Syndr* 2010 Feb;53(2):234-239. [doi: 10.1097/QAI.0b013e3181c2fc86] [Medline: 19901844]
9. Price MA, Rida W, Mwangome M, Mutua G, Middelkoop K, Roux S, et al. Identifying at-risk populations in Kenya and South Africa: HIV incidence in cohorts of men who report sex with men, sex workers, and youth. *J Acquir Immune Defic Syndr* 2012 Feb 01;59(2):185-193. [doi: 10.1097/QAI.0b013e31823d8693] [Medline: 22227488]
10. Phillips AN, Cambiano V, Nakagawa F, Brown AE, Lampe F, Rodger A, et al. Increased HIV incidence in men who have sex with men despite high levels of ART-induced viral suppression: analysis of an extensively documented epidemic. *PLoS One* 2013;8(2):e55312 [FREE Full text] [doi: 10.1371/journal.pone.0055312] [Medline: 23457467]
11. Beyrer C, Baral SD, Walker D, Wirtz AL, Johns B, Sifakis F. The expanding epidemics of HIV type 1 among men who have sex with men in low- and middle-income countries: diversity and consistency. *Epidemiol Rev* 2010;32:137-151. [doi: 10.1093/epirev/mxq011] [Medline: 20573756]
12. Beyrer C, Baral SD, van GF, 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]
13. Baral SD, 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.0000000000000037] [Medline: 24445371]
14. Sullivan PS, Jones JS, Baral SD. The global north: HIV epidemiology in high-income countries. *Curr Opin HIV AIDS* 2014 Mar;9(2):199-205. [doi: 10.1097/COH.0000000000000039] [Medline: 24445370]
15. Lou J, Blevins M, Ruan Y, Vermund SH, Tang S, Webb GF, et al. Modeling the impact on HIV incidence of combination prevention strategies among men who have sex with men in Beijing, China. *PLoS One* 2014;9(3):e90985. [doi: 10.1371/journal.pone.0090985] [Medline: 24626165]
16. 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]
17. Paz-Bailey G, Noble M, Salo K, Tregear SJ. Prevalence of HIV among U.S. female sex workers: systematic review and meta-analysis. *AIDS Behav* 2016 Oct;20(10):2318-2331 [FREE Full text] [doi: 10.1007/s10461-016-1332-y] [Medline: 26914165]
18. Baral S, Beyrer C, Muessig K, Poteat T, Wirtz AL, Decker MR, et al. Burden of HIV among female sex workers in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet Infect Dis* 2012 Jul;12(7):538-549. [doi: 10.1016/S1473-3099(12)70066-X] [Medline: 22424777]
19. Risher K, Mayer KH, Beyrer C. HIV treatment cascade in MSM, people who inject drugs, and sex workers. *Curr Opin HIV AIDS* 2015 Nov;10(6):420-429 [FREE Full text] [Medline: 26352393]
20. Knox J, Sandfort T, Yi H, Reddy V, Maimane S. Social vulnerability and HIV testing among South African men who have sex with men. *Int J STD AIDS* 2011 Dec;22(12):709-713 [FREE Full text] [doi: 10.1258/ijsa.2011.010350] [Medline: 22174050]
21. Hladik W, Barker J, Ssenkusu JM, Opio A, Tappero JW, Hakim A, et al. HIV infection among men who have sex with men in Kampala, Uganda--a respondent driven sampling survey. *PLoS One* 2012;7(5):e38143 [FREE Full text] [doi: 10.1371/journal.pone.0038143] [Medline: 22693590]

22. Hu Y, Lu H, Raymond HF, Sun Y, Sun J, Jia Y, et al. Measures of condom and safer sex social norms and stigma towards HIV/AIDS among Beijing MSM. *AIDS Behav* 2014 Jun;18(6):1068-1074 [[FREE Full text](#)] [doi: [10.1007/s10461-013-0609-7](https://doi.org/10.1007/s10461-013-0609-7)] [Medline: [24057931](#)]
23. Choi K, Hudes ES, Steward WT. Social discrimination, concurrent sexual partnerships, and HIV risk among men who have sex with men in Shanghai, China. *AIDS Behav* 2008 Jul;12(4 Suppl):S71-S77. [doi: [10.1007/s10461-008-9394-0](https://doi.org/10.1007/s10461-008-9394-0)] [Medline: [18427972](#)]
24. Schwartz SR, Nowak RG, Orazulike I, Keshinro B, Ake J, Kennedy S, et al. The immediate effect of the same-sex marriage prohibition act on stigma, discrimination, and engagement on HIV prevention and treatment services in men who have sex with men in Nigeria: analysis of prospective data from the TRUST cohort. *Lancet HIV* 2015 Jul;2(7):e299-e306. [doi: [10.1016/S2352-3018\(15\)00078-8](https://doi.org/10.1016/S2352-3018(15)00078-8)] [Medline: [26125047](#)]
25. Turan JM, Bukusi EA, Onono M, Holzemer WL, Miller S, Cohen CR. HIV/AIDS stigma and refusal of HIV testing among pregnant women in rural Kenya: results from the MAMAS Study. *AIDS Behav* 2011 Aug;15(6):1111-1120 [[FREE Full text](#)] [doi: [10.1007/s10461-010-9798-5](https://doi.org/10.1007/s10461-010-9798-5)] [Medline: [20827573](#)]
26. Logie CH, Newman PA, Weaver J, Roungraphon S, Tepjan S. HIV-related stigma and HIV prevention uptake among young men who have sex with men and transgender women in Thailand. *AIDS Patient Care STDS* 2016 Feb;30(2):92-100. [doi: [10.1089/apc.2015.0197](https://doi.org/10.1089/apc.2015.0197)] [Medline: [26788978](#)]
27. Pachankis JE, Hatzenbuehler ML, Hickson F, Weatherburn P, Berg RC, Marcus U, et al. Hidden from health: structural stigma, sexual orientation concealment, and HIV across 38 countries in the European MSM Internet Survey. *AIDS* 2015 Jun 19;29(10):1239-1246. [doi: [10.1097/QAD.0000000000000724](https://doi.org/10.1097/QAD.0000000000000724)] [Medline: [26035323](#)]
28. Ayala G, Makofane K, Santos G, Beck J, Do TD, Hebert P, et al. Access to basic HIV-related services and PrEP acceptability among men who have sex with men worldwide: barriers, facilitators, and implications for combination prevention. *J Sex Transm Dis* 2013;2013:953123 [[FREE Full text](#)] [doi: [10.1155/2013/953123](https://doi.org/10.1155/2013/953123)] [Medline: [26316968](#)]
29. Quinn K, Dickson-Gomez J. Homonegativity, religiosity, and the intersecting identities of young black men who have sex with men. *AIDS Behav* 2016 Jan;20(1):51-64 [[FREE Full text](#)] [doi: [10.1007/s10461-015-1200-1](https://doi.org/10.1007/s10461-015-1200-1)] [Medline: [26373283](#)]
30. Mburu G, Ram M, Siu G, Bitira D, Skovdal M, Holland P. Intersectionality of HIV stigma and masculinity in eastern Uganda: implications for involving men in HIV programmes. *BMC Public Health* 2014 Oct 11;14:1061 [[FREE Full text](#)] [doi: [10.1186/1471-2458-14-1061](https://doi.org/10.1186/1471-2458-14-1061)] [Medline: [25304035](#)]
31. Kennedy CE, Baral SD, Fielding-Miller R, Adams D, Dlodlu P, Sithole B, et al. "They are human beings, they are Swazi": intersecting stigmas and the positive health, dignity and prevention needs of HIV-positive men who have sex with men in Swaziland. *J Int AIDS Soc* 2013 Dec 02;16 Suppl 3:18749 [[FREE Full text](#)] [Medline: [24321112](#)]
32. Harcourt C, O'Connor J, Egger S, Fairley CK, Wand H, Chen MY, et al. The decriminalization of prostitution is associated with better coverage of health promotion programs for sex workers. *Aust N Z J Public Health* 2010 Oct;34(5):482-486. [doi: [10.1111/j.1753-6405.2010.00594.x](https://doi.org/10.1111/j.1753-6405.2010.00594.x)] [Medline: [21040176](#)]
33. Lazarus L, Deering KN, Nabess R, Gibson K, Tyndall MW, Shannon K. Occupational stigma as a primary barrier to health care for street-based sex workers in Canada. *Cult Health Sex* 2012;14(2):139-150 [[FREE Full text](#)] [doi: [10.1080/13691058.2011.628411](https://doi.org/10.1080/13691058.2011.628411)] [Medline: [22084992](#)]
34. Baral S, Holland CE, Shannon K, Logie C, Semugoma P, Sithole B, et al. Enhancing benefits or increasing harms: community responses for HIV among men who have sex with men, transgender women, female sex workers, and people who inject drugs. *J Acquir Immune Defic Syndr* 2014 Aug 15;66 Suppl 3:S319-S328. [doi: [10.1097/QAI.0000000000000233](https://doi.org/10.1097/QAI.0000000000000233)] [Medline: [25007203](#)]
35. World Health Organization. Guidelines: prevention and treatment of HIV and other sexually transmitted infections among men who have sex with men and transgender people: recommendations for a public health approach URL: http://www.who.int/hiv/pub/guidelines/msm_guidelines2011/en/ [accessed 2017-03-15] [[WebCite Cache ID 6nYpaeiwR](#)]
36. UNAIDS. Department of Health, New York. 2014. Action on social drivers to end AIDS and extreme poverty URL: https://www.health.ny.gov/diseases/aids/ending_the_epidemic/docs/key_resources/housing_and_supportive_services/social_drivers_to_end_aids.pdf [accessed 2017-03-15] [[WebCite Cache ID 6nYpjicS4](#)]
37. AIDS.gov. 2015. National HIV/AIDS Strategy for the United States: Updated to 2020 URL: <https://www.aids.gov/federal-resources/national-hiv-aids-strategy/overview/> [accessed 2017-01-16] [[WebCite Cache ID 6nYpmOKPE](#)]
38. Goffman E. *Stigma: Notes on the Management of Spoiled Identity*. New York: Simon and Schuster; Jun 15, 1986.
39. Link B, Phelan J. Conceptualizing stigma. *Annu Rev Sociol* 2001;27:363-385.
40. Green G. Attitudes towards people with HIV: are they as stigmatizing as people with HIV perceive them to be? *Soc Sci Med* 1995 Aug;41(4):557-568. [Medline: [7481950](#)]
41. Scambler G. Health-related stigma. *Sociol Health Illn* 2009 Apr;31(3):441-455 [[FREE Full text](#)] [doi: [10.1111/j.1467-9566.2009.01161.x](https://doi.org/10.1111/j.1467-9566.2009.01161.x)] [Medline: [19366430](#)]
42. Pryor JB, Bos AER. *Social Psychological Perspectives on Stigma: Advances in Theory and Research*. New York, NY: Routledge; 2016.
43. Nyblade LC. Measuring HIV stigma: existing knowledge and gaps. *Psychol Health Med* 2006 Aug;11(3):335-345. [doi: [10.1080/13548500600595178](https://doi.org/10.1080/13548500600595178)] [Medline: [17130069](#)]

44. Fay H, Baral SD, Trapence G, Motimedi F, Umar E, Ipinge S, et al. Stigma, health care access, and HIV knowledge among men who have sex with men in Malawi, Namibia, and Botswana. *AIDS Behav* 2011 Aug;15(6):1088-1097. [doi: [10.1007/s10461-010-9861-2](https://doi.org/10.1007/s10461-010-9861-2)] [Medline: [21153432](https://pubmed.ncbi.nlm.nih.gov/21153432/)]
45. Mak Winnie W S, Cheung Rebecca Y M, Law RW, Woo J, Li Patrick C K, Chung Rita W Y. Examining attribution model of self-stigma on social support and psychological well-being among people with HIV+/AIDS. *Soc Sci Med* 2007 Apr;64(8):1549-1559. [doi: [10.1016/j.socscimed.2006.12.003](https://doi.org/10.1016/j.socscimed.2006.12.003)] [Medline: [17223239](https://pubmed.ncbi.nlm.nih.gov/17223239/)]
46. Meyer IH. Minority stress and mental health in gay men. *J Health Soc Behav* 1995 Mar;36(1):38-56. [Medline: [7738327](https://pubmed.ncbi.nlm.nih.gov/7738327/)]
47. Dembosky A. KQED News. 2016. Orlando shooting raises questions about internalized homophobia and violence URL: <http://www2.kqed.org/stateofhealth/2016/06/17/orlando-shooting-raises-questions-about-internalized-homophobia-and-violence/> [WebCite Cache ID 6nYqFU04O]
48. Logie C, James L, Tharao W, Loutfy M. HIV, gender, race, sexual orientation, and sex work: a qualitative study of intersectional stigma experienced by HIV-positive women in Ontario, Canada. *PLoS Med* 2011 Nov;8(11):e1001124 [FREE Full text] [doi: [10.1371/journal.pmed.1001124](https://doi.org/10.1371/journal.pmed.1001124)] [Medline: [22131907](https://pubmed.ncbi.nlm.nih.gov/22131907/)]
49. Bauer GR. Incorporating intersectionality theory into population health research methodology: challenges and the potential to advance health equity. *Soc Sci Med* 2014 Jun;110:10-17 [FREE Full text] [doi: [10.1016/j.socscimed.2014.03.022](https://doi.org/10.1016/j.socscimed.2014.03.022)] [Medline: [24704889](https://pubmed.ncbi.nlm.nih.gov/24704889/)]
50. Kohler-Hausmann I. Oxford University Press. 2011. Discrimination URL: <http://www.oxfordbibliographies.com/view/document/obo-9780199756384/obo-9780199756384-0013.xml;jsessionid=B3F6D0D9C83EDC8E8CB0AF389CE3B27C> [accessed 2017-03-30] [WebCite Cache ID 6pLMpGxJv]
51. Deacon H. Towards a sustainable theory of health-related stigma: lessons from the HIV/AIDS literature. *J Community Appl Soc Psychol* 2006 Nov;16(6):418-425. [doi: [10.1002/casp.900](https://doi.org/10.1002/casp.900)]
52. Charurat ME, Emmanuel B, Akolo C, Keshinro B, Nowak RG, Kennedy S, et al. Uptake of treatment as prevention for HIV and continuum of care among HIV-positive men who have sex with men in Nigeria. *J Acquir Immune Defic Syndr* 2015 Mar 1;68 Suppl 2:S114-S123 [FREE Full text] [doi: [10.1097/QAI.0000000000000439](https://doi.org/10.1097/QAI.0000000000000439)] [Medline: [25723975](https://pubmed.ncbi.nlm.nih.gov/25723975/)]
53. McNaghten A, Kearns R, Siegler AJ, Phaswana-Mafuya N, Bekker L, Stephenson R, et al. Sibanye methods for prevention packages program project protocol: pilot study of HIV prevention interventions for men who have sex with men in South Africa. *JMIR Res Protoc* 2014;3(4):e55 [FREE Full text] [doi: [10.2196/resprot.3737](https://doi.org/10.2196/resprot.3737)] [Medline: [25325296](https://pubmed.ncbi.nlm.nih.gov/25325296/)]
54. van der Elst EM, Kombo B, Gichuru E, Omar A, Musyoki H, Graham SM, et al. The green shoots of a novel training programme: progress and identified key actions to providing services to MSM at Kenyan health facilities. *J Int AIDS Soc* 2015;18:20226 [FREE Full text] [Medline: [26493863](https://pubmed.ncbi.nlm.nih.gov/26493863/)]
55. Colchero MA, Bautista-Arredondo S, Cortés-Ortiz MA, Romero-Martinez M, Salas J, Sosa-Rubí SG, et al. Impact and economic evaluations of a combination prevention programme for men who have sex with men in Mexico. *AIDS* 2016 Jan;30(2):293-300. [doi: [10.1097/QAD.0000000000000925](https://doi.org/10.1097/QAD.0000000000000925)] [Medline: [26691549](https://pubmed.ncbi.nlm.nih.gov/26691549/)]
56. Henderson M. WHO. 2014. On the ground: Programmes serving the needs of key populations URL: http://apps.who.int/iris/bitstream/10665/145002/1/WHO_HIV_2014.50_eng.pdf [WebCite Cache ID 6nYqKvm7o]
57. Fitzgerald-Husek A, Wert M, Rosman L, Bowes A, Grosso A, Baral S. Measuring stigma related to sex worker identity and MSM identity: a systematic review. 2014 Presented at: APHA; November 18, 2014; New Orleans, LA.
58. Brown L, Macintyre K, Trujillo L. Interventions to reduce HIV/AIDS stigma: what have we learned? *AIDS Educ Prev* 2003 Feb;15(1):49-69. [Medline: [12627743](https://pubmed.ncbi.nlm.nih.gov/12627743/)]
59. Stangl AL, Lloyd JK, Brady LM, Holland CE, Baral S. A systematic review of interventions to reduce HIV-related stigma and discrimination from 2002 to 2013: how far have we come? *J Int AIDS Soc* 2013 Nov 13;16(3 Suppl 2):18734 [FREE Full text] [Medline: [24242268](https://pubmed.ncbi.nlm.nih.gov/24242268/)]
60. Shilts R. *And the Band Played On: Politics, People and the AIDS Epidemic*. New York, NY: Griffin; 1987.
61. The United Republic of Tanzania Prime Minister's Office, Tanzania Commission for AIDS (TACAIDS). Health Policy Initiative. 2013. National multi-sectoral HIV and AIDS stigma and discrimination reduction strategy URL: http://www.healthpolicyinitiative.com/Publications/Documents/1536_1_SD_Strategy_booklet_TACAIDS.pdf [accessed 2017-01-16] [WebCite Cache ID 6nYqPnvjZ]
62. Buehler JW. Surveillance. In: *Modern Epidemiology*. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.
63. Wirtz AL, Jumbe V, Trapence G, Kamba D, Umar E, Ketende S, et al. HIV among men who have sex with men in Malawi: elucidating HIV prevalence and correlates of infection to inform HIV prevention. *J Int AIDS Soc* 2013;16 Suppl 3:18742 [FREE Full text] [Medline: [24321110](https://pubmed.ncbi.nlm.nih.gov/24321110/)]
64. Risher K, Adams D, Sithole B, Ketende S, Kennedy C, Mnisi Z, et al. Sexual stigma and discrimination as barriers to seeking appropriate healthcare among men who have sex with men in Swaziland. *J Int AIDS Soc* 2013;16(3 Suppl 2):18715 [FREE Full text] [Medline: [24242263](https://pubmed.ncbi.nlm.nih.gov/24242263/)]
65. Baral S, Adams D, Lebona J, Kaibe B, Letsie P, Tshello R, et al. A cross-sectional assessment of population demographics, HIV risks and human rights contexts among men who have sex with men in Lesotho. *J Int AIDS Soc* 2011;14:36 [FREE Full text] [doi: [10.1186/1758-2652-14-36](https://doi.org/10.1186/1758-2652-14-36)] [Medline: [21726457](https://pubmed.ncbi.nlm.nih.gov/21726457/)]

66. Fay H, Baral SD, Trapence G, Motimedi F, Umar E, Ipinge S, et al. Stigma, health care access, and HIV knowledge among men who have sex with men in Malawi, Namibia, and Botswana. *AIDS Behav* 2011 Aug;15(6):1088-1097. [doi: [10.1007/s10461-010-9861-2](https://doi.org/10.1007/s10461-010-9861-2)] [Medline: [21153432](https://pubmed.ncbi.nlm.nih.gov/21153432/)]
67. Baral S, Trapence G, Motimedi F, Umar E, Ipinge S, Dausab F, et al. HIV prevalence, risks for HIV infection, and human rights among men who have sex with men (MSM) in Malawi, Namibia, and Botswana. *PLoS One* 2009;4(3):e4997 [FREE Full text] [doi: [10.1371/journal.pone.0004997](https://doi.org/10.1371/journal.pone.0004997)] [Medline: [19325707](https://pubmed.ncbi.nlm.nih.gov/19325707/)]
68. Ha H, Risser J, Ross M, Huynh N, Nguyen H. Homosexuality-related stigma and sexual risk behaviors among men who have sex with men in Hanoi, Vietnam. *Arch Sex Behav* 2015;44(2):349-356. [doi: [10.1007/s10508-014-0450-8](https://doi.org/10.1007/s10508-014-0450-8)]
69. Oldenburg CE, Biello KB, Colby D, Closson EF, Mai T, Nguyen T, et al. Stigma related to sex work among men who engage in transactional sex with men in Ho Chi Minh City, Vietnam. *Int J Public Health* 2014 Oct;59(5):833-840 [FREE Full text] [doi: [10.1007/s00038-014-0562-x](https://doi.org/10.1007/s00038-014-0562-x)] [Medline: [24858522](https://pubmed.ncbi.nlm.nih.gov/24858522/)]
70. Liu S, Srikrishnan AK, Zelaya CE, Solomon S, Celentano DD, Sherman SG. Measuring perceived stigma in female sex workers in Chennai, India. *AIDS Care* 2011 May;23(5):619-627. [doi: [10.1080/09540121.2010.525606](https://doi.org/10.1080/09540121.2010.525606)] [Medline: [21293991](https://pubmed.ncbi.nlm.nih.gov/21293991/)]
71. Yang Y, Wang J, Lin F, Zhang T, Yu F, Zhao Y, et al. Stigma against HIV/AIDS among female sex workers and general migrant women in eastern China. *BMC Womens Health* 2015 Jan 22;15:2. [doi: [10.1186/s12905-014-0160-3](https://doi.org/10.1186/s12905-014-0160-3)] [Medline: [25608866](https://pubmed.ncbi.nlm.nih.gov/25608866/)]
72. Stahlman S, Sanchez TH, Sullivan PS, Ketende S, Lyons C, Charurat ME, et al. The prevalence of sexual behavior stigma affecting gay men and other men who have sex with men across Sub-Saharan Africa and in the United States. *JMIR Public Health Surveill* 2016 Jul 26;2(2):e35 [FREE Full text] [doi: [10.2196/publichealth.5824](https://doi.org/10.2196/publichealth.5824)] [Medline: [27460627](https://pubmed.ncbi.nlm.nih.gov/27460627/)]
73. Dramé FM, Crawford EE, Diouf D, Beyrer C, Baral SD. A pilot cohort study to assess the feasibility of HIV prevention science research among men who have sex with men in Dakar, Senegal. *J Int AIDS Soc* 2013 Dec 02;16 Suppl 3:18753 [FREE Full text] [Medline: [24321115](https://pubmed.ncbi.nlm.nih.gov/24321115/)]
74. Herrick AL, Stall R, Chmiel JS, Guadamuz TE, Penniman T, Shoptaw S, et al. It gets better: resolution of internalized homophobia over time and associations with positive health outcomes among MSM. *AIDS Behav* 2013 May;17(4):1423-1430 [FREE Full text] [doi: [10.1007/s10461-012-0392-x](https://doi.org/10.1007/s10461-012-0392-x)] [Medline: [23283578](https://pubmed.ncbi.nlm.nih.gov/23283578/)]
75. Shannon K, Bright V, Allinott S, Alexson D, Gibson K, Tyndall MW, et al. Community-based HIV prevention research among substance-using women in survival sex work: the Maka Project Partnership. *Harm Reduct J* 2007 Dec 08;4:20 [FREE Full text] [doi: [10.1186/1477-7517-4-20](https://doi.org/10.1186/1477-7517-4-20)] [Medline: [18067670](https://pubmed.ncbi.nlm.nih.gov/18067670/)]
76. GNP+, ICW, UNAIDS. The PLHIV Stigma Index. 2016. URL: <http://www.stigmaindex.org> [accessed 2017-01-16] [WebCite Cache ID 6nYqVDpcA]
77. Joint United Nations Programme on HIV/AIDS (UNAIDS). 2015. Terminology Guidelines URL: http://www.unaids.org/sites/default/files/media_asset/2015_terminology_guidelines_en.pdf [WebCite Cache ID 6oZ9YGssU]
78. Stigma Index. 2013. Persons Living with HIV Stigma Index: Belize URL: <http://www.stigmaindex.org/belize> [accessed 2017-01-16] [WebCite Cache ID 6nYqaAGg9]
79. Trinidad A, Quinto D, Naldoza R. The experiences of external and internal stigma of HIV positive Filipinos. *Philippine Population Review* 2011;10(1):43-63.
80. Stigma Index. 2012. People living with HIV stigma index: Jamaica URL: <http://www.stigmaindex.org/jamaica> [accessed 2017-01-16] [WebCite Cache ID 6nYqdvZEd]
81. Stigma Index. 2014. Vietnam National Network of People Living with HIV (VNP+) URL: <http://www.stigmaindex.org/vietnam> [accessed 2017-01-16] [WebCite Cache ID 6nYqghuWN]
82. Stigma Index. 2013. All Ukrainian network of people living with HIV URL: <http://www.stigmaindex.org/ukraine> [accessed 2017-01-16] [WebCite Cache ID 6nYqit7nu]
83. Stigma Index. 2012. The people living with HIV stigma index URL: <http://www.stigmaindex.org/malawi> [accessed 2017-01-16] [WebCite Cache ID 6nYqkbbk0z]
84. Hargreaves JR, Stangl A, Bond V, Hoddinott G, Krishnaratne S, Mathema H, HPTN 071 (PopART) study team. HIV-related stigma and universal testing and treatment for HIV prevention and care: design of an implementation science evaluation nested in the HPTN 071 (PopART) cluster-randomized trial in Zambia and South Africa. *Health Policy Plan* 2016 Dec;31(10):1342-1354. [doi: [10.1093/heapol/czw071](https://doi.org/10.1093/heapol/czw071)] [Medline: [27375126](https://pubmed.ncbi.nlm.nih.gov/27375126/)]
85. Sanchez TH, Sineath RC, Kahle EM, Tregear SJ, Sullivan PS. The annual American men's internet survey of behaviors of men who have sex with men in the United States: protocol and key indicators report 2013. *JMIR Public Health Surveill* 2015;1(1):e3 [FREE Full text] [doi: [10.2196/publichealth.4314](https://doi.org/10.2196/publichealth.4314)] [Medline: [27227126](https://pubmed.ncbi.nlm.nih.gov/27227126/)]
86. Herek GM. Attitudes toward lesbians and gay men: a factor-analytic study. *J Homosex* 1984;10(1-2):39-51. [doi: [10.1300/J082v10n01_03](https://doi.org/10.1300/J082v10n01_03)] [Medline: [6520388](https://pubmed.ncbi.nlm.nih.gov/6520388/)]
87. Ross MW, Rosser BR. Measurement and correlates of internalized homophobia: a factor analytic study. *J Clin Psychol* 1996 Jan;52(1):15-21. [doi: [10.1002/\(SICI\)1097-4679\(199601\)52:1<15::AID-JCLP2>3.0.CO;2-V](https://doi.org/10.1002/(SICI)1097-4679(199601)52:1<15::AID-JCLP2>3.0.CO;2-V)] [Medline: [8682906](https://pubmed.ncbi.nlm.nih.gov/8682906/)]
88. Morrison MA, Morrison TG. Development and validation of a scale measuring modern prejudice toward gay men and lesbian women. *J Homosex* 2002;43(2):15-37. [Medline: [12739696](https://pubmed.ncbi.nlm.nih.gov/12739696/)]
89. Basow S, Campanile F. Attitudes toward prostitution as a function of attitudes toward feminism in college students. *Psychol Women Q* 1990;14(1):135-141.

90. Levin L, Peled E. The attitudes toward prostitutes and prostitution scale: a new tool for measuring public attitudes toward prostitutes and prostitution. *Res Soc Work Pract* 2011;21(5):582-593.
91. Indicator Registry. 2017. Indicators for Global AIDS Monitoring URL: <http://www.indicatorregistry.org/> [WebCite Cache ID 6oZ9rpkKI]
92. Moreno A, Herazo E, Oviedo H, Campo-Arias A. Measuring homonegativity: psychometric analysis of Herek's attitudes toward lesbians and gay men scale (ATLG) in Colombia, South America. *J Homosex* 2015;62(7):924-935. [doi: [10.1080/00918369.2014.1003014](https://doi.org/10.1080/00918369.2014.1003014)] [Medline: [25569818](https://pubmed.ncbi.nlm.nih.gov/25569818/)]
93. Papadaki V, Plotnikof K, Gioumidou M, Zisimou V, Papadaki E. A comparison of attitudes toward lesbians and gay men among students of helping professions in Crete, Greece: the cases of social work, psychology, medicine, and nursing. *J Homosex* 2015;62(6):735-762. [doi: [10.1080/00918369.2014.998956](https://doi.org/10.1080/00918369.2014.998956)] [Medline: [25530439](https://pubmed.ncbi.nlm.nih.gov/25530439/)]
94. Yu Y, Xiao S, Xiang Y. Application and testing the reliability and validity of a modified version of Herek's attitudes toward lesbians and gay men scale in China. *J Homosex* 2011;58(2):263-274 [FREE Full text] [doi: [10.1080/00918369.2011.540182](https://doi.org/10.1080/00918369.2011.540182)] [Medline: [21294029](https://pubmed.ncbi.nlm.nih.gov/21294029/)]
95. Strong KL, Folse VN. Assessing undergraduate nursing students' knowledge, attitudes, and cultural competence in caring for lesbian, gay, bisexual, and transgender patients. *J Nurs Educ* 2015 Jan;54(1):45-49. [doi: [10.3928/01484834-20141224-07](https://doi.org/10.3928/01484834-20141224-07)] [Medline: [25535762](https://pubmed.ncbi.nlm.nih.gov/25535762/)]
96. Norman LR, Carr R, Jiménez J. Sexual stigma and sympathy: attitudes toward persons living with HIV in Jamaica. *Cult Health Sex* 2006;8(5):423-433. [doi: [10.1080/13691050600855748](https://doi.org/10.1080/13691050600855748)] [Medline: [16923646](https://pubmed.ncbi.nlm.nih.gov/16923646/)]
97. Chan BT, Tsai AC. HIV stigma trends in the general population during antiretroviral treatment expansion: analysis of 31 countries in Sub-Saharan Africa, 2003-2013. *J Acquir Immune Defic Syndr* 2016 Aug 15;72(5):558-564. [doi: [10.1097/QAI.0000000000001011](https://doi.org/10.1097/QAI.0000000000001011)] [Medline: [27035888](https://pubmed.ncbi.nlm.nih.gov/27035888/)]
98. Nyblade L, MacQuarrie K, Phillip F, Kwesigabo G, Mbwambo J, Ndega J, et al. Measuring HIV stigma: Results of a field test in Tanzania. Washington, DC: Tanzania Stigma-Indicators Field Test Group; 2005.
99. Rogers SJ, Tureski K, Cushnie A, Brown A, Bailey A, Palmer Q. Layered stigma among health-care and social service providers toward key affected populations in Jamaica and The Bahamas. *AIDS Care* 2014;26(5):538-546. [doi: [10.1080/09540121.2013.844762](https://doi.org/10.1080/09540121.2013.844762)] [Medline: [24125067](https://pubmed.ncbi.nlm.nih.gov/24125067/)]
100. Nyblade L, Jain A, Benkirane M, Li L, Lohiniva A, McLean R, et al. A brief, standardized tool for measuring HIV-related stigma among health facility staff: results of field testing in China, Dominica, Egypt, Kenya, Puerto Rico and St. Christopher & Nevis. *J Int AIDS Soc* 2013 Nov 13;16(3 Suppl 2):18718 [FREE Full text] [Medline: [24242266](https://pubmed.ncbi.nlm.nih.gov/24242266/)]
101. Reynolds W. Development of reliable and valid short forms of the marlowe-crowne social desirability scale. *J Clin Psychol* 1982;38(1):119-125.
102. Castro R. Inconsistent respondents and sensitive questions. *Field Methods* 2012 Dec 11;25(3):283-298. [doi: [10.1177/1525822x12466988](https://doi.org/10.1177/1525822x12466988)]
103. Adebajo S, Obianwu O, Eluwa G, Vu L, Oginni A, Tun W, et al. Comparison of audio computer assisted self-interview and face-to-face interview methods in eliciting HIV-related risks among men who have sex with men and men who inject drugs in Nigeria. *PLoS One* 2014;9(1):e81981 [FREE Full text] [doi: [10.1371/journal.pone.0081981](https://doi.org/10.1371/journal.pone.0081981)] [Medline: [24416134](https://pubmed.ncbi.nlm.nih.gov/24416134/)]
104. Durso L, Meyer I. Patterns and predictors of disclosure of sexual orientation to healthcare providers among lesbians, gay men, and bisexuals. *Sex Res Social Policy* 2013;10(1):35-42. [Medline: [23463442](https://pubmed.ncbi.nlm.nih.gov/23463442/)]
105. JMIR. 2015. American men's internet survey (AMIS): online HIV behavioral survey of men who have sex with men URL: <https://publichealth.jmir.org/article/downloadSuppFile/4314/22192> [accessed 2017-03-30] [WebCite Cache ID 6pLO4YqBM]
106. Yu D, Calleja JM, Zhao J, Reddy A, Seguy N, Technical Consultation on Lessons Learnt from Size Estimation among Key Populations in Asian Countries. Estimating the size of key populations at higher risk of HIV infection: a summary of experiences and lessons presented during a technical meeting on size estimation among key populations in Asian countries. *Western Pac Surveill Response J* 2014;5(3):43-49 [FREE Full text] [doi: [10.5365/WPSAR.2014.5.2.008](https://doi.org/10.5365/WPSAR.2014.5.2.008)] [Medline: [25320676](https://pubmed.ncbi.nlm.nih.gov/25320676/)]
107. Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS* 2005 May;19 Suppl 2:S67-S72. [Medline: [15930843](https://pubmed.ncbi.nlm.nih.gov/15930843/)]
108. Stahlman S, Johnston LG, Yah C, Ketende S, Maziya S, Trapence G, et al. Respondent-driven sampling as a recruitment method for men who have sex with men in southern Sub-Saharan Africa: a cross-sectional analysis by wave. *Sex Transm Infect* 2016 Jun;92(4):292-298. [doi: [10.1136/sextrans-2015-052184](https://doi.org/10.1136/sextrans-2015-052184)] [Medline: [26424712](https://pubmed.ncbi.nlm.nih.gov/26424712/)]
109. Zahn R, Grosso A, Scheibe A, Bekker L, Ketende S, Dausab F, et al. Human rights violations among men who have sex with men in Southern Africa: comparisons between legal contexts. *PLoS One* 2016;11(1):e0147156 [FREE Full text] [doi: [10.1371/journal.pone.0147156](https://doi.org/10.1371/journal.pone.0147156)] [Medline: [26764467](https://pubmed.ncbi.nlm.nih.gov/26764467/)]
110. Pyun T, Santos G, Arreola S, Do T, Hebert P, Beck J, et al. Internalized homophobia and reduced HIV testing among men who have sex with men in China. *Asia Pac J Public Health* 2014 Mar;26(2):118-125. [doi: [10.1177/1010539514524434](https://doi.org/10.1177/1010539514524434)] [Medline: [24554493](https://pubmed.ncbi.nlm.nih.gov/24554493/)]
111. Quinn K, Dickson-Gomez J, DiFranceisco W, Kelly JA, St Lawrence JS, Amirkhanian YA, et al. Correlates of internalized homonegativity among black men who have sex with men. *AIDS Educ Prev* 2015 Jun;27(3):212-226 [FREE Full text] [doi: [10.1521/aeap.2015.27.3.212](https://doi.org/10.1521/aeap.2015.27.3.212)] [Medline: [26010313](https://pubmed.ncbi.nlm.nih.gov/26010313/)]

112. Ross M, Kajubi P, Mandel J, McFarland W, Raymond H. Internalized homonegativity/homophobia is associated with HIV-risk behaviours among Ugandan gay and bisexual men. *Int J STD AIDS* 2013 May;24(5):409-413. [doi: [10.1177/0956462412472793](https://doi.org/10.1177/0956462412472793)] [Medline: [23970711](https://pubmed.ncbi.nlm.nih.gov/23970711/)]
113. Peterson J, Bakeman R, Sullivan P, Millett G, Rosenberg E, Salazar L, et al. Social discrimination and resiliency are not associated with differences in prevalent HIV infection in black and white men who have sex with men. *J Acquir Immune Defic Syndr* 2014 Aug 15;66(5):538-543 [FREE Full text] [doi: [10.1097/QAI.000000000000203](https://doi.org/10.1097/QAI.000000000000203)] [Medline: [24820109](https://pubmed.ncbi.nlm.nih.gov/24820109/)]
114. Tsai A, Bangsberg D, Weiser S. Harnessing poverty alleviation to reduce the stigma of HIV in Sub-Saharan Africa. *PLoS Med* 2013 Nov;10(11):e1001557 [FREE Full text] [doi: [10.1371/journal.pmed.1001557](https://doi.org/10.1371/journal.pmed.1001557)] [Medline: [24319400](https://pubmed.ncbi.nlm.nih.gov/24319400/)]

Abbreviations

AIDS: acquired immunodeficiency syndrome
ART: antiretroviral therapy
HIV: human immunodeficiency virus
LGB: lesbian, gay, or bisexual
LGBT: lesbian, gay, bisexual, or transgender
MSM: men who have sex with men
PLHIV: people living with HIV
RDS: respondent-driven sampling
TLS: time-location sampling
UNAIDS: Joint United Nations Program on HIV and AIDS

Edited by K Sabin; submitted 16.01.17; peer-reviewed by V Bendaud, K Deuba; comments to author 19.02.17; revised version received 26.02.17; accepted 01.03.17; published 26.04.17.

Please cite as:

Stahlman S, Hargreaves JR, Sprague L, Stangl AL, Baral SD

Measuring Sexual Behavior Stigma to Inform Effective HIV Prevention and Treatment Programs for Key Populations

JMIR Public Health Surveill 2017;3(2):e23

URL: <http://publichealth.jmir.org/2017/2/e23/>

doi: [10.2196/publichealth.7334](https://doi.org/10.2196/publichealth.7334)

PMID: [28446420](https://pubmed.ncbi.nlm.nih.gov/28446420/)

©Shauna Stahlman, James R Hargreaves, Laurel Sprague, Anne L Stangl, Stefan D Baral. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 26.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Short Paper

The RSVP Project: Factors Related to Disengagement From Human Immunodeficiency Virus Care Among Persons in San Francisco

Susan Scheer¹, MPH, PhD; Miao-Jung Chen¹, MPH, PhD; Maree Kay Parisi¹, BA; Maya Yoshida-Cervantes¹, MPH; Erin Antunez², MS; Viva Delgado¹, MPH; Nicholas J Moss³, MPH, MD; Kate Buchacz⁴, MPH, PhD

¹HIV Epidemiology Section, San Francisco Department of Public Health, San Francisco, CA, United States

²Disease Prevention and Control, San Francisco Department of Public Health, San Francisco, CA, United States

³Division of Communicable Disease Control and Prevention, Alameda County Public Health Department, Oakland, CA, United States

⁴Divisions of HIV/AIDS Prevention, Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:

Susan Scheer, MPH, PhD

HIV Epidemiology Section

San Francisco Department of Public Health

25 Van Ness Avenue

Suite 500

San Francisco, CA, 94114

United States

Phone: 1 415 437 6259

Fax: 1 415 431 0353

Email: susan.scheer@sfdph.org

Abstract

Background: In the United States, an estimated two-thirds of persons with human immunodeficiency virus (HIV) infection do not achieve viral suppression, including those who have never engaged in HIV care and others who do not stay engaged in care. Persons with an unsuppressed HIV viral load might experience poor clinical outcomes and transmit HIV.

Objective: The goal of the Re-engaging Surveillance-identified Viremic Persons (RSVP) project in San Francisco, CA, was to use routine HIV surveillance databases to identify, contact, interview, and reengage in HIV care persons who appeared to be out of care because their last HIV viral load was unsuppressed. We aimed to interview participants about their HIV care and barriers to reengagement.

Methods: Using routinely collected HIV surveillance data, we identified persons with HIV who were out of care (no HIV viral load and CD4 laboratory reports during the previous 9-15 months) and with their last plasma HIV RNA viral load >200 copies/mL. We interviewed the located persons, at baseline and 3 months later, about whether and why they disengaged from HIV care and the barriers they faced to care reengagement. We offered them assistance with reengaging in HIV care from the San Francisco Department of Public Health linkage and navigation program (LINCS).

Results: Of 282 persons selected, we interviewed 75 (26.6%). Of these, 67 (89%) reported current health insurance coverage, 59 (79%) had ever been prescribed and 45 (60%) were currently taking HIV medications, 59 (79%) had seen an HIV provider in the past year, and 34 (45%) had missed an HIV appointment in the past year. Reasons for not seeing a provider included feeling healthy, using alcohol or drugs, not having enough money or health insurance, and not wanting to take HIV medicines. Services needed to get to an HIV medical care appointment included transportation assistance, stable living situation or housing, sound mental health, and organizational help and reminders about appointments. A total of 52 (69%) accepted a referral to LINCS. Additionally, 64 (85%) of the persons interviewed completed a follow-up interview 3 months later and, of these, 62 (97%) had health insurance coverage and 47 (73%) reported having had an HIV-related care appointment since the baseline interview.

Conclusions: Rather than being truly out of care, most participants reported intermittent HIV care, including recent HIV provider visits and health insurance coverage. Participants also frequently reported barriers to care and unmet needs. Health department assistance with HIV care reengagement was generally acceptable. Understanding why people previously in HIV care disengage from care and what might help them reengage is essential for optimizing HIV clinical and public health outcomes.

KEYWORDS

HIV care; HIV care continuum; engagement in HIV care; HIV surveillance; viral suppression; linkage and navigation to care

Introduction

In the United States, an estimated two-thirds of persons living with human immunodeficiency virus (HIV) infection do not achieve viral suppression, including both those who have never engaged in HIV care and those who linked to HIV care after diagnosis but did not stay engaged [1-3]. Persons with an unsuppressed HIV viral load might experience poor clinical outcomes and transmit HIV [4]. The National HIV/AIDS Strategy for the United States focuses on improving the HIV care continuum, including interventions that link, retain, and reengage persons in HIV care [5,6].

Retention in HIV care has been studied in surveillance registries [2,7-9], observational cohorts [10], health care databases and networks [11-14], and research populations [15,16]. In San Francisco, 19% of persons diagnosed with HIV in 2006-2007 were not adequately retained in care (did not have at least two laboratory measurements reported annually) [17]. Other US jurisdictions have documented reductions in retention after the initial HIV diagnosis and linkage to care, and HIV care disparities in population subgroups [8,18].

The goal of the Re-engaging Surveillance-identified Viremic Persons (RSVP) project in San Francisco was to use routine HIV surveillance databases to identify, contact, interview, and reengage persons living with HIV infection who appeared to be out of care because their last HIV viral load was unsuppressed. We interviewed participants about their HIV care patterns and barriers to reengagement. Understanding why people disengage from HIV care and what would help them reengage is essential for optimizing HIV clinical and public health outcomes.

Methods

RSVP began on April 20, 2012 for a 12-month period. Our project methods and implementation, including success in locating truly viremic out-of-care persons, have been previously described [19]. Briefly, persons with an unsuppressed plasma

HIV RNA viral load (>200 copies/mL) at their last measurement were eligible if they appeared to have disengaged from HIV care because they lacked HIV viral load and CD4 cell count laboratory reports during the 9 to 15 months prior to April 20, 2012 in the San Francisco Department of Public Health Enhanced HIV/AIDS Reporting System. We asked participants to complete a 30-minute interviewer-administered survey and offered them assistance with reengaging in HIV care from the San Francisco Department of Public Health linkage and navigation program (LINCS). Baseline interview questions included demographics, health insurance coverage, HIV testing and care experiences, treatment use, sexual activities, unmet services, and willingness to talk with LINCS. Participants were asked to complete a follow-up interview 3 months later that assessed changes since baseline and care reengagement. RSVP participants could meet with LINCS staff for health insurance assistance, appointments for HIV care reengagement, and referrals to ancillary services. The US Centers for Disease Control and Prevention and the University of California, San Francisco determined this to be a nonresearch activity; therefore, we did not require institutional review board approval.

Results

Baseline RSVP Interview

The characteristics of the 75 (26.6%) interviewed participants were broadly representative of the 282 eligible persons [19]. Most of the 75 interviewed participants (median age 45 years) were male (85%), born in the United States (77%), and current San Francisco residents (75%), and identified as gay or homosexual (69%) (Table 1). Participants were racially and ethnically diverse, one-third reported a college degree or higher, the majority were either unemployed or receiving disability benefits, and 1 in 5 reported being homeless or in unstable housing in the previous 12 months. A total of 64 (85%) were sexually active in the past year, and 19 (25%) reported having had condomless anal or vaginal sex with a person of HIV-negative or unknown status. Drug use and binge drinking were common as were symptoms of depression (Table 1).

Table 1. Baseline characteristics of interviewed Re-engaging Surveillance-identified Viremic Persons (RSVP) participants, San Francisco, 2012-2013 (n=75).

Characteristics	Median (IQR ^a) or n (%)
Demographics	
Age (years)	
Median (IQR)	45 (37-51)
<30, n (%)	5 (7)
30-44, n (%)	31 (41)
≥45, n (%)	39 (52)
Gender, n (%)	
Male	64 (85)
Female	10 (13)
Transgender	1 (1)
Sexual orientation, n (%)	
Gay or homosexual	52 (69)
Straight or heterosexual	12 (16)
Bisexual	9 (12)
Questioning	1 (1)
Queer	1 (1)
Country or territory of birth, n (%)	
United States	58 (77)
Puerto Rico	1 (1)
Mexico	5 (7)
Other	11 (15)
Current city of residence, n (%)	
San Francisco	56 (75)
Other	19 (25)
Race/ethnicity, n (%)	
Non-Hispanic white	32 (43)
Non-Hispanic black	20 (27)
Non-Hispanic Asian/Pacific Islander	2 (3)
Hispanic/Latino	20 (27)
Other	1 (1)
Education, n (%)	
High school, General Equivalency Diploma, or less	22 (29)
Some technical or college training	30 (40)
College degree or more	23 (31)
Current housing situation, n (%)	
Person's own house or apartment	43 (57)
Someone else's house or apartment	16 (21)
Single room, rented room, motel, single-room occupancy	11 (15)
All other (shelter, transitional housing, homeless)	5 (7)
Current work situation, n (%)	
Working full-time or part-time	25 (33)

Characteristics	Median (IQR ^a) or n (%)
Unemployed, laid off, looking for work	25 (33)
Disabled (receiving disability income)	16 (21)
Other	3 (4)
Missing data	6 (8)
In jail, detention, or prison in the past 12 months, n (%)	
Yes	2 (3)
No	73 (97)
Sexual risk, drug use, and other behaviors (past 12 months)	
Had any sex (anal, vaginal, or oral), n (%)	
Yes	64 (85)
No	11 (15)
Had anal or vaginal sex without a condom with a person of HIV^b-negative or unknown status, n (%)	
Yes	19 (25)
No	45 (60)
Not applicable (not sexually active)	11 (15)
Injected any nonprescription drugs, n (%)	
Yes	18 (24)
No	55 (73)
Missing data	2 (3)
Used the following drugs (not mutually exclusive), n (%)	
Powder cocaine	12 (16)
Crack cocaine	14 (19)
Poppers	20 (27)
Heroin	5 (7)
Methamphetamine	27 (36)
Prescription drugs or painkillers without a prescription	16 (21)
Binge drinking (≥5, if male, and ≥4, if female, alcoholic drinks in one sitting), n (%)	
Daily or weekly	15 (20)
Monthly or less	21 (28)
Never	39 (52)
Health and care utilization	
Has health insurance or other health care coverage, n (%)	
Yes	67 (89)
No	8 (11)
Has one place in particular where usually goes for medical care not related to HIV infection, n (%)	
Yes	44 (59)
No	30 (40)
Missing data	1 (1)
Years since first HIV-positive test, n (%)	
<5	24 (32)
5-20	38 (51)
>20	13 (17)

Characteristics	Median (IQR ^a) or n (%)
Has seen a provider for HIV medical care in the past 12 months, n (%)	
Yes	59 (79)
No	11 (15)
Missing data	5 (7)
Has had a CD4 cell count or HIV viral load test in the past 12 months, n (%)	
Yes	67 (89)
No	5 (7)
Missing data/don't know	3 (4)
Ever prescribed HIV medications by a doctor, n (%)	
Yes	59 (79)
No	16 (21)
Currently taking any medications to treat HIV infection, n (%)	
Yes	45 (60)
No	30 (40)
Missed any HIV medical care appointments in the past 12 months, n (%)	
Yes	34 (45)
No	39 (52)
Missing data	2 (3)
Ever told anyone (other than doctor, nurse, or health care worker) about being HIV-positive, n (%)	
Yes	71 (95)
No	4 (5)
Little interest or pleasure in doing things (in the past 2 weeks), n (%)	
Yes	34 (45)
No	41 (55)
Feeling down, depressed, or hopeless (in the past 2 weeks), n (%)	
Yes	40 (53)
No	35 (47)
Interested in talking with San Francisco Department of Public Health linkage and navigation program (LINCS) staff, n (%)	
Yes	52 (69)
No	21 (28)
Missing	2 (3)

^aIQR: interquartile range.

^bHIV: human immunodeficiency virus.

A total of 67 (89%) reported current health insurance or coverage, 59 (79%) had ever been prescribed HIV medications, and 45 (60%) reported current medication use. In the past year, 59 (79%) had seen an HIV provider for medical care, and 67 (89%) reported a CD4 cell count or viral load test. For the 11 (15%) participants who had not seen an HIV provider in the past year, the frequently reported reasons included feeling healthy (n=6), drinking alcohol or using drugs (n=5), not having enough money or health insurance (n=4), and not wanting to take HIV medicines (n=4).

Nearly half (45%) reported missing an HIV medical care appointment in the past year (Table 1). Among 59 participants

who reported seeing a provider in the past year, 27 (46%) reported missing an HIV medical care appointment in the past year. These 27 participants volunteered 1 or more of these responses to "What would help you the most to get to HIV medical care appointment within the next 3 months?": transportation assistance (n=8), a stable living situation or housing (n=3), sound mental health (n=3), help getting organized to track appointments (n=3), and more appointment reminders (phone calls, text messages, email, letters) (n=3).

The 30 (40%) participants who reported not currently taking medications to treat HIV were similar in their characteristics to the overall interviewed population: 24 were men, 25 had current

health insurance or coverage, 20 had seen an HIV provider for medical care in the past year, and 23 expressed interest in meeting with LINCS staff. Overall, the most frequently reported services used in the past year were HIV education or information, public benefits support, HIV case management,

financial assistance, and spiritual support (Table 2). The most frequently reported needed but not accessed services were dental services, mental health services, financial assistance, transportation assistance, and HIV case management (Table 2).

Table 2. Services used, and services that were needed but not used, among the Re-engaging Surveillance-identified Viremic Persons (RSVP) project participants at baseline, San Francisco, 2012-2013 (n=75).

Services in the past 12 months	Used, n (%)	Unmet need ^a : did not use the service but needed it, n (%)
Dental services	23 (31)	35 (67) ^b
Mental health services, including one-to-one counseling	19 (25)	28 (50) ^b
Transportation assistance	19 (25)	26 (46) ^b
Practical support (bills, buddy program, help with cleaning)	12 (16)	25 (40)
Financial assistance	28 (37) ^b	23 (49) ^b
Legal services	13 (17)	22 (35)
Shelter or housing services	8 (11)	21 (31)
HIV ^c case management services	29 (39) ^b	20 (43) ^b
Drug or alcohol counseling or treatment	13 (17)	18 (29)
Meal or food services	25 (33)	16 (32)
HIV peer group support	15 (20)	16 (27)
Public benefits, including Supplemental Security Income or Social Security Disability Insurance	34 (45) ^b	15 (37)
Spiritual support	27 (36) ^b	14 (29)
Domestic violence services	4 (5)	10 (14)
Education or information about HIV	39 (52) ^b	6 (17)
Interpreter services	0 (0)	1 (1)
Childcare services	1 (1)	0 (0)
Any other HIV-related services	6 (8)	3 (4)

^aOrdered by highest count; percentage calculated among those who did not use a service in past 12 months.

^bThe top 5 most frequently used services, and the top 5 most frequently needed services (unmet needs).

^cHIV: human immunodeficiency virus.

Of the participants, 52 (69%) accepted referral to LINCS; their linkage outcomes were previously published [19]. Among the 11 participants who reported not having seen an HIV medical provider in the previous 12 months, 10 (91%) agreed to talk to LINCS staff about reengagement or ancillary services. Of these, 8 (80%) did meet with LINCS staff and, of these, 5 (63%) were already in HIV care and 3 (38%) reengaged in care within 6 months with LINCS assistance.

Follow-Up Interview

Of the 75 participants, 64 (85%) completed the 3-month follow-up interview. Among the 64, 97% (n=62) had health insurance or coverage, 73% (n=47) reported having "seen a doctor, nurse or other health care provider for HIV medical care" since baseline interview, 81% (n=52) reported having a CD4 cell count test, 77% (n=49) had an HIV viral load test, and 47% (n=30) were prescribed HIV treatment. HIV risk behaviors

since the baseline visit included using "any drugs that were not prescribed by a doctor" (n=30, 47%) and injecting drugs that were not prescribed (n=9, 14%). A total of 33 (52%) reported vaginal or anal sex, including 7 (11%) who reported condomless vaginal or anal sex with an HIV-negative or unknown-status partner.

Discussion

The RSVP project sought to identify a high-priority population of persons living with HIV who appeared to be out of care and viremic based on HIV surveillance data. However, the majority of interviewed participants reported both having seen an HIV provider in the past 12 months and having health insurance, thus having means and opportunity to access the HIV care needed to achieve viral suppression. The frequent self-reported care engagement was corroborated by HIV surveillance data:

over 80% had at least one viral load or CD4 cell count test during the 12 months after they met RSVP project eligibility [19].

Nevertheless, RSVP participants identified personal and structural barriers to HIV care and many unmet needs. Almost half reported missing a scheduled medical appointment in the previous year, and 40% reported no current HIV treatment. Deficits in comprehensive health care and social support included unmet needs for legal services, dental care, transportation assistance, mental health services, financial assistance, and practical support. Notably, about one-fourth of participants reported engaging in condomless sex with negative or unknown HIV-status partners. Given that all RSVP participants had an unsuppressed HIV viral load at their last measurement, the risk of further HIV transmission was possible, indicating that we reached persons who would benefit from assistance remaining in HIV care.

Similar to our findings, a myriad of social, behavioral, and structural factors have previously been found to be correlated with poor retention in care [20,21], including substance and alcohol use, poor mental health, homelessness, and low socioeconomic status [8,10,22]; low care satisfaction, medical establishment distrust, stigma, and lack of social or ancillary

support services [15,23]; and early HIV disease or feeling well [8,17]. Patient perceptions of HIV care engagement, including when accessing care only sporadically, may also differ from standard care engagement metrics [24]. The multifaceted barriers to staying engaged in HIV care point to the need for comprehensive, ongoing, innovative, client-centered approaches to support retention (eg, case management, wraparound services, and contingency management). We were unable to analyze specific associations between barriers participants faced and the services they used, and their likelihood of reengaging in care, due to the heterogeneity of barriers and care patterns in our relatively small interviewed population.

In summary, among persons presumed to be out of HIV care, there were self-reported indicators of at least intermittent HIV care, as well as frequently reported barriers to care and unmet needs. As previously described, locating persons who are truly out of HIV care is difficult [19]. Nevertheless, surveillance-based public health efforts support the HIV care continuum, as most RSVP project participants returned for a follow-up interview, seemed willing to engage with health department staff, and accepted assistance with ancillary services and HIV care reengagement, a prerequisite for ongoing HIV treatment and viral suppression.

Acknowledgments

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the US Centers for Disease Control and Prevention.

Conflicts of Interest

None declared.

References

1. Hall HI, Frazier EL, Rhodes P, Holtgrave DR, Furlow-Parmley C, Tang T, et al. Differences in human immunodeficiency virus care and treatment among subpopulations in the United States. *JAMA Intern Med* 2013 Jul 22;173(14):1337-1344. [doi: [10.1001/jamainternmed.2013.6841](https://doi.org/10.1001/jamainternmed.2013.6841)] [Medline: [23780395](https://pubmed.ncbi.nlm.nih.gov/23780395/)]
2. Hall HI, Gray KM, Tang T, Li J, Shouse L, Mermin J. Retention in care of adults and adolescents living with HIV in 13 U.S. areas. *J Acquir Immune Defic Syndr* 2012 May 01;60(1):77-82. [doi: [10.1097/QAI.0b013e318249fe90](https://doi.org/10.1097/QAI.0b013e318249fe90)] [Medline: [22267016](https://pubmed.ncbi.nlm.nih.gov/22267016/)]
3. Bradley H, Hall HI, Wolitski RJ, Van Handel MM, Stone AE, LaFlam M, et al. Vital Signs: HIV diagnosis, care, and treatment among persons living with HIV--United States, 2011. *MMWR Morb Mortal Wkly Rep* 2014 Nov 28;63(47):1113-1117 [FREE Full text] [Medline: [25426654](https://pubmed.ncbi.nlm.nih.gov/25426654/)]
4. Skarbinski J, Rosenberg E, Paz-Bailey G, Hall HI, Rose CE, Viall AH, et al. Human immunodeficiency virus transmission at each step of the care continuum in the United States. *JAMA Intern Med* 2015 Apr;175(4):588-596. [doi: [10.1001/jamainternmed.2014.8180](https://doi.org/10.1001/jamainternmed.2014.8180)] [Medline: [25706928](https://pubmed.ncbi.nlm.nih.gov/25706928/)]
5. Office of National AIDS Policy. National HIV/AIDS Strategy for the United States: updated to 2020. Washington, DC: The White House; 2015 Jul. URL: <https://www.aids.gov/federal-resources/national-hiv-aids-strategy/nhas-update.pdf> [accessed 2017-01-13] [WebCite Cache ID 6nUbwFeYS]
6. Higa DH, Marks G, Crepaz N, Liau A, Lyles CM. Interventions to improve retention in HIV primary care: a systematic review of U.S. studies. *Curr HIV/AIDS Rep* 2012 Dec;9(4):313-325. [doi: [10.1007/s11904-012-0136-6](https://doi.org/10.1007/s11904-012-0136-6)] [Medline: [22996171](https://pubmed.ncbi.nlm.nih.gov/22996171/)]
7. Muthulingam D, Chin J, Hsu L, Scheer S, Schwarcz S. Disparities in engagement in care and viral suppression among persons with HIV. *J Acquir Immune Defic Syndr* 2013 May 01;63(1):112-119. [doi: [10.1097/QAI.0b013e3182894555](https://doi.org/10.1097/QAI.0b013e3182894555)] [Medline: [23392459](https://pubmed.ncbi.nlm.nih.gov/23392459/)]
8. Torian LV, Wiewel EW. Continuity of HIV-related medical care, New York City, 2005-2009: do patients who initiate care stay in care? *AIDS Patient Care STDS* 2011 Feb;25(2):79-88. [doi: [10.1089/apc.2010.0151](https://doi.org/10.1089/apc.2010.0151)] [Medline: [21284498](https://pubmed.ncbi.nlm.nih.gov/21284498/)]

9. 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/)]
10. Rebeiro P, Althoff KN, Buchacz K, Gill J, Horberg M, Krentz H, North American AIDS Cohort Collaboration on Research and Design. Retention among North American HIV-infected persons in clinical care, 2000-2008. *J Acquir Immune Defic Syndr* 2013 Mar 01;62(3):356-362 [FREE Full text] [doi: [10.1097/QAI.0b013e31827f578a](https://doi.org/10.1097/QAI.0b013e31827f578a)] [Medline: [23242158](https://pubmed.ncbi.nlm.nih.gov/23242158/)]
11. Magnus M, Herwehe J, Gruber D, Wilbright W, Shepard E, Abrams A, et al. Improved HIV-related outcomes associated with implementation of a novel public health information exchange. *Int J Med Inform* 2012 Oct;81(10):e30-e38. [doi: [10.1016/j.ijmedinf.2012.06.005](https://doi.org/10.1016/j.ijmedinf.2012.06.005)] [Medline: [22883431](https://pubmed.ncbi.nlm.nih.gov/22883431/)]
12. Horberg MA, Hurley LB, Silverberg MJ, Klein DB, Quesenberry CP, Mugavero MJ. Missed office visits and risk of mortality among HIV-infected subjects in a large healthcare system in the United States. *AIDS Patient Care STDS* 2013 Aug;27(8):442-449 [FREE Full text] [doi: [10.1089/apc.2013.0073](https://doi.org/10.1089/apc.2013.0073)] [Medline: [23869466](https://pubmed.ncbi.nlm.nih.gov/23869466/)]
13. Gardner EM, Daniloff E, Thrun MW, Reirden DH, Davidson AJ, Johnson SC, et al. Initial linkage and subsequent retention in HIV care for a newly diagnosed HIV-infected cohort in Denver, Colorado. *J Int Assoc Provid AIDS Care* 2013;12(6):384-390. [doi: [10.1177/2325957413500532](https://doi.org/10.1177/2325957413500532)] [Medline: [23962912](https://pubmed.ncbi.nlm.nih.gov/23962912/)]
14. Bove JM, Golden MR, Dhanireddy S, Harrington RD, Dombrowski JC. Outcomes of a clinic-based surveillance-informed intervention to re-link patients to HIV care. *J Acquir Immune Defic Syndr* 2015 Nov 01;70(3):262-268 [FREE Full text] [doi: [10.1097/QAI.0000000000000707](https://doi.org/10.1097/QAI.0000000000000707)] [Medline: [26068720](https://pubmed.ncbi.nlm.nih.gov/26068720/)]
15. Christopoulos KA, Massey AD, Lopez AM, Geng EH, Johnson MO, Pilcher CD, et al. "Taking a half day at a time:" patient perspectives and the HIV engagement in care continuum. *AIDS Patient Care STDS* 2013 Apr;27(4):223-230 [FREE Full text] [doi: [10.1089/apc.2012.0418](https://doi.org/10.1089/apc.2012.0418)] [Medline: [23565926](https://pubmed.ncbi.nlm.nih.gov/23565926/)]
16. Haley DF, Lucas J, Golin CE, Wang J, Hughes JP, Emel L, HPTN 064 Study Team. Retention strategies and factors associated with missed visits among low income women at increased risk of HIV acquisition in the US (HPTN 064). *AIDS Patient Care STDS* 2014 Apr;28(4):206-217 [FREE Full text] [doi: [10.1089/apc.2013.0366](https://doi.org/10.1089/apc.2013.0366)] [Medline: [24697160](https://pubmed.ncbi.nlm.nih.gov/24697160/)]
17. Hsu LC, Chen M, Kali J, Pipkin S, Scheer S, Schwarcz S. Assessing receipt of medical care and disparity among persons with HIV/AIDS in San Francisco, 2006-2007. *AIDS Care* 2011 Mar;23(3):383-392. [doi: [10.1080/09540121.2010.507740](https://doi.org/10.1080/09540121.2010.507740)] [Medline: [21347902](https://pubmed.ncbi.nlm.nih.gov/21347902/)]
18. Dombrowski JC, Kent JB, Buskin SE, Stekler JD, Golden MR. Population-based metrics for the timing of HIV diagnosis, engagement in HIV care, and virologic suppression. *AIDS* 2012 Jan 02;26(1):77-86 [FREE Full text] [doi: [10.1097/QAD.0b013e32834dcee9](https://doi.org/10.1097/QAD.0b013e32834dcee9)] [Medline: [22008656](https://pubmed.ncbi.nlm.nih.gov/22008656/)]
19. Buchacz K, Chen M, Parisi MK, Yoshida-Cervantes M, Antunez E, Delgado V, et al. Using HIV surveillance registry data to re-link persons to care: the RSVP Project in San Francisco. *PLoS One* 2015;10(3):e0118923 [FREE Full text] [doi: [10.1371/journal.pone.0118923](https://doi.org/10.1371/journal.pone.0118923)] [Medline: [25748668](https://pubmed.ncbi.nlm.nih.gov/25748668/)]
20. Horstmann E, Brown J, Islam F, Buck J, Agins BD. Retaining HIV-infected patients in care: Where are we? Where do we go from here? *Clin Infect Dis* 2010 Mar 1;50(5):752-761 [FREE Full text] [doi: [10.1086/649933](https://doi.org/10.1086/649933)] [Medline: [20121413](https://pubmed.ncbi.nlm.nih.gov/20121413/)]
21. Mugavero MJ, Norton WE, Saag MS. Health care system and policy factors influencing engagement in HIV medical care: piecing together the fragments of a fractured health care delivery system. *Clin Infect Dis* 2011 Jan 15;52(Suppl 2):S238-S246 [FREE Full text] [doi: [10.1093/cid/ciq048](https://doi.org/10.1093/cid/ciq048)] [Medline: [21342913](https://pubmed.ncbi.nlm.nih.gov/21342913/)]
22. Hessol NA, Weber KM, Holman S, Robison E, Goparaju L, Alden CB, et al. Retention and attendance of women enrolled in a large prospective study of HIV-1 in the United States. *J Womens Health (Larchmt)* 2009 Oct;18(10):1627-1637 [FREE Full text] [doi: [10.1089/jwh.2008.1337](https://doi.org/10.1089/jwh.2008.1337)] [Medline: [19788344](https://pubmed.ncbi.nlm.nih.gov/19788344/)]
23. Wohl AR, Carlos J, Tejero J, Dierst-Davies R, Daar ES, Khanlou H, et al. Barriers and unmet need for supportive services for HIV patients in care in Los Angeles County, California. *AIDS Patient Care STDS* 2011 Sep;25(9):525-532. [doi: [10.1089/apc.2011.0149](https://doi.org/10.1089/apc.2011.0149)] [Medline: [21774689](https://pubmed.ncbi.nlm.nih.gov/21774689/)]
24. Castel AD, Tang W, Peterson J, Mikre M, Parenti D, Elion R, et al. Sorting through the lost and found: are patient perceptions of engagement in care consistent with standard continuum of care measures? *J Acquir Immune Defic Syndr* 2015 May 01;69(Suppl 1):S44-S55 [FREE Full text] [doi: [10.1097/QAI.0000000000000575](https://doi.org/10.1097/QAI.0000000000000575)] [Medline: [25867778](https://pubmed.ncbi.nlm.nih.gov/25867778/)]

Abbreviations

HIV: human immunodeficiency virus

LINCS: San Francisco Department of Public Health linkage and navigation program

RSVP: Re-engaging Surveillance-identified Viremic Persons

Edited by T Sanchez; submitted 13.01.17; peer-reviewed by J Colasanti, N Benbow; comments to author 06.03.17; revised version received 25.03.17; accepted 06.04.17; published 04.05.17.

Please cite as:

Scheer S, Chen MJ, Parisi MK, Yoshida-Cervantes M, Antunez E, Delgado V, Moss NJ, Buchacz K

The RSVP Project: Factors Related to Disengagement From Human Immunodeficiency Virus Care Among Persons in San Francisco
JMIR Public Health Surveill 2017;3(2):e25

URL: <http://publichealth.jmir.org/2017/2/e25/>

doi: [10.2196/publichealth.7325](https://doi.org/10.2196/publichealth.7325)

PMID: [28473307](https://pubmed.ncbi.nlm.nih.gov/28473307/)

©Susan Scheer, Miao-Jung Chen, Maree Kay Parisi, Maya Yoshida-Cervantes, Erin Antunez, Viva Delgado, Nicholas J Moss, Kate Buchacz. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 04.05.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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

Physician's Perceptions of Telemedicine in HIV Care Provision: A Cross-Sectional Web-Based Survey

Kelly Anderson¹, MD, CCFP(EM); Troy Francis², MSc; Francisco Ibanez-Carrasco³, PhD; Jason Globberman³, MSc

¹St. Michael's Hospital Family Health Team, Department of Family and Community Medicine, Toronto, ON, Canada

²Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada

³Ontario HIV Treatment Network, Toronto, ON, Canada

Corresponding Author:

Kelly Anderson, MD, CCFP(EM)
St. Michael's Hospital Family Health Team
Department of Family and Community Medicine
St. Michael's Health Centre
80 Bond Street
Toronto, ON, M5G 1X2
Canada
Phone: 1 416 864 3011
Fax: 1 416 864 3099
Email: andersonk@smh.ca

Abstract

Background: Telemedicine, or electronic interactive health care consultation, offers a variety of benefits to both patients and primary care clinicians. However, little is known about the opinions of physicians using these modalities.

Objective: The aim of this study was to examine physician perceptions, including challenges, risks, and benefits of the use of telemedicine in human immunodeficiency virus (HIV) patient care.

Methods: A Web-based, self-administered, anonymous, cross-sectional survey was sent to physicians known to be providing medical care to patients living with HIV in Ontario, Canada. Descriptive statistics and frequencies were used to examine physician perceptions and characteristics of participants.

Results: Among the 51 invited participants, 48 (94%) completed the survey. Sixty-two percent (29/47) of respondents reported that they used some form of telemedicine to care for HIV patients in their practice. Of the respondents who identified as having used telemedicine in their practice, telephone (86%, 25/29), email (69%, 20/29), and teleconsultation (24%, 7/29) were listed as frequent modalities used. A significant number of physicians (83%, 38/46) agreed that an obstacle to adopting telemedicine is their perception that this modality does not allow for a comprehensive assessment of their patients' health. In addition, 65% (28/43) of physicians agreed that patients may not feel adequately connected to them as a provider if they used telemedicine. However, 85% (39/46) of respondents believed that telemedicine could improve access and timeliness to care along with increasing the number of times physicians can interact with their patients.

Conclusions: From the perceptions of physicians, telemedicine shows promise in the care of patients living with HIV. More than half of the respondents are already using telemedicine modalities. Whereas many physicians are concerned about their ability to fully assess the health of a patient via telemedicine, most physicians do see a need for it—to reduce patient travel times, reduce exposure to stigma, and improve efficiency and timely access to care. Challenges and risks such as technological gaps, confidentiality, and medicolegal concerns must be addressed for physicians to feel more comfortable using telemedicine.

(*JMIR Public Health Surveill* 2017;3(2):e31) doi:[10.2196/publichealth.6896](https://doi.org/10.2196/publichealth.6896)

KEYWORDS

HIV; AIDS; telemedicine; health surveys

Introduction

Telemedicine, or electronic interactive health care consultation, offers a variety of benefits to both patients and primary care clinicians [1]. Telemedicine encompasses a wide variety of health care services, and with the current advances in technology, these services are quickly evolving and becoming more affordable and accessible. Telemedicine models are also wide ranging: from live synchronous connections between two or more parties in a health encounter to asynchronous training modules; from programs delivered with a desktop in one place to just-in-time learning delivered via mobile devices [2-4]. In all of these cases, there are a variety of potential obstacles and challenges such as technical (eg, type of device and data plans), administrative (coordination and support), financial (costs of technology), and cultural (affinity with some media over other) [5]. The complexity of the illness is a barrier that also needs to be addressed when identifying which telemedicine model is most appropriate to use. Although telemedicine is promising in the provision of services to patients, the evidence remains limited and inconsistent [5].

Human immunodeficiency virus (HIV) is a highly complex chronic illness and should be thought of as a myriad of conditions and not one single issue. Many patients suffering from HIV are also afflicted with physical and psychological impairments, making the use of telemedicine in this population a strong compliment to current treatment options [6,7]. Although various telemedicine modalities are well received by patients—including people living with HIV—due to its convenience, ability to increase confidentiality, and reduce stigma, little is known about what physicians providing HIV care think about telemedicine [8-11]. The uptake of telemedicine services relies on system structures being embedded within medical practice and physicians playing a pivotal role in adopting the technology [12,13]. Physicians' acceptance of telemedicine is essential in its use, and thus understanding physicians' perceptions on the use of, and the skills required to use telemedicine to care for those living with HIV, may help identify the perceived challenges, risks, and benefits to the uptake of its many modalities. Additionally, identifying physician characteristics may shed light into how often HIV care providers use telemedicine and what impact it has on their practice and on the health of patients. Accordingly, the purpose of this study was to explore how physicians perceive the use of telemedicine in HIV care in the province of Ontario. This is a necessary entry point into understanding whether HIV disease itself can and must be treated using telemedicine or specialized physicians in general, who use telemedicine, need to be trained in the complexity of HIV to better serve those they see who happen to be HIV positive as well.

Methods

Participants

Physicians who were providing medical care to people living with HIV in Ontario, Canada were asked to complete a Web-based, self-administered, and anonymous survey regarding the use of telemedicine in HIV care.

Recruitment

A list of all registered infectious disease specialists in the province (n=218) was obtained via the College of Physicians and Surgeons of Ontario. Of these, 51 individuals were identified as providing HIV-specific care. A link to the survey was emailed to these individuals. In addition, 2 additional email reminders were circulated 1 week apart to get as many responses as possible.

Eligibility Criteria

Only physicians practicing in Ontario and who have HIV patients in their care were included in this study. All participants provided informed consent.

Data Collection

A total of 48 participants completed a cross-sectional survey hosted by FluidSurveys [14]. Questions were designed in consultation with HIV researchers, program evaluators, HIV educators, distance education specialists, community leaders living with HIV, and physicians with expertise in providing telemedicine services as well as HIV care.

The aims of the survey were to better understand current practices and perceived risks, benefits, and challenges in using telemedicine in HIV care. To do this, questions were organized into 5 sections: (1) general concepts of telemedicine, (2) perceived challenges to telemedicine use, (3) benefits of telemedicine, (4) perceived risks of telemedicine, and (5) information about the health care provider and their practice and patient population.

Analysis

Data were analyzed using descriptive statistics and frequencies to examine the characteristics of invited participants and their responses to the survey items. Data are presented as total counts and percentages. For simplicity, physician perceptions were dichotomized by aggregating the categories "agree" and "strongly agree" to "agree" in the text, while "disagree" and "strongly disagree" were combined to "disagree." In addition, "minor," "moderate," and "severe" barriers were consolidated to "barrier" in the text. Data synthesis and statistical analyses were performed using R 3.2.3 (R Core Team, Vienna Austria) [15].

Ethics

This study was approved by the Research Ethics Board at St Michael's Hospital in Toronto, Canada (REB#15-337). Per our study protocol, participation was completely voluntary. All questions were self-administered and anonymous. Participants were not compensated for completing the survey and were free to withdraw or not answer any questions they did not want to with no professional or other consequences.

Results

Recruitment and Participation

Of the 51 physicians who were invited to take part in the Web-based survey, 50 (98%) consented to participate in the study. Two physicians who consented were removed because

they did not answer any of the survey questions. All questions had 10% or fewer missing items, with most having only 4% missing.

Demographic Characteristics of Respondents

Physician characteristics are presented in Table 1. About half of the invited physicians were generalists (44% family physician, 2% internal medicine, and 2% pediatrics), whereas 36% of physicians were specialists in infectious diseases and 9% were psychiatrists focusing in HIV. The majority of respondents primarily practiced in the Greater Toronto Area (69%) and represented a broad range of years in practice with 78% of physicians having at least ten years of experience. The respondents each represented a varying caseload of HIV patients.

HIV Care Providers' Use of Telemedicine

More than half of the physicians (62%, 29/47) reported that they used some form of telemedicine to care for HIV patients

in their practice. Eighteen (38%, 18/47) respondents stated that they have never used telemedicine in their practice. Of the 29 respondents who were identified as having used telemedicine in their practice, telephone (86%, 25/29), email (69%, 20/29), and teleconsultation (24%, 7/29) were listed as frequent modalities used to care for patients. HIV care providers were more likely to use the telephone (69%, 20/29) to care for their patients as opposed to email (21%, 6/29) and teleconsultation systems (10%, 3/29; Table 1).

When requiring the assistance of a specialist in order to care for HIV patients, about half of the physicians (55%, 26/47) reported that they used telemedicine in order to consult with specialists. However, 21 respondents (46%, 21/47) stated that they have never used telemedicine to contact a specialist. The primary modalities used to consult with a specialist by the 26 respondents were email (58%, 15/26) and telephone (38%, 10/26).

Table 1. Characteristics of participants (N=48).

Participant characteristics	Number of respondents n (%)
Location of practice, n=48	
Toronto or Ottawa	37 (77)
Other locations in Ontario	11 (23)
Type of physician, n=45	
Family medicine	20 (44)
Infectious disease	16 (36)
Other (ie, internal medicine, psychiatrist, etc)	9 (20)
Years in practice, n=45	
<10	10 (22)
10-15	8 (18)
16-24	12 (27)
≥25	15 (33)
Percentage of patient case load that are people living with HIV^a, n=45	
< 25%	14 (31)
25-49%	12 (27)
50-74%	9 (20)
75-100%	10 (22)
Modalities of HIV telemedicine currently used, n=29	
Phone consultations	25 (86)
Email consultations	20 (69)
Teleconsultation systems	7 (24)
Other (ie, instant messaging software)	1 (3)
Current use of telemedicine services, n=48	
Never or rarely	29 (60)
Sometimes	11 (23)
Often	7 (15)

^aHIV: human immunodeficiency virus.

Twenty-six physicians (55%, 26/47) felt that there was a need to expand the use of HIV telemedicine services in Ontario; however, 20 physicians (43%, 20/47) were unsure of the need for an expansion of HIV telemedicine services, with 1 physician expressing a negative response to the development of telemedicine services in Ontario. Whereas only 26 physicians expressed interest in developing telemedicine services for HIV patients, many physicians (28/46, 61%) stated they served patients who could benefit from telemedicine. There were also a significant number of physicians who stated that they provided care to a number of patients who had difficulties traveling because of physical disabilities or mental health issues (82%, 37/45 and 72%, 34/47 respectively). Also, 74% (34/46) of those surveyed stated they had patients who traveled more than 100 kilometers for a visit.

Physicians Perceived Challenges to Telemedicine Use in Their HIV Patients

Physicians endorsed various challenges relating to the use of telemedicine when caring for HIV patients (Table 2). Most notably, 83% (38/46) of physicians felt that they could not adequately assess the health of a patient via telemedicine. Sixty two percent (28/45 respondents) reported that telemedicine took too much time, and 60% (27/45 respondents) felt they lacked the technology to use telemedicine in their practice. The majority of respondents (76%, 34/45) believed that their patients did not have access to the necessary equipment needed to use telemedicine services. Many physicians cited other challenges such as confidentiality (60%, 27/45), lack of remuneration (62%, 28/45), concerns that patients will abuse telemedicine services (71%, 32/45), and medicolegal concerns (51%, 23/45). However, an absence of patients that would benefit from telemedicine was not identified as a barrier; 75% (33/44) of physicians stated that lack of need did not prevent them from using telemedicine.

Table 2. Participants' perspectives on challenges to using telemedicine.

Perceived challenges to using telemedicine	No barrier n (%)	Minor barrier n (%)	Moderate barrier n (%)	Severe barrier n (%)
It takes too much time	17 (38)	20 (44)	6 (13)	2 (4)
I lack access to the necessary technology	18 (40)	9 (20)	15 (33)	3 (7)
I have no patients that require telemedicine services	33 (75)	7 (16)	1 (2)	3 (7)
Most of my patients do not have access to the necessary equipment for telemedicine	11 (24)	12 (27)	19 (42)	3 (7)
I have concerns about confidentiality	18 (40)	17 (38)	7 (16)	3 (7)
There is too much diversity in my practice to adopt telemedicine as a regular practice	26 (58)	13 (29)	5 (11)	1 (2)
I cannot adequately assess a patient using telemedicine	8 (17)	19 (41)	15 (33)	4 (9)
I have concerns about remuneration for telemedicine	17 (38)	12 (27)	10 (22)	6 (13)
I worry that patients will abuse the use of telemedicine to communicate with me	13 (29)	16 (36)	15 (33)	1 (2)
I have medicolegal and licensing concerns about using telemedicine	22 (49)	15 (33)	8 (18)	0

Perceived Risks of Using Telemedicine in Patients With HIV

Respondents expressed opinions regarding the perceived risks of using telemedicine for patients living with HIV (Table 3). Many respondents (65%, 28/43) agreed that patients may not "feel adequately connected" to them as a provider with the use of telemedicine. Fifty-eight percent (25/43 respondents) agreed that HIV patients would receive poorer quality assessments with the use of telemedicine. However, most physicians (77%, 33/43) disagreed with the statement that HIV patients would feel more social isolation with the use of telemedicine. Also, the majority of the respondents (67%, 29/43) disagreed with the statement that remote patients would lose the opportunity to come visit their practice.

Perceived Benefits of Using Telemedicine to Care for HIV Patients

The majority of physicians stated that HIV patients would benefit from the use of telemedicine (Table 4). There was unanimous agreement among respondents (100%) when assessing telemedicine's ability to reduce patients' travel times. Eighty-three percent (38/46) of respondents agreed with the premise that using telemedicine can reduce a patients' exposure to the stigma of having HIV. Additionally, 65% (30/46) of respondents agreed that the use of telemedicine can increase the patients' privacy. Many physicians also agreed that both the quality of care (61%, 28/46) and efficacy of patient care (67%, 31/46) could improve in HIV patients through the use of telemedicine. Eighty-five percent (39/46) of respondents agreed with the statement that telemedicine will be able to increase the number of times patients are able to interact with their physician as well as agreed with its ability to improve access and timeliness to care.

Table 3. Participants' perspectives on risks of telemedicine.

Perceived risks of using telemedicine	Strongly disagree n (%)	Disagree n (%)	Agree n (%)	Strongly agree n (%)
I am concerned that telemedicine will increase the social isolation experienced by people living with HIV ^a	4 (9)	29 (67)	7 (16)	3 (7)
Patients may not feel adequately connected to me as a health care provider	3 (7)	12 (28)	22 (51)	6 (14)
Patients will receive lesser quality assessments	3 (7)	15 (35)	20 (47)	5 (12)
Patients may not understand my instructions	3 (7)	20 (47)	19 (44)	1 (2)
Remote patients will lose the opportunity to come see me in person if they prefer it	3 (7)	26 (60)	13 (30)	1 (2)

^aHIV: human immunodeficiency virus.

Table 4. Participants' perspectives on benefits of telemedicine.

Perceived benefits of using telemedicine	Strongly disagree n (%)	Disagree n (%)	Agree n (%)	Strongly agree n (%)
Reduces their exposure to stigma (in rural areas, for example, acquaintances or family would not see them visit medical services)	0	8 (17)	34 (74)	4 (9)
Increases their privacy	0	16 (35)	26 (57)	4 (9)
Reduces travel time	0	0	15 (33)	31 (67)
Improves quality of patient care	1 (2)	17 (37)	21 (46)	7 (15)
Improves efficacy of patient care	0	15 (33)	22 (48)	9 (20)
Increases the number of times we can interact	0	7 (15)	29 (63)	10 (22)
Improves access and timeliness to care	0	7 (15)	28 (61)	11 (24)

Discussion

Principal Findings

The aim of this study was to help describe physicians' perceptions of the use of telemedicine to care for patients living with HIV and the perception they have of their patients' use of telemedicine; the barriers and facilitators. In terms of benefits, physicians agreed across the board that telemedicine services would decrease travel times and could likely reduce patient experience of HIV-related stigma at appointments. Physicians also felt that telemedicine would increase access and efficiency of care, while benefiting patients who have difficulty travelling due to physical and mental impairments.

We discovered some consistent trends when evaluating the perceived challenges physicians reported to using telemedicine in their practice. Physicians most notably reported apparent issues around lack of time for telemedicine, lack of necessary technology for the patient and the provider, as well as concerns about confidentiality, remuneration, and inability to adequately assess patients using the service. These results echo some of the current sentiments regarding the uptake of telemedicine into general medical practice [1-5]. Physicians were also concerned about patients not feeling adequately connected to them as a health care provider when using telemedicine, and to a lesser extent, believed that patients would receive lesser quality assessments.

Although there was consistent agreement around possible benefits of telemedicine for HIV care, our findings highlight many perceived challenges and risks that must be addressed

before HIV telemedicine is likely to expand dramatically in practice. Confidentiality, privacy, and remuneration were reported as key challenges to physicians adopting telemedicine in Ontario. These challenges may primarily be due to a lack of information on the physicians' part in regards to the telemedicine services at their disposal and have less to do with regulation limitations, or limitations of the technology. Previous research has shown that physicians are less likely to use telemedicine services on a regular basis if they are not adequately compensated for their time and effort [13,16]. Addressing these perceived barriers to the implementation of telemedicine services is a complex problem that requires assistance from many sources including health care institutions, policy makers, physicians, and patients alike.

Some of these data reflect current themes in distance education; the impact of in-person consultation is glamorized, whereas "Web-based presence" is misunderstood as impersonal. In many responses, it seems that physicians fear that the introduction of telemedicine into practice will be a replacement to in-person consultations and not a compliment to current practices. The nature of HIV as a complex syndrome of various clinical manifestations may also be contributing to the physicians' hesitation to endorse telemedicine. For example, people living with HIV may have complex social, financial, and psychological concerns that seem less amenable to telemedicine techniques when a provider is inexperienced with these remote modalities.

In Canada, the Canada Health Infoway, which is a national initiative, was created to expedite the development and adoption of telemedicine services while addressing reported barriers to

the implementation of eHealth systems. Through the use of provincial partners like the Ontario Telemedicine Network, the infrastructure is in place to provide telemedicine services to those receiving care for HIV. The use of telemedicine to treat other chief health complaints has shown positive results as evidenced by the success of Telehomecare, Telestroke, and Teledermatology programs in Ontario [17].

Telemedicine networks in Canada offer education for physicians and their staff around remuneration, the technology of telemedicine, as well as ways to incorporate telemedicine alongside in-person care. Physicians and health care providers may need ongoing training and support in the form of distance educational sessions to gain up-to-date and meaningful instruction on the benefit of telemedicine services and how to seamlessly integrate them into their practices. By providing physicians with evidence-based research on the growing need for innovative care and the benefits of implementing telemedicine services, their perspectives may change and allow for a greater adoption of the service in HIV care provision.

Limitations

In terms of collecting respondent characteristics, we collected few demographic identifiers in order to protect confidentiality

of northern providers that may be using telemedicine more often or more proficiently than tertiary providers. We also did not gather the ages or genders of physicians who responded in our survey, which could possibly be a predictor of telemedicine perceptions. We also were not able to gather opinions of all physicians practicing HIV care in Ontario, as there is not a full, up-to-date registry of these physicians.

Conclusions

From the perceptions of physicians, telemedicine shows promise in the care of patients living with HIV. More than half of the respondents are already using telemedicine. Whereas many physicians are concerned about their ability to fully assess the health of a patient via telemedicine, most physicians do see a need for it to reduce patient travel times, reduce exposure to stigma, and improve efficiency and timely access to care. Challenges and risks such as technological gaps, confidentiality, and medicolegal concerns must be addressed for physicians to feel more comfortable using telemedicine. Further research is warranted to determine the levels and needs for training of physicians and patients on various telemedicine modalities and technologies. Also, there is a need to compare and contrast the data collected with research evidence in telemedicine uptake in other health areas such as wound care, diabetes, and counseling.

Acknowledgments

This study was funded by a small grant from St Michael's Hospital family health team, and was supported in kind by staff time at the Ontario HIV Treatment Network. Neither organization had any influence in the content of the survey or resulting publication, nor review or approval of the paper prior to publication.

The authors would like to acknowledge the HIV primary care physician team at St Michael's Hospital, as well as Drs Sean Rourke and Rob Williams for their thorough review and edits of the survey.

Authors' Contributions

Authors KA, TF, FIC, and JG were extensively involved in the creation and distribution of the survey, analysis of results, and drafting of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Ontario HIV Treatment Network. Telemedicine and HIV health care URL: <http://www.ohtn.on.ca/wp-content/uploads/sites/9/2014/11/RR88-Telemedicine.pdf> [accessed 2016-10-25] [WebCite Cache ID 6lWscE78n]
2. Palalas A. Blended mobile learning: expanding learning spaces with mobile technologies. In: Tsinakos A, Ally M, editors. *Global Mobile Learning Implementations and Trends*. China: China Central Radio & TV University Press; 2013:86-104.
3. Waldura JF, Neff S, Goldschmidt RH. Teleconsultation for clinicians who provide human immunodeficiency virus care: experience of the national HIV telephone consultation service. *Telemed J E Health* 2011;17(6):472-477 [FREE Full text] [doi: [10.1089/tmj.2010.0210](https://doi.org/10.1089/tmj.2010.0210)] [Medline: [21612517](https://pubmed.ncbi.nlm.nih.gov/21612517/)]
4. Waldura JF, Neff S, Dehlendorf C, Goldschmidt RH. Teleconsultation improves primary care clinicians' confidence about caring for HIV. *J Gen Intern Med* 2013 Jun;28(6):793-800 [FREE Full text] [doi: [10.1007/s11606-013-2332-5](https://doi.org/10.1007/s11606-013-2332-5)] [Medline: [23371417](https://pubmed.ncbi.nlm.nih.gov/23371417/)]
5. Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: a systematic review of reviews. *Int J Med Inform* 2010 Nov;79(11):736-771. [doi: [10.1016/j.ijmedinf.2010.08.006](https://doi.org/10.1016/j.ijmedinf.2010.08.006)] [Medline: [20884286](https://pubmed.ncbi.nlm.nih.gov/20884286/)]
6. Underwood J, Winston A. Guidelines for evaluation and management of cognitive disorders in HIV-positive individuals. *Curr HIV/AIDS Rep* 2016 Oct;13(5):235-240 [FREE Full text] [doi: [10.1007/s11904-016-0324-x](https://doi.org/10.1007/s11904-016-0324-x)] [Medline: [27353598](https://pubmed.ncbi.nlm.nih.gov/27353598/)]
7. Carroll MB, Fields JH, Clerc PG. Rheumatoid arthritis in patients with HIV: management challenges. *Open Access Rheumatol* 2016;8:51-59 [FREE Full text] [doi: [10.2147/OARRR.S87312](https://doi.org/10.2147/OARRR.S87312)] [Medline: [27843370](https://pubmed.ncbi.nlm.nih.gov/27843370/)]

8. León A, Cáceres C, Fernández E, Chausa P, Martin M, Codina C, et al. A new multidisciplinary home care telemedicine system to monitor stable chronic human immunodeficiency virus-infected patients: a randomized study. *PLoS One* 2011 Jan 21;6(1):e14515 [FREE Full text] [doi: [10.1371/journal.pone.0014515](https://doi.org/10.1371/journal.pone.0014515)] [Medline: [21283736](https://pubmed.ncbi.nlm.nih.gov/21283736/)]
9. Ohl M, Dillon D, Moeckli J, Ono S, Waterbury N, Sissel J, et al. Mixed-methods evaluation of a telehealth collaborative care program for persons with HIV infection in a rural setting. *J Gen Intern Med* 2013 Sep;28(9):1165-1173 [FREE Full text] [doi: [10.1007/s11606-013-2385-5](https://doi.org/10.1007/s11606-013-2385-5)] [Medline: [23475640](https://pubmed.ncbi.nlm.nih.gov/23475640/)]
10. Saberi P, Yuan P, John M, Sheon N, Johnson MO. A pilot study to engage and counsel HIV-positive African American youth via telehealth technology. *AIDS Patient Care STDS* 2013 Sep;27(9):529-532 [FREE Full text] [doi: [10.1089/apc.2013.0185](https://doi.org/10.1089/apc.2013.0185)] [Medline: [23991691](https://pubmed.ncbi.nlm.nih.gov/23991691/)]
11. Saifu HN, Asch SM, Goetz MB, Smith JP, Graber CJ, Schaberg D, et al. Evaluation of human immunodeficiency virus and hepatitis C telemedicine clinics. *Am J Manag Care* 2012 Apr;18(4):207-212 [FREE Full text] [Medline: [22554009](https://pubmed.ncbi.nlm.nih.gov/22554009/)]
12. Villalba-Mora E, Casas I, Lupiañez-Villanueva F, Maghiros I. Adoption of health information technologies by physicians for clinical practice: the Andalusian case. *Int J Med Inform* 2015 Jul;84(7):477-485. [doi: [10.1016/j.ijmedinf.2015.03.002](https://doi.org/10.1016/j.ijmedinf.2015.03.002)] [Medline: [25823578](https://pubmed.ncbi.nlm.nih.gov/25823578/)]
13. de Grood C, Raissi A, Kwon Y, Santana MJ. Adoption of e-health technology by physicians: a scoping review. *J Multidiscip Healthc* 2016;9:335-344 [FREE Full text] [doi: [10.2147/JMDH.S103881](https://doi.org/10.2147/JMDH.S103881)] [Medline: [27536128](https://pubmed.ncbi.nlm.nih.gov/27536128/)]
14. Fluidsurveys. Online survey software URL: <http://fluidsurveys.com> [accessed 2016-09-30] [WebCite Cache ID 6kucRYNa4]
15. R Core Team. R-project. Vienna R: A language and environment for statistical computing URL: <https://www.r-project.org/> [accessed 2017-05-17] [WebCite Cache ID 6qWznspIA]
16. Pong R, Hogenbirk J. Reimbursing Physicians for telehealth practice: issues and policy options. *Health Law Review* 2000;9(1):3-13 [FREE Full text]
17. Canada's Health Informatics Association. Leccsct. 2015 Canadian TeleHealth Report URL: <https://www.leccsct.ca/wp-content/uploads/2012/10/2015-TeleHealth-Public-eBook-Final-10-9-15-secured.pdf> [accessed 2017-03-08] [WebCite Cache ID 6ooWbDnII]

Abbreviations

HIV: human immunodeficiency virus

Edited by G Eysenbach; submitted 26.10.16; peer-reviewed by K Claborn, I Prieto-Egido; comments to author 28.01.17; revised version received 10.03.17; accepted 14.04.17; published 30.05.17.

Please cite as:

Anderson K, Francis T, Ibanez-Carrasco F, Globerman J

Physician's Perceptions of Telemedicine in HIV Care Provision: A Cross-Sectional Web-Based Survey

JMIR Public Health Surveill 2017;3(2):e31

URL: <http://publichealth.jmir.org/2017/2/e31/>

doi: [10.2196/publichealth.6896](https://doi.org/10.2196/publichealth.6896)

PMID: [28559226](https://pubmed.ncbi.nlm.nih.gov/28559226/)

©Kelly Anderson, Troy Francis, Francisco Ibanez-Carrasco, Jason Globerman. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 30.05.2017. 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

Recruiting Young Gay and Bisexual Men for a Human Papillomavirus Vaccination Intervention Through Social Media: The Effects of Advertisement Content

Paul L Reiter¹, PhD; Mira L Katz², PhD; Jose A Bauermeister³, PhD; Abigail B Shoben², PhD; Electra D Paskett¹, PhD; Annie-Laurie McRee⁴, DrPH

¹College of Medicine, Ohio State University, Columbus, OH, United States

²College of Public Health, Ohio State University, Columbus, OH, United States

³Department of Family and Community Health, University of Pennsylvania School of Nursing, Philadelphia, PA, United States

⁴Division of General Pediatrics and Adolescent Health, University of Minnesota Medical School, Minneapolis, MN, United States

Corresponding Author:

Paul L Reiter, PhD

College of Medicine

Ohio State University

1590 North High Street, Suite 525

Columbus, OH, 43201

United States

Phone: 1 614 366 4265

Fax: 1 614 366 5454

Email: paul.reiter@osumc.edu

Abstract

Background: Web-based approaches, specifically social media sites, represent a promising approach for recruiting young gay and bisexual men for research studies. Little is known, however, about how the performance of social media advertisements (ads) used to recruit this population is affected by ad content (ie, image and text).

Objective: The aim of this study was to evaluate the effects of different images and text included in social media ads used to recruit young gay and bisexual men for the pilot test of a Web-based human papillomavirus (HPV) vaccination intervention.

Methods: In July and September 2016, we used paid Facebook advertisements to recruit men who were aged 18-25 years, self-identified as gay or bisexual, US resident, and had not received HPV vaccine. A 4x2x2 factorial experiment varied ad image (a single young adult male, a young adult male couple, a group of young adult men, or a young adult male talking to a doctor), content focus (text mentioning HPV or HPV vaccine), and disease framing (text mentioning cancer or a sexually transmitted disease [STD]). Poisson regression determined whether these experimental factors affected ad performance.

Results: The recruitment campaign reached a total of 35,646 users who viewed ads for 36,395 times. This resulted in an overall unique click-through rate of 2.01% (717/35,646) and an overall conversion rate of 0.66% (241/36,395). Reach was higher for ads that included an image of a couple (incidence rate ratio, IRR=4.91, 95% CI 2.68-8.97, $P<.001$) or a group (IRR=2.65, 95% CI 1.08-6.50, $P=.03$) compared with those that included an image of a single person. Ads that included an image of a couple also had a higher conversion rate (IRR=2.56, 95% CI 1.13-5.77, $P=.02$) than ads that included an image of a single person. Ads with text mentioning an STD had a higher unique click-through rate compared with ads with text mentioning cancer (IRR=1.34, 95% CI 1.06-1.69, $P=.01$). The campaign cost a total of US \$413.72 and resulted in 150 eligible and enrolled individuals (US \$2.76 per enrolled participant).

Conclusions: Facebook ads are a convenient and cost-efficient strategy for reaching and recruiting young gay and bisexual men for a Web-based HPV vaccination intervention. To help optimize ad performance among this population, researchers should consider the importance of the text and image included in the social media recruitment ads.

(*JMIR Public Health Surveill* 2017;3(2):e33) doi:[10.2196/publichealth.7545](https://doi.org/10.2196/publichealth.7545)

KEYWORDS

HPV; HPV vaccine; gay and bisexual men; social media

Introduction

Human papillomavirus (HPV) infection is the most common sexually transmitted infection in the United States [1]. Gay and bisexual men have high rates of HPV infection and several HPV-related diseases including anal cancer and genital warts [2-4]. The Advisory Committee on Immunization Practices (ACIP) currently recommends routine HPV vaccination for males aged 11-12 years in the United States, with catch-up vaccination for ages 13-21 years [5]. Importantly, the ACIP also recommends routine HPV vaccination for men who have sex with men, including those who identify as gay or bisexual or who intend to have sex with men, through the age of 26 years [5]. Two doses of HPV vaccine are now recommended if the vaccine series is initiated before turning 15 years old, whereas 3 doses are recommended if the vaccine series is initiated after turning 15 years old [5]. Despite recommendations, current HPV vaccine coverage remains modest among males in the United States, including among young gay and bisexual men [6-9]. Recent studies suggest that fewer than 15% of young gay and bisexual men have received any doses of the HPV vaccine [7-9].

Efforts are therefore needed to increase HPV vaccination among young gay and bisexual men. However, reaching and recruiting this population may pose a challenge to HPV vaccination interventions. Recruitment via community settings (eg, community events and organizations) has been a frequently used approach for recruiting gay and bisexual men into research studies, but it is an approach that may oversample individuals who are more involved with the lesbian, gay, bisexual, and transgender (LGBT) community [10]. Some individuals may also be less willing to self-identify as gay or bisexual (or disclose same-sex behavior or attraction) when recruited in-person compared with other approaches [11].

Web-based approaches, specifically social media sites, represent a newer and promising recruitment strategy [12,13]. Facebook is the most popular social media site among US adults, with an estimated 72% of adult Internet users reporting use of Facebook [14]. With more than 6 million sexual minority users in the United States [15], Facebook is a promising platform for recruiting young gay and bisexual men. Several recent efforts have successfully used Facebook to recruit young adults including sexual minority young adults, into research studies for a range of health-related topics [16-25]. Several of these studies reported that Facebook was a convenient and cost-efficient strategy for recruiting young adults [17,18,20,22].

The content of the recruitment ads likely plays a role in the success of a Facebook recruitment campaign. Indeed, past studies involving smoking cessation and mental health have shown that both the images and text included in Facebook ads affect ad performance and recruitment metrics (eg, number of Facebook users who click on an ad) [26,27]. Little is known, however, about how Facebook ad content (ie, image and text) affects the success of recruitment campaigns for sexual minority populations. We report the results of an experiment to determine how Facebook ad content affected the recruitment of young gay and bisexual men for the pilot test of an HPV vaccination

intervention. Results will be highly useful for planning and implementing future interventions among this population.

Methods

Study Overview

“Outsmart HPV” is a mobile-friendly Web-based HPV vaccination intervention developed for young gay and bisexual men. Men were eligible for this project if they (1) were in the age range of 18-25 years, (2) self-identified as gay or bisexual, (3) were a US resident, and (4) had not received any doses of HPV vaccine. Following intervention development, we conducted a pilot test of this project. Recruitment for the pilot test occurred in two waves, with the first wave occurring in July 2016 and the second in September 2016. All recruitment occurred via paid Facebook ads. This report includes recruitment results from the pilot test. The Institutional Review Board at The Ohio State University approved this study. The parent randomized trial is registered at ClinicalTrials.gov (identifier NCT02835755).

Facebook Advertisements

We created recruitment ads using Facebook’s Ads Manager program. All ads included a headline, main text, image, Outsmart HPV project logo, and weblink to the project website. Ads adhered to Facebook’s requirements at the time of ad development, including character limits (25 characters for the headline and 90 characters for the main text) and image restrictions (any image used could not include more than 20% text) [28]. We designed all ads to appear in the “News Feed” on Facebook, which is a streaming list of updates from the user’s connections (eg, friends) and advertisers. We focused solely on News Feed ads, as opposed to other Facebook ad locations (eg, right column), since News Feed ads are more effective in terms of recruitment metrics for research studies [26].

The headline (“Earn Up To \$95 Online!”), Outsmart HPV logo, and weblink to the project website were identical for all ads. The main text for all ads started with “Online HPV study for gay & bisexual men,” but the remaining main text and image varied across the ads using a 4x2x2 factorial experiment design. Experimental factors included image type, content focus, and disease framing. Image type (4 conditions) refers to the primary image featured in the ad. All ad images featured at least one male who appeared to be a young adult in the target age group, but ads varied on the number and type of people in the image: (1) a single man (“single person”), (2) a male couple (“couple”), (3) a group of more than 2 men (“group”), or (4) a man interacting with a health care provider (“doctor”). The people included in the images were from a range of racial or ethnic groups. Content focus (2 conditions) involved whether the ad’s remaining main text mentioned (1) HPV, or (2) HPV vaccine. Disease framing (2 conditions) involved whether the ad’s remaining main text mentioned (1) cancer, or (2) a sexually transmitted disease (STD). We created 2 ads for each of the 16 experimental conditions (ie, 2 variations of each image type), resulting in a total of 32 ads. [Figure 1](#) shows example ads across various experimental conditions. Facebook approved all ads before their use.

Figure 1. Example Facebook recruitment ads for Outsmart HPV (human papillomavirus). All ads were News Feed ads, and experimental factors included: image type (single person, couple, group, or doctor), content focus (HPV or HPV vaccine), and disease framing (cancer or sexually transmitted disease [STD]).



Image Type: Single Person
Content Focus: HPV
Disease Framing: Cancer



Image Type: Couple
Content Focus: HPV
Disease Framing: STD



Image Type: Group
Content Focus: HPV Vaccine
Disease Framing: Cancer



Image Type: Doctor
Content Focus: HPV Vaccine
Disease Framing: STD

Recruitment Campaign

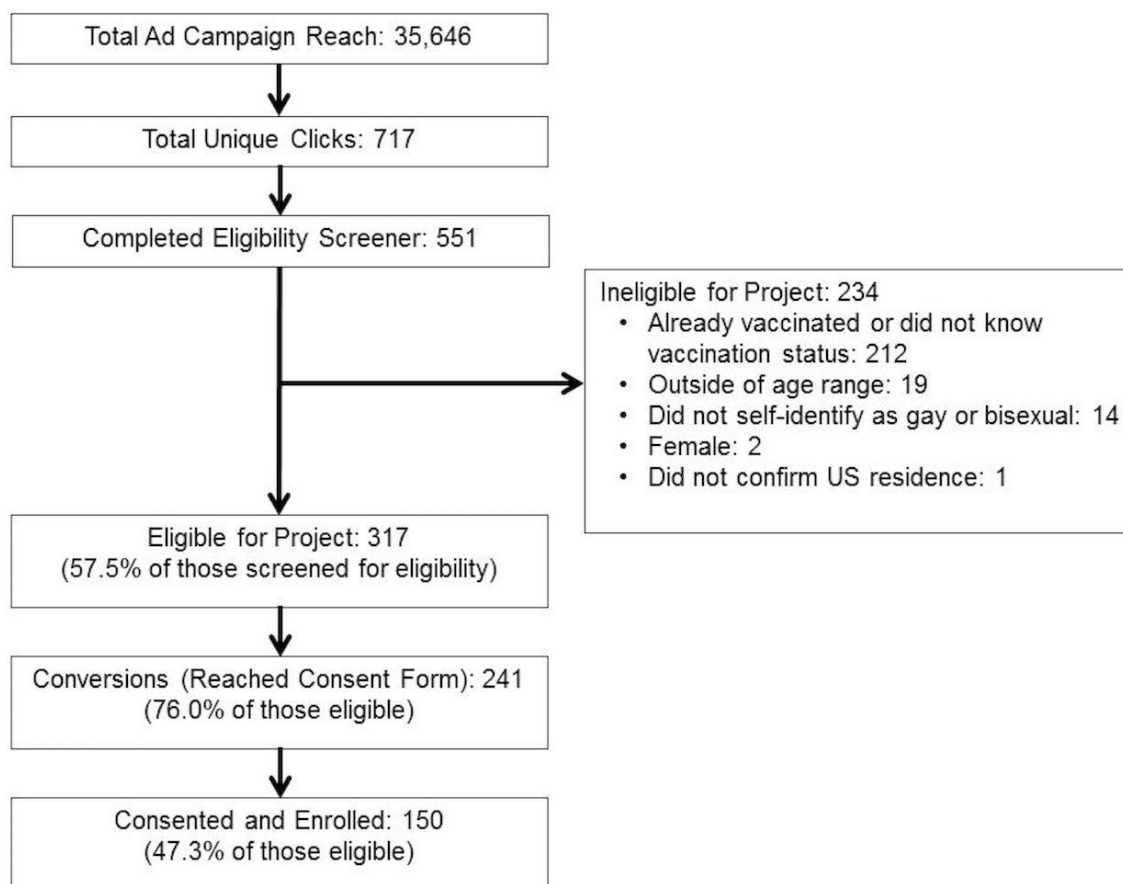
We targeted all ads to Facebook users by sex (male), age (18-25 years), location (United States), and language (English). We further targeted the ads on keywords generated from information that Facebook users add to their Timeline, pages they “like,” or ads they have previously clicked on. The keywords used in our campaign included “bisexuality,” “homosexuality,” “same-sex relationship,” “genderqueer,” “gay pride,” “LGBT community,” “LGBT culture,” or “rainbow flag (LGBT movement).” The chance that a given ad appeared on a user’s page was determined by a Facebook algorithm that considers several factors (eg, the spending limit specified for the ad, competition from other ads).

Project Enrollment

Ads linked potential participants to the project website or the project’s Facebook page (which in turn directed potential participants to the project website). We collaborated with the Center for Health Communications Research (CHCR) at the University of Michigan to develop and maintain the project website, which was a mobile-friendly website accessible by desktop or laptop, tablet computer, or smartphone (iOS and Android). Once on the project website, potential participants first completed a project eligibility screener. Those determined to be eligible were then asked to provide informed consent and create a project account. Following account creation, participants completed a preintervention Web-based survey, were randomly assigned to either the intervention or control arm, and then completed remaining study activities (eg, viewing intervention or control materials on the Web about HPV vaccine and completing 3 additional Web-based surveys over the course of

7 months). Participants could earn up to US \$95 in gift cards during the course of the project.

Figure 2. Facebook recruitment campaign results for Outsmart HPV (human papillomavirus).



Measures

Facebook provided standard data on several metrics related to ad performance. Similar to past studies involving Facebook recruitment campaigns [17,20,25-27], we examined 3 metrics as outcomes: (1) reach (ie, the number of Facebook users an ad was shown to), (2) unique click-through rate (UCTR; number of unique Facebook users who clicked on an ad divided by reach), and (3) conversion rate (number of unique Facebook users who reached the study consent form divided by the number of times an ad was viewed). We express UCTRs and conversion rates as percentages for the remainder of this paper. We also report information on the cost of the recruitment campaign, as well as the demographic and health-related characteristics of enrolled participants (using data from the project eligibility screener and preintervention survey).

Data Analysis

We calculated descriptive statistics for all Facebook metrics and survey data. We used multivariable Poisson models to determine the effects of our experimental factors (image type, content focus, and disease framing) on ad reach, UCTR, and conversion rate. The multivariable models included all three experimental factors and controlled for recruitment wave. Poisson models for UCTR and conversion rate included an offset or exposure term due to differences in the denominators

of these quantities. The analytic dataset used for Poisson analyses included 64 observations (32 created ads that were used in two recruitment waves). We report incidence rate ratios (IRRs) and 95% CI from the multivariable Poisson models. Robust sandwich standard errors were used in calculating the 95% CIs and *P* values. We conducted analyses using Stata version 12.0 (Stata Corp), and all statistical tests were 2-tailed with a critical alpha of .05.

Results

Recruitment Campaign Results

Our Facebook recruitment campaign reached a total of 35,646 users who viewed ads for 36,395 times (Figure 2). There were 717 unique clicks on the ads, resulting in an overall UCTR=2.01% (717/35,646). A total of 551 potential participants completed the project eligibility screener, of whom 317 (57.5%) were eligible for participation in the study. Among those found to be ineligible, reasons for ineligibility (more than one reason could apply per person) included: already received HPV vaccine or did not know vaccination status (90.5%, 212/234), outside of the project's age range (8.1%, 19/234), did not self-identify as gay or bisexual (6.0%, 14/234), female gender (0.9%, 2/234), and did not confirm US residence (0.4%, 1/234).

The recruitment campaign resulted in 241 conversions (76.0% of potential participants found to be eligible [241/317]) and an overall conversion rate of 0.66% (241/36,395). A total of 150 individuals provided consent and were enrolled into the project (47.3% of potential participants found to be eligible [150/317]). It took about 4 days in July 2016 (ie, recruitment wave 1) and about 2 days in September 2016 (ie, recruitment wave 2) to reach this enrollment goal. The total cost of the campaign was US \$413.72. This translates into a cost of US \$0.58 per unique click on an ad, US \$1.72 per conversion, and US \$2.76 per enrolled participant.

Factorial Ad Experiment

The image included in ads affected ad performance. Reach was higher for ads that included an image of a couple (IRR=4.91, 95% CI 2.68-8.97) or a group (IRR=2.65, 95% CI 1.08-6.50) compared with those that included an image of a single person (Table 1). Ads that included an image of a couple also had a higher conversion rate (IRR=2.56, 95% CI 1.13-5.77) and UCTR (OR=1.66, 95% CI 0.99-2.80) than ads that included an image of a single person, though the latter association was of borderline statistical significance in the multivariable model ($P=.06$). For disease framing, ads with main text mentioning an STD had a higher UCTR compared with ads with main text mentioning cancer (IRR=1.34, 95% CI 1.06-1.69). The content focus of ads did not affect ad performance.

Table 1. Results of Facebook advertisement factorial experiment. The analytic dataset included 64 observations (32 created ads that were used in two recruitment waves). Multivariable models included all variables in the table and controlled for recruitment wave.

Factor	Reach ^a				UCTR ^{b,c}			Conversion rate ^{c,d}		
	n	Mean	IRR ^e (95% CI)	P	Mean	IRR (95% CI)	P	Mean	IRR (95% CI)	P
Image type										
Single person	16	210.88	ref ^f		1.51	ref		0.35	ref	
Couple	16	1034.56	4.91 (2.68-8.97)	<.001	2.52	1.66 (0.99-2.80)	.06	0.85	2.56 (1.13-5.77)	.02
Group	16	557.94	2.65 (1.08-6.50)	.03	1.73	1.03 (0.62-1.71)	.91	0.59	1.76 (0.77-4.03)	.18
Doctor	16	424.50	2.01 (0.88-4.62)	.10	1.40	0.91 (0.50-1.67)	.77	0.45	1.26 (0.52-3.09)	.61
Content focus										
HPV ^g	32	558.34	ref		2.05	ref		0.62	ref	
HPV vaccine	32	555.59	1.00 (0.53-1.86)	.99	1.97	0.98 (0.79-1.22)	.86	0.71	1.15 (0.93-1.43)	.19
Disease framing										
Cancer	32	528.38	ref		1.79	ref		0.68	ref	
STD ^h	32	585.56	1.11 (0.59-2.07)	.75	2.21	1.34 (1.06-1.69)	.01	0.65	0.98 (0.79-1.22)	.85

^aReach was defined as number of Facebook users ad was shown to.

^bUnique click-through rate (UCTR) was defined as the number of unique users who clicked on an ad divided by reach.

^cDisplayed as a percentage.

^dConversion rate was defined as the number of unique users who reached the study consent form divided by the number of times an ad was viewed.

^eIRR: incidence rate ratio.

^fref: reference group.

^gHPV: human papillomavirus.

^hSTD: sexually transmitted disease.

Participant Characteristics

Enrolled participants were from 31 states and the District of Columbia. Most participants were in the age range of 22-25 years (58.7%, 88/150), non-Hispanic white (56.7%, 85/150), and not married or living with a partner (80.0%, 120/150; [Multimedia Appendix 1](#)). About 82.7% (124/150) of participants self-identified as gay. Most participants did not have a college degree (62.7%, 94/150) and reported a household income of less than US \$50,000 (76.0%, 114/150). Most participants had some form of health insurance (82.0%, 123/150), though fewer than half reported having a routine medical check-up in the last year (46.7%, 70/150). About half of the participants (51.3%, 77/150) reported they were younger than 18 years of age at

sexual debut (first vaginal, anal, or oral intercourse), and most reported having at least six male sexual partners during their lifetime (64.7%, 97/150). Few participants reported being positive for human immunodeficiency virus (HIV; 5.3%, 8/150) or a history of either genital warts (6.0%, 9/150) or another STD (18.0%, 27/150).

Discussion

Principal Findings

Reaching and recruiting young gay and bisexual men may be a challenge faced by interventions to increase HPV vaccination. Social media sites represent a potentially effective strategy for overcoming this challenge, yet little is known about the

effectiveness of this strategy in recruiting for such interventions. We successfully enrolled 150 eligible young gay and bisexual men via Facebook ads for a pilot test of Outsmart HPV. Recruitment occurred over a short time period and cost less than US \$500 total, translating into less than US \$3 spent on recruitment per enrolled participant. The recruitment metrics for our study (eg, UCTR and cost per enrolled participant) are similar to those from past studies that recruited via Facebook ads [16,17,25,29]. Our findings not only add to the growing body of literature that shows Facebook ads are a convenient and cost-efficient recruitment strategy for research studies but specifically suggest that they are an effective strategy for recruiting young gay and bisexual men for Web-based HPV vaccination interventions.

Findings from our study also identify potential strategies for optimizing the performance of Facebook ads. Similar to a past study [27], ad images had a meaningful effect on ad performance. Ads that featured an image of a young adult male couple, and to a lesser degree an image of a group of young adult men, performed better than ads that featured a single young adult male. Ads that included an image of a couple may have appealed to the importance of romantic relationships during adolescence and young adulthood [30]. Furthermore, these ads could have prompted potential participants to think about the health of their sexual partners, which may have been motivating since previous research suggests that young sexual minority men view the protection of their partners' health as an advantage of HPV vaccination [31]. Ads that included an image of a group of young adult men may have affected potential participants' perceived social norms about the topic and the study. Positive social norms have been previously associated with HPV vaccination behaviors among young gay and bisexual men [7]. Future studies that recruit through Facebook should use ads that feature multiple individuals in the target age range to help improve ad performance and recruitment. Furthermore, we recommend that recruitment ads in future research feature individuals who are racially or ethnically diverse, as we did in this study, so that the ads are relatable and appealing to a wide audience.

The text included in Facebook ads also affected ad performance in our study. Ads with main text mentioning an STD had a higher UCTR compared with ads mentioning cancer. Past research suggests that adolescents and young adults are less future-orientated (eg, tend to focus on potential short-term vs long-term consequences) than older adults [32], and potential participants likely perceived STDs to be a more relevant and immediate health outcome than cancer. Indeed, nearly half of all incident STDs in the United States occur among adolescents and young adults [33], whereas cancer is much less common among young adults than older ages [34]. Our finding is also consistent with past research examining the effects of message framing on HPV vaccine acceptability, as young adults tended to be more receptive to HPV vaccine when it was presented in the context of an STD or genital warts [35,36]. This is in contrast to older adults, who tended to be more receptive to HPV vaccine when it was presented as preventing cancer [37,38]. It is therefore important that future efforts consider framing HPV and HPV vaccine in the context of an STD when targeting

materials including study recruitment materials, toward young adults.

The characteristics of participants in our study were comparable with past samples of gay and bisexual men who were recruited through various other strategies. For example, the demographic characteristics of participants in our study (eg, age, race or ethnicity, education level) are highly similar to those of young sexual minority men from the National Survey of Family Growth and the National Longitudinal Study of Adolescent Health [39,40]. Furthermore, about 80% of participants in our study self-identified as gay, which closely resembles other national samples of gay and bisexual men [7,41-43]. Health-related characteristics (eg, lack of health insurance and receipt of routine medical check-up in the last year) were also similar between participants in our study and those from past studies [7,40,44,45]. These similarities should not be altogether surprising given the ubiquitous nature of Facebook use among young adults [46], and they suggest that Facebook ads can recruit diverse samples of young gay and bisexual men that are comparable with what would be obtained through more traditional recruitment methods. Future research should further explore these comparisons, as well as continue to monitor how samples of young gay and bisexual men recruited via Facebook compare with those recruited through other social media sites that may have different audiences (eg, Instagram, Grindr, and MiGente). Data from the American Men's Internet Survey suggest that characteristics of sexual minority men differ by the social media site used of recruitment [47], and it is important to continue to monitor these potential differences as social media evolve.

Strengths and Limitations

Study strengths include the use of an experimental design, recruitment of a national sample, and examining several metrics related to ad performance. Our study also has several limitations. Data were not available on the characteristics of Facebook users who were shown an ad but did not click on it. Similarly, we were not able to assess reasons why Facebook users did not click on an ad or clicked on an ad but did not complete the project eligibility screener. We were also not able to link ad performance data with data from study surveys to determine how ad performance may have differed across demographic groups. Future research should identify strategies for linking these data and explore potential differences in ad performance. Fraudulent accounts (eg, multiple accounts) are a growing concern for Web-based research [48], and we used several recommended strategies to reduce this possibility. We used a Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) during the project account creation process, inspected participants' account and contact information (eg, email address) for similarities between accounts, and inspected survey data for inconsistent or illogical responses [48]. Our pilot test had a somewhat modest sample size, and future efforts should examine how ad performance and metrics may change over time when recruiting large samples of young gay and bisexual men.

Conclusions

Facebook ads are a convenient and cost-efficient strategy for reaching and recruiting young gay and bisexual men for a Web-based HPV vaccination intervention. Importantly, the characteristics of young gay and bisexual men enrolled via this

recruitment strategy are similar to those enrolled via other recruitment strategies. Future Facebook recruitment efforts for this population should strongly consider the importance of ad content. Specifically, it may be beneficial to include ads that feature a young adult male couple and that frame HPV and HPV vaccine in the context of an STD.

Acknowledgments

This study was supported by the National Cancer Institute of the National Institutes of Health under Award Number R21CA194831. 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

PLR had received research grants from Merck Sharp & Dohme Corp. and Cervical Cancer-Free America, via an unrestricted educational grant from GlaxoSmithKline. EDP had received research grants from Merck Sharp & Dohme Corp. These funds were not used to support this research study.

Multimedia Appendix 1

Demographic and health-related characteristics of young gay and bisexual men enrolled in Outsmart HPV (human papillomavirus; n=150).

[[PDF File \(Adobe PDF File\), 49KB - publichealth_v3i2e33_app1.pdf](#)]

References

- Centers for Disease Control and Prevention (CDC). 2017. Genital HPV infection - fact sheet URL: <https://www.cdc.gov/std/HPV/STDFact-HPV.htm> [accessed 2017-02-15] [WebCite Cache ID 60fWwBya]
- Smith JS, Gilbert PA, Melendy A, Rana RK, Pimenta JM. Age-specific prevalence of human papillomavirus infection in males: a global review. *J Adolesc Health* 2011 Jun;48(6):540-552. [doi: [10.1016/j.jadohealth.2011.03.010](https://doi.org/10.1016/j.jadohealth.2011.03.010)] [Medline: [21575812](https://pubmed.ncbi.nlm.nih.gov/21575812/)]
- Daling JR, Weiss NS, Hislop TG, Maden C, Coates RJ, Sherman KJ, et al. Sexual practices, sexually transmitted diseases, and the incidence of anal cancer. *N Engl J Med* 1987 Oct 15;317(16):973-977. [doi: [10.1056/NEJM198710153171601](https://doi.org/10.1056/NEJM198710153171601)] [Medline: [2821396](https://pubmed.ncbi.nlm.nih.gov/2821396/)]
- Dinh T, Sternberg M, Dunne EF, Markowitz LE. Genital warts among 18- to 59-year-olds in the United States, national health and nutrition examination survey, 1999--2004. *Sex Transm Dis* 2008 Apr;35(4):357-360. [doi: [10.1097/OLQ.0b013e3181632d61](https://doi.org/10.1097/OLQ.0b013e3181632d61)] [Medline: [18360316](https://pubmed.ncbi.nlm.nih.gov/18360316/)]
- Meites E, Kempe A, Markowitz LE. Use of a 2-dose schedule for human papillomavirus vaccination - updated recommendations of the advisory committee on immunization practices. *MMWR Morb Mortal Wkly Rep* 2016 Dec 16;65(49):1405-1408 [FREE Full text] [doi: [10.15585/mmwr.mm6549a5](https://doi.org/10.15585/mmwr.mm6549a5)] [Medline: [27977643](https://pubmed.ncbi.nlm.nih.gov/27977643/)]
- Reagan-Steiner S, Yankey D, Jeyarajah J, Elam-Evans LD, Curtis CR, MacNeil J, et al. National, regional, state, and selected local area vaccination coverage among adolescents aged 13-17 years - United States, 2015. *MMWR Morb Mortal Wkly Rep* 2016 Aug 26;65(33):850-858 [FREE Full text] [doi: [10.15585/mmwr.mm6533a4](https://doi.org/10.15585/mmwr.mm6533a4)] [Medline: [27561081](https://pubmed.ncbi.nlm.nih.gov/27561081/)]
- Reiter PL, McRee A, Katz ML, Paskett ED. Human papillomavirus vaccination among young adult gay and bisexual men in the United States. *Am J Public Health* 2015 Jan;105(1):96-102. [doi: [10.2105/AJPH.2014.302095](https://doi.org/10.2105/AJPH.2014.302095)] [Medline: [25393178](https://pubmed.ncbi.nlm.nih.gov/25393178/)]
- Agénor M, Peitzmeier SM, Gordon AR, Charlton BM, Haneuse S, Potter J, et al. Sexual orientation identity disparities in human papillomavirus vaccination initiation and completion among young adult US women and men. *Cancer Causes Control* 2016 Oct;27(10):1187-1196. [doi: [10.1007/s10552-016-0796-4](https://doi.org/10.1007/s10552-016-0796-4)] [Medline: [27507284](https://pubmed.ncbi.nlm.nih.gov/27507284/)]
- Jones J, Poole A, Lasley-Bibbs V, Johnson M. LGBT health and vaccinations: findings from a community health survey of Lexington-Fayette County, Kentucky, USA. *Vaccine* 2016 Apr 07;34(16):1909-1914. [doi: [10.1016/j.vaccine.2016.02.054](https://doi.org/10.1016/j.vaccine.2016.02.054)] [Medline: [26930365](https://pubmed.ncbi.nlm.nih.gov/26930365/)]
- Meyer IH, Wilson PA. Sampling lesbian, gay, and bisexual populations. *J Couns Psychol* 2009;56(1):23-31. [doi: [10.1037/a0014587](https://doi.org/10.1037/a0014587)]
- Institute of Medicine (IOM). *The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding*. Washington, DC: The National Academies Press; 2011.
- Sanchez T, Smith A, Denson D, Dinunno E, Lansky A. Internet-based methods may reach higher-risk men who have sex with men not reached through venue-based sampling. *Open AIDS J* 2012;6:83-89 [FREE Full text] [doi: [10.2174/1874613601206010083](https://doi.org/10.2174/1874613601206010083)] [Medline: [23049657](https://pubmed.ncbi.nlm.nih.gov/23049657/)]

13. Hernandez-Romieu AC, Sullivan PS, Sanchez TH, Kelley CF, Peterson JL, Del RC, et al. The comparability of men who have sex with men recruited from venue-time-space sampling and facebook: a cohort study. *JMIR Res Protoc* 2014;3(3):e37 [FREE Full text] [doi: [10.2196/resprot.3342](https://doi.org/10.2196/resprot.3342)] [Medline: [25048694](https://pubmed.ncbi.nlm.nih.gov/25048694/)]
14. Duggan M. Pew Research Center. 2015. The demographics of social media users URL: <http://www.pewinternet.org/2015/08/19/the-demographics-of-social-media-users/> [accessed 2017-02-15] [WebCite Cache ID 6oIgaZeSI]
15. State B, Wernerfelt N. Facebook research. 2015. America's coming out on facebook URL: <https://research.fb.com/americas-coming-out-on-facebook/> [accessed 2017-02-15] [WebCite Cache ID 6oIh0X8cr]
16. Lane TS, Armin J, Gordon JS. Online recruitment methods for web-based and mobile health studies: a review of the literature. *J Med Internet Res* 2015;17(7):e183 [FREE Full text] [doi: [10.2196/jmir.4359](https://doi.org/10.2196/jmir.4359)] [Medline: [26202991](https://pubmed.ncbi.nlm.nih.gov/26202991/)]
17. Pedersen ER, Helmuth ED, Marshall GN, Schell TL, PunKay M, Kurz J. Using facebook to recruit young adult veterans: online mental health research. *JMIR Res Protoc* 2015;4(2):e63 [FREE Full text] [doi: [10.2196/resprot.3996](https://doi.org/10.2196/resprot.3996)] [Medline: [26033209](https://pubmed.ncbi.nlm.nih.gov/26033209/)]
18. Staffileno BA, Zschunke J, Weber M, Gross LE, Fogg L, Tangney CC. The feasibility of using facebook, craigslist, and other online strategies to recruit young African American women for a web-based healthy lifestyle behavior change intervention. *J Cardiovasc Nurs* 2016:- Epub ahead of print. [doi: [10.1097/JCN.0000000000000360](https://doi.org/10.1097/JCN.0000000000000360)] [Medline: [27428356](https://pubmed.ncbi.nlm.nih.gov/27428356/)]
19. Kaysen D, Davis K, Kilmer J. Use of social networking sites to sample lesbian and bisexual women. *The Addictions Newsletter* 2011;18(1):14-15 [FREE Full text]
20. Subasinghe AK, Nguyen M, Wark JD, Tabrizi SN, Garland SM. Targeted facebook advertising is a novel and effective method of recruiting participants into a human papillomavirus vaccine effectiveness study. *JMIR Res Protoc* 2016 Jul 22;5(3):e154 [FREE Full text] [doi: [10.2196/resprot.5679](https://doi.org/10.2196/resprot.5679)] [Medline: [27450586](https://pubmed.ncbi.nlm.nih.gov/27450586/)]
21. Raviotta JM, Nowalk MP, Lin CJ, Huang H, Zimmerman RK. Using facebook™ to recruit college-age men for a human papillomavirus vaccine trial. *Am J Mens Health* 2016;10(2):110-119. [doi: [10.1177/1557988314557563](https://doi.org/10.1177/1557988314557563)] [Medline: [25389213](https://pubmed.ncbi.nlm.nih.gov/25389213/)]
22. Ryan GS. Online social networks for patient involvement and recruitment in clinical research. *Nurse Res* 2013;21(1):35-39. [doi: [10.7748/nr2013.09.21.1.35.e302](https://doi.org/10.7748/nr2013.09.21.1.35.e302)] [Medline: [24004430](https://pubmed.ncbi.nlm.nih.gov/24004430/)]
23. Bauermeister JA, Pingel ES, Jadwin-Cakmak L, Harper GW, Horvath K, Weiss G, et al. Acceptability and preliminary efficacy of a tailored online HIV/STI testing intervention for young men who have sex with men: the Get Connected! program. *AIDS Behav* 2015 Oct;19(10):1860-1874 [FREE Full text] [doi: [10.1007/s10461-015-1009-y](https://doi.org/10.1007/s10461-015-1009-y)] [Medline: [25638038](https://pubmed.ncbi.nlm.nih.gov/25638038/)]
24. Bauermeister JA, Zimmerman MA, Johns MM, Glowacki P, Stoddard S, Volz E. Innovative recruitment using online networks: lessons learned from an online study of alcohol and other drug use utilizing a web-based, respondent-driven sampling (webRDS) strategy. *J Stud Alcohol Drugs* 2012 Sep;73(5):834-838 [FREE Full text] [Medline: [22846248](https://pubmed.ncbi.nlm.nih.gov/22846248/)]
25. Prescott TL, Phillips IG, DuBois LZ, Bull SS, Mustanski B, Ybarra ML. Reaching adolescent gay, bisexual, and queer men online: development and refinement of a National recruitment strategy. *J Med Internet Res* 2016;18(8):e200 [FREE Full text] [doi: [10.2196/jmir.5602](https://doi.org/10.2196/jmir.5602)] [Medline: [27492781](https://pubmed.ncbi.nlm.nih.gov/27492781/)]
26. Ramo DE, Rodriguez TM, Chavez K, Sommer MJ, Prochaska JJ. Facebook recruitment of young adult smokers for a cessation trial: methods, metrics, and lessons learned. *Internet Interv* 2014 Apr;1(2):58-64. [doi: [10.1016/j.invent.2014.05.001](https://doi.org/10.1016/j.invent.2014.05.001)] [Medline: [25045624](https://pubmed.ncbi.nlm.nih.gov/25045624/)]
27. Hui A, Wong PW, Fu K. Evaluation of an online campaign for promoting help-seeking attitudes for depression using a facebook advertisement: an online randomized controlled experiment. *JMIR Ment Health* 2015;2(1):e5 [FREE Full text] [doi: [10.2196/mental.3649](https://doi.org/10.2196/mental.3649)] [Medline: [26543911](https://pubmed.ncbi.nlm.nih.gov/26543911/)]
28. Facebook. 2016. Facebook ads URL: <https://www.facebook.com/business/products/ads> [accessed 2017-05-22] [WebCite Cache ID 6qeO9ac83]
29. Amon KL, Campbell AJ, Hawke C, Steinbeck K. Facebook as a recruitment tool for adolescent health research: a systematic review. *Acad Pediatr* 2014;14(5):439-47.e4. [doi: [10.1016/j.acap.2014.05.049](https://doi.org/10.1016/j.acap.2014.05.049)] [Medline: [25169155](https://pubmed.ncbi.nlm.nih.gov/25169155/)]
30. Meier A, Allen G. Romantic relationships from adolescence to young adulthood: evidence from the National longitudinal study of adolescent health. *Sociol Q* 2009;50(2):308-335 [FREE Full text] [doi: [10.1111/j.1533-8525.2009.01142.x](https://doi.org/10.1111/j.1533-8525.2009.01142.x)] [Medline: [25332511](https://pubmed.ncbi.nlm.nih.gov/25332511/)]
31. Wheldon CW, Daley EM, Walsh-Buhi ER, Baldwin JA, Nyitray AG, Giuliano AR. An integrative theoretical framework for HPV vaccine promotion among male sexual minorities. *Am J Mens Health* 2016 Jun 06:- Epub ahead of print. [doi: [10.1177/1557988316652937](https://doi.org/10.1177/1557988316652937)] [Medline: [27272241](https://pubmed.ncbi.nlm.nih.gov/27272241/)]
32. Steinberg L, Graham S, O'Brien L, Woolard J, Cauffman E, Banich M. Age differences in future orientation and delay discounting. *Child Dev* 2009;80(1):28-44. [doi: [10.1111/j.1467-8624.2008.01244.x](https://doi.org/10.1111/j.1467-8624.2008.01244.x)] [Medline: [19236391](https://pubmed.ncbi.nlm.nih.gov/19236391/)]
33. Satterwhite CL, Tortrone E, Meites E, Dunne EF, Mahajan R, Ocfemia MC, et al. Sexually transmitted infections among US women and men: prevalence and incidence estimates, 2008. *Sex Transm Dis* 2013 Mar;40(3):187-193. [doi: [10.1097/OLQ.0b013e318286bb53](https://doi.org/10.1097/OLQ.0b013e318286bb53)] [Medline: [23403598](https://pubmed.ncbi.nlm.nih.gov/23403598/)]
34. Bleyer A, Viny A, Barr R. Cancer in 15- to 29-year-olds by primary site. *Oncologist* 2006 Jun;11(6):590-601 [FREE Full text] [doi: [10.1634/theoncologist.11-6-590](https://doi.org/10.1634/theoncologist.11-6-590)] [Medline: [16794238](https://pubmed.ncbi.nlm.nih.gov/16794238/)]
35. Yang JZ, Pittman MM. The silver lining of shame: framing HPV to influence vaccination intentions. *Health Commun* 2016 Jul 27;32(8):987-994. [doi: [10.1080/10410236.2016.1196420](https://doi.org/10.1080/10410236.2016.1196420)] [Medline: [27463558](https://pubmed.ncbi.nlm.nih.gov/27463558/)]

36. Krieger JL, Sarge MA. A serial mediation model of message framing on intentions to receive the human papillomavirus (HPV) vaccine: revisiting the role of threat and efficacy perceptions. *Health Commun* 2013;28(1):5-19. [doi: [10.1080/10410236.2012.734914](https://doi.org/10.1080/10410236.2012.734914)] [Medline: [23330854](https://pubmed.ncbi.nlm.nih.gov/23330854/)]
37. McRee A, Reiter PL, Chantala K, Brewer NT. Does framing human papillomavirus vaccine as preventing cancer in men increase vaccine acceptability? *Cancer Epidemiol Biomarkers Prev* 2010 Aug;19(8):1937-1944 [FREE Full text] [doi: [10.1158/1055-9965.EPI-09-1287](https://doi.org/10.1158/1055-9965.EPI-09-1287)] [Medline: [20647398](https://pubmed.ncbi.nlm.nih.gov/20647398/)]
38. Leader AE, Weiner JL, Kelly BJ, Hornik RC, Cappella JN. Effects of information framing on human papillomavirus vaccination. *J Womens Health (Larchmt)* 2009 Feb;18(2):225-233 [FREE Full text] [doi: [10.1089/jwh.2007.0711](https://doi.org/10.1089/jwh.2007.0711)] [Medline: [19183094](https://pubmed.ncbi.nlm.nih.gov/19183094/)]
39. Fasula AM, Oraka E, Jeffries WL, Carry M, Bañez OM, Balaji AB, Workgroup on Adolescent Sexual and Reproductive Health. Young sexual minority males in the United States: sociodemographic characteristics and sexual attraction, identity and behavior. *Perspect Sex Reprod Health* 2016 Mar;48(1):3-8. [doi: [10.1363/48e7016](https://doi.org/10.1363/48e7016)] [Medline: [26742996](https://pubmed.ncbi.nlm.nih.gov/26742996/)]
40. Strutz KL, Herring AH, Halpern CT. Health disparities among young adult sexual minorities in the U.S. *Am J Prev Med* 2015 Jan;48(1):76-88 [FREE Full text] [doi: [10.1016/j.amepre.2014.07.038](https://doi.org/10.1016/j.amepre.2014.07.038)] [Medline: [25241194](https://pubmed.ncbi.nlm.nih.gov/25241194/)]
41. Gates G. UCLA. 2011. How many people are lesbian, gay, bisexual, and transgender? URL: <http://williamsinstitute.law.ucla.edu/wp-content/uploads/Gates-How-Many-People-LGBT-Apr-2011.pdf> [accessed 2017-02-15] [WebCite Cache ID [60lhJ02Yq](https://www.webcitation.org/60lhJ02Yq)]
42. Pew Research Center. 2013. A survey LGBT Americans URL: <http://www.pewsocialtrends.org/2013/06/13/a-survey-of-lgbt-americans/> [WebCite Cache ID [60lgIO2NN](https://www.webcitation.org/60lgIO2NN)]
43. Reiter PL, Brewer NT, McRee A, Gilbert P, Smith JS. Acceptability of HPV vaccine among a national sample of gay and bisexual men. *Sex Transm Dis* 2010 Mar;37(3):197-203 [FREE Full text] [doi: [10.1097/OLQ.0b013e3181bf542c](https://doi.org/10.1097/OLQ.0b013e3181bf542c)] [Medline: [20118831](https://pubmed.ncbi.nlm.nih.gov/20118831/)]
44. Kates J, Ranji U, Beamesderfer A, Salganicoff A, Dawson L. KFF. 2016. Health and access to care and coverage for lesbian, gay, bisexual, and transgender (LGBT) individuals in the U.S URL: <http://kff.org/report-section/health-and-access-to-care-and-coverage-for-lgbt-individuals-in-the-u-s-update-health-challenges/> [WebCite Cache ID [60lhTiOcb](https://www.webcitation.org/60lhTiOcb)]
45. Jackson CL, Agénor M, Johnson DA, Austin SB, Kawachi I. Sexual orientation identity disparities in health behaviors, outcomes, and services use among men and women in the United States: a cross-sectional study. *BMC Public Health* 2016 Aug 17;16(1):807 [FREE Full text] [doi: [10.1186/s12889-016-3467-1](https://doi.org/10.1186/s12889-016-3467-1)] [Medline: [27534616](https://pubmed.ncbi.nlm.nih.gov/27534616/)]
46. Pew Research Center. 2016. Social media update URL: <http://www.pewinternet.org/2016/11/11/social-media-update-2016/> [accessed 2017-02-15] [WebCite Cache ID [60IgtKKOC](https://www.webcitation.org/60IgtKKOC)]
47. Zlotorzynska M, Sullivan P, Sanchez T. The annual American men's internet survey of behaviors of men who have sex with men in the United States: 2015 key indicators report. *JMIR Public Health Surveill* 2017 Mar 25;3(1):e13 [FREE Full text] [doi: [10.2196/publichealth.7119](https://doi.org/10.2196/publichealth.7119)] [Medline: [28356240](https://pubmed.ncbi.nlm.nih.gov/28356240/)]
48. Teitcher JE, Bockting WO, Bauermeister JA, Hofer CJ, Miner MH, Klitzman RL. Detecting, preventing, and responding to “fraudsters” in internet research: ethics and tradeoffs. *J Law Med Ethics* 2015;43(1):116-133 [FREE Full text] [doi: [10.1111/jlme.12200](https://doi.org/10.1111/jlme.12200)] [Medline: [25846043](https://pubmed.ncbi.nlm.nih.gov/25846043/)]

Abbreviations

ACIP: Advisory Committee on Immunization Practices
HPV: human papillomavirus
CHCR: Center for Health Communications Research
IRR: incidence rate ratio
LGBT: lesbian, gay, bisexual, and transgender
STD: sexually transmitted disease
UCTR: unique click-through rate

Edited by G Eysenbach; submitted 23.02.17; peer-reviewed by C Wheldon, H Fontenot, M Zlotorzynska; comments to author 12.04.17; revised version received 20.04.17; accepted 25.04.17; published 02.06.17.

Please cite as:

Reiter PL, Katz ML, Bauermeister JA, Shoben AB, Paskett ED, McRee AL

Recruiting Young Gay and Bisexual Men for a Human Papillomavirus Vaccination Intervention Through Social Media: The Effects of Advertisement Content

JMIR Public Health Surveill 2017;3(2):e33

URL: <http://publichealth.jmir.org/2017/2/e33/>

doi: [10.2196/publichealth.7545](https://doi.org/10.2196/publichealth.7545)

PMID: [28576758](https://pubmed.ncbi.nlm.nih.gov/28576758/)

©Paul L Reiter, Mira L Katz, Jose A Bauermeister, Abigail B Shoben, Electra D Paskett, Annie-Laurie McRee. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 02.06.2017. 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

Perceptions of HIV Seriousness, Risk, and Threat Among Online Samples of HIV-Negative Men Who Have Sex With Men in Seven Countries

Anna N Chard^{1*}, MPH; Nicholas Metheny^{2,3*}, RN, MPH; Rob Stephenson^{2,3*}, PhD

¹Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, United States

²Department of Health Behavior and Biological Sciences, School of Nursing, University of Michigan, Ann Arbor, MI, United States

³Center for Sexuality and Health Disparities, School of Nursing, University of Michigan, Ann Arbor, MI, United States

* all authors contributed equally

Corresponding Author:

Rob Stephenson, PhD

Center for Sexuality and Health Disparities

School of Nursing

University of Michigan

Room 2236

400 North Ingalls

Ann Arbor, MI,

United States

Phone: 1 734 615 0149

Email: rbsteph@umich.edu

Abstract

Background: Rates of new HIV infections continue to increase worldwide among men who have sex with men (MSM). Despite effective prevention strategies such as condoms and pre-exposure prophylaxis (PrEP), low usage of both methods in many parts of the world hinder prevention efforts. An individual's perceptions of the risk of acquiring HIV and the seriousness they afford to seroconversion are important drivers of behavioral risk-taking. Understanding the behavioral factors suppressing the uptake of HIV prevention services is a critical step in informing strategies to improve interventions to combat the ongoing HIV pandemic among MSM.

Objective: The study aimed to examine cross-national perceptions of HIV/AIDS seriousness, risk, and threat and the association between these perceptions and sociodemographic characteristics, relationships, and high-risk sexual behaviors among MSM.

Methods: Participants in Australia, Brazil, Canada, Thailand, South Africa, the United Kingdom, and the United States were recruited for a self-administered survey via Facebook (N=1908). Respondents were asked to rate their perceived seriousness from 1 (not at all serious) to 5 (very serious) of contracting HIV, their perceived risk from 1 (no risk) to 10 (very high risk) of contracting HIV based on their current behavior, and their perception of the threat of HIV—measured as their confidence in being able to stay HIV-negative throughout their lifetimes—on a scale from 1 (will not have HIV by the end of his lifetime) to 5 (will have HIV by the end of his lifetime). Covariates included sociodemographic factors, sexual behavior, HIV testing, drug use, and relationship status. Three ordered logistic regression models, one for each outcome variable, were fit for each country.

Results: Contracting HIV was perceived as serious (mean=4.1-4.6), but perceptions of HIV risk (mean=2.7-3.8) and threat of HIV (mean=1.7-2.2) were relatively low across countries. Older age was associated with significantly lower perceived seriousness of acquiring HIV in five countries (Australia: odds ratio, OR 0.97, 95% CI 0.94-0.99; Brazil: OR 0.95, 95% CI 0.91-0.98; Canada: OR 0.96, 95% CI 0.93-0.98; South Africa: OR 0.96, 95% CI 0.94-0.98; United Kingdom: OR 0.95, 95% CI 0.92-0.98). Being in a male-male sexual relationship was associated with significantly lower perceived risk of HIV in four countries (Australia: OR 0.47, 95% CI 0.30-0.75; Canada: OR: 0.54, 95% CI 0.35-0.86; United Kingdom: OR 0.38, 95% CI 0.24-0.60; United States: OR 0.5, 95% CI 0.31-0.82). Drug use in the previous year was associated with greater threat of contracting HIV in two countries (Canada: OR 1.81, 95% CI 1.13-2.91; United Kingdom: OR 1.7, 95% CI 1.06-2.74).

Conclusions: Few measures of behavioral or sexual risk-taking were significantly associated with perceived HIV seriousness, risk, or threat across countries. Overall, low levels of reported risk were identified, and results illustrate important gaps in the understanding of risk among MSM across societies that could be addressed through culturally-tailored prevention messaging.

KEYWORDS

sexual minorities; HIV; risk; perception

Introduction

Although new HIV infections are stable or decreasing among other risk groups, incident infections continue to increase among men who have sex with men (MSM) in the United States [1], Western Europe, and Australia [2]. In low and middle-income countries, MSM have 19-fold greater odds of testing HIV sero-positive compared with the general adult male population [3]. The quality and coverage of HIV prevention services varies dramatically across the world. In many low- and middle-income countries, lack of MSM-centric HIV prevention programs, low rates of HIV knowledge, and low rates of condom use exist among MSM [4,5]. In these countries, low prevention method uptake is driven largely by societal stigma, homophobia, and constrained resources for MSM outreach [4]. In many resource-rich countries, the inverse is often a cause for low prevention uptake among MSM. Prevention fatigue, or safer sex fatigue, refers to the attitude that HIV prevention messaging and programs have become tiresome because of frequent and targeted interventions and social marketing campaigns regarding safe-sex behaviors since the beginning of the pandemic [6,7]. In both resource-rich and resource-poor settings, the advent of treatment as prevention (TasP) and pre-exposure prophylaxis (PrEP) provide a promising biomedical method for preventing incident HIV infections. Yet, large numbers of MSM in both resource-rich and resource-poor settings are unaware of their sero-status [8,9], and PrEP usage rates remain low, with studies reporting use at less than 3% among MSM in samples across resource-rich and resource-poor settings [10,11]. However, with a significant scale-up in the use of these prevention methods, incident rates could fall dramatically [12,13], with one model showing a 70% reduction in new HIV cases when 90% of those who are HIV-positive are virally suppressed and PrEP usage is at 65% among those who are eligible [14]. Understanding the behavioral factors suppressing the uptake of HIV prevention services is a critical step in informing strategies to scale up promising biomedical interventions.

Several theories focused on understanding HIV-related risk behaviors such as the Health Belief Model and the Theory of Planned Behavior, which assert that perceptions of HIV susceptibility severity drive HIV-negative individuals' motivations to use HIV prevention methods [15]. Therefore, in order to increase HIV prevention uptake, accurate assessment of HIV risk is important. Although recent work has found that many MSM underestimate their risk of contracting HIV [16-21], there is a scarcity of research addressing what drives perceptions of HIV seriousness, risk, and threat among HIV-negative MSM. This study fills gaps in this knowledge in two key ways. First, the authors found no previous studies in which MSM quantify their perceptions of HIV seriousness, risk, or threat, providing comparable, standardized data across multiple countries. Second, this is the first study to examine cross-national perceptions of HIV and AIDS seriousness, risk, and threat, and the association

between these perceptions and sociodemographic characteristics, relationships, and high-risk sexual behaviors among MSM. Results from this study may be influential in informing HIV prevention uptake strategies for HIV-negative MSM internationally.

Methods

Data

Participants were recruited online for a self-administered survey. Banner ads were placed on Facebook and displayed to men who indicated an interest in men only on their profiles and reported living in Australia, Brazil, Canada, South Africa, Thailand, the United Kingdom, or the United States. These countries were selected because of their large populations of self-identified gay and bisexual men on Facebook and to allow comparisons across socioeconomic and cultural contexts. Clicking on the advertisement led potential participants to information regarding the survey. After obtaining electronic informed consent, participants were presented with the eligibility screener that affirmed that the participants were born male, were over 18 years old, and self-reported anal or oral sex with a man in the previous year. Except for the Brazilian and Thai surveys, which were administered in Portuguese and Thai, respectively, all surveys were administered in English. Ads were displayed until 500 men in each country were recruited (range= 5 days (United States) to 14 days (Thailand)). The study was approved by the Institutional Review Board (IRB determination IRB00047677).

A total of 11,850 people across the seven sample countries clicked on the advertisement and were exposed to the eligibility screener. Of these, 58.01% (6874/11,850) began the eligibility screener; 13.09% (1551/11,850) did not meet eligibility criteria and were disqualified from completion. Among the 5323 eligible participants, 54.27% (2889/5323) men completed the survey. Of the men who completed the survey, 6.44% (186/2889) reported being HIV-positive and 27.52% (795/2889) had never been tested for HIV. Thus, they did not meet the criteria for this analysis of being known to be HIV-negative. A final convenience sample of 1908 participants who met all eligibility criteria provided data for all covariates of interest and was included in the analysis.

The anonymous survey took approximately 30 minutes to complete and collected information on participants' sociodemographic characteristics (including age; race, ethnicity, or ancestry; and level of education). Age of sexual debut with a man was self-reported as the age at which the respondent first had sex (type of sex not defined) with a man; drug use was assessed by asking whether respondents had used any nonprescription drugs in the previous 12 months. Relationship status was determined by asking respondents whether they were currently in a sexual relationship and whether that relationship was with a man. Participants self-reported whether they had ever had an HIV test; respondents answering affirmatively were

asked the date of their last HIV test. To evaluate the percentage of all condomless anal intercourse (CAI) engaged in by respondents in the previous year and the percentage of lifetime anal sex partners encountered in the previous year, respondents were asked the numbers of male anal sex partners and condomless male anal sex partners they had had intercourse with in the past 12 months and in their lifetimes.

The outcome variables were developed specifically for this survey from a review of the literature that highlighted three distinct domains of risk perception: perception of the seriousness of HIV, perception of the risk of acquiring HIV based on current behavior, and perception of the ability to remain HIV negative. The authors then developed simple scale items designed to address each of these domains. To measure perceptions of HIV seriousness, respondents were asked to rate their perception of the seriousness of contracting HIV (“How serious for you would it be if you contracted HIV?”) on a scale from 1 (not at all serious) to 5 (very serious), their perception of the risk of contracting HIV based on their current behavior (“How would you rate your risk for contracting HIV based on your current behavior?”) on a scale from 1 (no risk) to 10 (very high risk), and their perception of the threat of HIV (“How confident are you that you can stay HIV-negative in your lifetime?”) on a scale from 1 (will not have HIV by the end of his lifetime) to 5 (will have HIV by the end of his lifetime). These three questions provide three outcome variables measuring (1) the perceived seriousness of contracting HIV, (2) the perceived risk of contracting HIV, and (3) the perceived threat of HIV. Respondents were not compensated for completing the survey.

Analysis

Data were cleaned and analyzed using Stata version 12 (StataCorp, 2012). As respondents were asked to rate their perceived risk, seriousness, and threat of contracting HIV in the future, only participants reporting having ever been tested for HIV and having had a known HIV-negative status were included in the analysis. The key sociodemographic covariates included age; education, which was dichotomized as receiving a secondary school education or less (≤ 12 years) or receiving higher education (> 12 years); race, categorized as nonminority race or ethnicity (Australia: European, Brazil: white, Canada: European, South Africa: black, Thailand: Thai, United Kingdom: white, and United States: white) or minority race or ethnicity (Australia: other; Brazil: mixed, brown, or Other; Canada: Other; South Africa: white or other; Thailand: other; United Kingdom: black or other; United States: black or other), and drug use in the previous 12 months. Currently being in a sexual relationship was dichotomized as being in a relationship with a man versus being single.

Covariates measuring sexual behaviors included age of sexual debut with a man; percentage of condomless versus all anal intercourse partners encountered in the previous year, which was calculated by dividing respondents' reported number of CAI partners in the previous year by the reported number of all anal intercourse partners in the previous year; and percent of lifetime anal sex partners encountered in the previous year,

which was calculated by dividing the reported number of anal sex partners in the previous year by the total number of lifetime anal sex partners. Time since respondents' last HIV test was calculated in years by subtracting the month and year of survey completion from the reported month and year of respondents' last HIV test. Given the ordinal outcomes of perceived HIV seriousness, risk, and threat, the analysis employed ordered logistic regression models. In order to make cross-country comparisons of the outcome variables, separate models were fitted for each of the three outcomes in each of the seven countries. This required separate datasets for each country, fitting three models per dataset.

Results

Demographic Factors

Demographic characteristics of the sample, respondents' sexual behaviors, time since respondents' last HIV test, and perceived seriousness, risk, and threat of HIV seroconversion are summarized in [Table 1](#). The mean age of respondents ranged from 25.8 (Brazil) to 33.9 (Canada). The majority of respondents reported having 12 years or more of education, belonging to a majority race (except in South Africa), currently being in a sexual relationship with a man, and no drug use in the previous 12 months.

Across all the countries, between 50.2%-63.8% of respondents' anal sex partners in the previous year were condomless, and less than 1/3 of respondents' anal intercourse partners had been encountered in the previous year. Time since respondents' most recent HIV test ranged from 1.2 (South Africa) to 2.3 (Thailand and the United Kingdom) years.

Country-specific mean perceived seriousness of contracting HIV ranged from 4.1 (SD=0.1) in Thailand to 4.7 (SD=0.9) in the United States (see [Table 2](#)), mean perceived risk of contracting HIV ranged from 2.7 (SD=2.0) in Canada to 3.8 (SD=0.2) in Thailand (see [Table 3](#)), and mean perceived threat of HIV ranged from 1.7 (SD=0.9) in Canada and the United Kingdom to 2.2 (SD=0.1) in Thailand (see [Table 4](#)).

Respondent age was negatively associated with perceived HIV seriousness in all countries except Thailand and the United States. Age was negatively associated with perceived risk in the United States. There was a significant, positive association between education and perceived HIV seriousness, risk, and threat in two countries. Compared to respondents with less than or equal to 12 years of education, respondents with more than 12 years of education had significantly greater odds of perceiving a higher HIV seriousness in the United Kingdom and a higher HIV risk and threat in Canada. Moreover, in Canada, respondents belonging to minority races had significantly lower odds of perceiving a higher HIV seriousness than respondents belonging to nonminority races. South African respondents belonging to a minority race had significantly lower odds of perceiving a higher HIV risk and HIV threat than those belonging to a nonminority race.

Table 1. Respondent characteristics and mean minority stress scale index scores of 1891 internet-recruited men who have sex with men in seven countries.

Variable	Australia (n=274), μ (%)	Brazil (n=294), μ (%)	Canada (n=274), μ (%)	South Africa (n=386), μ (%)	Thailand (n=146), μ (%)	United Kingdom (n=280), μ (%)	United States (n=254), μ (%)
Age in years	30.4 (11.5)	25.8 (7.7)	33.9 (12.6)	33.2 (10.3)	32.3 (7.8)	30.6 (11.2)	30.9 (13.4)
Education							
≤12 years	103 (37.6)	97 (33.0)	64 (23.4)	118 (30.6)	22 (15.1)	62 (22.1)	71 (28.0)
>12 years	171 (62.4)	197 (67.0)	210 (76.6)	268 (69.4)	124 (84.9)	218 (77.9)	183 (72.1)
Race/Ethnicity^a							
Nonminority	168 (61.3)	189 (64.3)	219 (79.9)	56 (14.5)	138 (94.5)	266 (95.0)	207 (81.5)
Minority	106 (38.7)	105 (35.7)	55 (20.1)	330 (85.5)	8 (5.5)	14 (5.0)	47 (18.5)
Drug use in previous 12 months							
No	153 (55.8)	203 (69.1)	145 (52.9)	233 (60.4)	112 (76.7)	179 (63.9)	162 (63.8)
Yes	121 (44.2)	91 (30.9)	129 (47.1)	153 (39.6)	34 (23.3)	101 (36.1)	92 (36.2)
Relationship Status							
Single	130 (47.5)	150 (51.0)	107 (39.1)	165 (42.8)	60 (41.1)	122 (43.6)	118 (46.5)
In a relationship	144 (52.6)	144 (49.0)	167 (60.9)	221 (57.3)	86 (58.9)	158 (56.4)	136 (53.5)
Age of sexual debut with a man	17.5 (4.7)	15.2 (3.9)	17.4 (5.0)	17.1 (4.6)	17.9 (4.4)	17.5 (4.6)	17.1 (4.8)
Percentage of unprotected versus all anal intercourse partners in previous year	54.7 (42.8)	50.2 (41.2)	54.4 (43.9)	63.8 (42.5)	54.0 (43.7)	55.9 (43.2)	61.1 (43.7)
Percentage of lifetime anal sex partners encountered in previous year	26.5 (29.5)	30.1 (29.3)	25.6 (29.3)	26.3 (25.7)	35.0 (33.6)	28.1 (29.3)	31.8 (34.0)
Years since most recent HIV test	1.4 (2.5)	1.4 (1.9)	1.7 (2.6)	1.2 (1.9)	2.3 (2.7)	2.3 (4.1)	1.5 (2.7)
Perceived seriousness of contracting HIV (1-10)	4.6 (0.9)	4.6 (0.9)	4.6 (1.0)	4.3 (1.1)	4.1 (0.1)	4.6 (0.9)	4.7 (0.9)
Perceived risk of contracting HIV (1-5)	2.9 (2.1)	3.2 (2.4)	2.7 (2.1)	3.0 (2.4)	3.8 (0.2)	2.8 (2.1)	3.0 (2.3)
Perceived inability to stay HIV-negative throughout life (1-10)	1.8 (1.0)	2.1 (1.1)	1.7 (0.9)	1.8 (1.0)	2.2 (0.1)	1.7 (0.9)	1.8 (1.0)

^aNon-minority race/ethnicity was categorized as follows: Australia: European, Brazil: white, Canada: European, South Africa: black, Thailand: Thai, United Kingdom: white, United States: white. Minority race/ethnicity was categorized as follows: Australia: Other; Brazil: Mixed, brown, and Other; Canada: Other; South Africa: white and Other; Thailand: Other; United Kingdom: black and Other; United States: black and Other.

Table 2. Adjusted Odds Ratios and 95% CIs between perceived HIV seriousness and respondent characteristics.

Variable	Australia (n=274), aOR (95% CI)	Brazil (n=294), aOR (95% CI)	Canada (n=274), aOR (95% CI)	South Africa (n=386), aOR (95% CI)	Thailand (n=146), aOR (95% CI)	United Kingdom (n=280), aOR (95% CI)	United States (n=254), aOR (95% CI)
Age in years	0.97 (0.94-0.99) ^a	0.95 (0.91-0.98) ^a	0.96 (0.93-0.98) ^a	0.96 (0.94-0.98) ^a	0.97 (0.93-1.01)	0.95 (0.92-0.98) ^a	0.98 (0.95-1.01)
≤12 years of education (ref: >12 years)	1.21 (0.65-2.28)	0.95 (0.53-1.71)	0.64 (0.29-1.43)	0.85 (0.52-1.39)	1.84 (0.71-4.77)	1.94 (1.02-3.71) ^a	0.94 (0.45-1.92)
Minority race (ref: nonminority race)	1.11 (0.59-2.09)	1.15 (0.65-2.02)	0.4 (0.19-0.82) ^a	1.3 (0.67-2.51)	0.4 (0.09-1.72)	1.3 (0.34-4.97)	0.79 (0.34-1.83)
Drug use in previous 12 months (ref: no drug use)	0.88 (0.47-1.66)	0.84 (0.48-1.45)	1.44 (0.76-2.72)	1.07 (0.68-1.69)	0.9 (0.41-1.97)	0.49 (0.28-0.87) ^a	0.86 (0.44-1.67)
In a relationship (ref: single)	1.17 (0.61-2.24)	0.8 (0.47-1.34)	1.03 (0.52-2.02)	0.64 (0.4-1.03)	0.48 (0.24-0.98) ^a	0.84 (0.47-1.51)	1.26 (0.62-2.56)
Age of sexual debut with a man	1.08 (1.01-1.16) ^a	1.02 (0.95-1.08)	1.09 (1.02, 1.17) ^a	1.02 (0.97-1.07)	1.01 (0.93-1.1)	1.04 (0.98-1.11)	0.98 (0.92-1.04)
Percentage of unprotected versus all anal intercourse partners in previous year	0.87 (0.4-1.88)	0.92 (0.49-1.75)	0.93 (0.44-1.96)	1.76 (1.02-3.05) ^a	1.65 (0.74-3.7)	1.5 (0.76-2.94)	0.63 (0.28-1.45)
Percentage of lifetime anal sex partners encountered in previous year	1.8 (0.45-7.25)	0.76 (0.32-1.8)	4.38 (0.85-22.56)	1.33 (0.5-3.58)	1.73 (0.57-5.28)	1.62 (0.47-5.56)	1.3 (0.41-4.13)
Years since most recent HIV test	1 (0.9-1.12)	1.08 (0.94-1.24)	1.04 (0.91-1.18)	0.96 (0.87-1.07)	0.97 (0.87-1.09)	1.08 (1.01-1.15) ^a	0.93 (0.83-1.05)

^aStatistical significance, alpha=.05.

Table 3. Adjusted Odds Ratios and 95% CIs between perceived HIV risk and respondent characteristics.

Variable	Australia (n=274), aOR (95% CI)	Brazil (n=294), aOR (95% CI)	Canada (n=274), aOR (95% CI)	South Africa (n=386), aOR (95% CI)	Thailand (n=146), aOR (95% CI)	United Kingdom (n=280), aOR (95% CI)	United States (n=254), aOR (95% CI)
Age in years	0.98 (0.96-1)	1 (0.97-1.03)	1 (0.98, 1.02)	1.01 (0.99-1.03)	0.97 (0.94-1.01)	1 (0.97-1.03)	0.97 (0.95-0.99) ^a
≤12 years of education (ref: >12 years)	0.75 (0.48-1.18)	1.17 (0.74-1.84)	1.83 (1.06, 3.16) ^a	1.3 (0.87-1.96)	0.77 (0.34-1.73)	1.31 (0.74-2.33)	1.55 (0.92-2.62)
Minority race (ref: nonminority race)	1.24 (0.78-1.95)	1.18 (0.76-1.81)	1.37 (0.79, 2.39)	0.53 (0.3-0.91) ^a	0.49 (0.14-1.71)	0.69 (0.26-1.85)	0.75 (0.4-1.37)
Drug use in previous 12 months (ref: no drug use)	1.36 (0.86-2.14)	1.39 (0.89-2.16)	2.6 (1.67, 4.05) ^a	1.78 (1.22-2.58) ^a	1.77 (0.88-3.55)	2.23 (1.41-3.52) ^a	1.85 (1.16-2.95) ^a
In a relationship (ref: single)	0.47 (0.3-0.75) ^a	0.73 (0.48-1.11)	0.54 (0.35, 0.86) ^a	0.75 (0.51-1.1)	1.31 (0.7-2.42)	0.38 (0.24-0.6) ^a	0.5 (0.31-0.82) ^a
Age of sexual debut with a man	0.95 (0.9-0.99) ^a	0.94 (0.9-1) ^a	1 (0.96, 1.05)	0.99 (0.95-1.03)	1.06 (0.98-1.15)	1 (0.95-1.05)	1 (0.95-1.05)
Percentage of unprotected versus all anal intercourse partners in previous year	0.65 (0.37-1.12)	0.5 (0.3-0.84) ^a	0.9 (0.53, 1.53)	0.69 (0.44-1.1)	0.98 (0.48-2.01)	0.67 (0.4-1.12)	1.07 (0.6-1.89)
Percentage of lifetime anal sex partners encountered in previous year	1.05 (0.45-2.43)	1.33 (0.67-2.66)	1.28 (0.56, 2.92)	1.99 (0.91-4.34)	0.86 (0.35-2.09)	1.65 (0.71-3.8)	1.04 (0.48-2.27)
Years since most recent HIV test	0.93 (0.83-1.04)	1.06 (0.95-1.19)	0.95 (0.87, 1.05)	0.94 (0.86-1.04)	1.18 (1.05-1.31) ^a	0.92 (0.87-0.98) ^a	1 (0.9-1.1)

^aStatistical significance at alpha=.05.

Behavioral Factors

There was a significant negative association between drug use and perceived HIV seriousness in the United Kingdom. Conversely, there was a significant positive association between drug use and perceived HIV risk in Canada, South Africa, the United Kingdom, and the United States and between drug use and threat of HIV in Canada and the United Kingdom.

Reporting a current sexual relationship with a man was negatively associated with perceived HIV seriousness (Thailand), risk (Australia, Canada, the United Kingdom, and the United States), and threat (Australia) in several countries. As compared with those not in a male-male relationship, respondents reporting currently being in a sexual relationship had significantly lower odds of perceiving a higher HIV seriousness in Thailand; significantly lower odds of perceiving a higher risk of HIV in Australia, Brazil, Canada, the United Kingdom, and the United States; and significantly lower odds of perceiving a higher threat of HIV in Australia.

Age of sexual debut with a man was positively associated with perceived seriousness of HIV in two countries and negatively

associated with perceived risk of HIV in two countries. As age of sexual debut with a man increased, the odds of perceiving a higher seriousness of HIV increased significantly in Australia and Canada, while the odds of perceiving a higher HIV risk significantly decreased in Australia and Brazil. As the percentage of condomless anal intercourse partners in the previous year increased, the odds of perceiving a higher HIV risk significantly decreased in Brazil and the odds of perceiving a higher HIV threat significantly decreased in Brazil and South Africa. The percent of lifetime anal sex partners encountered in the previous year was associated with only one outcome (HIV threat) in one country. In the United States, the odds of perceiving a higher HIV threat significantly decreased as the percentage of lifetime anal sex partners encountered in the previous year increased. Time since last HIV test was positively associated with perceived HIV seriousness in the United Kingdom. The association between time since last HIV test and perceived HIV risk yielded mixed results. As time since respondents' last HIV test increased, odds of perceiving a higher HIV risk increased significantly among participants in Thailand, but decreased significantly among participants in the United Kingdom.

Table 4. Adjusted Odds Ratios and 95% CIs between perceived HIV threat and respondent characteristics.

Variable	Australia (n=274), aOR (95% CI)	Brazil (n=294), aOR (95% CI)	Canada (n=274), aOR (95% CI)	South Africa (n=386), aOR (95% CI)	Thailand (n=146), aOR (95% CI)	United Kingdom (n=280), aOR (95% CI)	United States (n=254), aOR (95% CI)
Age in years	0.99 (0.96-1.01)	1.02 (0.98-1.05)	1 (0.98-1.03)	1.01 (0.99-1.03)	0.99 (0.95-1.03)	1.02 (0.99-1.04)	0.99 (0.97-1.01)
≤12 years of education (ref: >12 years)	0.87 (0.54-1.42)	1.29 (0.81-2.05)	2.27 (1.23-4.19) ^a	1.37 (0.89-2.1)	0.46 (0.19-1.11)	1.06 (0.58-1.94)	1.31 (0.75-2.27)
Minority race (ref: nonminority race)	1.2 (0.74-1.94)	0.74 (0.47-1.16)	0.83 (0.45-1.54)	0.53 (0.3-0.95) ^a	2.17 (0.59-7.88)	1.06 (0.37-3.01)	1.55 (0.82-2.92)
Drug use in previous 12 months (ref: no drug use)	1.42 (0.87-2.31)	1.33 (0.84-2.12)	1.81 (1.13-2.91) ^a	1.4 (0.94-2.08)	1.22 (0.6-2.51)	1.7 (1.06-2.74) ^a	0.9 (0.55-1.48)
In a relationship (ref: single)	0.55 (0.34-0.89) ^a	0.75 (0.49-1.16)	0.84 (0.51-1.38)	0.94 (0.63-1.4)	1.15 (0.6-2.2)	0.7 (0.44-1.11)	0.78 (0.46-1.32)
Age of sexual debut with a man	0.98 (0.93-1.04)	0.98 (0.93-1.03)	0.98 (0.93-1.03)	1 (0.96-1.04)	1.05 (0.97-1.13)	1 (0.95-1.05)	1 (0.95-1.06)
Percentage of unprotected versus all anal intercourse partners in previous year	0.74 (0.41-1.34)	0.54 (0.32-0.93) ^a	0.89 (0.5-1.58)	0.52 (0.32-0.84) ^a	1.4 (0.67-2.94)	0.81 (0.47-1.41)	0.55 (0.3-1.02)
Percentage of lifetime anal sex partners encountered in previous year	0.65 (0.26-1.61)	1.21 (0.59-2.49)	1.11 (0.43-2.87)	1.2 (0.52-2.76)	1.17 (0.45-2.99)	0.96 (0.38-2.42)	0.36 (0.15-0.87) ^a
Years since most recent HIV test	1.03 (0.92-1.15)	1 (0.89-1.13)	0.98 (0.89-1.08)	0.97 (0.88-1.07)	1.08 (0.97-1.21)	0.94 (0.88-1.01)	0.92 (0.82-1.03)

^aStatistical significance at alpha=.05.

Discussion

Principal Findings

The results of this study point to the potential influence of drug use, relationships, and age on individual reporting of perceived HIV seriousness, risk of HIV seroconversion, and lifetime threat of HIV seroconversion among MSM across culturally and economically diverse settings. While some cross-national trends were found, the variations across countries point to the importance of cultural context in understanding the drivers of perceived risk of HIV.

Respondents reporting currently being in a sexual relationship with a man perceived a lower risk of contracting HIV compared to respondents not in a relationship. This may be the result of messaging to MSM regarding HIV risk. Since the beginning of the epidemic, HIV prevention messaging has focused on the risks associated with having multiple concurrent sexual partners. However, recent evidence suggests that between one-third [22] to two-thirds [23] of incident HIV infections among MSM result from a main partner, a pattern driven by more frequent anal intercourse and more infrequent condom use between main partners relative to casual partners [23,24]. Recent research also shows that US MSM with main partners are less likely to test regularly for HIV than single MSM [25]. Central to understanding the differences in HIV risk behaviors between single and partnered men is the idea that partnered men perceive being in a relationship as protective, regardless of their outside sexual behavior or their knowledge of their partners' serostatus

[25]. The findings of this study reinforce this idea: partnered men in a range of social and cultural contexts perceived themselves to be at less risk of HIV acquisition. The models did not control for the presence of a sexual agreement in the relationship: it may be that the partners are monogamous and have knowledge of each other's serostatus and are thus correct in their estimation of a lowered risk of HIV. Despite this, the results suggest the need for ongoing research attention to understand how partnered men perceived their HIV risks and the extent to which this varies by context, as well as greater programmatic attention to the unique needs of partnered men. For example, Couples HIV Counseling and Testing (CHCT) allows partnered men to examine HIV prevention as a dyad [26-28].

Older respondents reported a significantly lower perceived seriousness (five countries) and risk (US) of contracting HIV or AIDS than younger respondents. These results may reflect some degree of treatment optimism; largely as a result of highly active anti-retroviral treatment (HAART), HIV-positive MSM are living longer than they were twenty years ago. Although the sample is of HIV-negative men, these men may have observed the changes in the impact of HIV on those living with it, —and the transformation of HIV into manageable chronic condition. This may allow them to perceive the threat of HIV as less serious now than in the past. The prevalence of HIV and AIDS among older MSM is higher than among young MSM in many places [29,30], although incidence is increasing among younger MSM. Several studies have identified having older MSM sex partners to be an independent risk factor for HIV

transmission [30,31] because older MSM have had more lifetime sexual partners and sexual encounters, and thus more opportunities for HIV seroconversion. Our findings suggest that decreased perceptions of the seriousness and risk of HIV acquisition might be an issue for older MSM in the countries sampled and point to the need for HIV prevention messaging and services to be adapted to specifically reach older MSM who may be underestimating their HIV risk and need for prevention.

Few measures of recent sexual behavior were significantly associated with perceived HIV seriousness, risk, or threat across multiple countries. An older age of sexual debut with a man was associated with a higher perceived HIV seriousness (Australia and Canada) and lower perceived HIV risk (Australia and Brazil). It is possible that a higher perceived seriousness of HIV actually delayed participants' sexual debut. An earlier age of sexual debut has been associated with engaging in sexual risk behaviours and positive HIV status [32,33]. We have observed the inverse; perhaps a later age of sexual debut may in turn be associated with more conservative sexual behaviors or fewer sexual partners, leading to participants' perception of a lower risk of contracting HIV. Additionally, despite a higher perceived seriousness of HIV among men with a greater percentage of CAI partners in the previous year in South Africa, perceived risk (Australia and Brazil) and threat (Brazil and South Africa) were lower among participants in some countries. The discordance between sexual risk taking and perceptions of HIV seriousness, risk, and threat may reflect some degree of cognitive dissonance or HIV fatalism among these participants [34,35]. The knowledge of treatment as prevention (TasP) may be another reason for a reduced risk perception among men with a higher percentage of CAI, with the assumption that an undetectable viral load equates to an inability to transmit HIV. While being on PrEP or engaging in CAI with a partner on PrEP may also reduce risk perception, the survey was administered before the widespread availability of PrEP in these countries. Therefore, its use does not likely factor into decreased HIV risk and threat perceptions. In contrast, in multiple countries, participants reporting drug use in the previous year perceived their risk and threat of HIV to be higher than non-drug using respondents. It is possible that MSM using drugs identify the increased risks of HIV acquisition associated with drug use (through needle sharing or drug-linked sexual risk-taking). It is also possible that drug use is a maladaptive response to social pressure and discrimination, which acts to increase both the use of drugs and the perception of increased risk of HIV acquisition. In some contexts, such as environments where homosexuality is highly stigmatized, drug use may be a strategy to mitigate stress associated with internalized homonegativity and to normalize same-sex thoughts, feelings, and behaviours and connect with others in the gay community [36-38].

Limitations

There are limitations to the present study, most of which result from the Web-based convenience sampling design. In all countries, the survey was advertised only to men who were registered users of Facebook and had a profile indicating an interest in men, leading to selection bias towards men who may be more public about their sexual identity or behavior. A recent study reported that Web-based recruitment and venue-based recruitment yields similar samples of MSM in the United States [39]; however, whether this holds true in international settings is yet to be established. An additional limitation is that a significant proportion of those who clicked on the banner ads did not complete the survey; we do not have data on their characteristics to establish the extent of this selectivity bias. The use of a one-year recall period may not be ideal for assessing the number of sex partners and could lead to biased responses. Another limitation is the lack of a country-specific context to this study. While we focus on the broad, international patterns of HIV seriousness, risk, and threat in this publication, future research at the country level should include a discussion of the cultural, political, and historical factors that may affect these perceptions among MSM. Despite these limitations, however, this study demonstrated the usefulness of a Web-based survey tool in reaching MSM across countries and collecting standardized data across economically and culturally-diverse settings with large samples of men who have sex with men.

Conclusions

Perceived risks of HIV infection were generally low across MSM in this international sample. Across countries, older MSM and those in sexual relationships perceived themselves to have lower risk and threat of HIV, while drug using MSM perceived their HIV risk as higher. Few measures of sexual behavior were associated with perceptions of HIV seriousness, risk, or threat. Our results point to the need for more nuanced research into the constructs of risk perception among MSM. Currently, little is known about how risk perception varies with age, stage of life, or the availability of new biomedical preventative options such as PrEP. Among MSM internationally, varying sociopolitical climates and the relative strength of structural minority stressors such as heteronormative social pressure and homophobic discrimination may also have important effects. A more cogent understanding of the constructs associated with HIV risk perception may lead to more tailored and effective HIV prevention interventions for MSM. The results also suggest that, while existing HIV prevention messages continue to focus on the HIV risks of condomless sex, such messages may need to be packaged differently for men in relationships and for older men. Additionally, novel strategies to help MSM understand the true risks of unsafe sex and to surmount the effects of behavioral discordance are still needed.

Conflicts of Interest

None declared.

References

1. CDC. 2015. HIV in the United States: at a glance internet URL: <https://www.cdc.gov/hiv/statistics/overview/ataglance.html> [accessed 2017-06-05] [WebCite Cache ID 6qzzQJSBC]
2. Sullivan PS, Hamouda O, Delpech V, Geduld JE, Prejean J, Semaille C, Annecy MSM Epidemiology Study Group. Reemergence of the HIV epidemic among men who have sex with men in North America, Western Europe, and Australia, 1996-2005. *Ann Epidemiol* 2009 Jun;19(6):423-431. [doi: [10.1016/j.annepidem.2009.03.004](https://doi.org/10.1016/j.annepidem.2009.03.004)] [Medline: [19460672](https://pubmed.ncbi.nlm.nih.gov/19460672/)]
3. United Nations Population Fund, Global Forum on MSM & HIV, World Health Organization, United States Agency for International Development, World Bank. UNFPA. New York: United Nations Population Fund; 2015. Implementing comprehensive HIV and STI prevention programmes for men who have sex with men URL: <http://www.unfpa.org/publications/implementing-comprehensive-hiv-and-sti-programmes-men-who-have-sex-men> [accessed 2017-06-05] [WebCite Cache ID 6qzso2JPD]
4. Adam PC, de Wit JB, Toskin I, Mathers BM, Nashkoev M, Zablotska I, et al. Estimating levels of HIV testing, HIV prevention coverage, HIV knowledge, and condom use among men who have sex with men (MSM) in low-income and middle-income countries. *J Acquir Immune Defic Syndr* 2009 Dec;52(Suppl 2):S143-S151. [doi: [10.1097/QAI.0b013e3181baf111](https://doi.org/10.1097/QAI.0b013e3181baf111)] [Medline: [19901627](https://pubmed.ncbi.nlm.nih.gov/19901627/)]
5. Sullivan PS, Carballo-Diéguez A, Coates T, Goodreau SM, McGowan I, Sanders EJ, et al. Successes and challenges of HIV prevention in men who have sex with men. *Lancet* 2012 Jul 28;380(9839):388-399 [FREE Full text] [doi: [10.1016/S0140-6736\(12\)60955-6](https://doi.org/10.1016/S0140-6736(12)60955-6)] [Medline: [22819659](https://pubmed.ncbi.nlm.nih.gov/22819659/)]
6. Ostrow DG, Silverberg MJ, Cook RL, Chmiel JS, Johnson L, Li X, multicenter AIDS cohort study. Prospective study of attitudinal and relationship predictors of sexual risk in the multicenter AIDS cohort study. *AIDS Behav* 2008 Jan;12(1):127-138. [doi: [10.1007/s10461-007-9223-x](https://doi.org/10.1007/s10461-007-9223-x)] [Medline: [17410419](https://pubmed.ncbi.nlm.nih.gov/17410419/)]
7. Stockman JK, Schwarcz SK, Butler LM, de Jong B, Chen SY, Delgado V, et al. HIV prevention fatigue among high-risk populations in San Francisco. *J Acquir Immune Defic Syndr* 2004 Apr 01;35(4):432-434. [Medline: [15097163](https://pubmed.ncbi.nlm.nih.gov/15097163/)]
8. Castro R, Ribeiro-Alves M, Corrêa RG, Derrico M, Lemos K, Grangeiro JR, et al. The men who have sex with men HIV care cascade in Rio de Janeiro, Brazil. *PLoS One* 2016;11(6):e0157309 [FREE Full text] [doi: [10.1371/journal.pone.0157309](https://doi.org/10.1371/journal.pone.0157309)] [Medline: [27299524](https://pubmed.ncbi.nlm.nih.gov/27299524/)]
9. Kerr LR, Mota RS, Kendall C, Pinho A, Mello MB, Guimarães MD, et al. HIV among MSM in a large middle-income country. *AIDS* 2013 Jan 28;27(3):427-435. [doi: [10.1097/QAD.0b013e32835ad504](https://doi.org/10.1097/QAD.0b013e32835ad504)] [Medline: [23291540](https://pubmed.ncbi.nlm.nih.gov/23291540/)]
10. Krakower DS, Mimiaga MJ, Rosenberger JG, Novak DS, Mitty JA, White JM, et al. Limited awareness and low immediate uptake of pre-exposure prophylaxis among men who have sex with men using an internet social networking site. *PLoS One* 2012 Mar;7(3):e33119 [FREE Full text] [doi: [10.1371/journal.pone.0033119](https://doi.org/10.1371/journal.pone.0033119)] [Medline: [22470438](https://pubmed.ncbi.nlm.nih.gov/22470438/)]
11. Mayer KH, Oldenburg CE, Novak DS, Elsesser SA, Krakower DS, Mimiaga MJ. Early adopters: correlates of HIV chemoprophylaxis use in recent online samples of us men who have sex with men. *AIDS Behav* 2016 Jul;20(7):1489-1498. [doi: [10.1007/s10461-015-1237-1](https://doi.org/10.1007/s10461-015-1237-1)] [Medline: [26530863](https://pubmed.ncbi.nlm.nih.gov/26530863/)]
12. Arnold EA, Steward WT. HIV transmission in MSM: considerations for PrEP scale-up. *Lancet HIV* 2016 Feb;3(2):e62-e63. [doi: [10.1016/S2352-3018\(15\)00060-0](https://doi.org/10.1016/S2352-3018(15)00060-0)] [Medline: [26847227](https://pubmed.ncbi.nlm.nih.gov/26847227/)]
13. Punyacharoensin N, Edmunds W, De AD, Delpech V, Hart G, Elford J, et al. Effect of pre-exposure prophylaxis and combination HIV prevention for men who have sex with men in the UK: a mathematical modelling study. *Lancet HIV* 2016 Feb;3(2):e94-e104. [doi: [10.1016/S2352-3018\(15\)00056-9](https://doi.org/10.1016/S2352-3018(15)00056-9)] [Medline: [26847231](https://pubmed.ncbi.nlm.nih.gov/26847231/)]
14. Grant R. Scale-up of preexposure prophylaxis in San Francisco to impact HIV incidence. 2015 Presented at: Conference on Retroviruses and Opportunistic Infections; 2015; Seattle, Washington.
15. DiClemente R, Crosby R, Kegler M. Emerging theories in health promotion practice and research: strategies for improving public health. San Francisco: Jossey-Bass; 2002.
16. Clifton S, Nardone A, Field N, Mercer CH, Tanton C, Macdowall W, et al. HIV testing, risk perception, and behaviour in the British population. *AIDS* 2016 Mar 27;30(6):943-952 [FREE Full text] [doi: [10.1097/QAD.0000000000001006](https://doi.org/10.1097/QAD.0000000000001006)] [Medline: [26963528](https://pubmed.ncbi.nlm.nih.gov/26963528/)]
17. Darbes LA, Chakravarty D, Beougher SC, Neilands TB, Hoff CC. Partner-provided social support influences choice of risk reduction strategies in gay male couples. *AIDS Behav* 2012 Jan;16(1):159-167 [FREE Full text] [doi: [10.1007/s10461-010-9868-8](https://doi.org/10.1007/s10461-010-9868-8)] [Medline: [21221756](https://pubmed.ncbi.nlm.nih.gov/21221756/)]
18. Gallagher T, Link L, Ramos M, Bottger E, Aberg J, Daskalakis D. Self-perception of HIV risk and candidacy for pre-exposure prophylaxis among men who have sex with men testing for HIV at commercial sex venues in New York city. *LGBT Health* 2014 Sep;1(3):218-224. [doi: [10.1089/lgbt.2013.0046](https://doi.org/10.1089/lgbt.2013.0046)] [Medline: [26789715](https://pubmed.ncbi.nlm.nih.gov/26789715/)]
19. Hoffmann M, MacCarthy S, Batson A, Crawford-Roberts A, Rasanathan J, Nunn A, et al. Barriers along the care cascade of HIV-infected men in a large urban center of Brazil. *AIDS Care* 2016;28(1):57-62 [FREE Full text] [doi: [10.1080/09540121.2015.1062462](https://doi.org/10.1080/09540121.2015.1062462)] [Medline: [26291264](https://pubmed.ncbi.nlm.nih.gov/26291264/)]
20. Mitchell JW, Petroll AE. HIV testing rates and factors associated with recent HIV testing among male couples. *Sex Transm Dis* 2012 May;39(5):379-381 [FREE Full text] [doi: [10.1097/OLQ.0b013e3182479108](https://doi.org/10.1097/OLQ.0b013e3182479108)] [Medline: [22504604](https://pubmed.ncbi.nlm.nih.gov/22504604/)]
21. Stephenson R, White D, Mitchell JW. Sexual agreements and perception of HIV prevalence among an online sample of partnered men who have sex with men. *Arch Sex Behav* 2015 Oct;44(7):1813-1819. [doi: [10.1007/s10508-015-0532-2](https://doi.org/10.1007/s10508-015-0532-2)] [Medline: [26048482](https://pubmed.ncbi.nlm.nih.gov/26048482/)]

22. Goodreau S, Carnegie N, Vittinghoff E, Lama J, Sanchez J, Grinsztejn B, et al. What drives the US and Peruvian HIV epidemics in men who have sex with men (MSM)? *PLoS One* 2012;7(11):e50522 [FREE Full text] [doi: [10.1371/journal.pone.0050522](https://doi.org/10.1371/journal.pone.0050522)] [Medline: [23209768](https://pubmed.ncbi.nlm.nih.gov/23209768/)]
23. 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 01;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/)]
24. Gass K, Hoff CC, Stephenson R, Sullivan PS. Sexual agreements in the partnerships of internet-using men who have sex with men. *AIDS Care* 2012;24(10):1255-1263 [FREE Full text] [doi: [10.1080/09540121.2012.656571](https://doi.org/10.1080/09540121.2012.656571)] [Medline: [22375729](https://pubmed.ncbi.nlm.nih.gov/22375729/)]
25. Stephenson R, White D, Darbes L, Hoff C, Sullivan P. HIV testing behaviors and perceptions of risk of HIV infection among MSM with main partners. *AIDS Behav* 2015 Mar;19(3):553-560 [FREE Full text] [doi: [10.1007/s10461-014-0862-4](https://doi.org/10.1007/s10461-014-0862-4)] [Medline: [25081599](https://pubmed.ncbi.nlm.nih.gov/25081599/)]
26. Stephenson R, Rentsch C, Sullivan P. High levels of acceptability of couples-based HIV testing among MSM in South Africa. *AIDS Care* 2012;24(4):529-535 [FREE Full text] [doi: [10.1080/09540121.2011.617413](https://doi.org/10.1080/09540121.2011.617413)] [Medline: [22007940](https://pubmed.ncbi.nlm.nih.gov/22007940/)]
27. Wagenaar BH, Christiansen-Lindquist L, Khosropour C, Salazar LF, Benbow N, Prachand N, et al. Willingness of US men who have sex with men (MSM) to participate in couples HIV voluntary counseling and testing (CVCT). *PLoS One* 2012;7(8):e42953 [FREE Full text] [doi: [10.1371/journal.pone.0042953](https://doi.org/10.1371/journal.pone.0042953)] [Medline: [22905191](https://pubmed.ncbi.nlm.nih.gov/22905191/)]
28. Stephenson R, Sullivan P, Salazar L, Gratz B, Allen S, Seelbach E. Attitudes towards couples-based HIV testing among MSM in three US cities. *AIDS Behav* 2011 Apr;15(Suppl 1):S80-S87 [FREE Full text] [doi: [10.1007/s10461-011-9893-2](https://doi.org/10.1007/s10461-011-9893-2)] [Medline: [21336607](https://pubmed.ncbi.nlm.nih.gov/21336607/)]
29. Li A, Varangrat A, Wimonsate W, Chemnasiri T, Sinthuwattanawibool C, Phanuphak P, et al. Sexual behavior and risk factors for HIV infection among homosexual and bisexual men in Thailand. *AIDS Behav* 2009 Apr;13(2):318-327. [doi: [10.1007/s10461-008-9448-3](https://doi.org/10.1007/s10461-008-9448-3)] [Medline: [18758936](https://pubmed.ncbi.nlm.nih.gov/18758936/)]
30. Joseph HA, Marks G, Belcher L, Millett GA, Stueve A, Bingham TA, et al. Older partner selection, sexual risk behaviour and unrecognised HIV infection among black and Latino men who have sex with men. *Sex Transm Infect* 2011 Aug;87(5):442-447. [doi: [10.1136/sextrans-2011-050010](https://doi.org/10.1136/sextrans-2011-050010)] [Medline: [21705378](https://pubmed.ncbi.nlm.nih.gov/21705378/)]
31. Hurt CB, Matthews DD, Calabria MS, Green KA, Adimora AA, Golin CE, et al. Sex with older partners is associated with primary HIV infection among men who have sex with men in North Carolina. *J Acquir Immune Defic Syndr* 2010 Jun;54(2):185-190 [FREE Full text] [doi: [10.1097/QAI.0b013e3181c99114](https://doi.org/10.1097/QAI.0b013e3181c99114)] [Medline: [20057320](https://pubmed.ncbi.nlm.nih.gov/20057320/)]
32. Cheung DH, Suharlim C, Guadamuz TE, Lim SH, Koe S, Wei C. Earlier anal sexarche and co-occurring sexual risk are associated with current HIV-related risk behaviors among an online sample of men who have sex with men in Asia. *AIDS Behav* 2014 Dec;18(12):2423-2431 [FREE Full text] [doi: [10.1007/s10461-014-0821-0](https://doi.org/10.1007/s10461-014-0821-0)] [Medline: [24920344](https://pubmed.ncbi.nlm.nih.gov/24920344/)]
33. Nelson KM, Gamarel KE, Pantalone DW, Carey MP, Simoni JM. Sexual debut and HIV-related sexual risk-taking by birth cohort among men who have sex with men in the United States. *AIDS Behav* 2016 Oct;20(10):2286-2295. [doi: [10.1007/s10461-016-1330-0](https://doi.org/10.1007/s10461-016-1330-0)] [Medline: [26860630](https://pubmed.ncbi.nlm.nih.gov/26860630/)]
34. Frasca T, Ventuneac A, Balan I, Carballo-Diéguez A. Inner contradictions among men who bareback. *Qual Health Res* 2012 Jul;22(7):946-956 [FREE Full text] [doi: [10.1177/1049732312443592](https://doi.org/10.1177/1049732312443592)] [Medline: [22645221](https://pubmed.ncbi.nlm.nih.gov/22645221/)]
35. Nemeroff CJ, Hoyt MA, Huebner DM, Proescholdbell RJ. The cognitive escape scale: measuring HIV-related thought avoidance. *AIDS Behav* 2008 Mar;12(2):305-320. [doi: [10.1007/s10461-007-9345-1](https://doi.org/10.1007/s10461-007-9345-1)] [Medline: [18188691](https://pubmed.ncbi.nlm.nih.gov/18188691/)]
36. Bauermeister JA. It's all about connecting: reasons for drug use among Latino gay men living in the San Francisco bay area. *J Ethn Subst Abuse* 2007;6(1):109-129. [doi: [10.1300/J233v06n01_07](https://doi.org/10.1300/J233v06n01_07)] [Medline: [17430819](https://pubmed.ncbi.nlm.nih.gov/17430819/)]
37. Ha H, Risser JM, Ross MW, Huynh NT, Nguyen HT. Homosexuality-related stigma and sexual risk behaviors among men who have sex with men in Hanoi, Vietnam. *Arch Sex Behav* 2015 Feb;44(2):349-356. [doi: [10.1007/s10508-014-0450-8](https://doi.org/10.1007/s10508-014-0450-8)] [Medline: [25617010](https://pubmed.ncbi.nlm.nih.gov/25617010/)]
38. Peacock E, Andrinopoulos K, Hembling J. Binge drinking among men who have sex with men and transgender women in San Salvador: correlates and sexual health implications. *J Urban Health* 2015 Aug;92(4):701-716 [FREE Full text] [doi: [10.1007/s11524-014-9930-3](https://doi.org/10.1007/s11524-014-9930-3)] [Medline: [25591660](https://pubmed.ncbi.nlm.nih.gov/25591660/)]
39. Hernandez-Romieu AC, Sullivan PS, Sanchez TH, Kelley CF, Peterson JL, Del RC, et al. The comparability of men who have sex with men recruited from venue-time-space sampling and facebook: a cohort study. *JMIR Res Protoc* 2014 Jul 17;3(3):e37 [FREE Full text] [doi: [10.2196/resprot.3342](https://doi.org/10.2196/resprot.3342)] [Medline: [25048694](https://pubmed.ncbi.nlm.nih.gov/25048694/)]

Abbreviations

- AIDS:** acquired immune deficiency syndrome
- CAI:** condomless anal intercourse
- CHCT:** couples HIV counseling and testing
- HAART:** highly active anti-retroviral treatment
- HIV:** human immunodeficiency virus
- MSM:** men who have sex with men
- aOR:** adjusted Odds Ratio

Edited by T Sanchez; submitted 21.02.17; peer-reviewed by B Koblin, J Mitchell; comments to author 20.03.17; revised version received 29.03.17; accepted 25.04.17; published 20.06.17.

Please cite as:

Chard AN, Metheny N, Stephenson R

Perceptions of HIV Seriousness, Risk, and Threat Among Online Samples of HIV-Negative Men Who Have Sex With Men in Seven Countries

JMIR Public Health Surveill 2017;3(2):e37

URL: <http://publichealth.jmir.org/2017/2/e37/>

doi: [10.2196/publichealth.7546](https://doi.org/10.2196/publichealth.7546)

PMID: [28634155](https://pubmed.ncbi.nlm.nih.gov/28634155/)

©Anna N Chard, Nicholas Metheny, Rob Stephenson. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 20.06.2017. 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

TwiMed: Twitter and PubMed Comparable Corpus of Drugs, Diseases, Symptoms, and Their Relations

Nestor Alvaro^{1,2}, PhD; Yusuke Miyao^{1,2}, PhD; Nigel Collier³, PhD

¹National Institute of Informatics, Department of Informatics, Tokyo, Japan

²The Graduate University for Advanced Studies (SOKENDAI), Kanagawa, Japan

³Faculty of Modern & Medieval Languages, Department of Theoretical and Applied Linguistics, University of Cambridge, Cambridge, United Kingdom

Corresponding Author:

Nestor Alvaro, PhD

National Institute of Informatics

Department of Informatics

2-1-2 Hitotsubashi, Chiyoda-ku

Tokyo, 101-8430

Japan

Phone: 81 3 4212 2164

Fax: 81 3 4212 2150

Email: nestoralvaro@gmail.com

Abstract

Background: Work on pharmacovigilance systems using texts from PubMed and Twitter typically target at different elements and use different annotation guidelines resulting in a scenario where there is no comparable set of documents from both Twitter and PubMed annotated in the same manner.

Objective: This study aimed to provide a comparable corpus of texts from PubMed and Twitter that can be used to study drug reports from these two sources of information, allowing researchers in the area of pharmacovigilance using natural language processing (NLP) to perform experiments to better understand the similarities and differences between drug reports in Twitter and PubMed.

Methods: We produced a corpus comprising 1000 tweets and 1000 PubMed sentences selected using the same strategy and annotated at entity level by the same experts (pharmacists) using the same set of guidelines.

Results: The resulting corpus, annotated by two pharmacists, comprises semantically correct annotations for a set of drugs, diseases, and symptoms. This corpus contains the annotations for 3144 entities, 2749 relations, and 5003 attributes.

Conclusions: We present a corpus that is unique in its characteristics as this is the first corpus for pharmacovigilance curated from Twitter messages and PubMed sentences using the same data selection and annotation strategies. We believe this corpus will be of particular interest for researchers willing to compare results from pharmacovigilance systems (eg, classifiers and named entity recognition systems) when using data from Twitter and from PubMed. We hope that given the comprehensive set of drug names and the annotated entities and relations, this corpus becomes a standard resource to compare results from different pharmacovigilance studies in the area of NLP.

(*JMIR Public Health Surveill* 2017;3(2):e24) doi:[10.2196/publichealth.6396](https://doi.org/10.2196/publichealth.6396)

KEYWORDS

Twitter; PubMed; corpus; pharmacovigilance; natural language processing; text mining; annotation

Introduction

Corpora annotated for adverse drug events are becoming important in order to train computers to automatically build adverse drug reaction profiles for post marketing surveillance.

Researchers are typically interested in understanding the accuracy of their systems [1-8], whereas at the same time only a limited number of corpora exist [9-12].

Pharmacovigilance (drug safety) systems using texts obtained from the scientific literature have received attention for many years [1,6-8] and since recently researchers started exploring Twitter and other nonscientific texts where patients describe

diseases and symptoms [2-4]. However, there is currently no way to systematically compare systems performance across text types.

In this paper we provide a benchmark corpus composed of semantically correct annotations that can be used in natural language processing (NLP) studies and show our approach to produce a comparable corpora using texts from Twitter and PubMed, explaining our strategy for controlling external variables that may affect the sample.

Social media texts are known for containing a high proportion of ungrammatical constructions out of vocabulary words, abbreviations, and metaphoric usage [13], whereas scientific texts are known for the use of specialized vocabulary and well-formed sentences. Secondary key factors involved in a direct comparison are the data selection methods and the topicality [14].

Existing corpora from PubMed and Twitter cannot be directly compared and the goal of this research is to produce a comparable corpus of drug-related sentences targeting at the same set of drugs.

To date, most of the curated corpora for pharmacovigilance come from scientific formal texts obtained from PubMed [15,16], although datasets curated from other scientific resources, such as the Khresmoi project [17], are also available [18].

Since a few years ago, corpora obtained from social media texts started emerging. At first, researchers focused on blogs and forums [19,20] and then on Twitter's data [3,21,22] due to the high volume of the information it provides, with 310 million monthly active users [23] generating over 500 million tweets per day [24] and also motivated by its "realtime" information, allowing health researchers to potentially investigate and identify new adverse drug event (ADE) types faster than traditional methods such as physician reports.

Researchers have assessed the number of scientific works in PubMed where Twitter data was used [25], finding that the interest in Twitter is growing within the medical domain. Pharmacovigilance sees Twitter as a useful resource in different areas. Messages found in Twitter tweets can help researchers to understand temporal patterns on the drugs usage [26], can provide a good resource for obtaining first-hand experience reports on the drugs use [21], and can be useful in the early detection of prescription medication abuse [27] and adverse events [3,28].

Recent examples where researchers used Twitter to build different corpora are: (1) the corpus built to understand patient experiences at health care facilities [29], (2) the corpus built to measure the public interest and concerns about different diseases [30], and (3) the corpus used to assess the positive or negative attitude toward specific treatments [31]. Although texts written in English have been used very frequently in NLP for pharmacovigilance, texts in Spanish extracted from social media [32] and French clinical texts [33] have been also used.

The work on Twitter and PubMed is an ongoing effort resulting in promising NLP studies on the automatic recognition of

medications and adverse events [34,35] and also on the medical question answering [36,37]. By releasing our corpus, we hope other researchers can benefit from it and continue advancing in this area.

Methods

Data Selection

For our study, we selected a set of 30 different drugs used in other pharmacovigilance studies [3,5,21,38,39]. Including these drugs allowed us to cover different interests in the research community and also allowed us to account for drugs used to treat very different conditions such as cancer [3], attention deficit disorders [5,21,39], schizophrenia [3,38], or depression [5,21,38,39].

We employed Twitter's application programming interface (API) to download messages mentioning any of those drug names or their synonyms by running our script from September 7, 2015 to October 10, 2015, obtaining 165,489 tweets. In the case of PubMed, we obtained the list of articles about those drugs by using EuropePMC RESTful Web Services [40], issuing our query on October 21, 2015 to search for texts containing the same keywords that we used when collecting tweets. Once we had the list of PubMed articles, we processed them to extract the sentences containing the drug mentions obtaining 29,435 sentences.

From these sentences, we removed all non-ASCII (American standard code for information interchange) characters (eg, emojis), replaced all user name mentions with "`__username__`," all email addresses with "`__email__`," and all numbers with "`__number__`." We also reduced characters elongation by removing the repetition of a character after the second occurrence eg, "greeeeeat" would become "great"), and lowercased all sentences.

Using the preprocessed sentences and aiming at maximizing the informativeness and the variability of the texts, we limited the number of tweets any user could contribute to 5 and discarded sentences shorter than 20 characters in length, retweets, tweets not written in English, sentences containing keywords related to marketing campaigns (for this we created a list built heuristically using 5 words commonly related to marketing campaigns: "buy," "cheap," "online," "pharmacy," "price"), and also discarded sentences including URLs.

To discard possibly duplicated sentences, we stored 40-character long substrings appearing in the chosen sentences and searched for these substrings in the candidate sentences keeping only the messages not containing them. For each chosen sentence, we only stored one substring composed of a maximum of 40 characters (less for sentences shorter than 60 characters in length), extracted from the character in position 20th onwards. This decision was driven by the observation that there were a number of tweets conveying the same information using minor rewording for the sentences, making them unique. In this scenario, discarding the sentences replicating information contributes to increase the information diversity.

This strategy aims at further increasing the variability of the texts by filtering out similar messages, and in case of selecting the message “Lisinopril is used for treating high blood pressure alone or with other medicines. Other names for this medication. Acecomb, Acelisino” by extracting the characters in position 20 to position 40 (“or treating high blood pressure alone or”), we are able to discard possible duplicated sentences such as “lisinopril and hctz 20 mg 25 mg—national institutes of...lisinopril is used for treating high blood pressure alone or with other,” and similarly the system is also able to discard the sentence “Jun 29, 2015...Active Ingredient: Lisinopril. Prinivil is used for treating high blood pressure alone or with other medicines. Other.” This strategy also showed its usefulness when applied to PubMed sentences as observed in the substring “mg oral granules are bioequivalent to s” appearing twice in the same article, first in the abstract (“Sandoz montelukast 4 mg oral granules are bioequivalent to Singulair 4 mg mini oral granules, with a similar safety profile .”), and also in the discussion (“The current study has clearly demonstrated that Sandoz montelukast 4 mg oral granules are bioequivalent to Singulair mini 4 mg oral granules in terms of the rate and extent of absorption of each formulation .”), thereby showing that this method can help in reducing the amount of duplicated information.

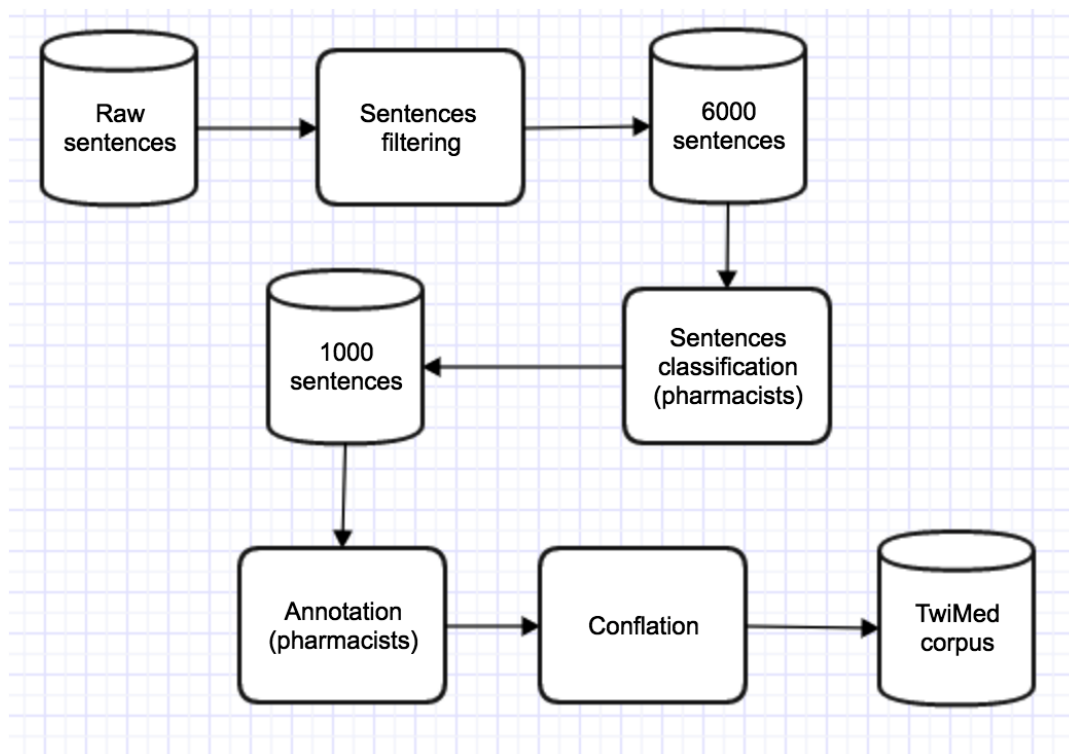
Out of the resulting sentences, we automatically selected 6000 sentences each for both Twitter and PubMed, which we extracted in a round-robin fashion aiming at a balanced sample of the drug mentions.

We were interested in finding which sentences would be of interest, for which we divided the main task in two phases. During the first phase, both annotators were requested to perform a sentence level annotation to extract 1000 positive sentences (ie, the sentences mentioning drugs, symptoms, and diseases related to the drug effects in humans) out of the 6000 sentences. In the second phase, the annotators would use the annotation guidelines to identify the entities and relations appearing in the 1000 sentences identified during the first phase.

The aim of this pipeline is to filter the most informative sentences, discarding those sentences that are prone to include information that is not of high relevance for pharmacovigilance studies.

Figure 1 shows the pipeline used to filter, classify, and annotate the sentences. Despite the difference in the initial number of raw sentences we had from Twitter and PubMed (165,489 tweets and 29,435 PubMed sentences), the steps described in the figure provided the same number of sentences at the end of each process.

Figure 1. Annotation pipeline. The initial number of raw sentences differed between twitter (165,489 tweets) and PubMed (29,435 sentences).



Selecting the Annotators

We identified 6 people who were willing to contribute to the task and prepared a test to identify the best candidates. We provided them with 20 sentences from Twitter and 20 sentences from PubMed obtained from the 6000 sentences we had previously filtered, and their task was to identify the sentences

containing a mention to a drug and a related disease or symptom. The 6 candidate annotators had different backgrounds: one of them was a native English speaker, three of them were pharmacists, and the last two of them were active social media users. Except for the native English speaker, the rest of the annotators were native Spanish speakers able to read English

texts. Although the pharmacists are not referred to as “active social media users,” the three of them were millennials who used social media networks to some extent.

Annotation

The annotation guidelines were prepared after reviewing existing guidelines used in other pharmacovigilance projects. In the ADE corpus guidelines [10], the researchers annotated the drugs, adverse effects, dosages, and the existing relations between these elements, whereas in the meta-knowledge annotation of bio-events [11], the researchers followed a slightly different approach and focused on different “dimensions” of the biomedical events. Those dimensions can be thought as attributes of those events as these dimensions are the knowledge type, the level of certainty, the polarity, the manner, and the source of the annotated event.

Similarly, to the annotation of the ADE corpus, the Arizona disease corpus (AZDC) annotation guidelines [41] focused on the annotation of the diseases, also covering syndromes, illnesses, and disorders. Another document consulted to prepare the first draft of our annotation guidelines was the shared

annotated resources (ShARe) or Conference and Labs of the Evaluation Forum (CLEF) eHealth 2013 shared task I [12], where the authors annotated disorders using the concept unique identifier (CU), also clarifying that a disorder is understood as “any span of text that can be mapped to a concept in the SNOMED-CT terminology, which belongs to the Disorder semantic group,” clarifying that the Disorder semantic group should include “congenital abnormalities,” “diseases or syndromes,” and “signs and symptoms” among others.

Other supporting document used to prepare the annotation guidelines was the annotation guidelines for the drug-drug interaction (DDI) corpus [9], where the authors focused on the annotation of a number of entities such as drugs approved for human use, brand names for approved drugs, drugs that have not been approved for human use, and different drug groups. These guidelines also describe the annotation for different types of relations existing between the entities: “advice,” “effect,” “mechanism,” or “other.”

The details on the resulting corpora produced by the researchers using the aforementioned annotation guidelines can be found in [Table 1](#).

Table 1. Details on the resulting corpora produced by the researchers who used the guidelines we reviewed.

Corpus name	DDI ^a corpus	ADE ^b corpus	AZDC ^c corpus	ShARe ^d or CLEF ^e eHealth 2013 Task I
Annotated entities	Pharmacological substances	Drug, adverse effects, dosages	Diseases	Disorders
Annotated relations	Drug-drug interactions	Drug-adverse effect, drug-dosage	-	-
Texts origin	DrugBank and MEDLINE	MEDLINE	PubMed abstracts	Clinical notes
Number of documents	1025	2972	794 (2775 sentences)	200
Number of annotators	2	3 (after automatic annotation)	2 (after automatic annotation)	2
Availability	Free use for academic research	Free	Free	Upon request
Annotation Tool	Brat	Knowtator	In-house tool	Knowtator

^aDDI: drug-drug interaction.

^bADE: adverse drug event.

^cAZDC: Arizona disease corpus.

^dShARe: shared annotated resources.

^eCLEF: Conference and Labs of the Evaluation Forum.

As shown in [Table 1](#), for the annotation of these corpora, researchers have used tools such as Brat [42] and Knowtator [43]. In our case, Brat tool was chosen after taking into account that it is a Web-based annotation tool that eases key elements of the annotation process.

The use of the mentioned guidelines eased the task of generating the first draft of our guidelines, and allowed us to identify which were the entities, relations, and attributes to be annotated. This first draft was then used by three external annotators with a background in computer science who annotated a small set of PubMed and Twitter sentences. During that first annotation period, we had daily meetings after each annotation session and refined the guidelines upon the discrepancies we found and the questions raised by the annotators. Those comments and question in combination with the information we found in the

existing guidelines was used to provide the annotators with an updated version of the guidelines for the next annotation session.

After 2 weeks and 6 annotation sessions, the number of discrepancies was reduced to a minimum and no more questions were raised, leading us to agree on freezing the guidelines so that these would be used as they were.

The final version of the guidelines used in our study includes three different entities: (1) Drug: any of the marketed medicines that appears in the SIDER database [44], which is also listed in the closed set of drugs we provided to the annotators, (2) symptom: any sign or symptom contained in MedDRA [45] ontology, and (3) disease: any disease contained in MedDRA ontology.

The use of SIDER, which contains information on marketed medicines extracted from public documents and package inserts,

and MedDRA, a medical terminology dictionary aimed at easing the annotators' task by providing them two well-known resources to provide the annotated entities with a standardized concept identifier. We believe the fact that those resources are of common use in the research community, and combined with the current trend to map concepts in these databases to concepts in other resources [46,47], provides an important element toward TwiMed corpus reuse.

- **Polarity:** Used to indicate whether the entity was negated or not. The negation had to be a linguistic negation (“not,” “don’t”...).
- **Person:** Used to indicate whether the entity was affecting the “1st,” “2nd,” “3rd” person, or whether there was no information. This attribute was based on the original sender.
- **Modality:** Used to indicate whether the entity was stated in an “actual,” “hedged,” “hypothetical,” or “generic” way.
- **Exemplification:** Used to indicate whether the entity was presented using an example or a description. This attribute was only to be used when the entity was presented through an exemplification.
- **Duration:** Used to indicate whether the entity’s lasting span was “intermittent,” “regular,” “irregular,” or not stated. In the case of drugs, this attribute referred to the time span when the drug had been taken.
- **Severity:** Used to indicate whether the seriousness of an entity was “mild,” “severe,” or not stated. This was the only attribute that did not apply to drugs.
- **Status:** Used to indicate whether the duration of the entity was “complete,” “continuing,” or not stated. In the case of drugs, this attribute referred to the time span when the drug was perceived as having effect.
- **Sentiment:** Used to indicate whether the entity was perceived as “positive,” “negative,” or “neutral.”
- **Entity identifier:** Used to indicate the CUI for that entity. This was the only attribute that had to be filled for all annotated entities. For this attribute we provided a list of allowed values, and used the value “-1” (not found) for entities whose CUI would not be present in the list.

The list of attributes was decided based on the combination of elements noticed to be annotated in pharmacovigilance studies using formal texts (eg, “duration” or “modality”), as well as in pharmacovigilance studies using informal texts (eg, “polarity” or “sentiment”).

- **Reason-to-use:** Used to represent the relation appearing when a symptom or disease leads to the use of some drug.
- **Outcome-positive:** Used to represent the relation between a drug, and an expected or unexpected symptom or disease appearing after the drug consumption. The outcome had to be positive.
- **Outcome-negative:** Used to represent the relation between a drug, and an expected or unexpected symptom or disease appearing after the drug consumption. The outcome had to be negative.

These elements are further explained in the annotations guidelines that are shared in the [Multimedia Appendix 2](#).

Once the guidelines were ready and the annotators were chosen, we preprocessed the sentences before presenting them to the annotators by replacing the existing emojis with a string describing each character, and discarded other non-ASCII characters. We also decided not to lower case the sentences as we thought that would ease the annotator’s task to detect some sentiments and disambiguate acronyms. Besides these changes, the preprocessing strategy is the same we described in the “data selection” section.

To compare the annotations produced by the experts, we focused on both the “type” assigned to the entity (ie, disease, drug, or symptom) and also on the offsets for that entity. Taking that into account, we decided to compute the results when using relaxed constraints and strict constraints. In the case of using relaxed constraints, we say that the entity annotated by both annotators is a match if the type for the entity matches between annotations and the spans of those annotations have some overlap. In the case of using strict constraints, the match would happen if the type in both annotations matches and the spans for the annotated entities have the same offsets. Discontinuous annotations were allowed and taken into account when computing the matches, which means that in case of using strict constraints, all the spans taking part on the entity’s annotation should be the same.

We measured the level of agreement between the annotations produced by our experts following the inter annotator agreement (IAA) measure in the CLEF corpus [48]. This IAA metric is reported to approximate the kappa score [48], and to be more suited for this case [49]:

$$\text{IAA} = \frac{\text{matches}}{\text{matches} + \text{nonmatches}}$$

In our case matches accounts for the total number of token matches for which both annotators agreed, and matches + nonmatches counts all annotations performed by the annotator being evaluated.

Results

Data Selection

Out of the 6000 selected sentences each for both Twitter and PubMed that we extracted, we observed differing sample frequencies of each drug. In both Twitter and PubMed, some drugs attracted more attention than others, although in the case of Twitter, temporal variability is a known fact [26] that has to be taken into consideration.

We found that the frequency of the drugs in the extracted sample had no correlation between Twitter and PubMed (Spearman $\rho = .03$), as shown in [Table 2](#).

Table 2. Total number of sentences for each drug name in Twitter and PubMed.

Drug name	# Tweets	# Sentences in PubMed
Bevacizumab	69	239
Buprenorphine	363	244
Carbamazepine	74	239
Ciprofloxacin	81	250
Citalopram	331	251
Cortisone	344	231
Destroamphetamine sulphate	373	19
Docetaxel	34	246
Duloxetine	242	241
Fluoxetine	344	238
Fluvoxamine maleate	13	204
Lamotrigine	168	242
Lisdexamfetamine	348	84
Lisinopril	56	147
Melphalan	2	234
Methylphenidate hydrochloride	349	112
Modafinil	287	10
Montelukast	71	239
Olanzapine	190	248
Paroxetine	365	249
Prednisone	350	249
Quetiapine	339	247
Rupatadine	1	45
Sertraline	343	236
Tamoxifen	122	238
Topiramate	133	231
Trazodone	206	70
Triamcinolone acetonide	14	253
Venlafaxine	326	238
Ziprasidone	62	226

Selecting the Annotators

To evaluate the annotator's performance, we used a gold standard set of labels that we generated obtaining the majority vote from the results we received from the 6 annotators and the annotations produce by the first author of the paper, also giving more weight to the pharmacists' annotations in PubMed and to

social media users' annotations in Twitter. That is, when there were clear differences between the annotations provided by the contributors with the higher weights and the rest of the annotators, we took the former annotations into account.

As can be seen in [Table 3](#), one pharmacist scored the best result, 87.5% agreement with the gold standard data (35 out of 40 sentences were correctly labelled).

Table 3. Agreement with gold standard data during the annotator selection phase. We compared the results from 2 very active social media users, one native English speaker and 3 pharmacists. We indicate between brackets the time it took to complete the annotation for that dataset (time in min).

Annotator	Twitter (min)	PubMed (min)	Total (min)
Social1	0.70 (9)	0.80 (10)	0.75 (19)
Social2	1.00 (8)	0.70 (7)	0.85 (15)
Native speaker	0.85 (6)	0.50 (6)	0.67 (12)
Pharmacist1	0.90 (8)	0.85 (7)	0.87 (15)
Pharmacist2	0.70 (11)	0.80 (9)	0.75 (20)
Pharmacist3	0.50 (15)	0.70 (15)	0.60 (30)

Those results were in line with our expectations as social media users got the best scores in social media texts, and the best scores in PubMed texts were obtained by the pharmacists. However, we were very surprised by the low scores obtained by Pharmacist3 and the native English speaker. We followed up with them discovering that Pharmacist3 had some trouble understanding the samples because of those being written in English language (it was also evidenced in the time it took her to complete the task). In the case of the native English speaker, he reported that he was not an active social media user and requested further information on the set of tweets as he found those texts to be hard to understand. Overall, we discovered the native English speaker was too cautious when indicating which sentences were positive cases as he annotated 7 sentences as positive out of the 40 sentences (the gold standard data had 16 sentences tagged as positive sentences), whereas the rest of the

annotators indicated 13-18 sentences were positive (Pharmacist3, who obtained the lowest score, was above that range as she annotated 24 sentences as positive).

We decided to hire Pharmacist1 as she scored the best results, and out of Social1, Social2, and Pharmacist2, we decided to hire Pharmacist2 taking into account that the resulting corpus would require annotation at entity level for which Pharmacist2's in-domain knowledge would be very valuable.

Annotation

Once the 2 pharmacists completed the annotation at sentence level, we focused on the entity level annotation targeting at the diseases, drugs, and symptoms. The results for Twitter and PubMed are shown in Tables 4 and 5, using the relaxed constraints and strict constraints strategy described in the Methods section.

Table 4. Detail of annotations in Twitter. The first column shows the element being evaluated. Columns 2-5 show the inter annotator agreement scores of pharmacist 1 (Ph1) and pharmacist 2 (Ph2) using relaxed and strict constraints. Columns 6 and 7 show the number of elements annotated by each pharmacist. Columns 8 and 9 show the number of matching elements between pharmacist's annotations using relaxed and strict constraints.

Annotated element	Ph1 (relaxed constraints)	Ph2 (relaxed constraints)	Ph1 (strict constraints)	Ph2 (strict constraints)	#Ph1	#Ph2	#Matches (relaxed constraints)	#Matches (strict constraints)
Drug	97.39	98.72	93.52	94.80	1111	1096	1082	1039
Disease	50.86	91.47	46.12	82.95	464	258	236	214
Symptom	77.23	76.71	54.21	53.84	1164	1172	899	631
Outcome-negative	63.27	75.19	43.02	51.12	795	669	503	342
Outcome-positive	11.01	40.00	8.26	30.00	109	30	12	9
Reason-to-use	55.82	60.18	44.66	48.14	842	781	470	376
Duration	46.37	8.96	39.11	7.56	248	1283	115	97
Exemplification	10.11	64.77	3.37	21.59	564	88	57	19
Modality	56.92	30.58	49.57	26.63	585	1089	333	290
Person	72.56	58.55	60.21	48.58	1709	2118	1240	1029
Polarity	76.06	52.43	53.52	36.89	71	103	54	38
Sentiment	72.48	19.46	60.92	16.36	476	1773	345	290
Severity	64.18	19.59	44.03	13.44	134	439	86	59
Status	59.41	22.07	45.94	17.07	542	1459	322	249

Table 5. Detail of annotations in PubMed. The first column shows the element being evaluated. Columns 2-5 show the inter annotator agreement scores of pharmacist 1 (Ph1) and pharmacist 2 (Ph2) using relaxed and strict constraints. Columns 6 and 7 show the number of elements annotated by each pharmacist. Columns 8 and 9 show the number of matching elements between pharmacist's annotations using relaxed and strict constraints.

Annotated element	Ph1 (relaxed constraints)	Ph2 (relaxed constraints)	Ph1 (strict constraints)	Ph2 (strict constraints)	#Ph1	#Ph2	#Matches (relaxed constraints)	#Matches (strict constraints)
Drug	95.20	97.90	86.23	88.67	1271	1236	1210	1096
Disease	64.18	95.22	53.41	79.23	1086	732	697	580
Symptom	85.13	60.59	70.61	50.26	558	784	475	394
Outcome-negative	60.97	64.86	50.35	53.56	433	407	264	218
Outcome-positive	56.25	32.73	43.75	25.45	32	55	18	14
Reason-to-use	62.87	77.39	47.10	57.98	1535	1247	965	723
Duration	52.17	9.38	48.70	8.75	115	640	60	56
Exemplification	0.64	50.00	0.32	25.00	311	4	2	1
Modality	74.23	50.52	64.60	43.96	1370	2013	1017	885
Person	63.93	77.18	56.08	67.70	1439	1192	920	807
Polarity	25.00	22.22	25.00	22.22	16	18	4	4
Sentiment	33.33	1.96	22.22	1.31	9	153	3	2
Severity	42.22	33.33	37.78	29.82	45	57	19	17
Status	53.85	2.52	53.85	2.52	26	555	14	14

By focusing on the results appearing in [Tables 4](#) and [5](#), we see that the agreement for the drugs in Twitter and PubMed is very high, which was expected, given our sampling strategy, although for diseases and symptoms the agreement score decreases noticeably in both Twitter and PubMed.

When comparing the results for the relations (outcome-negative, outcome-positive, and reason-to-use), we saw low levels of agreement, having Twitter lower results in all cases. Analyzing the number of annotations it was clear that the use of outcome-positive relation varied considerably between annotators, contributing to the low scores.

The attributes "person" (in PubMed), "modality" (in both PubMed and Twitter), "polarity," and "sentiment" (in Twitter) were the ones obtaining the best scores. On the other hand, the attribute "exemplification" (in both PubMed and Twitter), "sentiment," "polarity" (in PubMed), and "duration" (in Twitter) were very prone to disagreements as these scores were the lowest in [Tables 4](#) and [5](#).

By analyzing the discrepancies in the annotations, we discovered that the distinction between disease and symptom entities, although theoretically clear, was hard to disambiguate in a number of sentences. We can see that in the tweet "Is steroid induced psychosis a thing? (Like short term prednisone tx)" (see [Figure 2](#)), psychosis could be identified as a symptom [50,51] or as a disease [52]. Similarly, the entity comorbid obesity found in the PubMed sentence "The present case report of topiramate's effect on comorbid obesity," could be also understood as both a symptom and a disease [53]. Interestingly, we observed in those examples that even if the chosen type of entity (disease and symptom) was different, the annotators agreed on the chosen CUI.

In the case of relations, we discovered that both outcome-positive relation and reason-to-use relation were confounded in some cases. One example from Twitter is the sentence "How about trazodone, so I can just feel a little funny and then knock out and have the best sleep of my life," where the drug, trazodone, and the symptom, the best sleep of my life, were annotated as such by both annotators, although one annotator indicated the relation between these entities was an outcome-positive relation whereas the other annotator marked it as a reason-to-use relation. The same observation was seen in PubMed sentences as in "Because fatigue is a frequent symptom of depression and there is some evidence that treatment with an antidepressant improves fatigue in patients with fibromyalgia, we hypothesized that the antidepressant fluvoxamine might improve fatigue related to PBC and PSC." In this sentence, the drug (fluvoxamine) and the symptom (fatigue) were correctly identified, same as the existing relation between the entities, but the chosen type of relation was different. This observation, combined with the fact that outcome-positive relation was the least used type of relation, helps in understanding the causes for the low inter annotator agreement score.

Given the similarities between those concepts and the disagreements that we detected, we evaluated the inter annotator agreement score when conflating the concepts disease and symptom under "disease or symptom" concept. We also grouped together outcome-positive and reason-to-use relations under "benefit" relation. The use of those categories produced a noticeable improvement in the IAA scores. This strategy also improved the agreement scores for most of the attributes as can be seen in [Table 6](#) (for Twitter), and [Table 7](#) (for PubMed).

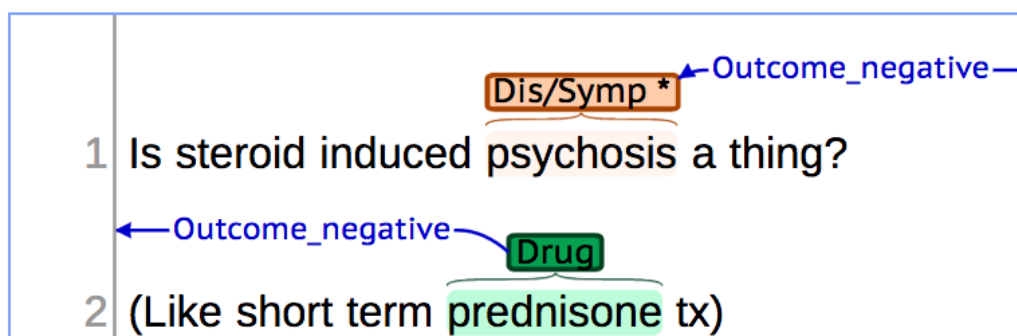
Table 6. Detail of annotations in Twitter using the conflation strategy. The first column shows the element being evaluated. Columns 2-5 show the inter annotator agreement scores of pharmacist 1 (Ph1) and pharmacist 2 (Ph2) using relaxed and strict constraints. Columns 6 and 7 show the number of elements annotated by each pharmacist. Columns 8 and 9 show the number of matching elements between pharmacist's annotations using relaxed and strict constraints.

Annotated element	Ph1 (relaxed constraints)	Ph2 (relaxed constraints)	Ph1 (strict constraints)	Ph2 (strict constraints)	#Ph1	#Ph2	#Matches (relaxed constraints)	#Matches (strict constraints)
Drug	97.39	98.72	93.52	94.80	1111	1096	1082	1039
Disease or symptom	82.25	93.64	61.36	69.86	1628	1430	1339	999
Outcome-negative	67.30	79.97	46.29	55.01	795	669	535	368
Benefit	68.14	79.90	52.37	61.41	951	811	648	498
Duration	50.00	9.66	41.94	8.11	248	1283	124	104
Exemplification	10.11	64.77	3.37	21.59	564	88	57	19
Modality	64.44	34.62	54.53	29.29	585	1089	377	319
Person	77.30	62.37	63.96	51.61	1709	2118	1321	1093
Polarity	80.28	55.34	57.75	39.81	71	103	57	41
Sentiment	75.00	20.14	62.61	16.81	476	1773	357	298
Severity	67.16	20.50	47.01	14.35	134	439	90	63
Status	61.81	22.96	48.15	17.89	542	1459	335	261

Table 7. Detail of annotations in PubMed using the conflation strategy. The first column shows the element being evaluated. Columns 2-5 show the inter annotator agreement scores of pharmacist 1 (Ph1) and pharmacist 2 (Ph2) using relaxed and strict constraints. Columns 6 and 7 show the number of elements annotated by each pharmacist. Columns 8 and 9 show the number of matching elements between pharmacist's annotations using relaxed and strict constraints.

Annotated element	Ph1 (relaxed constraints)	Ph2 (relaxed constraints)	Ph1 (strict constraints)	Ph2 (strict constraints)	#Ph1	#Ph2	#Matches (relaxed constraints)	#Matches (strict constraints)
Drug	95.20	97.90	86.23	88.67	1271	1236	1210	1096
Disease or symptom	91.91	99.67	74.21	80.47	1644	1516	1511	1220
Outcome-negative	81.52	86.73	65.82	70.02	433	407	353	285
Benefit	77.41	93.16	56.86	68.43	1567	1302	1213	891
Duration	53.91	9.69	50.43	9.06	115	640	62	58
Exemplification	0.64	50.00	0.32	25.00	311	4	2	1
Modality	83.43	56.78	71.39	48.58	1370	2013	1143	978
Person	71.58	86.41	62.13	75.00	1439	1192	1030	894
Polarity	43.75	38.89	43.75	38.89	16	18	7	7
Sentiment	33.33	1.96	22.22	1.31	9	153	3	2
Severity	53.33	42.11	46.67	36.84	45	57	24	21
Status	53.85	2.52	53.85	2.52	26	555	14	14

Figure 2. Sample with the annotation of a drug, a disease and the relation between these concepts in a sentence from Twitter.



Discussion

Principal Findings

We produced a corpus of documents obtained using a very similar pipeline for both Twitter and PubMed. All the documents were filtered and double annotated by the same experts (pharmacists).

TwiMed corpus shows that the drugs appearing less often in Twitter are those used to treat cancer (“docetaxel,” “bevacizumab,” “tamoxifen”) and epilepsy (“topiramate,” “carbamazepine”). On the other hand, the drugs used to treat attention deficit hyperactivity disorder (“modafinil,” “destroamphetamine sulphate,” “lisdexamfetamine”) are the drugs having the most mentions in Twitter, whereas at the same time those are the drugs having the least mentions in PubMed sentences. This fact evidences that public concerns are not always aligned with the interests of the scientific community, and potential areas of research may emerge from those findings to understand the reasons and related outcomes.

In our analysis, we observed a much higher balance for the drug mentions in PubMed than in Twitter. Factors such as the demographics of the user base in Twitter [54], or the time when the messages were gathered [26] are elements that should be studied in detail to measure their correlation with the different distribution of drug mentions.

Given our main goal was to create a corpus covering mentions of drugs used to treat a number of different conditions, we applied a set of controlling mechanisms to extract the data from Twitter and from PubMed, and researchers should be aware that the drugs in this corpus have different distribution in the original sources of information than the distributions presented here, and following quantitative studies may be needed to understand those differences.

We believe the reason for the high agreement in the annotation of drugs is the use of a closed set of drug names that both annotators knew beforehand. We can see, however, that there is a lower level of agreement for the annotation of symptoms and diseases and the main reason would be that the annotators had to identify these mentions, and there is an open list of entities that can be found in the texts, not to mention that these entities could be presented using an exemplification. In addition,

in some contexts a disease can be considered as a symptom, and the short nature of the sentences can act as a factor in confounding the nature of these entities. Similarly, more subjective concepts such as the duration attribute or the exemplification attribute show a low level of agreement probably because the annotators had to interpret these elements by themselves and in some cases a certain level of subjectivity led the decision.

Researchers should be aware that this corpus is not devised to capture everything about the selected set of 30 drugs, and there are a number of drug names appearing in the selected set of sentences which were not annotated because these were not included in the target set of drugs. Similarly, there are DDIs and other relations that the corpus does not include because of our constraints. However, we believe the provided sample can help in training NLP systems to capture more information. Nonetheless, we provide a set of semantically correct annotations that can be used in NLP studies.

Our annotation also confirmed that there is lower agreement in the annotation of tweets than in the annotation of PubMed sentences, showing the noisy nature of Twitter [13]. Moreover, when applying our conflation strategy aimed at resolving disagreements, we observed that these differences still remained.

We noticed that a number of those disagreements were caused when confounding “diseases” and “symptoms.” Similarly, acronyms appearing in documents from PubMed tend to be explained the first time they are presented, which does not necessarily have to be when the drugs and related symptoms and diseases are discussed. In our case we allowed the annotators to access the full articles during the annotation process to reduce the impact of this problem. Nonetheless, we believe the use of acronyms is a potential source of confusion in texts where the context is scarce, and this potential problem should be handled.

Additionally, other noticeable finding when using the same guidelines is the fact that disagreements appear in similar categories for both Twitter and PubMed. Figure 3 shows an example where the string “eyelids are itchy” was annotated with the duration of “regular” by one annotator (to indicate that there is a continued lasting span), whereas the other annotator chose “irregular” for the duration attribute (to indicate that there is no pattern in the lasting span).

Figure 3 also shows an example where the annotation for the attribute “exemplification” differs between annotators as the string “eyelids are itchy” was annotated as an exemplification by only one of the two annotators.

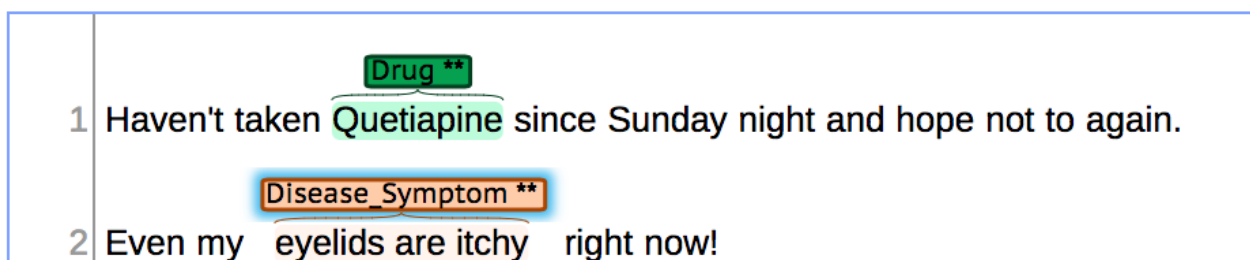
Besides PubMed-Twitter comparative studies, our corpus is of potential interest for researchers aiming at finding sentences containing information on the drugs, symptoms, and diseases. We believe this corpus can become a useful resource to discern informational sentences in the area of pharmacovigilance as other researchers can use the sentences we included in this dataset to create classifiers targeting at the correct identification of sentences reporting drug-use.

This dataset shows that for similar events coming from very different data sources, the way in which people communicate the same messages has noticeable differences. This corpus can

provide useful insights to science communicators and public institutions for adapting their messages when addressing the general public so that the information can attract more attention. One of such examples would be the use of social media by official health institutions, where most of the messages are more formal than average social media messages, as a mean to reach a wider audience during health promotion and disease prevention campaigns as the wording may affect the impact of the messages.

We believe combining the information contained in scientific reports, of high quality and very trustworthy, together with the information coming from social media messages, which is global, has a high volume, and is up-to-date, should be taken into account when building pharmacovigilance systems. We hope this corpus can help researchers interested in combining the potential of those data sources.

Figure 3. Sample of an annotation where “duration” and “exemplification” attributes are used.



Conclusions

We have presented a pharmacovigilance corpus that, to our knowledge, is the first corpus that allows researchers to perform direct comparative studies toward understanding the differences between drug reports in Twitter and PubMed.

Our corpus contains annotations for drugs, symptoms, and diseases; their attributes (polarity, person, modality, exemplification, duration, severity, status, sentiment); and the relations between the annotated entities (reason-to-use, outcome-negative, and outcome-positive).

We also identified the source of a number of disagreements for the annotated entities and relations, and proposed a conflation strategy to resolve those discrepancies. That approach resulted in higher agreement scores for most entities and relations.

We hope that given the comprehensive set of drug names and the annotated entities and relations included in this corpus, it

can become a standard resource to compare results from different pharmacovigilance studies, especially in the area of NLP as it can help in training to recognize the entities and relations in the texts. Similarly, this corpus can help in comparing the performance of NLP tools across the 2 different linguistic registers (formal and informal).

In summary, we present a comparable corpus for pharmacovigilance studies and the annotation scheme we devised. This work is presented to the research community in the belief that such resources can help in this rapidly growing area.

The corpus we release, available as [Multimedia Appendix 1](#), contains the annotations for all the entities, relations, and attributes where both annotators agreed. Additionally, we provide the tools to obtain the raw tweets used in the annotation, to comply with Twitter’s terms of service [55], and we also provide the tools to preprocess the raw sentences from both Twitter and PubMed to reuse the released annotations.

Acknowledgments

This research project was supported by a grant from the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT). We would also like to thank Ana Salto Hurtado and Marta Gómez Márquez de Prado for their help in annotating the sentences. We want to thank Ramiro Aparicio Gallardo for his technical support. Finally, we also want to thank Dr Nut Limsopatham and Dr Yuka Tateisi for their suggestions while developing the guidelines.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Twimed corpus.

[[ZIP File \(Zip Archive\), 2MB - jmir_v3i2e24_app1.zip](#)]

Multimedia Appendix 2

Annotation guidelines used to prepare Twimed corpus.

[[PDF File \(Adobe PDF File\), 139KB - jmir_v3i2e24_app2.pdf](#)]

References

1. Hakala K, Van Landeghem S, Salakoski T, Van de Peer Y, Ginter F. Application of the EVEX resource to event extraction and network construction: Shared Task entry and result analysis. In: BMC Bioinformatics. 2015 Presented at: BioNLP Shared Task 2013; August 9, 2013; Sofia, Bulgaria p. S3 URL: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-16-S16-S3> [doi: [10.1186/1471-2105-16-S16-S3](https://doi.org/10.1186/1471-2105-16-S16-S3)]
2. Freifeld CC, Brownstein JS, Menone CM, Bao W, Filice R, Kass-Hout T, et al. Digital drug safety surveillance: monitoring pharmaceutical products in twitter. In: Drug Safety. New York, NY: Springer; 2014:343-350.
3. Bian J, Topaloglu U, Yu F. Towards large-scale twitter mining for drug-related adverse events. 2012 Presented at: Proceedings of the 2012 international workshop on Smart health and wellbeing; October 29, 2012; Maui, HI p. 25-32. [doi: [10.1145/2389707.2389713](https://doi.org/10.1145/2389707.2389713)]
4. 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/)]
5. 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/)]
6. Björne J, Salakoski T. Biomedical Event Extraction for Diverse Corpora. BMC Bioinformatics 2015;16(Suppl 16):S4. [doi: [10.1186/1471-2105-16-S16-S4](https://doi.org/10.1186/1471-2105-16-S16-S4)]
7. Miwa M, Thompson P, McNaught J, Kell DB, Ananiadou S. Extracting semantically enriched events from biomedical literature. BMC Bioinformatics 2012 May 23;13:108 [FREE Full text] [doi: [10.1186/1471-2105-13-108](https://doi.org/10.1186/1471-2105-13-108)]
8. Miwa M, Ananiadou S. NaCTeM EventMine for BioNLP 2013 CG and PC tasks. 2013 Presented at: Proceedings of the BioNLP Shared Task 2013 Workshop; 2013; Sofia, Bulgaria p. 94-98.
9. Herrero-Zazo M, Segura-Bedmar I, Martínez P, Declerck T. The DDI corpus: an annotated corpus with pharmacological substances and drug-drug interactions. J Biomed Inform 2013 Oct;46(5):914-920 [FREE Full text] [doi: [10.1016/j.jbi.2013.07.011](https://doi.org/10.1016/j.jbi.2013.07.011)] [Medline: [23906817](https://pubmed.ncbi.nlm.nih.gov/23906817/)]
10. Gurulingappa H, Rajput AM, Roberts A, Fluck J, Hofmann-Apitius M, Toldo L. Development of a benchmark corpus to support the automatic extraction of drug-related adverse effects from medical case reports. J Biomed Inform 2012 Oct;45(5):885-892 [FREE Full text] [doi: [10.1016/j.jbi.2012.04.008](https://doi.org/10.1016/j.jbi.2012.04.008)] [Medline: [22554702](https://pubmed.ncbi.nlm.nih.gov/22554702/)]
11. Nawaz R, Thompson P, McNaught J, Ananiadou S. Meta-knowledge annotation of bio-events. 2010 Presented at: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010); May 17-23, 2010; Valletta, Malta p. 2498-2505.
12. Elhadad N, Chapman W, Savova G. CHPC. ShARe/CLEF eHealth 2013 shared task: guidelines for the annotation of disorders in clinical notes URL: http://blulab.chpc.utah.edu/sites/default/files/ShARe_Guidelines_CLEF_2013.pdf [accessed 2017-04-24] [WebCite Cache ID [6pxg52DD9](https://www.webcitation.org/6pxg52DD9)]
13. Baldwin T, Cook P, Lui M, MacKinlay A, Wang L. how noisy social media text, how diffrent social media sourcess? 2013 Presented at: International Joint Conference on Natural Language Processing; October 14-18, 2013; Nagoya, Japan p. 356-364.
14. Lippincott T, Séaghdha D, Korhonen A. Exploring subdomain variation in biomedical language. BMC Bioinformatics 2011 May 27;12:212 [FREE Full text] [doi: [10.1186/1471-2105-12-212](https://doi.org/10.1186/1471-2105-12-212)] [Medline: [21619603](https://pubmed.ncbi.nlm.nih.gov/21619603/)]
15. 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/)]
16. Kim J, Ohta T, Tateisi Y, Tsujii J. GENIA corpus-a semantically annotated corpus for bio-textmining. Bioinformatics 2003 Jul 10;19(Suppl 1):i180-i182. [doi: [10.1093/bioinformatics/btg1023](https://doi.org/10.1093/bioinformatics/btg1023)]
17. Khresmoi. Khresmoi - medical information analysis and retrieval URL: <http://www.khresmoi.eu> [accessed 2011-11-06] [WebCite Cache ID [62zteCKmX](https://www.webcitation.org/62zteCKmX)]
18. Goeuriot L, Kelly L, Jones G, Zuccon G, Suominen H, Hanbury A, et al. Creation of a new evaluation benchmark for information retrieval targeting patient information needs. 2013 Presented at: The Fifth International Workshop on Evaluating Information Access; June 18, 2013; Tokyo, Japan.

19. Frost J, Okun S, Vaughan T, Heywood J, Wicks P. Patient-reported outcomes as a source of evidence in off-label prescribing: analysis of data from PatientsLikeMe. *J Med Internet Res* 2011 Jan 21;13(1):e6 [FREE Full text] [doi: [10.2196/jmir.1643](https://doi.org/10.2196/jmir.1643)] [Medline: [21252034](https://pubmed.ncbi.nlm.nih.gov/21252034/)]
20. Benton A, Ungar L, Hill S, Hennessy S, Mao J, Chung A, et al. Identifying potential adverse effects using the web: a new approach to medical hypothesis generation. *J Biomed Inform* 2011 Dec;44(6):989-996 [FREE Full text] [doi: [10.1016/j.jbi.2011.07.005](https://doi.org/10.1016/j.jbi.2011.07.005)] [Medline: [21820083](https://pubmed.ncbi.nlm.nih.gov/21820083/)]
21. Alvaro N, Conway M, Doan S, Lofi C, Overington J, Collier N. Crowdsourcing Twitter annotations to identify first-hand experiences of prescription drug use. *J Biomed Inform* 2015 Dec;58:280-287 [FREE Full text] [doi: [10.1016/j.jbi.2015.11.004](https://doi.org/10.1016/j.jbi.2015.11.004)] [Medline: [26556646](https://pubmed.ncbi.nlm.nih.gov/26556646/)]
22. Yin Z, Fabbri D, Rosenbloom ST, Malin B. A scalable framework to detect personal health mentions on twitter. *J Med Internet Res* 2015 Jun 05;17(6):e138 [FREE Full text] [doi: [10.2196/jmir.4305](https://doi.org/10.2196/jmir.4305)] [Medline: [26048075](https://pubmed.ncbi.nlm.nih.gov/26048075/)]
23. Twitter. About company URL: <http://about.twitter.com/company> [accessed 2015-04-28] [WebCite Cache ID 6Y8Gpanwp]
24. Internet-Live-Stats. Twitter usage statistics URL: <http://www.internetlivestats.com/twitter-statistics/> [accessed 2016-10-13] [WebCite Cache ID 6IEjp8QRn]
25. Williams SA, Terras M, Warwick C. How twitter is studied in the medical professions: a classification of twitter papers indexed in PubMed. *Med* 2013;2(2):e2 [FREE Full text] [doi: [10.2196/med20.2269](https://doi.org/10.2196/med20.2269)] [Medline: [25075237](https://pubmed.ncbi.nlm.nih.gov/25075237/)]
26. Hanson CL, Burton SH, Giraud-Carrier C, West JH, Barnes MD, Hansen B. Tweaking and tweeting: exploring Twitter for nonmedical use of a psychostimulant drug (Adderall) among college students. *J Med Internet Res* 2013 Apr 17;15(4):e62 [FREE Full text] [doi: [10.2196/jmir.2503](https://doi.org/10.2196/jmir.2503)] [Medline: [23594933](https://pubmed.ncbi.nlm.nih.gov/23594933/)]
27. Sarker A, O'Connor K, Ginn R, Scotch M, Smith K, Malone D, et al. Social media mining for toxicovigilance: automatic monitoring of prescription medication abuse from twitter. *Drug safety* 2016;231-240. [doi: [10.1007/s40264-015-0379-4](https://doi.org/10.1007/s40264-015-0379-4)]
28. Plachouras V, Leidner JL, Garrow AG. Quantifying self-reported adverse drug events on twitter: signal and topic analysis. 2016 Presented at: International Conference on Social Media & Society; 2016; London. [doi: [10.1145/2930971.2930977](https://doi.org/10.1145/2930971.2930977)]
29. Rastegar-Mojarad M, Ye Z, Wall D, Murali N, Lin S. Collecting and analyzing patient experiences of health care from social media. *J Med Internet Res* 2015 Jul 02;4(3):e78 [FREE Full text] [doi: [10.2196/resprot.3433](https://doi.org/10.2196/resprot.3433)] [Medline: [26137885](https://pubmed.ncbi.nlm.nih.gov/26137885/)]
30. Weeg C, Schwartz HA, Hill S, Merchant RM, Arango C, Ungar L. Using twitter to measure public discussion of diseases: a case study. *JMIR Public Health Surveill* 2015 Jun 26;1(1):e6. [doi: [10.2196/publichealth.3953](https://doi.org/10.2196/publichealth.3953)]
31. Du L, Rachul C, Guo Z. Gordie Howe's "Miraculous Treatment": case study of twitter users' reactions to a sport celebrity's stem cell treatment. *JMIR Public Health Surveill* 2016;2(1):e8. [doi: [10.2196/publichealth.5264](https://doi.org/10.2196/publichealth.5264)]
32. Segura-Bedmar I, Martínez P, Revert R, Moreno-Schneider J. Exploring Spanish health social media for detecting drug effects. *BMC Med Inform Decis Mak* 2015 Jun;15 Suppl 2:S6 [FREE Full text] [doi: [10.1186/1472-6947-15-S2-S6](https://doi.org/10.1186/1472-6947-15-S2-S6)] [Medline: [26100267](https://pubmed.ncbi.nlm.nih.gov/26100267/)]
33. Deléger L, Grouin C, Zweigenbaum P. Extracting medication information from French clinical texts. *Stud Health Technol Inform* 2010;160(Pt 2):949-953. [Medline: [20841824](https://pubmed.ncbi.nlm.nih.gov/20841824/)]
34. Polepalli RB, Belknap SM, Li Z, Frid N, West DP, Yu H. Automatically recognizing medication and adverse event information from food and drug administration's adverse event reporting system narratives. *JMIR Med Inform* 2014 Jun 27;2(1):e10 [FREE Full text] [doi: [10.2196/medinform.3022](https://doi.org/10.2196/medinform.3022)] [Medline: [25600332](https://pubmed.ncbi.nlm.nih.gov/25600332/)]
35. Hoffman KB, Dimbil M, Tatonetti NP, Kyle RF. A pharmacovigilance signaling system based on FDA regulatory action and post-marketing adverse event reports. *Drug safety* 2016 Jun;39(6):561-575. [doi: [10.1007/s40264-016-0409-x](https://doi.org/10.1007/s40264-016-0409-x)] [Medline: [26946292](https://pubmed.ncbi.nlm.nih.gov/26946292/)]
36. Wongchaisuwat P, Klabjan D, Jonnalagadda SR. A semi-supervised learning approach to enhance health care community-based question answering: a case study in alcoholism. *JMIR Med Inform* 2016 Aug 02;4(3):e24 [FREE Full text] [doi: [10.2196/medinform.5490](https://doi.org/10.2196/medinform.5490)] [Medline: [27485666](https://pubmed.ncbi.nlm.nih.gov/27485666/)]
37. Abacha AB, Zweigenbaum P. MEANS: A medical question-answering system combining NLP techniques and semantic Web technologies. *Inf Process Manage* 2015;570-594. [doi: [10.1016/j.ipm.2015.04.006](https://doi.org/10.1016/j.ipm.2015.04.006)]
38. 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. 2010 Presented at: Proceedings of the 2010 Workshop on Biomedical Natural Language Processing; July 15, 2010; Uppsala, Sweden p. 117-125.
39. Sampathkumar H, Chen X, Luo B. Mining adverse drug reactions from online healthcare forums using hidden Markov model. *BMC Med Inform Decis Mak* 2014 Oct 23;14:91 [FREE Full text] [doi: [10.1186/1472-6947-14-91](https://doi.org/10.1186/1472-6947-14-91)] [Medline: [25341686](https://pubmed.ncbi.nlm.nih.gov/25341686/)]
40. Europe PMC. RESTful Web Service URL: <http://europepmc.org/RestfulWebService> [accessed 2017-04-24] [WebCite Cache ID 6pxubKVDG]
41. ASU. Arizona Disease Corpus Annotation Guidelines URL: http://diego.asu.edu/downloads/AZDCAnnotationGuidelines_v013.pdf [accessed 2017-04-24] [WebCite Cache ID 6pxufeemu]
42. Stenertorp P, Pyysalo S, Topic G, Ohta T, Ananiadou S, Tsujii J. BRAT: a web-based tool for NLP-assisted text annotation. 2012 Presented at: Conference of the European Chapter of the Association for Computational Linguistics; April 23-27, 2012; Avignon, France p. 102-107.

43. Ogren PV. Knowtator: a plug-in for creating training and evaluation data sets for biomedical natural language systems. 2006 Presented at: International Protégé Conference; July 23-26, 2006; Stanford, California p. 73-76.
44. SIDER. SIDER Side Effect Resource URL: <http://sideeffects.embl.de> [accessed 2013-01-23] [WebCite Cache ID 6DsJnvdb1]
45. BioPortal. Medical dictionary for regulatory activities URL: <http://bioportal.bioontology.org/ontologies/MEDDRA> [accessed 2017-04-24] [WebCite Cache ID 6pxvlyd8g]
46. Dhombres F, Bodenreider O. Interoperability between phenotypes in research and healthcare terminologies--investigating partial mappings between HPO and SNOMED CT. *J Biomed Semantics* 2016;7:3 [FREE Full text] [doi: [10.1186/s13326-016-0047-3](https://doi.org/10.1186/s13326-016-0047-3)] [Medline: [26865946](https://pubmed.ncbi.nlm.nih.gov/26865946/)]
47. Mizuno S, Ogishima S, Nishigori H, Jamieson DG, Verspoor K, Tanaka H, et al. The pre-eclampsia ontology: a disease ontology representing the domain knowledge specific to pre-eclampsia. *PLoS One* 2016;11(10):e0162828 [FREE Full text] [doi: [10.1371/journal.pone.0162828](https://doi.org/10.1371/journal.pone.0162828)] [Medline: [27788142](https://pubmed.ncbi.nlm.nih.gov/27788142/)]
48. Roberts A, Gaizauskas R, Hepple M, Demetriou G, Guo Y, Setzer A, et al. Semantic annotation of clinical text: the CLEF corpus. 2008 Presented at: LREC; May 26, 2008; Marrakech, Morocco p. 19-26.
49. Hripcsak G, Rothschild AS. Agreement, the f-measure, and reliability in information retrieval. *J Am Med Inform Assoc* 2005 Jan;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/)]
50. Tamminga CA, Carlsson A. Partial dopamine agonists and dopaminergic stabilizers, in the treatment of psychosis. *Curr Drug Targets CNS Neurol Disord* 2002 Apr;1(2):141-147. [Medline: [12769623](https://pubmed.ncbi.nlm.nih.gov/12769623/)]
51. Hunter A, Grealish A, Dowling M. Improving quality of life for adolescents with psychosis. *Mental Health Practice* 2010;32-35. [doi: [10.7748/mhp2010.04.13.7.32.c7650](https://doi.org/10.7748/mhp2010.04.13.7.32.c7650)]
52. Perlman G, Kotov R, Fu J, Bromet EJ, Fochtmann LJ, Medeiros H, Genomic Psychiatry Cohort Consortium, et al. Symptoms of psychosis in schizophrenia, schizoaffective disorder, and bipolar disorder: a comparison of African Americans and Caucasians in the Genomic Psychiatry Cohort. *Am J Med Genet B Neuropsychiatr Genet* 2016 Jun;171(4):546-555. [doi: [10.1002/ajmg.b.32409](https://doi.org/10.1002/ajmg.b.32409)] [Medline: [26663585](https://pubmed.ncbi.nlm.nih.gov/26663585/)]
53. Khaodhiar L, McCowen KC, Blackburn GL. Obesity and its comorbid conditions. *Clin Cornerstone* 1999;2(3):17-31. [Medline: [10696282](https://pubmed.ncbi.nlm.nih.gov/10696282/)]
54. Duggan M. Pew Research Center. Mobile messaging and social media 2015 URL: <http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/> [accessed 2016-03-22] [WebCite Cache ID 6gCZOB0Lz]
55. Twitter. Twitter Developer Documentation URL: <https://dev.twitter.com/overview/terms/agreement-and-policy> [accessed 2016-10-04] [WebCite Cache ID 6l0jyN5el]

Abbreviations

- ADE:** adverse drug events
ASCII: American standard code for information interchange
CLEF: Conference and Labs of the Evaluation Forum
CUI: concept unique identifier
DDI: drug-drug interaction
IAA: inter annotator agreement
NLP: natural language processing
ShARe: shared annotated resources

Edited by G Eysenbach; submitted 28.07.16; peer-reviewed by A Dunn, L Lafrado, J Kang, AC Tan, I Segura Bedmar; comments to author 22.08.16; revised version received 24.11.16; accepted 20.03.17; published 03.05.17.

Please cite as:

Alvaro N, Miyao Y, Collier N

TwiMed: Twitter and PubMed Comparable Corpus of Drugs, Diseases, Symptoms, and Their Relations

JMIR Public Health Surveill 2017;3(2):e24

URL: <http://publichealth.jmir.org/2017/2/e24/>

doi: [10.2196/publichealth.6396](https://doi.org/10.2196/publichealth.6396)

PMID: [28468748](https://pubmed.ncbi.nlm.nih.gov/28468748/)

©Nestor Alvaro, Yusuke Miyao, Nigel Collier. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 03.05.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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

Filtering Entities to Optimize Identification of Adverse Drug Reaction From Social Media: How Can the Number of Words Between Entities in the Messages Help?

Redhouane Abdellaoui^{1,2}, MSc; Stéphane Schüeck², MSc, MD; Nathalie Texier², PharmD; Anita Burgun^{1,3}, MD, PhD

¹INSERM, UMRS 1138 Team 22, Université Pierre et Marie Curie, Paris, France

²Kappa Santé, Innovation, Paris, France

³Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpital Européen Georges-Pompidou (HEGP), Medical Informatics, Paris, France

Corresponding Author:

Redhouane Abdellaoui, MSc

INSERM

UMRS 1138 Team 22

Université Pierre et Marie Curie

4 rue de Cléry

Paris, 75002

France

Phone: 33 1 44 82 74 74

Fax: 33 1 44 82 74 75

Email: redhouane.abdellaoui@kappasante.com

Abstract

Background: With the increasing popularity of Web 2.0 applications, social media has made it possible for individuals to post messages on adverse drug reactions. In such online conversations, patients discuss their symptoms, medical history, and diseases. These disorders may correspond to adverse drug reactions (ADRs) or any other medical condition. Therefore, methods must be developed to distinguish between false positives and true ADR declarations.

Objective: The aim of this study was to investigate a method for filtering out disorder terms that did not correspond to adverse events by using the distance (as number of words) between the drug term and the disorder or symptom term in the post. We hypothesized that the shorter the distance between the disorder name and the drug, the higher the probability to be an ADR.

Methods: We analyzed a corpus of 648 messages corresponding to a total of 1654 (drug and disorder) pairs from 5 French forums using Gaussian mixture models and an expectation-maximization (EM) algorithm.

Results: The distribution of the distances between the drug term and the disorder term enabled the filtering of 50.03% (733/1465) of the disorders that were not ADRs. Our filtering strategy achieved a precision of 95.8% and a recall of 50.0%.

Conclusions: This study suggests that such distance between terms can be used for identifying false positives, thereby improving ADR detection in social media.

(*JMIR Public Health Surveill* 2017;3(2):e36) doi:[10.2196/publichealth.6577](https://doi.org/10.2196/publichealth.6577)

KEYWORDS

pharmacovigilance; social media; text mining; Gaussian mixture model; EM algorithm; clustering; density estimation

Introduction

Background

Adverse drug reactions (ADRs) cause millions of injuries worldwide each year and require billions of Euros in associated costs [1,2]. Drug safety surveillance targets the detection, assessment, and prevention of ADRs in the postapproval period. A promise of augmenting drug safety with patient-generated

data drawn from the Internet was called for by several scientific committees related to pharmacovigilance in the United States and in Europe [3,4].

There are now sites for consumers that enable patients to report ADRs. Patients who experience ADRs want to contribute drug safety content, share their experience, and obtain information and support from other Internet users [5-8].

Three recently published review articles showed that the use of social media data for ADR monitoring was increasing. Sarker et al analyzed 22 studies that used social media data. They observed that publicly available annotated data remained scarce, thus making system performance comparisons difficult [9]. Golder et al analyzed 51 studies based on a total of 174 social media sites, most of which had discussion forums (71%). They used broad selection criteria and considered several types of social media including messages, social networks, patient forums, Twitter, blogs, and Facebook [10]. Ninety percent (45/51) of the papers looked for any adverse events, and 10% (5/51) focused on specific adverse events (eg, fatal skin reactions or hypersensitivity). The overall prevalence of adverse event reports in social media varied from 0.2% to 8% of the posts. There was general agreement that a high frequency of mild adverse events was identified but that the more serious events and laboratory-based ADRs were under-represented in social media. Lardon et al explored methods for identifying and extracting target data and evaluating the quality of medical information from social media. Most studies used supervised classification techniques to detect posts containing ADR mentions and lexicon-based approaches to extract ADR mentions from texts [9,11].

When the methods relied on the development of lexicons, these studies were generally limited in the number of drugs studied or the number of target ADRs. For example, Benton et al focused on 4 drugs [12]; Yang et al focused on 10 drugs and 5 ADRs [13]; Yates et al focused on breast cancer-associated ADRs [14]; Jiang et al focused on 5 drugs [15]; and Sarker and Gonzalez focused on various drugs prescribed in chronic diseases, such as type 2 diabetes [16].

Other authors focused on detecting user posts mentioning potential ADRs. Some of them combined social media with other knowledge sources such as Medline [17]. The binary classification of text into ADR versus non-ADR categories has been typically performed in previous research work using three supervised classification approaches: (1) Naïve Bayes (NB), (2) support vector machine (SVM), and (3) maximum entropy (ME). Among those, SVMs are the most popular for text classification tasks [18], including ADR text.

In 13 studies using automatic processing based on data mining to analyze patient declarations, 7 studies aimed at identifying the relationships between disease entities and drug names. Five of these studies used machine learning methods. Qualitative analyses of forums and mailing list posts show that it may be used to identify rare and serious ADRs (eg, [11,19,20]) and the unexpected frequency of known ADRs. However, the use of social media for data source pharmacovigilance must be validated [10].

Therefore, the main challenge lay in identifying a combination of methods that could reduce the overall number of misclassifications of potential ADRs from patient's posts. In all such studies, the authors analyzed messages that contained references to both a drug and a disorder or symptom. ADRMine, a machine learning-based concept extraction system [21] that uses conditional random fields (CRFs), achieved an *F* measure of 0.82 in the ADR extraction task.

However, ADR messages from social media are not only factual descriptions about adverse events [10]. The messages may also include contextual information (the patient's condition and comorbidities) and opinions and feelings about treatments and drugs (eg, providing personal experience about a treatment, discussing new research, explaining documentation and drug monograph to a peer, and exchanging information relevant to patient's daily lives).

Before robust conclusions can be drawn from social media regarding ADRs, the biggest problem with automated or semiautomated methods is distinguishing between genuine ADRs and other types of cooccurrence (eg, treatments and context) between drugs and diseases in messages. To quote Golder [10], "the purported adverse events may not be adverse events at all. Terms used to describe adverse events can also be used for indications of the condition being treated (eg, confounding by indication), beneficial effects (ie, sleepiness can be a beneficial effect for someone with insomnia), or may not have been experienced by a patient." This notion can be illustrated by an article published by Benton et al [12]. The authors analyzed social media to identify adverse events that were associated with the most commonly used drugs to treat breast cancer. In their study, "uterine cancer" cooccurred 374 times with tamoxifen. However, most of the messages involved anxiety about taking tamoxifen because of a possible adverse event (uterine cancer) that could potentially occur in the future.

These examples indicate that methods are required to eliminate such false positives. The Detec't project developed by Kappa Santé [22] is an adverse drug reaction monitoring program based on data mining and statistical analysis techniques using social media texts. Our intent at this point was to distinguish between potential ADRs and non-ADRs among the disorders associated with a drug in messages from social media. In this paper, we investigate whether the distance between the terms representing drugs and disorders in the messages may help distinguish between ADRs and false positives.

Related Work

The current technological challenges include the difficulty for text mining algorithms to interpret patient lay vocabulary [23].

After the review of multiple approaches, Sarker et al [9] concluded that following data collection, filtering was a real challenge. Filtering methods are likely to aid in the ADR detection process by removing most irrelevant information. Based on our review of prior research, two types of filtering methods can be used: semantic approaches and statistical approaches.

Semantic filtering relies on semantic information, for example, negation rules and vocabularies, to identify messages not corresponding to an ADR declaration. Liu and al [24] developed negation rules and incorporated linguistic and medical knowledge bases in their algorithms to filter out negated ADRs, then remove drug indications and non- and unreported cases on FAERS (FDA's Adverse Event Reporting System) database. In their use case of 1822 discussions about beta blockers, 71% of the related medical events were adverse drug events, 20%

were drug indications, and 9% were negated adverse drug events.

Powell et al [25] developed “Social Media Listening,” a tool to augment postmarketing safety. This tool consisted on the removal of questionable Internet pharmacy advertisements (named “Junk”), posts in which a drug was discussed (named “mention”), posts in which a potential event was discussed (called “Proto-AE”), and any type of medical interaction description (called “Health System Interaction”). Their study revealed that only 26% of the considered posts contained relevant information. The distribution of post classifications by social media source varied considerably among drugs. Between 11% (7/63) and 50.5% (100/198) of the posts contained Proto-AEs (between 3.2% (4/123) and 33.64% (726/2158) for over-the-counter products). The final step was a manual evaluation.

The second type of filtering was based on statistical approaches using the topic models method [26]. Yang et al [27] used latent Dirichlet allocation probabilistic modeling [28] to filter topics and thereby reduce the dataset to a cluster of posts to evoke an ADR declaration. This method was evaluated by the comparison of 4 benchmark methods (example adaption for text categorization [EAT], positive examples and negative examples labeling heuristics [PNLH], active semisupervised clustering based two-stage text classification [ACTC], and Laplacian SVM) and the calculation of F scores (the harmonic mean of precision and recall) on ADRs posts. These 4 methods were improved by the use of this approach. The F score gains fluctuated between 1.94% and 6.14%. Sarker and Gonzalez [16] improved their ADR detection method by using different features for filtering. These multiple features were selected by the use of leave-one-out classification scores and were evaluated with accuracy and F scores. These features were based on n -grams (accuracy 82.6%, F score 0.654), computing the Tf-idf values for the semantic types (accuracy 82.6%, F score 0.652), polarity of sentences (accuracy 84.0%, F score 0.669), the positive or negative outcome (accuracy 83.9%, F score 0.665), ADR lexicon match (accuracy 83.5%, F score 0.659), sentiment analysis in posts (accuracy 82.0%), and filtering by topics (accuracy 83.7%, F score 0.670) for filtering posts without mention of ADRs. The use of all features for the filtering process provided an accuracy of 83.6% and an F score of 0.678. Bian et al [29] utilized SVM to filter the noise in tweets. Their motivation for classifying tweets arose from the fact that most posts were not associated with ADRs; thus, filtering out nonrelevant posts was crucial.

Wei and al [30] performed an automatic chemical-diseases relation extraction on a corpus of PubMed articles. Their process was divided in two subtasks. The first one was a disease named entity recognition (DNER) subtask based on the 1500 PubMed

titles and abstracts. The second subtask was a chemical-induced disease (CID) relation extraction (on the same corpus as DNER subtask). Chemicals and diseases were described utilizing the medical subject headings (MeSH) controlled vocabulary. They evaluated several approaches and obtained an average precision, recall, and standard F score of 78.99%, 74.81%, and 76.03%, respectively for DNER step and an average of 43.37% of F score with the CID step. The best result for CID step was obtained by combining two SVM approaches.

Objective

We propose adding a filter based on Gaussian mixtures models to reduce the burden of other entities, that is, disorders that are mentioned in the messages but are not ADRs. The objective was to optimize ADR detection by reducing the number of false positives. We hypothesized that the shorter the distance between the disorder name and the drug, the higher the probability to be an ADR. The approach was applied to the Detec't corpus.

Methods

Materials

Detec't Database

We used a version of the Detec't database that contained 17,703,834 messages corresponding to 350 drugs. The messages were extracted from 20 general health forums, all in French, using a custom Web crawler to browse the selected forums and scrape messages. The forums scraped do not restrict users with a limited number of characters in the message. Detec't contains the messages extracted and associated metadata, namely users' aliases and dates.

The Detec't database was created in 2012 by Kappa Santé [22], a contract research organization founded in 2003 that specialized in post marketing studies and pharmacoepidemiology. Kappa Santé developed Detec't to achieve this goal.

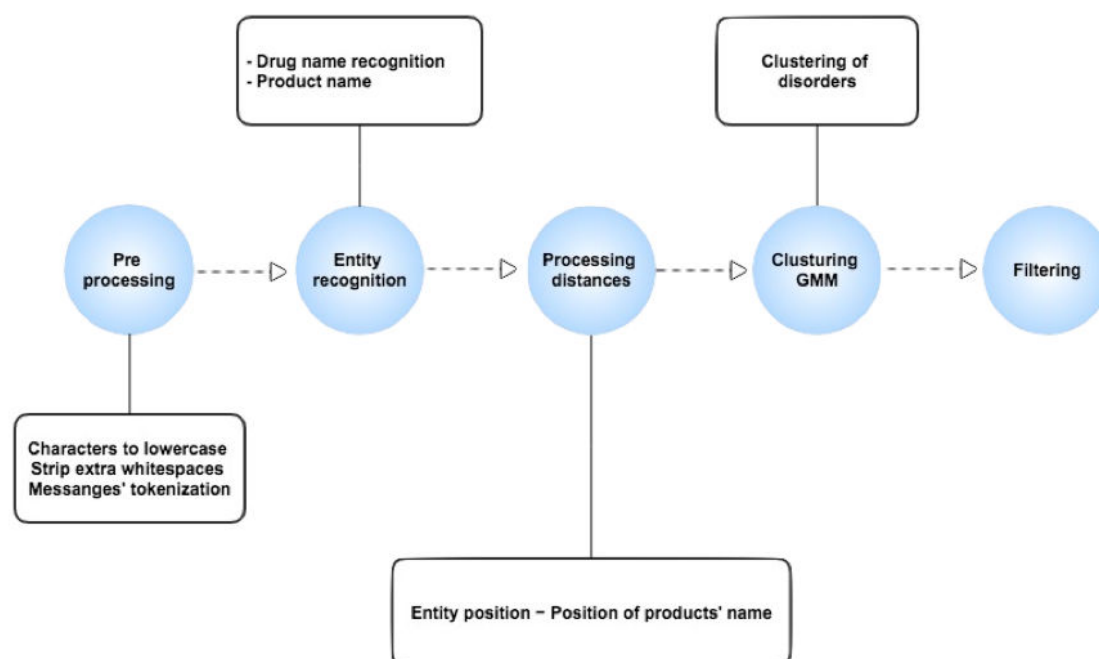
The messages that constitute our dataset came from (1) doctissimo, (2) atoute.org, (3) e-santé, (4) santé médecine, and (5) aufeminin. These are popular general forums dedicated to health with an average of 89,987 unique visitors a day in 2016. Users must register to be able to post a message in these forums.

Dataset Constitution

We randomly extracted 700 messages from the Detec't database related to 3 drugs from 3 different therapeutic classes: Teriflunomide, Insulin Glargine, and Zolpidem.

Of these, 52 messages did not contain any disease entity and were removed from the list. The remaining 648 messages were both manually annotated and automatically processed. Processing was performed in 5 steps; the method is summarized in Figure 1.

Figure 1. Summary diagram.



Medical Terminology

Regarding disorders, we used the Medical Dictionary for Regulatory Activities dictionary (MedDRA), which is the international medical terminology developed under the auspices of the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). The MedDRA dictionary is organized by system organ class (SOC) and divided into high-level group terms (HLGT), high-level terms (HLT), preferred terms (PT), and lowest-level terms (LLT). Synonymous LLT are grouped under a unique identifier labeled as preferred terms (PT).

We used a lexicon built in-house by Kappa Santé. The lexicon was derived from the French version of MedDRA 15.0 and was extended by adding more lay medical vocabulary. A fuzzy grouping technique was used to group commonly misspelled words or closely spelled words under one term. The grouping was performed at the MedDRA LLT level. The fuzzy grouping algorithm temporarily strips all vowels (except the first one), strips double or triple consonants from extracted words, and then compares them to see if they are the same. For example, “modeling” and “modelling” would be grouped together [31-34]. The original 15.0 release of MedDRA contains 19,550 PT and 70,177 LLT. Our lexicon contained a total of 19,530 PT and 63,392 LLT. Among them, 259 additional LLT were added by Kappa Santé, for example, “mal au crâne” (French familiar broadly used expression for headache) as a synonym of “mal de tête” (headache). Although not pure synonyms, as “crâne” is not equivalent to “head,” “mal de crâne is a familiar broadly used expression for headache. The decrease in the number of terms was caused by the removal of some PTs that were beyond the scope of ADRs, such as “married.” Moreover, the lexicon was manually reduced by grouping terms with similar meanings,

for example, the PTs for “alcoholic” and “alcoholism.” Our final version for disorders contained a total of 63,392 terms (LLT level), including both original MedDRA LLT terms and nonstandard terms. We used the most specific (LLT) level to search for disorder entities in the posts.

Manual Annotation

An ADR is a sign or symptom caused by taking a medication. ADRs may occur following a single dose or prolonged administration of a drug or result from the combination of 2 or more drugs.

A disorder concept corresponds to a sign or symptom, a disease, or any pathological finding. In the context of a message, a disorder may:

- Either play the role of an adverse event, (ADR) for example, “I took aspirin, it gave me a terrible headache.” These are considered “true ADRs.”
- Or correspond to a condition that is not reported by the patient as an ADR, for example, “I had a headache so I took aspirin.”

With the objective of distinguish between ADRs and disorders that were not ADRs, 2 experts manually annotated the messages to identify true ADRs.

The annotators labeled each disorder entity in the messages as (1) « ADR » if the patient reported the disorder as a possible ADR in his or her message, or (2) « other entity » if the disorder was not reported as an ADR in the message.

This annotated dataset was used as a gold standard.

Analysis Phases

Data Preprocessing

The standardization of our approach required preprocessing the dataset to avoid some cases of poor data quality. Figure 2 presents these preprocessing steps.

The character separation method involved inserting whitespaces around every punctuation character. This separation was necessary due to the poor data quality to optimize disorder identification.

Because we used the R software (a language and environment for statistical computing provide by the R core team in Vienna)

Figure 2. Data preprocessing steps.



Named Entity Recognition

The objective of the named entity recognition module was to identify 2 types of entities in a patient's post: drug names and disorders.

As the extended lexicon for disorders that we developed contained colloquial terms as well as expressions with spelling and/or grammatical errors, lexicon matching was performed using exact match methods after stemming of both messages and expressions in the lexicon.

Drug names in the messages were automatically identified using fuzzy matching and stored in the Detect database as message metadata. To minimize the impact of misspelled words, each word was first stemmed using a Porter stemmer, an algorithm meant to remove inflection from a word [12,16]. Savoy [36] demonstrates that the use of Porter stemmer algorithm improved information retrieval by 30.5% with French language.

All of the other terms in the messages were mapped to our extended version of MedDRA, which includes colloquialisms, abbreviations, and words with spelling errors. Lexicon matching was implemented as string matching using regular expressions. The granularity of the disorder concepts extracted from the messages corresponded to the LLT level of MedDRA.

Processing Distance Between Entities

After the preprocessing step, the position of each entity in the message was calculated. We defined a word as a continuous series of characters between 2 whitespaces. The distance

to process and analyze data, and given that R discriminates between lowercase and uppercase words, we used the "tm" Package (a text mining framework for R software) to convert the document text to lower case and remove extra whitespaces [35]. We did not remove stop words because our hypothesis was based on the number of words separating drug names and disorder terms. The stop words removing could have impacted the distances distribution.

The last step was the tokenization of messages. Word segmentation provides a list of words in each message and their positions in the post.

between a drug "a" and a disorder "b" in a message was defined as the number of words separating the two entities:

$$\text{Distance (a,b)} = (\text{position of b}) - (\text{position of a})$$

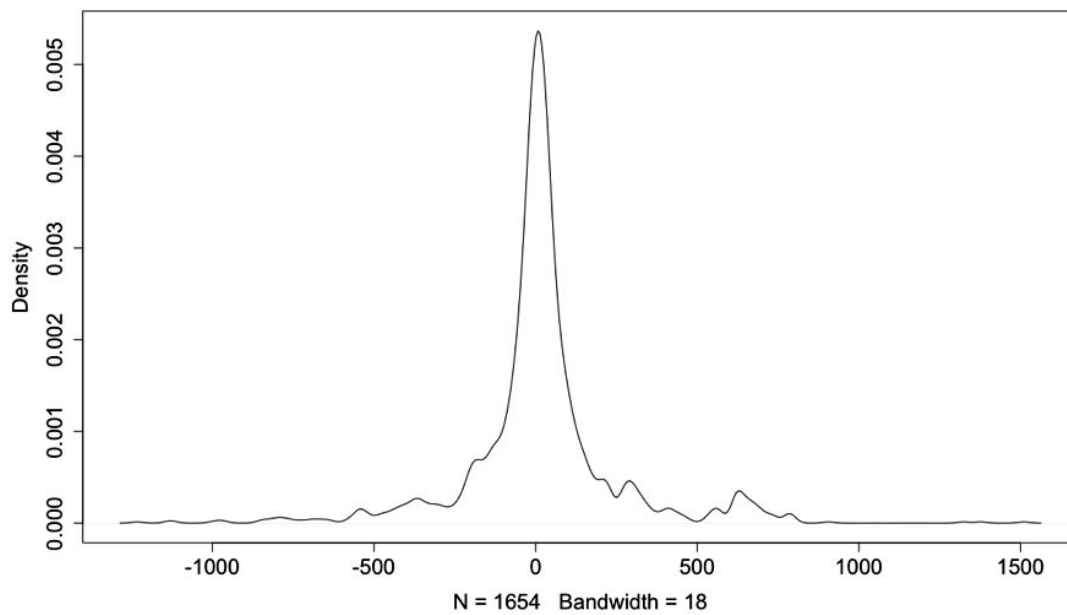
The following data were automatically collected:

- The disorder name (corresponding to b) detected in the message and the corresponding LLT.
- The MedDRA preferred term associated to the disorder term.
- The overall position of the disorder term in the message.
- Relative position of the detected disorder to the product's name (before or after).
- The distance between the disorder term and the drug name.
- Length of the message (expressed in number of words)

When the product name appears several times in a message, the algorithm evaluates the distances between a disorder and all drug name occurrences. The pairs identified are deduplicated. The only pair considered is the one that minimizes the distance to the drug name.

Clustering Method

We used Gaussian mixture models for the disorder clustering using "mclust" R package [37]. We modeled ADRs and other entities as normal distributions mixed on one. The global distribution is obtained by modeling distances calculated for each disorder (Figure 3). EM algorithm is used for model fitting. The affiliation of each type of entity is established by the use of likelihood maximization.

Figure 3. Observed density of distances between disorder terms and drug names.

Results

Descriptive Analysis

We processed a total of 648 messages from 5 French forums written from 2002 to 2013. The named entity recognition module automatically identified 320 unique disorders corresponding to 268 PTs (see [Multimedia Appendix 1](#) for an exhaustive list of disorders identified). Among the 648 messages, 40.9% (265/648) contained drug names but no disorder term. The automatic processing was able to extract 1654 (drug and disorder) pairs from the 383 messages containing at least one disorder term. [Figure 4](#) shows the number of messages consisting of (n_1 , n_2) words. Nine messages contain more than 1000 words.

All 1654 of the identified disorders were manually annotated as true ADRs or not. Among them, the experts identified 11.42%

(189/1654) of ADRs and 88.57% (1465/1654) of other entities. [Figure 5](#) shows the disorders found in the messages grouped at SOC level of MedDRA.

Analysis of Disorder Entity Distribution

As shown in [Figure 3](#), the distribution of the distances between disorder terms and drug names in the messages seems to follow a Gaussian distribution. However, a Shapiro-Wilk normality test significantly rejected the null-hypothesis with a P value of less than $2.2e-16$.

QQ-plot in [Figure 6](#) shows that the data are heterogeneous and can be a mixture of multiple Gaussian distributions [38,39].

The clustering method clusters the detected disorder concepts based on their distances (expressed as a number of words) to the product name in each message. To achieve this goal, we used Gaussian mixture models and EM algorithm [40-42].

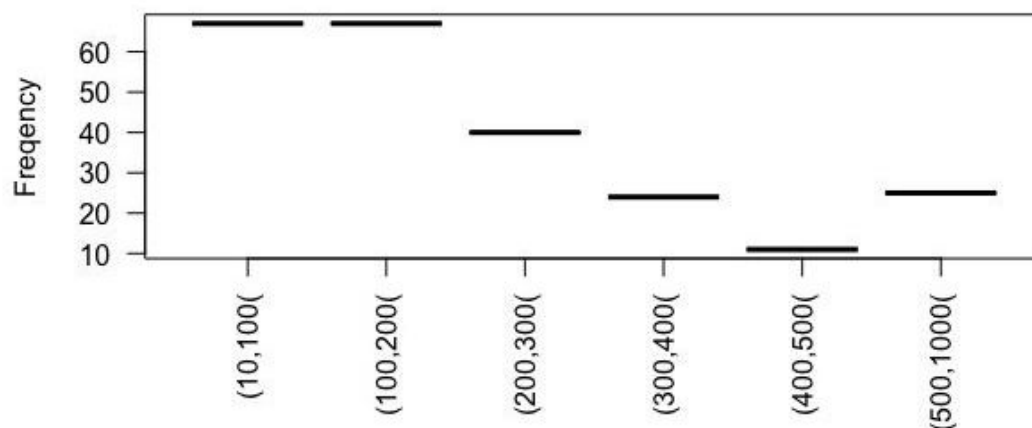
Figure 4. Documents under review found consisting of (n) words.

Figure 5. Disorders automatically identifies (MedDRA system organ class [SOC] level).

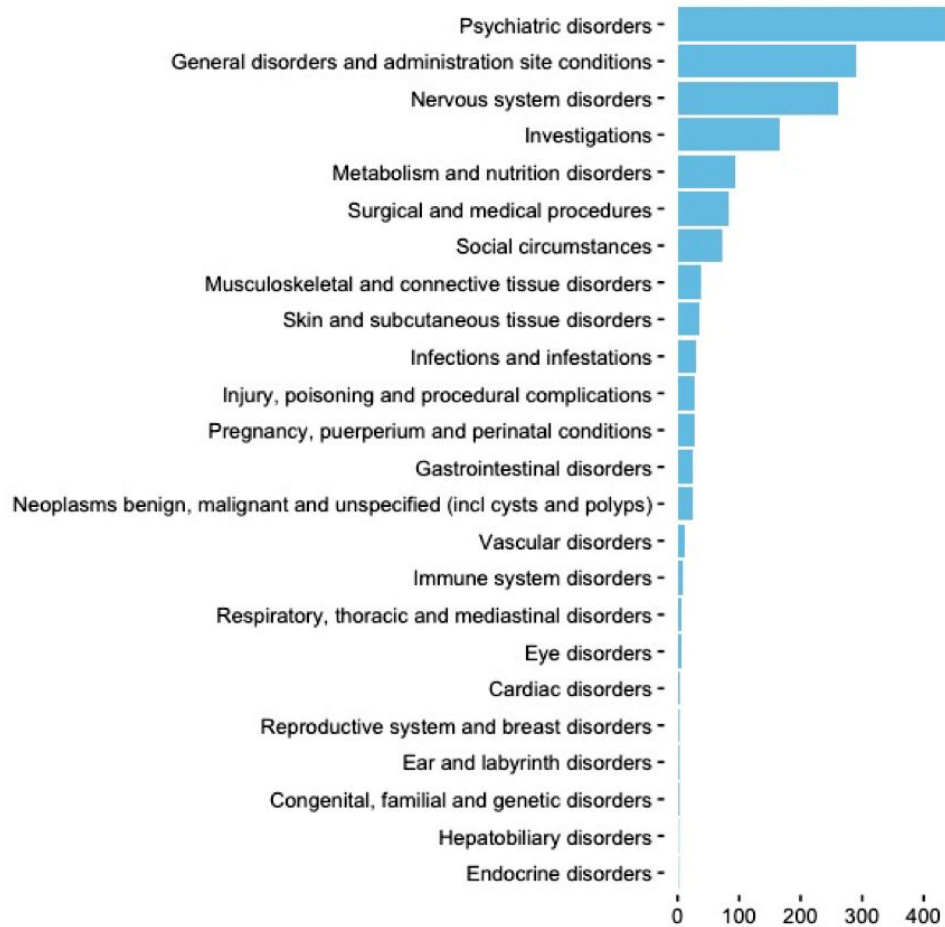


Figure 6. Normal Q-Q plot.

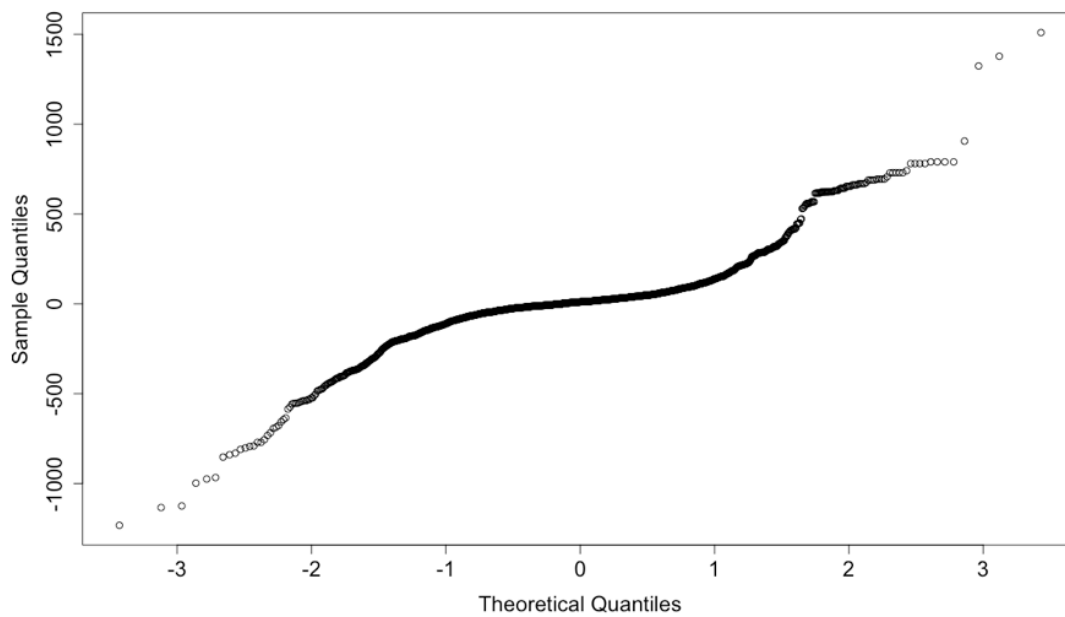
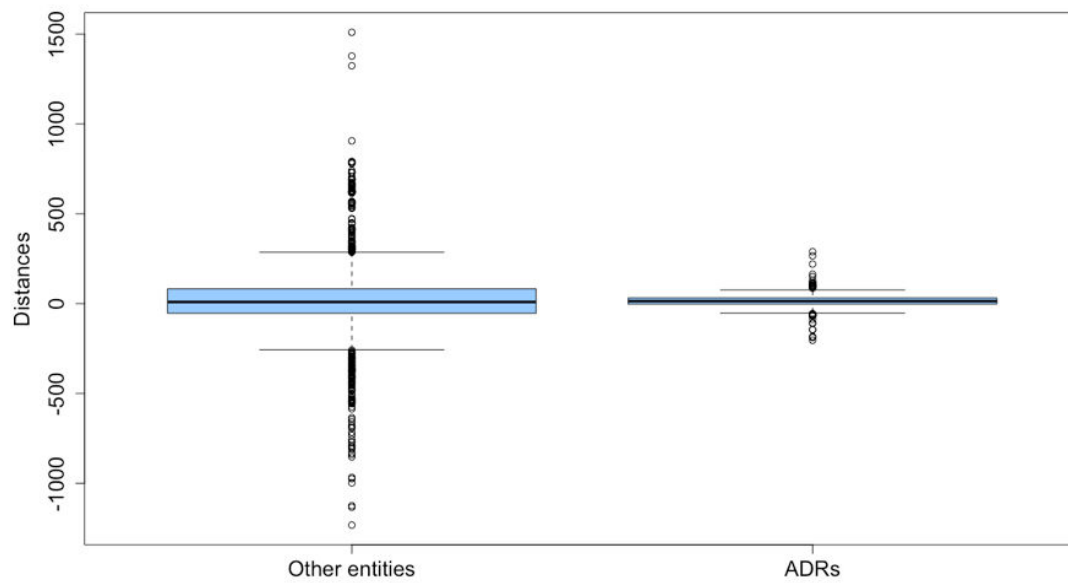
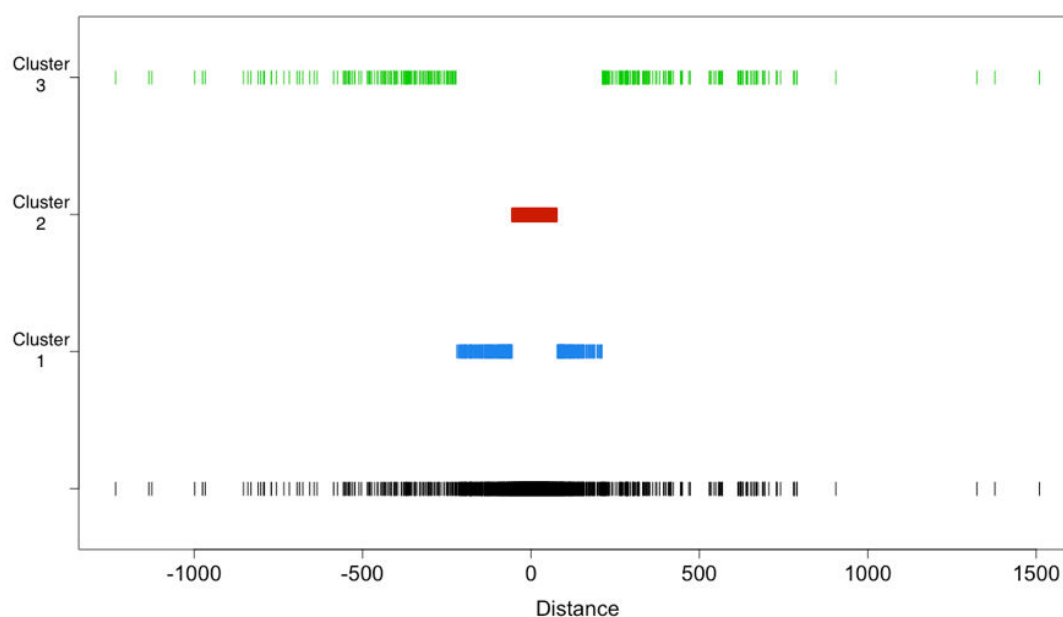


Figure 7. Repartition of true adverse drug reactions (ADRs) and other entities by distances.

Distance distribution also varies greatly with a viewing averaging 20.32 and a median value of 11.0. The distances vary between 1233 before the drug name and 1510 after. [Figure 7](#) shows the concentration of ADRs in a short interval around drug names. The ADRs are contained in an interval between 204 words before product names and 289 after words around drug names. All disorders located beyond 289 words are not ADRs (false positives).

Clusters Analysis

We applied a supervised clustering method with three fixed clusters ([Figure 8](#)).

Figure 8. Supervised clustering results.

Cluster 1 corresponds to distances in the $(-220, -57)$ union $(+78, +211)$ interval, that is, between 220 and 57 words before the drug name or in the interval between 78 and 211 words after the drug name. Cluster 1 contains 441 disorders. Among them, 6.6% (29/441) of the disorders found in cluster, are true ADRs.

Cluster 2 corresponds to distances in the $(-56, -1)$ union $(+2, +77)$ interval (ie, between 56 and 1 words before the drug name or between 2 and 77 words after the drug name). Cluster 2 contains 889 disorders. In cluster 2, 17.7% (157/889) of the disorders are true ADRs.

Table 1. Supervised clustering contingency table.

Clusters	Other entities (%)	ADRs ^a (%)	Total
Cluster 1	412 (93.4)	29 (6.6)	441
Cluster 2	732 (82.3)	157 (17.7)	889
Cluster 3	321 (99.1)	3 (0.9)	324
Total	1465	189	1654

^aADRs: adverse drug reactions.

Cluster 3 corresponds to distances between 1233 and 222 words before the drug name or between 212 and 1510 words after. Cluster 3 contains 324 disorders. Among them, 0.9% (3/324) are ADRs and 321 are other entities.

Filtering Strategies

We tested two filtering strategies. The objective was to filter out the entities that were not ADRs. Table 1 shows supervised clustering contingency.

In the first filtering strategy, we merged clusters 1 and 3 (Table 2). The disorders in clusters 1 and 3 (Table 1) are in the (−1233, −57) union (+78, +1510) interval. The objective of this strategy was to maximize the number of disorders that are not ADRs (412 in cluster 1 and 321 in cluster 3) in one cluster for filtering. The union of these 2 clusters contained only 4.2% (32/765) of ADRs. As shown in Table 2, 95.8% (733/765) of the disorders that are present in the union of clusters 1 and 3 correspond to disorders that are not ADRs (733 disorders).

Table 2. Filtering by merging of clusters 1 and 3.

Clusters	Other entities (%)	ADRs ^a (%)	Total
Clusters 1 and 3	733 (95.8)	32 (4.2)	765
Cluster 2	732 (82.3)	157 (17.7)	889
Total	1465	189	1654

^aADRs: adverse drug reactions.

In the context of ADR detection, the use of this approach to remove disorders of clusters 1 and 3 induces a 50.03% reduction of potential false positives.

The ability to detect false ADRs achieved a precision score of 95.8% and a recall of 50.0%. In other terms, almost all (>95%) of the pairs that were filtered out were not true ADRs, but the system detected only 50.03% (733/1465) of the false positives.

A second filtering strategy involved merging clusters 1 and 2 (Table 3). The main objective of this strategy was to aggregate as many ADRs as possible. We used the union of clusters 1 and 2 (412 disorders in cluster 1 and 732 disorders in cluster 2) and then filtered out the disorders from cluster 3.

Table 3. Filtering by merging clusters 1 and 2.

Clusters	Other entities (%)	ADRs ^a (%)	Total
Cluster 3	321 (99.1)	3 (0.9)	324
Clusters 1 and 2	1144 (86.0)	186 (14.0)	1330
Total	1465	189	1654

^aADRs: adverse drug reactions.

The union of clusters 1 and 2 contains 98.4% (186/189) of the true ADRs present in the dataset. Given that cluster 3 contains only 1.6% (3/189) of the ADRs in our study, exclusion of cluster 3 leads to erroneously ignoring only three relevant adverse events.

Using this filtering strategy, our detection of disorders that are not ADRs achieved a precision of 99.07% and a recall of 21.9%. In other terms, 99.07% (321/324) of the pairs that were filtered out were false positive, but the system detected only 21.91% (321/1465) of the non-ADRs.

Discussion

Principal Findings

We demonstrated that the meaning of a disorder term in a message varies considerably based on its distance to the drug name. Noticeably, before any filtering strategy, cluster 3 contained only three ADRs. The higher the distance between the disorder and the drug name is, the lower the probability that the disorder might be an ADR. Specifically, in cluster 3, 99.1% (321/324) of the disorder terms did not correspond to ADRs. Our approach based on distance measurement enabled us to filter out other (non-ADRs) entities from the detected disorders. The first strategy enabled us to automatically filter out 49.96%

(732/1465) of the disorders that were not ADRs. The second strategy filtered out 78.08% (1144/1465) of the disorders that were not ADRs. Consequently, we obtained a significant improvement in identifying non-ADRs (false positives) in messages. Such filtering can be used as a first step to optimize the screening of ADRs by reducing the false positive rate.

Comparison With Prior Work

Patient's adverse drug event discussions in forums are more informal and colloquial than biomedical literature and clinical notes. When messages in social media are mined to detect ADRs declarations, these informal chats lead to many noisy false positives. The use of filtering methods improves ADR detection in the huge data source that is social media [16,25,27]. Powell et al [25] showed that only 26% of such posts contain relevant information. Even when a message contains both a drug name and a disorder term, the latter may play a role other than an ADR. In our dataset, only 11.42% (189/1654) of the disorder terms corresponded to potential ADRs.

However, the use of distance (as number of words) has not been used for ADR detection, and the usage of this type of information for ADR classification is novel. Sarker and Gonzalez [16] used a leave-one-out classification to evaluate different features for a filtering approach. One of these approaches is based on the n-gram method (accuracy 80.7%), and another approach is based on topic evaluation (accuracy 86.1%). Our approach is different and can be combined with other filtering methods.

One challenge is the comparison of the different filtering methods and their evaluation on equivalent datasets. We evaluated our method on a corpus that was not specific to an adverse event. We relied on MedDRA, which encompasses the complete spectrum of possible ADRs.

Limitations and Future Work

Some limitations regarding the effectiveness of our filtering method should be noted. The main limitation is that our classification process is less efficient when the disorder term is closer to the drug name in the message. Another limitation is that the distance approach has been developed and tested on a French corpus and must be adapted to different languages. Finally, this approach is based on the number of words between drug names and disorder entities in messages and is therefore not applicable to some forms of social media such as Twitter because a tweet would not contain a sufficient number of words to satisfy a sufficient disparity of the disorders detected. The insufficient disparity would not allow our filter to effectively classify the disorders.

Many patients express sentiments when posting about drug associated events in social media, and (quoting Sarker and Gonzalez in [16]) "the sentiments generally correlate strongly with the reactions associated with the drugs they are taking." Combining lexical features from research areas such as sentiment analysis or polarity classification with methods that detect ADRs can improve the automatic classification of ADR mentions from social media text. Moreover, it can help analyzing consumer's perceptions and their changes in time, for example, following media coverage.

Conclusions

We have demonstrated that the distance between the disorder and the drug in a message influences the probability of a disorder to be a genuine ADR. The use of distance between entities on patient posts from social media enabled us to filter out false positives from the detected disorders, and thus, to optimize ADR screening.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Exhaustive list of disorders found at MedDRA preferred terms (PT) and system organ class (SOC) levels.

[[XLSX File \(Microsoft Excel File\), 53KB - publichealth_v3i2e36_app1.xlsx](#)]

References

1. Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, Servi D, et al. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group. JAMA 1995 Jul 05;274(1):29-34. [Medline: [7791255](#)]
2. Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality. JAMA 1997;277(4):301-306. [Medline: [9002492](#)]
3. Harpaz R, DuMouchel W, Shah N, Madigan D, Ryan P, Friedman C. Novel data-mining methodologies for adverse drug event discovery and analysis. Clin Pharmacol Ther 2012 Jun;91(6):1010-1021 [FREE Full text] [doi: [10.1038/clpt.2012.50](#)] [Medline: [22549283](#)]
4. 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. USA: Association for Computational Linguistics; 2010 Presented at: BioNLP '10 Proceedings of the 2010 Workshop on Biomedical Natural Language Processing; July 15, 2010; Uppsala, Sweden p. 117-125 URL: <http://dl.acm.org/citation.cfm?id=1869961.1869976>

5. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, Lau Annie Y S. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. *J Am Med Inform Assoc* 2015 Jan;22(1):243-256. [doi: [10.1136/amiajnl-2014-002841](https://doi.org/10.1136/amiajnl-2014-002841)] [Medline: [25005606](https://pubmed.ncbi.nlm.nih.gov/25005606/)]
6. van Hunsel F, van der Welle C, Passier A, van Puijenbroek E, van Grootheest K. Motives for reporting adverse drug reactions by patient-reporters in the Netherlands. *Eur J Clin Pharmacol* 2010 Nov;66(11):1143-1150 [FREE Full text] [doi: [10.1007/s00228-010-0865-7](https://doi.org/10.1007/s00228-010-0865-7)] [Medline: [20658130](https://pubmed.ncbi.nlm.nih.gov/20658130/)]
7. van Grootheest K, de Graaf L, de Jong-van den Berg LT. Consumer adverse drug reaction reporting: a new step in pharmacovigilance? *Drug Saf* 2003;26(4):211-217. [Medline: [12608885](https://pubmed.ncbi.nlm.nih.gov/12608885/)]
8. Gittelman S, Lange V, Gotway CC, Okoro CA, Lieb E, Dhingra SS, et al. A new source of data for public health surveillance: facebook likes. *J Med Internet Res* 2015;17(4):e98 [FREE Full text] [doi: [10.2196/jmir.3970](https://doi.org/10.2196/jmir.3970)] [Medline: [25895907](https://pubmed.ncbi.nlm.nih.gov/25895907/)]
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. Golder S, Norman G, Loke YK. Systematic review on the prevalence, frequency and comparative value of adverse events data in social media. *Br J Clin Pharmacol* 2015 Oct;80(4):878-888. [doi: [10.1111/bcp.12746](https://doi.org/10.1111/bcp.12746)] [Medline: [26271492](https://pubmed.ncbi.nlm.nih.gov/26271492/)]
11. Lardon J, Abdellaoui R, Bellet F, Asfari H, Souvignet J, Texier N, et al. Adverse drug reaction identification and extraction in social media: a scoping review. *J Med Internet Res* 2015 Jul 10;17(7):e171 [FREE Full text] [doi: [10.2196/jmir.4304](https://doi.org/10.2196/jmir.4304)] [Medline: [26163365](https://pubmed.ncbi.nlm.nih.gov/26163365/)]
12. Benton A, Ungar L, Hill S, Hennessy S, Mao J, Chung A, et al. Identifying potential adverse effects using the web: a new approach to medical hypothesis generation. *J Biomed Inform* 2011 Dec;44(6):989-996 [FREE Full text] [doi: [10.1016/j.jbi.2011.07.005](https://doi.org/10.1016/j.jbi.2011.07.005)] [Medline: [21820083](https://pubmed.ncbi.nlm.nih.gov/21820083/)]
13. Yang CC, Yang H, Jiang L, Zhang M. Social media mining for drug safety signal detection. 2012 Presented at: Proc Int Workshop Smart Health Wellbeing; October 29, 2012; Maui, HI p. 33-40.
14. Yates A, Goharian N. ADTrace: detecting expected and unexpected adverse drug reactions from user reviews on social media sites. In: *Advances in Information Retrieval*. Berlin, Heidelberg: Springer; 2013:816-819.
15. Jiang K, Zheng Y. Mining twitter data for potential drug effects. In: Motoda H, Wu Z, Cao L, Zaiane O, Yao M, Wang W, editors. *Advanced Data Mining and Applications*. Berlin, Heidelberg: Springer; 2013:434-443.
16. Sarker A, Gonzalez G. Portable automatic text classification for adverse drug reaction detection via multi-corpus training. *J Biomed Inform* 2015 Feb;53:196-207 [FREE Full text] [doi: [10.1016/j.jbi.2014.11.002](https://doi.org/10.1016/j.jbi.2014.11.002)] [Medline: [25451103](https://pubmed.ncbi.nlm.nih.gov/25451103/)]
17. Yeleswarapu S, Rao A, Joseph T, Saipradeep VG, Srinivasan R. A pipeline to extract drug-adverse event pairs from multiple data sources. *BMC Med Inform Decis Mak* 2014;14:13 [FREE Full text] [doi: [10.1186/1472-6947-14-13](https://doi.org/10.1186/1472-6947-14-13)] [Medline: [24559132](https://pubmed.ncbi.nlm.nih.gov/24559132/)]
18. Joachims T. Text categorization with support vector machines: learning with many relevant features. In: Nédellec C, Rouveirol C, editors. *Machine Learning: ECML-98*. Berlin, Heidelberg: Springer; 1998:137-142.
19. Abou Taam M, Rossard C, Cantaloube L, Bouscaren N, Roche G, Pochard L, et al. Analysis of patients' narratives posted on social media websites on benfluorex's (Mediator®) withdrawal in France. *J Clin Pharm Ther* 2014 Feb;39(1):53-55. [doi: [10.1111/jcpt.12103](https://doi.org/10.1111/jcpt.12103)] [Medline: [24304185](https://pubmed.ncbi.nlm.nih.gov/24304185/)]
20. Harpaz R, Callahan A, Tamang S, Low Y, Odgers D, Finlayson S, et al. Text mining for adverse drug events: the promise, challenges, and state of the art. *Drug Saf* 2014 Oct;37(10):777-790 [FREE Full text] [doi: [10.1007/s40264-014-0218-z](https://doi.org/10.1007/s40264-014-0218-z)] [Medline: [25151493](https://pubmed.ncbi.nlm.nih.gov/25151493/)]
21. 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/)]
22. Kappa Santé. URL: <https://www.kappasante.com/> [accessed 2017-06-03] [WebCite Cache ID 6qwWipp9j]
23. Sloane R, Osanlou O, Lewis D, Bollegala D, Maskell S, Pirmohamed M. Social media and pharmacovigilance: a review of the opportunities and challenges. *Br J Clin Pharmacol* 2015 Oct;80(4):910-920 [FREE Full text] [doi: [10.1111/bcp.12717](https://doi.org/10.1111/bcp.12717)] [Medline: [26147850](https://pubmed.ncbi.nlm.nih.gov/26147850/)]
24. Liu X, Chen H. A research framework for pharmacovigilance in health social media: identification and evaluation of patient adverse drug event reports. *J Biomed Inform* 2015 Dec;58:268-279 [FREE Full text] [doi: [10.1016/j.jbi.2015.10.011](https://doi.org/10.1016/j.jbi.2015.10.011)] [Medline: [26518315](https://pubmed.ncbi.nlm.nih.gov/26518315/)]
25. Powell GE, Seifert HA, Reblin T, Burstein PJ, Blowers J, Menius JA, et al. Social media listening for routine post-marketing safety surveillance. *Drug Saf* 2016 May;39(5):443-454. [doi: [10.1007/s40264-015-0385-6](https://doi.org/10.1007/s40264-015-0385-6)] [Medline: [26798054](https://pubmed.ncbi.nlm.nih.gov/26798054/)]
26. Blei D. Probabilistic topic models. *Commun ACM* 2012;55(4):84.
27. Yang M, Kiang M, Shang W. Filtering big data from social media--building an early warning system for adverse drug reactions. *J Biomed Inform* 2015 Apr;54:230-240 [FREE Full text] [doi: [10.1016/j.jbi.2015.01.011](https://doi.org/10.1016/j.jbi.2015.01.011)] [Medline: [25688695](https://pubmed.ncbi.nlm.nih.gov/25688695/)]
28. Canini K, Shi L, Griffiths T. MLR. 2009. Online inference of topics with latent dirichlet allocation URL: <http://proceedings.mlr.press/v5/canini09a/canini09a.pdf> [accessed 2017-06-09] [WebCite Cache ID 6r5kzf1P]
29. Bian J, Topaloglu U, Yu F. Towards Large-scale Twitter Mining for Drug-related Adverse Events. 2012 Presented at: Proc Int Workshop Smart Health Wellbeing; October 29 - 29, 2012; Maui, Hawaii, USA p. 25-32.

30. Wei CH, Peng Y, Leaman R, Davis AP, Mattingly CJ, Li J, et al. Assessing the state of the art in biomedical relation extraction: overview of the BioCreative V chemical-disease relation (CDR) task. Database (Oxford) 2016;2016:baw032 [FREE Full text] [doi: [10.1093/database/baw032](https://doi.org/10.1093/database/baw032)] [Medline: [26994911](https://pubmed.ncbi.nlm.nih.gov/26994911/)]
31. Gupta V, Lehal G. A survey of text mining techniques and applications. JETWI 2009 Aug 1;1(1):60-76 [FREE Full text]
32. Porter MF. Toronto.edu. 1980. An algorithm for suffix stripping URL: http://www.cs.toronto.edu/~frank/csc2501/Readings/R2_Porter/Porter-1980.pdf [accessed 2017-06-09] [WebCite Cache ID 6r5IONIWv]
33. Drazic M, Kukolj D, Vitas M, Pokric M, Manojlovic S, Tekic Z. Technology Matching of the Patent Documents Using Clustering Algorithms. New York: Ieee; 2013.
34. Iorio F, Bosotti R, Scacheri E, Belcastro V, Mithbaokar P, Ferriero R. Discovery of drug mode of action and drug repositioning from transcriptional responses. PNAS 2010;107(33):14621-14626. [doi: [10.1073/pnas.1000138107](https://doi.org/10.1073/pnas.1000138107)]
35. Feinerer I, Hornik K, Meyer D. Text mining infrastructure in R. J Stat Softw 2008;25(5):54. [doi: [10.18637/jss.v025.i05](https://doi.org/10.18637/jss.v025.i05)]
36. Savoy J. Light stemming approaches for the French, Portuguese, German and Hungarian languages. USA: ACM; 2006 Presented at: SAC '06 Proceedings of the 2006 ACM symposium on Applied computing; April 23-27, 2006; Dijon, France p. 1031-1035. [doi: [10.1145/1141277.1141523](https://doi.org/10.1145/1141277.1141523)]
37. Fraley C, Raftery A. Model-based methods of classification: using the mclust software in chemometrics. J Stat Softw 2007;18(6):1-13. [doi: [10.18637/jss.v018.i06](https://doi.org/10.18637/jss.v018.i06)]
38. Choi B. A graphical method to assess goodness-of-fit for inverse gaussian distribution. Korean J Appl Stat 2013;26(1):37-47.
39. Grego J, Yates P. Point and standard error estimation for quantiles of mixed flood distributions. J Hydrol 2010 Sep;391(3-4):289-301. [doi: [10.1016/j.jhydrol.2010.07.027](https://doi.org/10.1016/j.jhydrol.2010.07.027)]
40. Torres-Carrasquillo P, Reynolds D, Deller JJ. Language identification using Gaussian mixture model tokenization. : IEEE; 2002 Presented at: IEEE Int Conf Acoust Speech Signal Process ICASSP; May 13-17, 2002; Orlando, FL. [doi: [10.1109/ICASSP.2002.5743828](https://doi.org/10.1109/ICASSP.2002.5743828)]
41. Torres-Carrasquillo P, Singer E, Kohler M, Greene R, Reynolds D, Deller JJ. Approaches to language identification using Gaussian mixture models and shifted delta cepstral features. 2002 Presented at: 7th International Conference on Spoken Language Processing, ICSLP 2002; September 16, 2002; Denver, CO p. 89-92.
42. Dempster A, Laird N, Rubin D. Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc Ser B Methodol 1977;1:38.

Abbreviations

ACTC: active semisupervised clustering based two-stage text classification

ADR: adverse drug reaction

AE: adverse event

EAT: example adaption for text categorization

EM: expectation-maximization

FAERS: FDA's Adverse Event Reporting System

HLGT: high-level group terms

HLT: high-level terms

LLT: lowest-level terms

ME: Maximum Entropy

MedDRA: Medical Dictionary for Regulatory Activities

MeSH: medical subject headings

NB: Naïve Bayes

PNLH: positive examples and negative examples labeling heuristics

PT: preferred terms

SOC: system organ class

SVM: support vector machine

Edited by G Eysenbach; submitted 01.09.16; peer-reviewed by GE Powell, N Dasgupta, S Sahay, L Wang, Y Peng; comments to author 05.12.16; revised version received 24.01.17; accepted 15.02.17; published 22.06.17.

Please cite as:

Abdellaoui R, Schück S, Texier N, Burgun A

Filtering Entities to Optimize Identification of Adverse Drug Reaction From Social Media: How Can the Number of Words Between Entities in the Messages Help?

JMIR Public Health Surveill 2017;3(2):e36

URL: <http://publichealth.jmir.org/2017/2/e36/>

doi: [10.2196/publichealth.6577](https://doi.org/10.2196/publichealth.6577)

PMID: [28642212](https://pubmed.ncbi.nlm.nih.gov/28642212/)

©Redhouane Abdellaoui, Stéphane Schück, Nathalie Texier, Anita Burgun. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 22.06.2017. 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 and Sexual Behavior Among Adolescents: Is there a link?

Megan Landry^{1*}, MPH, DrPH; Monique Turner^{1*}, PhD; Amita Vyas^{1*}, PhD; Susan Wood^{2*}, PhD

¹Milken Institute School of Public Health, Department of Prevention and Community Health, The George Washington University, Washington, DC, United States

²Milken Institute School of Public Health, Department of Health Policy, The George Washington University, Washington, DC, United States

* all authors contributed equally

Corresponding Author:

Megan Landry, MPH, DrPH

Milken Institute School of Public Health

Department of Prevention and Community Health

The George Washington University

3rd Floor

950 New Hampshire Ave NW

Washington, DC, 20052

United States

Phone: 1 2029943695

Email: mmlandry@gwu.edu

Abstract

Background: Adolescent sexual risk taking and its consequences remain a global public health concern. Empirical evidence on the impact that social media has on sexual health behaviors among youth is sparse.

Objective: The study aimed to examine the relationship between social media and the change in sexual risk over time and whether parental monitoring moderates this relationship.

Methods: This study comprised a sample of 555 Latino youth aged 13-19 years from Maryland, United States completing baseline and follow-up surveys. Mixed-effects linear regression was used to examine the relationship between social media and the change in sexual risk over time and whether parental monitoring moderated the relationship.

Results: Sexual risk behaviors significantly increased between baseline (T1) and follow up (T2) (mean=0.432 vs mean=0.734, $P<.001$). Youth sending more than 100 text messages per day had significantly higher sexual risk scores (beta=1.008, $P<.001$) but significantly larger declines in sexual risk scores for higher levels of parental monitoring (beta=-.237, $P=.009$).

Conclusions: Although adolescents exchange SMS at high rates, parental monitoring remains vital to parent-child relationships and can moderate SMS frequency and sexual risk behaviors, despite parental influence diminishing and peer pressure and social influences increasing during adolescence.

(*JMIR Public Health Surveill* 2017;3(2):e28) doi:[10.2196/publichealth.7149](https://doi.org/10.2196/publichealth.7149)

KEYWORDS

text messaging; social media; parent-child relations; sexual behavior; adolescent

Introduction

Adolescent sexual risk taking and its consequences remain a global public health concern. Risky sexual behaviors may lead to increased likelihood of sexually transmitted infections (STIs) and unintended pregnancies [1-3]. Adolescence, defined as 13-19 years of age, is a phase of rapid physical, emotional, and cognitive development [4]. This period is marked by an increased importance on social relationships when youth are

focused on developing a sense of self and personal identity [5]. Given the importance of social relationships and the inability to fully control impulsive behaviors during adolescence [6], there may be some concern about the role social media has in adolescents' lives.

Adolescents use mobile phones and the Web to interact with both known and unknown peers to establish and maintain social connections [7]. Such communication platforms are free of locale and time and are relatively easy to use, making these

interactions a new way to foster the development of adolescents' identity, self-expression, intimate relationships, and social well-being [8-10]. Compared with adults older than 25 years, youth between 12 and 24 years of age are the most extensive users of new technology and are more likely to be connected to the virtual world, regardless of socioeconomic status (SES), race, or ethnicity [11,12]. Social media platforms such as short message service (SMS, or texting) and social networking sites (SNS) allow self-expression, intimacy, and privacy for adolescents [10,13]. Users are able to set their own preferences to convey messages about their social identity, in the same manner that face-to-face interactions allow, but on a global scale and in contexts that are not always monitored by adults [14].

Indeed, social media is a promising channel to deliver health information, including health promotion and disease prevention messages [15,16]. However, others suggest that Internet and social media platforms might also have negative health consequences due to a false belief of privacy leading to more provocative behavior and discussion around drinking, sex, violence, suicide ideation, and bullying, coupled with less parental monitoring [9,17]. The American Academy of Pediatrics Council on Communications and Media has also argued that although social media may facilitate socialization and communication, enhance learning opportunities, and increase access to health information, it may also lead to cyber bullying or harassment, sexting, and depression [18].

Yet, there is little empirical evidence on the impact social media use has on sexual health behaviors. Landry et al [19] concluded that Latino adolescents who sent or received more than 100 SMS per day were significantly more likely to ever have vaginal sex and adolescents who logged in to a social networking account at least once per day were significantly more likely to ever have vaginal sex. Their findings are consistent with Frank [20], who reported a relationship between excessive technology use among teens and increased health risk behaviors and poorer perceived health. Teens who hyper-text (ie, send or receive more than 120 messages per day) and hyper-network (ie, 3 or more hours on social sites per day) were much more likely to be involved with unhealthy uses of technology. In fact, Frank [20] reported that 75.8% of texters and 72% of social networkers sending messages or photos that they would not want their parents to see, while 56.4% admitted to using texting or social networking to find a place to gather without parental supervision, for example, to drink alcohol (41.5%) or to meet for sex (27.4%). Additionally, Frank [20] reported that minorities, children of parents with less education, and teenagers from homes without a father were more likely to engage in hypertexting and hyper networking.

One protective factor in reducing sexual risk behaviors during adolescence is parental monitoring [21-24]. Wight, Williamson, and Henderson's [25] longitudinal study suggested that low parental monitoring predicted early sexual activity for both sexes, and, for females, it also predicted more sexual partners and less condom use. Higher levels of parental knowledge of adolescent's whereabouts delayed sexual onset, especially among girls [26]. Still, other investigators have linked higher parental support to a delay in sexual debut for both girls and boys [27].

Although youth are the most extensive users of new technology and are more likely to be virtually connected, regardless of SES, race, or ethnicity, there are racial and ethnic disparities in the prevalence of sexual risky behavior. For example, Latino youth often engage in riskier sexual behaviors than their White counterparts. Compared to non-Latino Whites, Latino youth have higher rates of STIs, including chlamydia, gonorrhea, and syphilis [28]. Although condom use at most recent sexual intercourse has been on the rise among adolescents (ie, from 46% in 1991 to 59% in 2013), sexually active Latino adolescents were more likely than both White and Black adolescents to not have used a condom or birth control during their last sexual intercourse [29].

Additionally, Latino youth are just as likely as their White and Black counterparts to be extensive social media users [11,30,31], so we would expect an association between social media and sexual behavior among the larger population of youth. However, the Latino youth population are unique in that they face higher sexual risk behaviors and additional risk factors such as poverty, acculturative stress, and familial and cultural barriers that potentially segregate them from the larger society and put them at even greater risk of negative behaviors [12,32,33]. Given the limited but growing body of research surrounding the impact of social media on sexual health, investigating these relationships is important for public health practice and reducing negative health outcomes.

This study used two rounds of data from a longitudinal study of Latino youth to investigate sexual risk behavior over time. We questioned whether the rate of change in sexual behavior was related to social media utilization and frequency of use over the same time. Also, it was hypothesized that parental monitoring moderates the relationship between sexual risk behavior and social media utilization, such that when parental monitoring is higher, the association between social media utilization and sexual risk behavior will be weaker, relative to when parental monitoring is lower.

Methods

Participants

The data for this study were derived from a sample of self-identifying Latino adolescents aged 13-19 years (mean=15.73, SD=1.03). Participants were recruited from 12 public high schools in Maryland, United States. Participants completed baseline (T1) and 16-month follow-up (T2) surveys conducted as part of a program evaluation of the *Empowering Latino Youth Project* (ELYP) between spring 2012 and fall 2013. ELYP is a 5-year cluster-randomized controlled trial of a teen pregnancy prevention program. All participants provided parental consent and youth assent to participate in ELYP. Due to the data being from an intervention study, final analyses statistically controlled for participation in the intervention or control group. The analytic sample for this study is a subsample of the entire study sample (555/873) enrolled between spring 2012 and fall 2013. Participants who completed the baseline but no follow-up survey were 25.7% (224/873) and those completing only the follow-up survey were 10.8% (94/873). Participants lost to follow up were more likely to be male,

slightly older, born outside of the United States, and completed the baseline survey in Spanish. Participants lost to follow up also had statistically significantly higher sexual risk scores at baseline. Excluding participants lost to follow up from the analytic sample is underestimating the sexual risk score over time, therefore, providing a more conservative estimate.

Data Collection

To ensure privacy and reduce reporting bias, surveys were administered via individual laptops with audio capability for youth with low-literacy levels. Study participants chose to complete the survey in English or Spanish and were given US \$10 gift cards for completing the baseline survey and US \$20 gift cards for the 16-month survey. The surveys were translated and back translated by native Spanish speakers affiliated with the partner community organization then pretested for readability and accuracy. Upon survey completion, the data were stored in an encrypted file to be read only by the survey design software, SNAP surveys.

Measures

Social Media Use

Social media use includes SMS, Internet, and social media questions adapted from Pew Internet Project's Teen Survey [34]. Participants were asked if they have a mobile phone, if they use SMS, and the frequency of SMS per day (high SMS: >100 per day; low SMS: ≤100 per day). Participants with a mobile phone reported the following behaviors using their phone: send or receive email, take pictures, play music, send or receive instant messages, record videos, play games, or access Internet. SMS frequency was dichotomized at 100 based on Pew data that suggest the median number of SMS per day for Hispanic adolescents is 100 [7,34]. Additionally, results from one-way ANOVA suggested the two highest texting categories (ie, 101-200 per day and more than 200 per day) had significantly different sexual risk scores compared with four lower categories ($F_3=10.36, P<.001$). Participants reported how often they exchange SMS with friends, parents, and a boyfriend or girlfriend (1 = less often or never, 2 = a few times a week, 3=at least once a day, 4=several times a day), which was dichotomized into at least once per day versus less often.

Participants were asked if they use the Internet, for what purposes, and how often (0=never, 6=several times a day). Finally, participants were asked if they had accounts on any of the following SNS and a count variable was created, such as Facebook, MySpace, Twitter, Yahoo, YouTube, Instagram, Tumblr, Google Buzz, Flickr, Ustream, and other. Those with any account were asked about their frequency of logging in, which was dichotomized into daily login versus less frequent.

Parental Monitoring

The parental monitoring scale was adapted from Silverberg and Small [35] and validated in a Positive Youth Development Survey for Latinos [36]. The scale consisted of five items measuring participants' perceptions of their parent/guardian's knowledge of their whereabouts, making decisions that affect them, and seeking help or encouragement from a parent/guardian. The scale was measured on a 4-point

Likert-type scale (1=never, 2=some of the time, 3=most of the time, 4=all of the time). Internal consistency was high at both survey time points (Cronbach alpha = .85 at baseline and .86 at follow-up).

Sexual Risk Behavior

The sexual risk behavior measure used in this study was constructed based on the Centers for Disease Control's (CDC) description of sexual risk taking [29]. The sexual behavior questions were adapted from validated measures from the US Department of Health and Human Services, Office of Adolescent Health [37] and the CDC Youth Risk Behavior Surveillance Survey [29]. The six measured variables were as follows: ever had vaginal sex (0=no; 1=yes), condom use in the last 3 months (0=used condom; 1=did not use condom), contraception use in the last 3 months (0=used contraception; 1=did not use contraception), 2 or more sexual partners in the last 3 months (0=less than 2 partners; 1=2 or more partners), individual alcohol use with sex in the last 3 months (0=no individual alcohol use; 1=any individual alcohol use), and partner alcohol use with sex in the last 3 months (0=no partner alcohol use; 1=any partner alcohol use). The individual sexual behaviors were summed to create a composite score of sexual risk behavior (baseline mean=0.425, SD=0.044, range 0-6), where higher scores indicated increased risk behavior.

Statistical Analyses

Preliminary analyses examined frequencies and distributions of the variables of interest for the analytic sample (n=555). We first examined bivariate relationships between sexual risk behavior and mobile phone and social media use. In bivariate analyses (data not shown), there was no relationship between sexual risk behavior and access to a mobile phone, SMS with friends, or logging in to SNS once per day; so these variables were dropped from multivariate models.

Mixed-effects linear regression was used to examine the relationship between social media variables and the change in sexual risk between baseline and follow-up, while adjusting for time-varying and time-invariant covariates and allowing random effects for within and between subjects. Additionally, we examined if the relationship between social media and sexual risk was moderated at different levels of parental monitoring by entering an interaction term into the multivariate model. Final models were adjusted for gender, age, survey language, and intervention group. All analyses were conducted in STATA 12.0 (StataCorp). This study was reviewed and approved by the George Washington University Internal Review Board (IRB# 011217).

Results

Main Findings

Table 1 lists self-reported baseline demographic characteristics of the study sample. Participants were on average 15 years of age (mean=15.73, SD=1.03) and there were slightly more female participants (58.6%) than male participants (41.4%). The majority of respondents were in ninth grade at baseline (72.8%) and completed the survey in English (70.8%), while less than half were born in the United States (48%). On a scale from 1

to 4, participants reported above average levels of parental monitoring (mean=3.06, SD =0.748, range 1-4).

Sexual risk behaviors significantly increased between baseline (T1) and follow up (T2) (mean=0.432 vs mean=0.734, $P<.001$; Table 2). In terms of social media utilization, nearly 90% of participants used a mobile phone, SMS, Internet, and SNS at

both T1 and T2. The mean number of activities on a mobile phone significantly increased between surveys (mean=5.90 vs mean=6.30, $P<.001$), while high-frequency SMS (ie, >100 per day) significantly decreased (34.0% vs 27.9%, $P=.02$). In terms of SNS accounts, Facebook use decreased over time (82.9% vs 73.0%, $P<.001$) and YouTube and Instagram use significantly increased.

Table 1. Study sample characteristics at baseline.

Characteristics	(N=555) n (%)	Mean (SD)
Age		15.73 (1.0)
Gender		
Male	230 (41.4)	
Female	325 (58.6)	
Grade		
9th	404 (72.8)	
10th	151 (27.2)	
Survey language		
English	393 (70.8)	
Spanish	162 (29.2)	
Length of time in the United States		
US born (0)	268 (48.3)	
0-3 years (1)	136 (24.5)	
4-10 years (2)	102 (18.4)	
10+ years (3)	27 (4.9)	
Missing	22 (3.9)	
Parental Monitoring (alpha =.85) range 1-4		3.058 (0.7)

Mixed-Effect Models for Sexual Risk From Baseline to Follow-Up

The results of mixed-effects regression analyses are presented in Table 3. The unconditional means model was estimated to calculate intraclass correlation (ICC). Results showed a statistically significant sexual risk behavior score ($\gamma_{00}=.584$, $P<.001$), and participants' mean sexual risk scores (ie, the average score across both assessments) significantly varied around the mean sexual risk score (between subjects $\tau_{00}=.771$, $P<.001$), as well as significant differences between each participants' observed and predicted scores over time ($\sigma^2=.737$, $P<.001$). Further, ICC calculations suggest that 49.4% of sexual risk behavior scores varied across students. This ICC value is consistent with research that suggests ICC values exceeding .40 are common in longitudinal social science studies [38]. The crude estimated sexual risk score at baseline ($\gamma_{00}=.432$, $P<.001$) and the change in sexual risk score over the 16-month time period ($\gamma_{10}=.303$, $P<.001$) are presented in Model 1 of Table 3. In the first multivariate model (Model 2), only variables that were significantly associated with sexual risk in bivariate

analyses were included. Results from Model 2 indicate that high SMS (beta=.384, $P<.001$) and SMS to a boyfriend or girlfriend once per day (beta=.160, $P=.01$) were associated with increased sexual risk scores, controlling for age, gender, survey language, and intervention group. Higher levels of parental monitoring were significantly associated with lower reported sexual risk behavior over time (beta=-.140, $P=.01$).

In the second set of multivariate analyses (Model 3), we extended Model 2 to include an interaction term for high SMS and parental monitoring, along with significant predictors from Model 2. Results from Model 3 indicated that, on average, sexual risk behaviors increased over time and were significantly higher for males and older youth. Further, parental monitoring interacted with high SMS. The negative interaction was graphed (Figure 1) and indicated that higher levels of parental monitoring were related to a weaker association between high SMS and sexual risk. Youth that sent more than 100 SMS per day had significantly higher sexual risk scores (beta=1.008, $P<.001$), but also experienced significantly larger declines in sexual risk scores for higher levels of parental monitoring (beta=-.237, $P=.009$).

Table 2. Unadjusted changes in sexual risk, social media utilization, and parental monitoring at baseline (T1) and 16-month follow up (T2).

Variables of interest and response categories	T1 % (n)	T2 % (n)	Δ T1, T2
Sexual risk composite score mean (SD) (range 0-6)	0.425 (0.97)	0.733 (1.13)	0.31 ($P<.001$)
Social media use and behavior			
Mobile phone access (yes)	88.7 (488)	92.0 (494)	
Mean number of activities on mobile phone (SD)	5.90 (1.61)	6.30 (1.26)	0.4 ($P<.001$)
SMS use			
More than 100 SMS per day	95.7 (467)	98.6 (486)	2.88 ($P=.007$)
SMS Parents at least once per day	34.2 (149)	27.9 (135)	-6.13 ($P=.02$)
SMS Parents at least once per day	53.4 (239)	54.1 (260)	
SMS Friends at least once per day	87.6 (403)	86.0 (418)	
SMS boy/girlfriend at least once per day	64.8 (283)	68.7 (322)	
Internet use			
Using Internet once per day	96.9 (533)	98.5 (529)	
Social networking account			
Facebook	82.9 (458)	73.0 (390)	-9.94 ($P<.001$)
Twitter	58.7 (324)	55.8 (298)	
YouTube	61.8 (341)	66.9 (357)	
Instagram	14.8 (82)	27.4 (119)	12.65 ($P<.001$)
Mean number of activities on SNS (SD)			
Logging in to SNS one or more times per day	5.08 (0.08)	5.02 (0.09)	
Logging in to SNS one or more times per day	79.1 (417)	82.1 (426)	
Mean parental monitoring score (SD)	3.05 (0.75)	3.00 (0.75)	

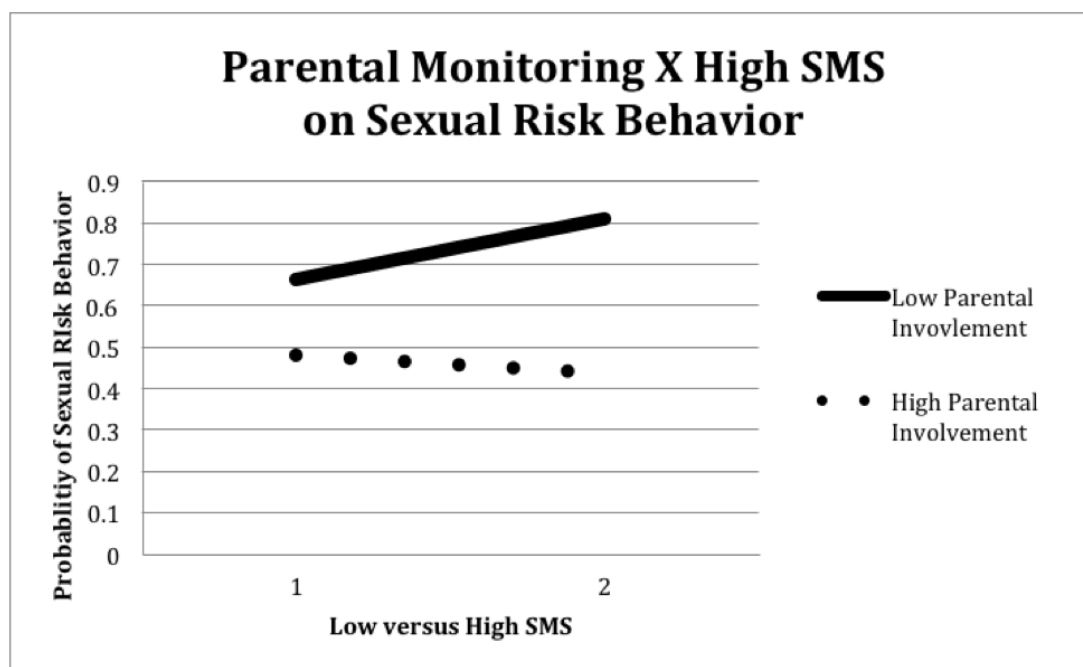
Figure 1. Interaction of parental monitoring and high SMS (ie, >100/day) on sexual risk behavior.

Table 3. Parameter estimates from mixed-effects models for change in sexual risk from T1 to T2.

Variables	Unconditional model Beta (SE)	Model 1: level 1 Beta (SE)	Model 2 ^a : main effects model Adjusted beta (SE)	Model 3 ^a : interaction model Adjusted beta (SE)
Fixed effect				
Intercept (γ_{00})	0.584 (0.040) ($P<.001$)	0.432 (0.045) ($P<.001$)	0.801 (0.263) ($P=.004$)	0.758 (0.256) ($P=.003$)
Time (γ_{10})		0.303 (0.046) ($P<.001$)	0.256 (0.061) ($P<.001$)	0.252 (0.058) ($P<.001$)
More than 100 SMS per day			0.384 (0.080) ($P<.001$)	1.088 (0.300) ($P<.001$)
SMS to whom				
Parents			-0.084 (0.074)	-
Boyfriend or girlfriend			0.160 (0.080) ($P=.01$)	0.148 (0.079)
Number of SNS accounts			0.018 (0.024)	-
Number of SNS activities performed			0.028 (0.024)	-
Parental monitoring			-0.140 (0.053) ($P=.006$)	-0.077 (0.064)
100 SMS \times parental monitoring				-0.237 (0.098) ($P=.009$)
Female			-0.219 (0.090) ($P=.01$)	-0.218 (0.089) ($P=.008$)
Age			0.177 (0.046) ($P=.001$)	0.187 (0.046) ($P<.001$)
Random effect				
Participant	0.771 (0.035)	0.785 (0.034)	0.729 (0.044)	0.746 (0.041)
Residual σ^2	0.737 (0.022)	0.705 (0.021)	0.616 (0.034)	0.612 (0.032)
LL ratio	-1547.279	-1523.1006	-813.11182	-853.34291

^a Models 2 and 3 control for age, gender, survey language, and intervention type. Age and gender were statistically significant so these variables are presented in the tables. The other control variables were not statistically significant. SMS use and having an SNS account were omitted from the final model due to collinearity with high SMS and number of SNS accounts.

Discussion

Principal Findings

Understanding predictors of sexual risk behavior is imperative for health and economic well-being over the life span, especially for underserved populations such as the Latino community. A plethora of studies have focused on sexual risk-taking behaviors, but with the proliferation of mobile technology and connectedness over the past decade, it is becoming clearer that social media utilization is also part of this relationship. Yet, there are still gaps in the literature with respect to social media use and sexual risk behaviors among adolescents in general. To our knowledge, this is the first study to longitudinally examine social media and sexual risk and the moderating effects of parental monitoring.

This study found a statistically significant positive association between high-frequency SMS and increased sexual risk behaviors over a 16-month period. Social media provides a context in which adolescents, who have a need for social acceptance and gratification and are still developing self-regulation skills, may find themselves vulnerable to pressures or unanticipated risk opportunities. Social media has the potential to expand and amplify existing peer relationships, which are well documented as influencing risk behaviors [39,40]. Social media may also provide increased access to partners that are more experienced, leading to increased communication about sex because of the perceived privacy of

social media [41]. Thus, those who are more active on social media could partake in more risky behaviors because of a larger peer network influencing their attitudes and social norms. Although these findings indicate a decrease in high frequency SMS and Facebook use between baseline and follow up, this does not necessarily imply a reduction in overall use. Adolescents are turning to newly developed software applications (eg, apps) that allow for communication within the app. We observe this as a result of the sharp increase in a newer app such as Instagram. Other research suggests similar results of a decline in Facebook use among US youth [42].

Although this study found a statistically significant association between increased sexual risk behaviors and high frequency SMS use over a 16-month period, parental monitoring was suggested to be a protective factor in this study. Results suggest increased sexual risk among higher SMS users, but higher levels of parental monitoring moderated this relationship in the hypothesized direction. Thus, parental monitoring was associated with lower levels of reported sexual risk behaviors despite high frequency SMS. Previous studies have documented that increased communication between parents and adolescents and the greater the parent knowledge of the adolescents' whereabouts (eg, parental monitoring), the lower the likelihood that adolescents will engage in health risk behaviors [43-46]. Specifically within the Latino community, Dittus and Jaccard [27] discovered that parental monitoring influenced delays in sexual intercourse, while Huebner and Howell [23] found parental monitoring and parenting style impacted having only

one sexual partner and to use a condom. The significant finding on the impact of parental monitoring on the relationship between high SMS use and sexual risk behavior compels parents to stay involved and maintain relationships with adolescents, despite adolescence being a period of time when parental influence is diminishing and peers/social influences are increasing [47].

While social media platforms are traditionally less monitored by adults, other research has suggested that mobile phones are one mechanism for parents to maintain a relationship with their adolescents while still affording them the autonomy and self-discovery they seek during this time [48]. It is important to highlight that the parental monitoring construct measured in this study was not specific to online monitoring; rather, it assessed adolescents' perceptions about whether their parents generally know their whereabouts and if they can go to their parents for support. This is important because in the aforementioned study, greater frequency of parental calls was associated with less adolescent-reported truthfulness, and parents calling when upset was associated with less parental knowledge and poorer family relations. Similarly, another study reported that parental social and technology supervision increased risky online activities [49]. Thus, it is important for parents to be involved in their adolescents' lives rather than solicit information via social media platforms solely to gain control of their lives.

The findings in this study are not intended to negate the substantial benefits of using social media in public health programs. However, our findings compel practitioners, parents and youth to be practical about the risks of high frequency SMS and other connections to expansive networks, and devise strategies for harnessing social media for good. For example, social media provides an excellent platform to strengthen supportive bonds and reach underserved youth to deliver health-related content [31], but public health professionals, policymakers, and parents should also embrace programs such as those that have encouraged adolescents to remove sexual content from their social networking profiles [50]. Other potential ways for parents to get involved is using for themselves or suggesting to their children, innovative technology, such as apps around sexual health and health-seeking behaviors that are being developed at a rapid pace. For example, Hablemos is a technology-based program attempting to close the communication gap between Latino parents and children through a parent-centered tool that is culturally appropriate and aims to empower Latino parents to have discussions about sexuality and contraception [51]. Specifically for youth, the team, An

Instant Gratification Situation, aims to develop content, stories, and messaging for social media platforms targeted toward youth for obtaining reproductive health services [51]. These platforms can be useful in engaging youth where they are—on mobile phones using social media—and continue to promote positive uses of social media platforms.

Limitations

There are several limitations specific to this study. First, results are based on self-reported, personal data that could be subject to response bias due to social desirability resulting from participants completing surveys in their school/program environment. This was attenuated by research assistants, not affiliated with the program, administering the surveys. Additionally, the use of personal laptop computers and audio capability increases data dependability [4,28].

A benefit of longitudinal data is the ability to control for individual heterogeneity and measure effects that are not detectable in pure cross-sectional studies, but these types of data are also sensitive to loss to follow up [52,53]. Strategies to reduce attrition permitted an overall 4% response rate at the 16-month follow-up survey, which is an acceptable response rate [54].

Despite this study utilizing longitudinal data, it is still limited in that we did not measure or test every variable theorized to influence sexual risk behavior. The relationships in this study may also be based on confounding variables that were not measured, such as peer behaviors [55], older sexual partners [43], parent-youth discussions about sexual health behaviors [56], relationship status, or having experienced dating violence [57]. Finally, although social media crosses cultural boundaries, this study was limited to a Latino population in Maryland; so generalizability is cautioned. However, it remains an important study of an emerging topic surrounding adolescent's social media use and how it affects their lives.

Conclusions

Although adolescents exchange SMS at high rates, parental monitoring remains vital to parent-child relationships and can moderate the relationship between adolescents' SMS frequency and sexual risk behaviors. When parents are involved in the lives of their adolescents, the extent of SMS sent or received does not influence risky sexual behaviors and suggests SMS is simply another form of communication that poses no additional threat.

Acknowledgments

This publication was made possible by Grant Number TP2AH000015 from the US Department of Health and Human Services, Office of Adolescent Health. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Office of Adolescent Health or the US Department of Health and Human Services. This research would not have been possible without the help of our community partners, Identity Inc. and Mary's Center, and the generous participation of the youth participants who chose to share their experiences. We are grateful for their participation and efforts with this research.

Conflicts of Interest

None declared.

References

1. Johnston L, O'Malley P, Bachman J, Schulenberg J. Monitoring the Future: National Survey Results on Drug Use-. Washington, DC: National Institute on Drug Abuse (NIDA); 1975:06-5883.
2. Kaestle CE, Halpern CT, Miller WC, Ford CA. Young age at first sexual intercourse and sexually transmitted infections in adolescents and young adults. *Am J Epidemiol* 2005 Apr 15;161(8):774-780. [doi: [10.1093/aje/kwi095](https://doi.org/10.1093/aje/kwi095)] [Medline: [15800270](https://pubmed.ncbi.nlm.nih.gov/15800270/)]
3. Lee J, Hahm HC. Acculturation and sexual risk behaviors among Latina adolescents transitioning to young adulthood. *J Youth Adolesc* 2010 Apr;39(4):414-427. [doi: [10.1007/s10964-009-9495-8](https://doi.org/10.1007/s10964-009-9495-8)] [Medline: [20020189](https://pubmed.ncbi.nlm.nih.gov/20020189/)]
4. Erikson E. Eight ages of man. *Klassiekers van de kinder-en jeugdpsychiatrie II* 1967:258.
5. Erikson E. Identity and the life cycle : selected papers. New York: International Universities Press; 1959.
6. Strauch B. The primal teen: What the new discoveries about the teenage brain tell us about our kids. New York, NY: Knopf Doubleday Publishing Group; 2007.
7. Lenhart A. Pew Research Center. 2012. Teens, smartphones & texting URL: <http://www.pewinternet.org/2012/03/19/teens-smartphones-texting/> [accessed 2017-04-04] [WebCite Cache ID 6pSrabT84]
8. Brown JD, Halpern CT, L'Engle KL. Mass media as a sexual super peer for early maturing girls. *J Adolesc Health* 2005 May;36(5):420-427. [doi: [10.1016/j.jadohealth.2004.06.003](https://doi.org/10.1016/j.jadohealth.2004.06.003)] [Medline: [15837346](https://pubmed.ncbi.nlm.nih.gov/15837346/)]
9. Houlihan D, Houlihan M. Adolescents and the social media: the coming storm. *J Child Adolesc Behav* 2014;2:e105. [doi: [10.4172/jcalb.1000e105](https://doi.org/10.4172/jcalb.1000e105)]
10. Moreno M. Social networking sites and adolescent health: new opportunities and new challenges. *ISJLP* 2011;7:57.
11. Landry M, Vyas A, Turner M, Glick S, Wood S. Evaluation of social media utilization by Latino adolescents: implications for mobile health interventions. *JMIR Mhealth Uhealth* 2015 Sep 29;3(3):e89 [FREE Full text] [doi: [10.2196/mhealth.4374](https://doi.org/10.2196/mhealth.4374)] [Medline: [26420553](https://pubmed.ncbi.nlm.nih.gov/26420553/)]
12. Berry J. Immigration, acculturation, and adaptation. *Appl Psychol* 1997 Jan;46(1):5-34. [doi: [10.1111/j.1464-0597.1997.tb01087.x](https://doi.org/10.1111/j.1464-0597.1997.tb01087.x)]
13. Livingstone S. Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression. *New Media & Soc* 2008;10(3):393-411.
14. Moreno MA, Kolb J. Social networking sites and adolescent health. *Pediatr Clin North Am* 2012 Jun;59(3):601-12, vii. [doi: [10.1016/j.pcl.2012.03.023](https://doi.org/10.1016/j.pcl.2012.03.023)] [Medline: [22643167](https://pubmed.ncbi.nlm.nih.gov/22643167/)]
15. Stockwell MS, Kharbanda EO, Martinez RA, Lara M, Vawdrey D, Natarajan K, et al. Text4Health: impact of text message reminder-recalls for pediatric and adolescent immunizations. *Am J Public Health* 2012 Feb;102(2):e15-e21. [doi: [10.2105/AJPH.2011.300331](https://doi.org/10.2105/AJPH.2011.300331)] [Medline: [22390457](https://pubmed.ncbi.nlm.nih.gov/22390457/)]
16. Wong CA, Merchant RM, Moreno MA. Using social media to engage adolescents and young adults with their health. *InHealthcare* 2014 Dec 31;2(4):220-224.
17. Borzekowski Dina L G. Adolescents' use of the internet: a controversial, coming-of-age resource. *Adolesc Med Clin* 2006 Feb;17(1):205-216. [doi: [10.1016/j.admecli.2005.10.006](https://doi.org/10.1016/j.admecli.2005.10.006)] [Medline: [16473301](https://pubmed.ncbi.nlm.nih.gov/16473301/)]
18. O'Keeffe GS, Clarke-Pearson K, Council OC. The impact of social media on children, adolescents, and families. *Pediatrics* 2011 Apr;127(4):800-804. [doi: [10.1542/peds.2011-0054](https://doi.org/10.1542/peds.2011-0054)] [Medline: [21444588](https://pubmed.ncbi.nlm.nih.gov/21444588/)]
19. Landry M, Gonzales FA, Wood S, Vyas A. New media use and sexual behavior among Latino adolescents. *Am J Health Behav* 2013 May;37(3):422-430. [doi: [10.5993/AJHB.37.3.15](https://doi.org/10.5993/AJHB.37.3.15)] [Medline: [23985189](https://pubmed.ncbi.nlm.nih.gov/23985189/)]
20. Frank S. Hyper-texting and hyper-networking examining why too much texting and social networking is associated with teen risk behavior. 2013 Nov 05 Presented at: 141st APHA Annual Meeting and Exposition; November 2-6, 2013; Boston, MA.
21. Deptula DP, Henry DB, Schoeny ME. How can parents make a difference? longitudinal associations with adolescent sexual behavior. *J Fam Psychol* 2010 Dec;24(6):731-739. [doi: [10.1037/a0021760](https://doi.org/10.1037/a0021760)] [Medline: [21171771](https://pubmed.ncbi.nlm.nih.gov/21171771/)]
22. DiClemente RJ, Wingood GM, Crosby RA, Sionean C, Cobb BK, Harrington K, et al. Sexual risk behaviors associated with having older sex partners: a study of black adolescent females. *Sex Transm Dis* 2002 Jan;29(1):20-24. [Medline: [11773874](https://pubmed.ncbi.nlm.nih.gov/11773874/)]
23. Huebner AJ, Howell LW. Examining the relationship between adolescent sexual risk-taking and perceptions of monitoring, communication, and parenting styles. *J Adolesc Health* 2003 Aug;33(2):71-78. [Medline: [12890597](https://pubmed.ncbi.nlm.nih.gov/12890597/)]
24. Killoren SE, Deutsch AR. A longitudinal examination of parenting processes and Latino youth's risky sexual behaviors. *J Youth Adolesc* 2014 Dec;43(12):1982-1993. [doi: [10.1007/s10964-013-0053-z](https://doi.org/10.1007/s10964-013-0053-z)] [Medline: [24186702](https://pubmed.ncbi.nlm.nih.gov/24186702/)]
25. Wight D, Williamson L, Henderson M. Parental influences on young people's sexual behaviour: a longitudinal analysis. *J Adolesc* 2006 Aug;29(4):473-494. [doi: [10.1016/j.adolescence.2005.08.007](https://doi.org/10.1016/j.adolescence.2005.08.007)] [Medline: [16213580](https://pubmed.ncbi.nlm.nih.gov/16213580/)]
26. de Graaf H, Vanwesenbeeck I, Woertman L, Meeus W. Parenting and adolescents' sexual development in western societies. *European Psychologist* 2011 Jan;16(1):21-31. [doi: [10.1027/1016-9040/a000031](https://doi.org/10.1027/1016-9040/a000031)]
27. Dittus PJ, Jaccard J. Adolescents' perceptions of maternal disapproval of sex: relationship to sexual outcomes. *J Adolesc Health* 2000 Apr;26(4):268-278. [Medline: [10734274](https://pubmed.ncbi.nlm.nih.gov/10734274/)]
28. Eaton D, Kann L, Kinchen S, Shanklin S, Flint KH, Hawkins J, Centers for Disease Control Prevention (CDC). Youth risk behavior surveillance - United States, 2011. *MMWR Surveill Summ* 2012 Jun 08;61(4):1-162 [FREE Full text] [Medline: [22673000](https://pubmed.ncbi.nlm.nih.gov/22673000/)]

29. Eaton DK, Kann L, Kinchen S, Shanklin S, Flint KH, Hawkins J, et al. Youth risk behavior surveillance-United States, 2011. *MMWR Morb Mortal Wkly Rep* 2012;61(4):1-162.
30. Livingston G, Lopez M. Pew Research Center. 2010. How young Latinos communicate with friends in the digital age URL: <http://www.pewhispanic.org/2010/07/28/how-young-latinos-communicate-with-friends-in-the-digital-age/> [accessed 2017-04-19] [WebCite Cache ID 6pq21c4PC]
31. Vyas AN, Landry M, Schnider M, Rojas AM, Wood SF. Public health interventions: reaching Latino adolescents via short message service and social media. *J Med Internet Res* 2012 Jul 12;14(4):e99 [FREE Full text] [doi: [10.2196/jmir.2178](https://doi.org/10.2196/jmir.2178)] [Medline: [22789678](https://pubmed.ncbi.nlm.nih.gov/22789678/)]
32. Gil A, Vega W, Dimas J. Acculturative stress and personal adjustment among Hispanic adolescent boys. *J Community Psychol* 1994;22(1):43-54. [doi: [10.1002/1520-6629\(199401\)22:1<43::AID-JCOP2290220106>3.0.CO;2-T](https://doi.org/10.1002/1520-6629(199401)22:1<43::AID-JCOP2290220106>3.0.CO;2-T)]
33. Romero AJ, Martinez D, Carvajal SC. Bicultural stress and adolescent risk behaviors in a community sample of Latinos and non-Latino European Americans. *Ethn Health* 2007 Nov;12(5):443-463. [doi: [10.1080/13557850701616854](https://doi.org/10.1080/13557850701616854)] [Medline: [17978943](https://pubmed.ncbi.nlm.nih.gov/17978943/)]
34. Lenhart A, Purcell K, Smith A, Zickuhr K. Pew Research Center. 2010. Social media and young adults URL: <http://www.pewinternet.org/2010/02/03/social-media-and-young-adults/> [accessed 2017-04-19] [WebCite Cache ID 6pq2NdgEW]
35. Silverberg S, Small S. Parental monitoring, family structure and adolescent substance use. 1991 Presented at: The Meeting of the Society of Research in Child Development; April, 1991; Seattle, Washington, DC.
36. Kerr M, Beck K, Shattuck TD, Kattar C, Uriburu D. Family involvement, problem and prosocial behavior outcomes of Latino youth. *Am J Health Behav* 2003;27 Suppl 1:S55-S65. [Medline: [12751647](https://pubmed.ncbi.nlm.nih.gov/12751647/)]
37. Kappeler EM, Farb AF. Historical context for the creation of the Office of Adolescent Health and the Teen Pregnancy Prevention program. *J Adolesc Health* 2014 Mar;54(3 Suppl):S3-S9 [FREE Full text] [doi: [10.1016/j.jadohealth.2013.11.020](https://doi.org/10.1016/j.jadohealth.2013.11.020)] [Medline: [24560074](https://pubmed.ncbi.nlm.nih.gov/24560074/)]
38. Spybrook J, Raudenbush S, Liu X, Congdon R. Optimal Design for Longitudinal and Multilevel Research: Documentation for the "Optimal Design" Software. 2006. URL: <http://www.rmcs.buu.ac.th/statcenter/HLM.pdf> [accessed 2017-04-19] [WebCite Cache ID 6pq3xQWWp]
39. Lyons H, Giordano P, Manning W, Longmore M. Identity, peer relationships, and adolescent girls' sexual behavior: an exploration of the contemporary double standard. *J Sex Res* 2011 Sep;48(5):437-449 [FREE Full text] [doi: [10.1080/00224499.2010.506679](https://doi.org/10.1080/00224499.2010.506679)] [Medline: [20818574](https://pubmed.ncbi.nlm.nih.gov/20818574/)]
40. Pfeifer J, Masten C, Moore W, Oswald T, Mazziotta J, Iacoboni M, et al. Entering adolescence: resistance to peer influence, risky behavior, and neural changes in emotion reactivity. *Neuron* 2011 Mar 10;69(5):1029-1036 [FREE Full text] [doi: [10.1016/j.neuron.2011.02.019](https://doi.org/10.1016/j.neuron.2011.02.019)] [Medline: [21382560](https://pubmed.ncbi.nlm.nih.gov/21382560/)]
41. Parks M, Floyd K. Making friends in cyberspace. *J Communication* 1996 Mar 01;46(1):80-97. [doi: [10.1111/j.1460-2466.1996.tb01462.x](https://doi.org/10.1111/j.1460-2466.1996.tb01462.x)]
42. Madden M, Lenhart A, Cortesi S, Gasser U, Duggan M, Smith A, et al. Pew Research Center. 2013. Teens, social media, and privacy URL: <http://www.pewinternet.org/2013/05/21/teens-social-media-and-privacy/> [accessed 2017-04-19] [WebCite Cache ID 6pq4IppIQ]
43. DiClemente R, Wingood G, Crosby R, Sionean C, Cobb B, Harrington K, et al. Parental monitoring: association with adolescents' risk behaviors. *Pediatrics* 2001 Jun;107(6):1363-1368. [Medline: [11389258](https://pubmed.ncbi.nlm.nih.gov/11389258/)]
44. Li X, Feigelman S, Stanton B. Perceived parental monitoring and health risk behaviors among urban low-income African-American children and adolescents. *J Adolesc Health* 2000 Jul;27(1):43-48. [Medline: [10867351](https://pubmed.ncbi.nlm.nih.gov/10867351/)]
45. Li X, Stanton B, Feigelman S. Impact of perceived parental monitoring on adolescent risk behavior over 4 years. *J Adolesc Health* 2000 Jul;27(1):49-56. [Medline: [10867352](https://pubmed.ncbi.nlm.nih.gov/10867352/)]
46. Romer D, Stanton B, Galbraith J, Feigelman S, Black MM, Li X. Parental influence on adolescent sexual behavior in high-poverty settings. *Arch Pediatr Adolesc Med* 1999 Oct;153(10):1055-1062. [Medline: [10520613](https://pubmed.ncbi.nlm.nih.gov/10520613/)]
47. Giordano PC. The wider circle of friends in adolescence. *Am J Sociol* 1995;101(3):661-697.
48. Weisskirch RS. Parenting by cell phone: parental monitoring of adolescents and family relations. *J Youth Adolesc* 2009 Sep;38(8):1123-1139. [doi: [10.1007/s10964-008-9374-8](https://doi.org/10.1007/s10964-008-9374-8)] [Medline: [19636776](https://pubmed.ncbi.nlm.nih.gov/19636776/)]
49. Sasson H, Mesch G. Parental mediation, peer norms and risky online behavior among adolescents. *Comput Human Behav* 2014 Apr;33:32-38. [doi: [10.1016/j.chb.2013.12.025](https://doi.org/10.1016/j.chb.2013.12.025)]
50. Moreno M, Vanderstoep A, Parks M, Zimmerman F, Kurth A, Christakis D. Reducing at-risk adolescents' display of risk behavior on a social networking web site: a randomized controlled pilot intervention trial. *Arch Pediatr Adolesc Med* 2009 Jan;163(1):35-41. [doi: [10.1001/archpediatrics.2008.502](https://doi.org/10.1001/archpediatrics.2008.502)] [Medline: [19124701](https://pubmed.ncbi.nlm.nih.gov/19124701/)]
51. The National Campaign to Prevent Teen & Unplanned Pregnancy. Innovation Next. 2016. Meet the winners URL: <http://innovationnext.org/winners> [accessed 2017-04-19] [WebCite Cache ID 6pq4y6j6n]
52. Baayen R, Davidson D, Bates D. Mixed-effects modeling with crossed random effects for subjects and items. *J Mem Lang* 2008 Nov;59(4):390-412. [doi: [10.1016/j.jml.2007.12.005](https://doi.org/10.1016/j.jml.2007.12.005)]
53. Wolfinger R, O'connell M. Generalized linear mixed models a pseudo-likelihood approach. *J Stat Comput Simul* 1993 Dec;48(3-4):233-243. [doi: [10.1080/00949659308811554](https://doi.org/10.1080/00949659308811554)]
54. Babbie ER. The basics of social research. Boston, United States: Cengage Learning; 1998.

55. Roberts ME, Gibbons FX, Gerrard M, Weng C, Murry VM, Simons LG, et al. From racial discrimination to risky sex: prospective relations involving peers and parents. *Dev Psychol* 2012 Jan;48(1):89-102 [[FREE Full text](#)] [doi: [10.1037/a0025430](https://doi.org/10.1037/a0025430)] [Medline: [21942666](#)]
56. Whitaker D, Miller K. Parent-adolescent discussions about sex and condoms impact on peer influences of sexual risk behavior. *J Adolesc Res* 2000;15(2):251-273.
57. Silverman JG, Raj A, Mucci LA, Hathaway JE. Dating violence against adolescent girls and associated substance use, unhealthy weight control, sexual risk behavior, pregnancy, and suicidality. *JAMA* 2001 Aug 01;286(5):572-579. [Medline: [11476659](#)]

Abbreviations

CDC: Centers for Disease Control and Prevention

SES: socioeconomic status

SMS: short message service

SNS: social networking sites

STIs: sexually transmitted infections

Edited by G Eysenbach; submitted 12.12.16; peer-reviewed by M Broom, N Allen; comments to author 05.01.17; revised version received 10.02.17; accepted 07.03.17; published 19.05.17.

Please cite as:

Landry M, Turner M, Vyas A, Wood S

Social Media and Sexual Behavior Among Adolescents: Is there a link?

JMIR Public Health Surveill 2017;3(2):e28

URL: <http://publichealth.jmir.org/2017/2/e28/>

doi: [10.2196/publichealth.7149](https://doi.org/10.2196/publichealth.7149)

PMID: [28526670](https://pubmed.ncbi.nlm.nih.gov/28526670/)

©Megan Landry, Monique Turner, Amita Vyas, Susan Wood. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 19.05.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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

Understanding Health Information Seeking on the Internet Among Sexual Minority People: Cross-Sectional Analysis From the Health Information National Trends Survey

Jennifer M Jabson^{1*}, MPH, PhD; Joanne G Patterson^{1*}, MSW, MPH; Charles Kamen², MPH, PhD

¹University of Tennessee, Department of Public Health, Knoxville, TN, United States

²University of Rochester, Cancer Control Unit, Department of Surgery, Rochester, NY, United States

*these authors contributed equally

Corresponding Author:

Jennifer M Jabson, MPH, PhD

University of Tennessee

Department of Public Health

390 HPER

1914 Andy Holt Ave

Knoxville, TN, 37996

United States

Phone: 1 865 974 0796

Fax: 1 865 974 6439

Email: jjabson@utk.edu

Abstract

Background: Individuals who face barriers to health care are more likely to access the Internet to seek health information. Pervasive stigma and heterosexism in the health care setting are barriers to health care for sexual minority people (SMP, ie, lesbian, gay, and bisexual people); therefore, SMP may be more likely to use the Internet as a source of health information compared to heterosexual people.

Objective: Currently, there is a dearth of published empirical evidence concerning health information seeking on the Internet among SMP; the current project addresses this gap.

Methods: Data from the 2015 Health Information National Trends Survey Food and Drug Administration Cycle were used to describe and summarize health information seeking among SMP (n=105) and heterosexual people (n=3405).

Results: Almost all of the SMP in this sample reported having access to the Internet (92.4%, 97/105). SMP were equally as likely as heterosexual people to seek health information on the Internet (adjusted odds ratio [aOR] 0.94, 95% CI 0.56-1.66) and to report incidental exposure to health information online (aOR 1.02, 95% CI 0.66-1.60). SMP were 58% more likely to watch a health-related video on YouTube than heterosexual people (aOR 1.58, 95% CI 1.00-2.47). Incidental exposure to health information was associated with seeking health information for oneself (aOR 3.87, 95% CI 1.16-14.13) and for someone else (aOR 6.30, 95% CI 2.40-17.82) among SMP.

Conclusions: SMP access the Internet at high rates and seek out health information online. Their incidental exposure could be associated with seeking information for self or others. This suggests that online interventions could be valuable for delivering or promoting health information for SMP.

(*JMIR Public Health Surveill* 2017;3(2):e39) doi:[10.2196/publichealth.7526](https://doi.org/10.2196/publichealth.7526)

KEYWORDS

sexual orientation; Internet-based health information seeking; internet access; sexual minority; homosexuality; bisexuality

Introduction

The general population increasingly uses the Internet as a source of health information. According to the Pew Internet and

American Life Project, 88% of the US population has access to the Internet [1] and 61% of adults use the Internet to seek health information [2]. Individuals who engage in online health information seeking for self and others are also more likely to experience incidental exposure to health information [3,4].

Incidental exposure to health information—“information gathered incidentally from sources in the environment” [5]—is worth consideration because it is associated both with increased health knowledge independent of active health information seeking behavior [4], and with positive health behaviors, such as fruit/vegetable intake, exercise, smoking, and cancer screening [5]. According to the Health Belief Model [6], both seeking health information and incidental exposure to health information on the Internet could increase an individual’s perceived susceptibility and perceived threat of disease, which could result in subsequent changes in health behavior and seeking of health care from a provider.

Not all subgroups of the general population use the Internet to seek health information at the same rate nor in the same way [7-10]. Individuals who face educational, economic, and cultural barriers to contact a health care professional are more likely to use the Internet to seek health information and to inform their health care decision making [8,9]. Sexual minority people (SMP; ie, lesbian, gay, and bisexual people) are a group of individuals who face multiple barriers to contact with health professionals. Therefore, it is possible that they may access the Internet at higher rates than heterosexual people (HP) to seek health information.

Pervasive heterosexism and stigma in the form of minority stressors [11,12] have been documented in the health care setting [13-17], and they produce multiple interpersonal and structural barriers for SMPs in accessing care and contacting health care professionals. For example, compared to HP, SMP are 50% less likely to receive needed preventive care including preventive screenings [18,19] and 85% more likely to leave a health care encounter with unmet needs [20]. SMP also report feeling disrespected by their medical providers. Among sexual minority men, 15% report not having enough time with providers, compared to 7% of heterosexual men [21]. In addition, providers also express discomfort and deny the importance of patient’s disclosure of sexual orientation, with 44-63% of health care providers reporting being unaware of SMP in their health care practice [14,22-24]. Health care providers’ medical education does not typically include thorough training in the unique health care needs of SMP, and providers often lack knowledge about health disparities experienced by SMP [25,26].

The Institute of Medicine [27], Healthy People 2020 [28], and the National Institute on Minority Health and Health Disparities [29] have all called for innovative, multilevel, public health solutions to reduce and eliminate health disparities among SMP. Promoting health information access or delivering behavioral health interventions on the Internet targeted to SMP may have multiple health-related benefits. Individuals who seek health information on the Internet are more likely to seek health care professionals for necessary treatments, make informed health care decisions, have positive feelings about information received from a health care professional, and report reductions in risky health behaviors [30,31]. Thus, the Internet has been identified as a promising channel for intervention delivery in the general population, and if SMPs are using the Internet to seek health information, the Internet could also be a useful channel for delivering innovative, disparities-reducing, public health interventions to SMP.

Seeking health information on the Internet may be useful for SMP who fear stigma and discrimination in the health care setting. However, seeking health information on the Internet may also have negative consequences. Health information on the Internet may be inaccurate, may produce anxiety, distress, and fear, and may result in further exposure to stigma for SMP. For example, individuals seeking health information on the Internet via a discussion board concerning human immunodeficiency virus could be exposed to discriminatory and other negative comments about SMP.

Currently there is a dearth of published empirical evidence concerning health information seeking and exposure to health information on the Internet among SMP as compared to HP. To the best of our knowledge, there is currently only one published, empirical article that describes health information seeking on the Internet in relation to sexual orientation [32]. Dahlhamer et al [32] used the National Health Interview Survey to estimate patterns in health information seeking on the Internet. In their study, a larger proportion of SMP than HP sought health information on the Internet.

The current project adds to this area by describing the rates at which SMP use the Internet to seek health information; describing incidental media exposure; estimating how SMP use the Internet for seeking health information compared to HP; and associating incidental and seeking health information on the Internet with health behavior, including seeking a health care provider. We hypothesized that (1) SMP would report higher use of the Internet for health information seeking than HP and that SMP incidental exposure would be similar to HP, and (2) incidental exposure to health information would be associated with seeking health information.

Methods

Survey Data

Health Information National Trends Survey (HINTS) Food and Drug Administration (FDA) Cycle 2015 data were used for this project [33]. HINTS is a nationally representative survey of adults aged 18 and older in the civilian, non-institutionalized population, administered by the National Cancer Institute [33]. HINTS was created to monitor changes in health communication and to understand how adults use communication channels to obtain health information. The response rate for HINTS FDA Cycle 4 was 33.04% [33]. Comprehensive information concerning HINTS FDA Cycle 4 methodology is available through HINTS [33].

In the HINTS FDA Cycle 4, 3738 participated and 3510 provided their sexual orientation. A total of 67 people identified as gay or lesbian, 38 as bisexual, and 3405 as heterosexual. Sexual orientation was missing for 228 participants. Due to small sample sizes and to facilitate analyses, sexual orientation was dichotomized where all gay men, lesbians, and bisexual individuals were combined into one group (SMP; n=105) and all heterosexual individuals were a second group (HP; n=3405). We believe this is a valid solution for small sample sizes for underrepresented populations, although we also acknowledge

the limitations this may introduce based on documented differences between sexual minority subgroups [27].

Measures

Dependent Variables

Sources and frequency of Internet access were measured with two questions: “Do you ever go online to access the Internet” (yes/no) and “How often do you access the Internet through each of the following: computer at home, computer at work, computer at school in a public place, computer at school in a private place, on a mobile device, on gaming device, Other” (daily, sometimes, never, n/a). Reasons for Internet use were measured with one question: “Sometimes people use the Internet for health-related reasons. Have you used the Internet for any of the following reasons in the past 12 months?” Sample reasons in the survey include “looked for health information for yourself, looked for health information for someone else, looked for information about quitting smoking.” Responses were dichotomous (yes/no).

Independent Variable

Sexual orientation was measured with a single item, “Do you think of yourself as heterosexual or straight, homosexual or gay/lesbian, bisexual, something else?” Sexual orientation was dichotomized (sexual minority=1, heterosexual=0). Respondents’ qualitative descriptions of “something else” did not indicate a sexual minority identity. Many of the responses included statements such as “normal,” “god’s child,” “human.” Therefore, respondents who selected “something else” (n=57) were excluded from the analysis.

Frequency and reading of incidental health information exposure online were measured with two questions: “Some people notice information about health on the Internet, even when they are not trying to find out about a health concern they have or someone in their family has. Have you read such health information on the Internet in the past 12 months?” (yes/no) and “About how often have you read this sort of information in the past 12 months?” (once a month or more, less than once a month).

Demographic Characteristics and Covariates

Demographic characteristics included age at time of survey, race/ethnicity (Hispanic, non-Hispanic Black/African-American, non-Hispanic white, other; ref=non-Hispanic white), highest level of education achieved (less than high school, high school/General Equivalency Diploma, some college/Associate

degree, college graduate/above; ref=less than high school), gender (male/female), insurance coverage (insured/uninsured), and income.

Time spent using the Internet for personal reasons was measured with two variables. First, “On a typical weekday, about how many hours do you use the Internet for personal reasons?” Respondents self-reported the number of hours on a typical weekday. Second, “During a typical weekend, about how many hours do you use the Internet for personal reasons?” Respondents self-reported the number of hours on a typical weekend.

Analyses

Descriptive statistics were calculated to describe and summarize demographic characteristics and covariates, sources and frequency of Internet access, and reasons for health information seeking among SMP. Chi-square tests were calculated for categorical variables for differences by sexual orientation. The *t* test was calculated to test for difference in continuous variables including age and average number of hours individuals accessed the Internet on weekdays and the weekend, by sexual orientation. Demographic characteristics that varied significantly by sexual orientation were applied as adjustment variables for multivariable tests. Multiple logistic regressions were calculated to test for association between sexual orientation and dependent variables. All of the multivariable logistic regression analyses were adjusted for age (continuous), education (categorical), and race/ethnicity (categorical); dummy variables were created for each categorical variable. Models were adjusted to account for possible confounded association between sexual orientation and health information seeking on the Internet. All adjustment variables were entered into each of the multivariable logistic regression models simultaneously. Multivariable logistic regression models concerning incidental exposure were also adjusted for average number of hours spent using the Internet on typical weekdays and weekends. Analyses were conducted with SPSS 14.0. This secondary analysis did not include human subjects and did not require a human subjects review.

Results

The sample’s demographic characteristics are stratified by sexual orientation and summarized in Table 1. SMP were more likely to be Hispanic ($\chi^2_3=13.91$, $P=.004$), younger ($t_1=4.39$, $P<.001$), and to have achieved higher levels of education ($\chi^2_3=8.10$, $P=.040$) than HP.

Table 1. Demographic characteristics by sexual orientation.

Characteristics	n (%) or mean (SD)			χ^2/t	df	P
	Full	Sexual minority	Heterosexual			
Total	3738 (100.0)	105 (3.0)	3405 (97.0)			
Sex				2.50	1	.13
Male	1497 (42.6)	52 (50.5)	1394 (42.7)			
Female	2018 (57.4)	51 (49.5)	49.5 (57.3)			
Race/Ethnicity				13.91	3	<.01
Hispanic	241 (7.2)	16 (15.2)	201 (6.4)			
White, Non-Hispanic	2633 (78.2)	72 (68.6)	2479 (79.1)			
Black or African-American, Non-Hispanic	232 (6.9)	6 (5.7)	212 (6.8)			
Other, Non-Hispanic	260 (7.7)	7 (6.7)	243 (7.8)			
Age	61.1 (17.5)	47.9 (15.6)	56.8 (16.6)	5.41	3508	<.001
Education				8.10	3	.04
<High school	237 (6.5)	5 (4.8)	194 (5.8)			
High school/GED	727 (19.8)	11 (10.5)	664 (19.5)			
Some college/AA degree	1132 (30.8)	30 (28.6)	1045 (31.0)			
College graduate or above	1578 (43.0)	58 (55.2)	1466 (43.1)			
Household income				1.97	4	.74
<\$20,000	664 (20.1)	16 (15.7)	591 (19.1)			
\$20,000-\$34,999	506 (15.3)	13 (12.7)	473 (13.9)			
\$35,000-\$49,999	415 (12.6)	14 (13.7)	392 (12.7)			
\$50,000-\$74,999	605 (18.3)	23 (22.5)	573 (18.6)			
>\$75,000	1112 (33.7)	36 (35.3)	1061 (34.3)			
Insurance coverage				0.01	1	1.00
Uninsured	207 (5.7)	6 (5.7)	186 (5.5)			
Insured	3444 (94.3)	99 (94.3)	3197 (94.5)			
Smoker				1.83	2	.41
Never	2041 (55.6)	51 (50.5)	1849 (55.3)			
Current	495 (13.5)	18 (17.8)	448 (13.4)			
Former	1132 (30.9)	32 (31.7)	1049 (31.4)			

Unadjusted descriptive statistics for sources and frequency of Internet access by sexual orientation are presented in [Table 2](#). SMP (92.4%, 97/105) were more likely to report Internet access than HP (79.4%, 2702/3738; $\chi^2_1=10.64$, $P=.001$). More SMP (26.1%, 24/105) reported accessing the Internet on a computer at school in a public location than HP (15.5%, 396/3738; $\chi^2_1=7.37$, $P=.009$).

[Table 2](#) also presents the unadjusted frequencies and descriptive reasons for seeking health information on the Internet, stratified

by sexual orientation. SMP most frequently used the Internet to seek health information for themselves (88.1%, 77/105), followed by seeking health information for someone else (60.0%, 57/105), and keeping track of personal health information (54.3%, 51/105). Watching health-related videos on YouTube was the only difference between SMP and HP in reasons for seeking health information on the Internet. More SMP watched health-related videos on YouTube (37.2%, 35/105) than HP (22.5%, 593/3405; $\chi^2_1=11.21$, $P<.001$).

Table 2. Sources, frequency, reasons, and unintentional exposure to health information via Internet access and use by sexual orientation.

Sources and frequency of Internet access	n (%) or mean (SD)		χ^2/t	df	P
	Sexual minority	Heterosexual			
Accesses Internet	97 (92.4)	2702 (79.4)	10.64	1	<.001
Average number of hours accessed Internet					
Weekday	4.18 (4.32)	2.79 (3.52)	-3.21	105	<.01
Weekend	4.92 (3.90)	3.69 (4.37)	-2.81	3380	<.01
Where accesses Internet					
Computer at home	90 (95.7)	2417 (92.5)	1.42	1	.32
Mobile device	75 (81.5)	1984 (76.5)	1.26	1	.32
Computer at work	54 (58.7)	1269 (49.8)	2.82	1	.11
Computer at school, public space	24 (26.1)	396 (15.5)	7.37	1	.01
Gaming device	21 (22.8)	397 (15.5)	3.55	1	.08
Computer at school, private space	12 (13.3)	201 (8.0)	3.26	1	.08
Other	7 (9.2)	105 (4.9)	2.88	1	.10
Noticed and read health information on Internet	62 (59.0)	1562 (45.9)	1.63	1	.21
How often read health information online					
Once a month	40 (65.6)	1000 (64.9)	0.01	1	1.00
Less than once a month	21 (34.4)	541 (35.1)	0.01	1	1.00
Used Internet last time sought health information	74 (89.2)	1669 (68.6)	15.97	1	<.001
Contacted doctor or health care provider last time sought health information	3 (3.6)	426 (17.5)	10.95	1	<.001
Reasons for Internet use					
Seeking health information for self	77 (88.1)	2076 (78.3)	0.41	1	.53
Seeking health information for someone else	57 (60.0)	1728 (65.3)	1.12	1	.32
Kept track of personal health information	51 (54.3)	1257 (47.5)	1.68	1	.21
Exchanged support about health concerns with family/friends	49 (52.1)	1237 (46.7)	1.06	1	.34
Used website to help with diet, weight, or physical activity	47 (49.5)	1163 (44.0)	1.13	1	.30
Seeking health care provider	42 (44.7)	943 (35.9)	3.02	1	.10
Watched a health-related video on YouTube	35 (37.2)	593 (22.5)	11.21	1	.001
Shared health information on social media sites	17 (18.1)	391 (14.8)	0.79	1	.38
Downloaded health information	14 (14.7)	397 (15.0)	0.00	1	1.00
Seeking information about quitting smoking	9 (9.6)	160 (6.1)	1.93	1	.18
Participation in online forum/support group	4 (4.2)	135 (5.1)	0.15	1	.82
Frequency of unintentional health information seeking					
Unintentionally noticed health information	62 (59.0)	1562 (58.0)	1.63	1	.21
How often read this type of information					
Once a month or more	40 (65.6)	1000 (64.9)	0.01	1	1.00
Less than once a month	21 (34.4)	541 (35.1)	0.01	1	1.00

Frequency of incidental exposure to health information (SMP 59% vs HP 58%) and frequency of reading such information did not vary by sexual orientation (Table 2). On a typical weekday, SMP and HP used the Internet for personal reasons on average 4.18 (SD 4.32) and 2.79 (SD 3.52) hours respectively ($t_{105}=-3.21$, $P<.01$). Internet use for personal reasons on a typical weekend also varied by sexual orientation. On average, SMP

reported 4.92 (SD 3.90) and HP 3.69 (SD 4.37) hours ($t_{3380}=-2.81$, $P<.01$).

Table 3 presents the adjusted associations between access to Internet and reasons for seeking health information on the Internet. In analyses adjusted for age, education, and race/ethnicity, SMP were 38% less likely than HP to report

seeking health information on the Internet for someone else (adjusted odds ratio [OR] 0.62, 95% CI 0.40-0.97; $P=.03$). After adjusting for age, race/ethnicity, and education, SMP were 58%

more likely to report watching health-related videos on YouTube compared to HP (aOR 1.58, 95% CI 1.00-2.47; $P=.04$).

Table 3. Associations^a between sexual orientation and health information seeking behaviors.

	Sexual minority (n=105) (heterosexual ref)	
	aOR (95% CI)	P
Accessed Internet	1.84 (0.85-4.56)	.15
Noticed health information (incidental exposure) ^b	1.02 (0.66-1.60)	.92
Read health information once a month or more (incidental exposure) ^b	0.83 (0.48-1.47)	.51
Seeking health information for self	0.94 (0.56-1.66)	.83
Seeking health information for someone else	0.62 (0.40-0.97)	.03
Kept track of personal health information	1.28 (0.84-1.96)	.25
Exchanged support about health concerns with family/friends	1.17 (0.77-1.78)	.47
Used website to help with diet, weight, or physical activity	0.95 (0.62-1.47)	.82
Seeking health care provider	1.14 (0.74-1.76)	.54
Watched a health-related video on YouTube	1.58 (1.003-2.47)	.04
Shared health information on social media sites	1.01 (0.56-1.72)	.97
Downloaded health information	0.76 (0.40-1.33)	.36
Seeking information about quitting smoking	1.37 (0.62-2.60)	.39
Participation in online forum/ support group	0.70 (0.21-1.72)	.50

^aAll models adjusted for age, education, and race/ethnicity.

^bModel adjusted for age, education, race/ethnicity, and average number of weekday and weekend hours accessed Internet.

Associations between incidental exposure to health information on the Internet and health information seeking behaviors were calculated with adjustment for age, race/ethnicity, and education, and stratified by sexual orientation (data available on request). Among HP, incidental exposure to health information on the Internet was associated with all health information seeking behaviors. Among SMP, small sample sizes made it impossible to calculate adjusted analyses for several models including seeking a health care provider, watching a health-related video, downloading health information, and seeking information about quitting smoking. For the models that could be calculated, incidental exposure to information online was associated with three times greater odds of seeking health information for self (aOR 3.87, 95% CI 1.16-14.13; $P<.05$) and six times greater odds of seeking health information for someone else (aOR 6.30, 95% CI 6.30-17.82; $P \leq .001$), relative to those who did not report incidental exposure to health information on the Internet.

Discussion

Principal Findings

The purpose of this project was to describe SMP's access to the Internet and to investigate SMP health information seeking and incidental exposure to health information on the Internet. Our results indicate that 94.4% of SMP respondents are accessing the Internet, and adjusted analyses indicate that this was not significantly different from heterosexual respondents.

We found that 88.1% of SMP and 78.3% of HP are seeking health information on the Internet for themselves. These findings are similar to those published by Dahlhamer et al [32]. From data provided by the National Health Interview Survey, Dahlhamer et al reported that 62.3% of sexual minority men and 65.8% of sexual minority women sought health information on the Internet compared to 42.3% of heterosexual men and 56.2% of heterosexual women. SMP have access to the Internet and are using the Internet to seek health information. This evidence is especially valuable for public health practitioners and researchers interested in testing and disseminating Internet-based interventions for improving health and reducing disparities among SMP [34].

SMP sought health information on the Internet for reasons that were largely the same as HP. One notable difference was in access to health-related videos on YouTube, where SMP were more likely to report viewing health-related videos than HP. Social media and YouTube have been used by some public health interventionists as a mechanism to reach gay and bisexual men for sexual health promotion [35,36]. In their review of social media for sexual health promotion, Gabarron and Wynn [36] found eight projects that determined that YouTube would be an effective means of delivering intervention content to SMP. Many of these projects reported thousands of online views, including "Queer as F**ck," which delivered sexual health promotion to gay and bisexual men via short "webisodes" [35]. Such webisodes were highly popular and commanded over 30,000 YouTube views. These and other forms of sexual health

promotion underscore the potential utility of YouTube for promoting health messages among SMP. It is also possible that SMP may be using health-related videos in response to experienced or anticipated barriers in accessing health care. Given the breadth of evidence documenting SMPs' experiences of heterosexism and homophobia in the health care setting [13-17], it may be that in the absence of consistently accessible, culturally competent care, SMP are more likely to seek health information from online sources, such as YouTube, that reflect their specific sexual orientation group or target relevant health concerns [37].

There are also methodological issues that could have influenced the results involving incidental exposure. The sample of SMP was small, and it may have been too small to detect an association. This is something that can be addressed only by increasing sample sizes of SMP in health surveillance. With respect to the significant associations identified between incidental exposure and health information seeking activities, ours was a cross-sectional, secondary study. Therefore, it cannot be known if incidental exposure came before, or after, seeking health information for self and seeking health information for someone else. It is plausible that in the process of seeking health information for themselves or someone else, SMP were incidentally exposed to health information online. This is a research question to be addressed by future research efforts that can test temporality.

It is useful to know that SMP are being exposed to incidental health information for the purposes of future health information programs that aim to target SMP. It could be useful for future Internet-based health information programs to know the types of health information that SMP are the most likely to read, and future studies should assess the health information to which SMP are exposed with greater nuance.

Strengths and Limitations

HINTS is a population-based data source focused on Internet access and health information seeking on the Internet that also includes sexual orientation questions. As a result, this project represents some of the first population-based evidence on Internet access among and health information seeking on the Internet among SMP. A very small number of individuals did not answer the sexual orientation question, suggesting that the population is increasingly comfortable reporting their sexual

orientation in health surveillance. Sexual minority samples sizes are notoriously small, with some population-based findings based on 50 or fewer SMP. Although still relatively small, HINTS provided a sample of more than 100 SMP, equally male and female.

This project involved a small sample of SMP. Small sample sizes result in underpowered statistical analyses and thereby make it difficult to detect statistically meaningful differences if differences between SMP and HP are present. Small sample size is a persistent problem when using health surveillance data sources to investigate SMP health. In order to capture the best possible approximation of a representative sample, the Institute of Medicine [27] and other national organizations strongly encourage the use of health surveillance data sources, such as HINTS, to investigate sexual minority health. However, this approach often produces very small samples where only 3-5% of the total sample comprising SMP, and sometimes much less. One possible solution for this in the future could be rigorous oversampling of SMP in health surveillance. The HINTS response rate is low (33.04% [33]) and may reflect the healthy volunteer effect where only the healthiest individuals participate [38]. Finally, our statistical methods that involved multiple tests of associations may have also put the findings at risk of a type I error in which we reported associations that were produced by chance. In addition, the confidence intervals and *P* values for "seeking health information for self" and "accessed the Internet" are very near and barely including 1.0 with *P* values of .06 and .05. It is possible that these confidence intervals and *P* values are an artifact of our statistical software package and small sample sizes. We believe that replicating these analyses with larger samples would aid in clarifying these limitations and associations.

Conclusion

It is often difficult to locate sexual minority people for health-related, disparities-reducing interventions, but our findings suggest that the Internet is a promising tool for delivering health interventions to this group. SMP use the Internet and are using it to access health information at high rates. This is valuable given the popularity and promise of Internet-based interventions for SMP. We now have evidence that the Internet is a promising delivery method for health-related information for SMP.

Conflicts of Interest

None declared.

References

1. Pew Research Center, Internet/Broadband Fact Sheet. 2017. URL: <http://www.pewinternet.org/fact-sheet/internet-broadband/> [accessed 2017-06-09] [WebCite Cache ID 6r5rV8z04]
2. Fox S, Jones S. The Social Life of Health Information. Washington, DC: Pew Internet and American Life Project; 2009 Jun. URL: http://www.pewinternet.org/files/old-media//Files/Reports/2009/PIP_Health_2009.pdf [accessed 2017-06-09] [WebCite Cache ID 6r5rqFKLG]
3. Cline R, Haynes KM. Consumer health information seeking on the Internet: the state of the art. *Health Educ Res* 2001 Dec;16(6):671-692. [Medline: [11780707](https://pubmed.ncbi.nlm.nih.gov/11780707/)]

4. Tian Y, Robinson JD. Incidental health information use on the Internet. *Health Commun* 2009 Jan;24(1):41-49. [doi: [10.1080/10410230802606984](https://doi.org/10.1080/10410230802606984)] [Medline: [19204857](https://pubmed.ncbi.nlm.nih.gov/19204857/)]
5. Shim M, Kelly B, Hornik R. Cancer information scanning and seeking behavior is associated with knowledge, lifestyle choices, and screening. *J Health Commun* 2006;11 Suppl 1:157-172. [doi: [10.1080/10810730600637475](https://doi.org/10.1080/10810730600637475)] [Medline: [16641081](https://pubmed.ncbi.nlm.nih.gov/16641081/)]
6. Janz N, Becker MH. The Health Belief Model: a decade later. *Health Educ Q* 1984;11(1):1-47. [doi: [10.1177/109019818401100101](https://doi.org/10.1177/109019818401100101)] [Medline: [6392204](https://pubmed.ncbi.nlm.nih.gov/6392204/)]
7. Anker A, Reinhart AM, Feeley TH. Health information seeking: a review of measures and methods. *Patient Educ Couns* 2011 Mar;82(3):346-354. [doi: [10.1016/j.pec.2010.12.008](https://doi.org/10.1016/j.pec.2010.12.008)] [Medline: [21239134](https://pubmed.ncbi.nlm.nih.gov/21239134/)]
8. Amante D, Hogan TP, Pagoto SL, English TM, Lapane KL. Access to care and use of the Internet to search for health information: results from the US National Health Interview Survey. *J Med Internet Res* 2015 Apr 29;17(4):e106 [FREE Full text] [doi: [10.2196/jmir.4126](https://doi.org/10.2196/jmir.4126)] [Medline: [25925943](https://pubmed.ncbi.nlm.nih.gov/25925943/)]
9. Perez S, Kravitz RL, Bell RA, Chan MS, Paterniti DA. Characterizing internet health information seeking strategies by socioeconomic status: a mixed methods approach. *BMC Med Inform Decis Mak* 2016 Aug 09;16:107 [FREE Full text] [doi: [10.1186/s12911-016-0344-x](https://doi.org/10.1186/s12911-016-0344-x)] [Medline: [27506607](https://pubmed.ncbi.nlm.nih.gov/27506607/)]
10. Bundorf M, Wagner TH, Singer SJ, Baker LC. Who searches the internet for health information? *Health Serv Res* 2006 Jun;41(3 Pt 1):819-836 [FREE Full text] [doi: [10.1111/j.1475-6773.2006.00510.x](https://doi.org/10.1111/j.1475-6773.2006.00510.x)] [Medline: [16704514](https://pubmed.ncbi.nlm.nih.gov/16704514/)]
11. Meyer I. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. *Psychol Bull* 2003;129(5):674-697. [doi: [10.1037/0033-2909.129.5.674](https://doi.org/10.1037/0033-2909.129.5.674)]
12. O'Hanlan K, Cabaj R, Schatz B, Lock J, Nemrow P. A Review of the Medical Consequences of Homophobia with Suggestions for Resolution. *Journal of the Gay and Lesbian Medical Association* 1997 Mar;1(1):25-39. [doi: [10.1023/B:JOLA.0000007009.83600.ae](https://doi.org/10.1023/B:JOLA.0000007009.83600.ae)]
13. Jabson J, Mitchell JW, Doty S. Associations between non-discrimination and training policies and physicians' attitudes and knowledge about sexual and gender minority patients: a comparison of physicians from two hospitals. *BMC Public Health* 2016;16(1):a. [doi: [10.1186/s12889-016-2927-y](https://doi.org/10.1186/s12889-016-2927-y)]
14. Schatz B, O'Hanlan R. *Anti-gay discrimination in medicine: Results of a national survey of lesbian, gay and bisexual physicians*. San Francisco, CA: American Association of Physicians for Human Rights; 1994.
15. Neville S, Henrickson M. Perceptions of lesbian, gay and bisexual people of primary healthcare services. *J Adv Nurs* 2006 Aug;55(4):407-415. [doi: [10.1111/j.1365-2648.2006.03944.x](https://doi.org/10.1111/j.1365-2648.2006.03944.x)] [Medline: [16866836](https://pubmed.ncbi.nlm.nih.gov/16866836/)]
16. Katz A. Gay and lesbian patients with cancer. *Oncol Nurs Forum* 2009;36(2):203-207. [Medline: [19273409](https://pubmed.ncbi.nlm.nih.gov/19273409/)]
17. Harris M, Nightengale J, Owen N. Health Care Professionals' Experience, Knowledge, and Attitudes Concerning Homosexuality. *Journal of Gay & Lesbian Social Services* 1995 May 11;2(2):91-108. [doi: [10.1300/J041v02n02_06](https://doi.org/10.1300/J041v02n02_06)]
18. Conron K, Mimiaga MJ, Landers SJ. A population-based study of sexual orientation identity and gender differences in adult health. *Am J Public Health* 2010 Oct;100(10):1953-1960. [doi: [10.2105/AJPH.2009.174169](https://doi.org/10.2105/AJPH.2009.174169)] [Medline: [20516373](https://pubmed.ncbi.nlm.nih.gov/20516373/)]
19. Blosnich J, Farmer GW, Lee JGL, Silenzio VMB, Bowen DJ. Health inequalities among sexual minority adults: evidence from ten U.S. states, 2010. *Am J Prev Med* 2014 Apr;46(4):337-349 [FREE Full text] [doi: [10.1016/j.amepre.2013.11.010](https://doi.org/10.1016/j.amepre.2013.11.010)] [Medline: [24650836](https://pubmed.ncbi.nlm.nih.gov/24650836/)]
20. McNair R, Szalacha LA, Hughes TL. Health status, health service use, and satisfaction according to sexual identity of young Australian women. *Womens Health Issues* 2011;21(1):40-47. [doi: [10.1016/j.whi.2010.08.002](https://doi.org/10.1016/j.whi.2010.08.002)] [Medline: [21185989](https://pubmed.ncbi.nlm.nih.gov/21185989/)]
21. Clift J, Kirby J. Health care access and perceptions of provider care among individuals in same-sex couples: findings from the Medical Expenditure Panel Survey (MEPS). *J Homosex* 2012;59(6):839-850. [doi: [10.1080/00918369.2012.694766](https://doi.org/10.1080/00918369.2012.694766)] [Medline: [22853183](https://pubmed.ncbi.nlm.nih.gov/22853183/)]
22. Dahan R, Feldman R, Hermoni D. Is patients' sexual orientation a blind spot of family physicians? *J Homosex* 2008;55(3):524-532. [doi: [10.1080/00918360802345321](https://doi.org/10.1080/00918360802345321)] [Medline: [19042284](https://pubmed.ncbi.nlm.nih.gov/19042284/)]
23. Westerståhl A, Segesten K, Björkelund C. GPs and lesbian women in the consultation: issues of awareness and knowledge. *Scand J Prim Health Care* 2002 Dec;20(4):203-207. [Medline: [12564570](https://pubmed.ncbi.nlm.nih.gov/12564570/)]
24. Eliason M. *Who cares? Institutional barriers to health care for lesbian, gay, and bisexual persons*. New York, NY: NLN Press; 1996.
25. Tesar C, Rovi SL. Survey of curriculum on homosexuality/bisexuality in departments of family medicine. *Fam Med* 1998 Apr;30(4):283-287. [Medline: [9568499](https://pubmed.ncbi.nlm.nih.gov/9568499/)]
26. Association of American Medical Colleges Group on Student Affairs, Association of American Medical Colleges Organization of Student Representatives. *AAMC Recommendations: Institutional Programs and Educational Activities to Address the Needs of Gay, Lesbian, Bisexual and Transgender (GLBT) Students and Patients*. 2007. URL: <https://www.aamc.org/download/54768/data/umassbestpractices.pdf> [accessed 2017-06-08] [WebCite Cache ID 6r5sffHhD]
27. Institute of Medicine. *The Health of Lesbian, Gay, Bisexual, and Transgender (LGBT) People: Building a Foundation for Better Understanding*. Washington, DC: National Academies Press; 2011.
28. US Department of Health and Human Services. *Healthy People 2020*. 2012. URL: <https://www.healthypeople.gov/2020/topics-objectives/topic/lesbian-gay-bisexual-and-transgender-health> [accessed 2017-06-09] [WebCite Cache ID 6r5sveOvM]

29. National Institute of Minority Health and Health Disparities, Mission and Vision. 2016. URL: <https://www.nimhd.nih.gov/about/overview/mission-vision.html> [accessed 2017-06-09] [WebCite Cache ID 6r5t4TSrT]
30. Ybarra M, Suman M. Help seeking behavior and the Internet: a national survey. *Int J Med Inform* 2006 Jan;75(1):29-41. [doi: [10.1016/j.ijmedinf.2005.07.029](https://doi.org/10.1016/j.ijmedinf.2005.07.029)] [Medline: [16129659](https://pubmed.ncbi.nlm.nih.gov/16129659/)]
31. Rice R. Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. *Int J Med Inform* 2006 Jan;75(1):8-28. [doi: [10.1016/j.ijmedinf.2005.07.032](https://doi.org/10.1016/j.ijmedinf.2005.07.032)] [Medline: [16125453](https://pubmed.ncbi.nlm.nih.gov/16125453/)]
32. Dahlhamer J, Galinsky AM, Joestl SS, Ward BW. Sexual Orientation and Health Information Technology Use: A Nationally Representative Study of U.S. Adults. *LGBT Health* 2017 Apr;4(2):121-129. [doi: [10.1089/lgbt.2016.0199](https://doi.org/10.1089/lgbt.2016.0199)] [Medline: [28287875](https://pubmed.ncbi.nlm.nih.gov/28287875/)]
33. Westat. Health Information National Trends Survey 4 (HINTS 4). 2015. URL: https://hints.cancer.gov/docs/HINTS_4_Cycle2_Methods_Report.pdf [accessed 2017-06-09] [WebCite Cache ID 6r5tKDI4e]
34. Bowen D, Jabson J, Kamen C. mHealth: an avenue for promoting health among sexual and gender minority populations? *Mhealth* 2016;2:36 [FREE Full text] [doi: [10.21037/mhealth.2016.09.01](https://doi.org/10.21037/mhealth.2016.09.01)] [Medline: [28293609](https://pubmed.ncbi.nlm.nih.gov/28293609/)]
35. Pedrana A, Hellard M, Gold J, Ata N, Chang S, Howard S, et al. Queer as F**k: reaching and engaging gay men in sexual health promotion through social networking sites. *J Med Internet Res* 2013 Feb 07;15(2):e25 [FREE Full text] [doi: [10.2196/jmir.2334](https://doi.org/10.2196/jmir.2334)] [Medline: [23391459](https://pubmed.ncbi.nlm.nih.gov/23391459/)]
36. Gabarron E, Wynn R. Use of social media for sexual health promotion: a scoping review. *Glob Health Action* 2016;9:a. [doi: [10.3402/gha.v9.32193](https://doi.org/10.3402/gha.v9.32193)]
37. Polonijo A, Hollister B. Normalcy, Boundaries, and Heterosexism: An Exploration of Online Lesbian Health Queries. *Journal of Gay & Lesbian Social Services* 2011 Apr;23(2):165-187. [doi: [10.1080/10538720.2011.560099](https://doi.org/10.1080/10538720.2011.560099)]
38. Pinsky PF, Miller A, Kramer BS, Church T, Reding D, Prorok P, et al. Evidence of a healthy volunteer effect in the prostate, lung, colorectal, and ovarian cancer screening trial. *Am J Epidemiol* 2007 Apr 15;165(8):874-881. [doi: [10.1093/aje/kwk075](https://doi.org/10.1093/aje/kwk075)] [Medline: [17244633](https://pubmed.ncbi.nlm.nih.gov/17244633/)]

Abbreviations

aOR: adjusted odds ratio

HINTS-FDA: Health Information National Trends Survey Food and Drug Administration

HP: heterosexual people

SMP: sexual minority people

Edited by T Sanchez; submitted 16.02.17; peer-reviewed by J Magee, T English; comments to author 13.03.17; revised version received 10.04.17; accepted 30.05.17; published 19.06.17.

Please cite as:

Jabson JM, Patterson JG, Kamen C

Understanding Health Information Seeking on the Internet Among Sexual Minority People: Cross-Sectional Analysis From the Health Information National Trends Survey

JMIR Public Health Surveill 2017;3(2):e39

URL: <http://publichealth.jmir.org/2017/2/e39/>

doi: [10.2196/publichealth.7526](https://doi.org/10.2196/publichealth.7526)

PMID: [28630036](https://pubmed.ncbi.nlm.nih.gov/28630036/)

©Jennifer M Jabson, Joanne G Patterson, Charles Kamen. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 19.06.2017. 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

GapMap: Enabling Comprehensive Autism Resource Epidemiology

Nikhila Albert^{1,2,3}, BS; Jena Daniels^{1,3}, BS; Jessey Schwartz^{1,3}, BA; Michael Du^{1,3}, BA; Dennis P Wall^{1,3}, PhD

¹Department of Pediatrics, Division of Systems Medicine, Stanford University, Stanford, CA, United States

²Department of Computer Science, Princeton University, Princeton, NJ, United States

³Department of Biomedical Data Science, Stanford University, Stanford, CA, United States

Corresponding Author:

Dennis P Wall, PhD

Department of Pediatrics

Division of Systems Medicine

Stanford University

1265 Welch Rd

Suite X141

Stanford, CA,

United States

Phone: 1 650 497 0921

Email: dpwall@stanford.edu

Abstract

Background: For individuals with autism spectrum disorder (ASD), finding resources can be a lengthy and difficult process. The difficulty in obtaining global, fine-grained autism epidemiological data hinders researchers from quickly and efficiently studying large-scale correlations among ASD, environmental factors, and geographical and cultural factors.

Objective: The objective of this study was to define resource load and resource availability for families affected by autism and subsequently create a platform to enable a more accurate representation of prevalence rates and resource epidemiology.

Methods: We created a mobile application, GapMap, to collect locational, diagnostic, and resource use information from individuals with autism to compute accurate prevalence rates and better understand autism resource epidemiology. GapMap is hosted on AWS S3, running on a React and Redux front-end framework. The backend framework is comprised of an AWS API Gateway and Lambda Function setup, with secure and scalable end points for retrieving prevalence and resource data, and for submitting participant data. Measures of autism resource scarcity, including resource load, resource availability, and resource gaps were defined and preliminarily computed using simulated or scraped data.

Results: The average distance from an individual in the United States to the nearest diagnostic center is approximately 182 km (50 miles), with a standard deviation of 235 km (146 miles). The average distance from an individual with ASD to the nearest diagnostic center, however, is only 32 km (20 miles), suggesting that individuals who live closer to diagnostic services are more likely to be diagnosed.

Conclusions: This study confirmed that individuals closer to diagnostic services are more likely to be diagnosed and proposes GapMap, a means to measure and enable the alleviation of increasingly overburdened diagnostic centers and resource-poor areas where parents are unable to diagnose their children as quickly and easily as needed. GapMap will collect information that will provide more accurate data for computing resource loads and availability, uncovering the impact of resource epidemiology on age and likelihood of diagnosis, and gathering localized autism prevalence rates.

(*JMIR Public Health Surveill* 2017;3(2):e27) doi:[10.2196/publichealth.7150](https://doi.org/10.2196/publichealth.7150)

KEYWORDS

autism; autism spectrum disorder; crowdsourcing; prevalence; resources; epidemiology

Introduction

Background

Autism spectrum disorder (ASD) has been attracting more interest recently, as a result of skyrocketing prevalence rates. Centers for Disease Control and Prevention (CDC) estimates in 2000 suggested that the condition affected 1 in 150 children; 10 years later, the rate rose to 1 in 68 children [1]. However, many present epidemiological studies suffer from small sample sizes and regional focuses [2-5]. For example, the most recent CDC prevalence estimates were based on a limited number of communities in only 11 states [6]. However, studies that estimate prevalence on broader geographic scales often rely on aggregate data that omit valuable regional prevalence rates. Although families report concerns as early as 1 year of age and reliable diagnoses can be made during early childhood [7], the average age of diagnosis for ASD in the United States is more than 4 years and an estimated quarter of children remain undiagnosed at 8 years of age [8-9]. Despite the considerable number of undiagnosed individuals who meet the criteria for autism [10], studies that calculate prevalence rates often do not include undiagnosed individuals [11]. This means that individuals without access to diagnostic centers for socioeconomic or geographic reasons are not reported, resulting in underrepresented statistics [12,13]. Instead, an ideal dataset would span a large geographic region (such as the United States) and maintain high specificity (to the city level).

Comprehensive regional autism prevalence rates would be extremely helpful for determining the true prevalence of autism and correlating genetic and environmental factors with higher standards of significance. In particular, comparing geographic trends in prevalence rates with autism resource epidemiology would be invaluable in revealing patient care deficits.

Whereas much effort has been focused on measuring autism incidence and prevalence rates, few have explored resource epidemiology. Yet, studies reveal alarming indicators of significant resource shortages. In metropolitan areas, parents may wait up to 18 months after initial suspicions for a final diagnosis [14]. For families of minority populations or lower socioeconomic status, this process may take even longer [15]. Children with autism in rural or resource-poor areas are diagnosed much later than their peers in resource-rich areas (if they are diagnosed at all) and face difficulties finding appropriate post diagnosis services [12].

Finding these resource gaps, regions in which there are limited diagnostic or treatment resources with respect to the demand, can support pushes for congressional change with hard data, allocate resources more efficiently, and provide information to emerging organizations and businesses to let them know where their services are most needed. These efforts can help reduce age at the first diagnosis and ensure that speech and behavioral therapies are started during critical periods when they are maximally impactful [16,17].

Aims of This Study

The specific aims of this study were to (1) obtain an early approximation of the disconnect between autism resources and

diagnosed individuals by determining the average distance between an individual with autism and the nearest diagnostic center based on public census data in the United States and the United Kingdom, (2) define useful metrics that can be used to determine whether a center is overloaded or whether a region is underserved, and (3) propose an online prototype to collect information pertaining to geographic variations of autism prevalence and the geographic resource utilization of autism resources.

Our proposed app, GapMap [18], will contribute to the growing field of participatory surveillance and epidemiology [19-21] by prospectively collecting locational, diagnostic, and resource use information from individuals with autism in order to compute accurate prevalence rates and better understand autism prevalence rates and resource epidemiology.

Methods

Computing the Average Distance From an Individual to the Nearest Diagnostic Center

A simulation of the average distance from an individual in the United States to the nearest diagnostic center was constructed using county-level 2015 Census Data [22] and a list of 840 developmental medical centers that were collected through Autism Speaks [23] and Autism Source [24]. County populations and land area were used to model randomly uniform distributions of individuals within county borders; individuals were then assigned to the nearest diagnostic center. The average and standard deviation were then computed for the distance from an individual to the nearest diagnostic center.

Computing the Average Distance From an Individual With ASD to the Nearest Diagnostic Center

A simulation of the average distance from an individual with an ASD to the nearest diagnostic center was constructed for the United States and the United Kingdom. Data from the United Kingdom are comprehensively collected and publicly available, which allowed us to conduct a proof-of-concept study with a more complete dataset to highlight disparities between families with a clinical diagnosis of autism and autism-related resources. City-level geographic approximations of individuals with autism were obtained by screen scraping information from the Internet and social networks. Lists of 840 and 135 diagnostic centers in the United States and United Kingdom, respectively, were collected through Autism Speaks [23], Autism Source [24], and the National Autistic Society (in the United Kingdom) [25]. Individuals were similarly assigned to the nearest diagnostic center, and statistics were computed. Maps of the individuals and centers were generated for both countries.

Computing Resource Load and Availability

Resource load was computed for 840 diagnostic centers in the United States using county-level 2015 Census Data [22-24]. Individuals were randomly assigned to a center near them, where nearby centers were defined as all the centers within a radius of 25, 100, 500, or 3000 km (the smallest radius was chosen so that at least one center existed within that distance). The

maximum and average distance an individual traveled to each center was calculated in addition to the resource load.

Resource load was calculated based on an approximation of the number of caseloads per clinician in a year and an estimation of the number of clinicians at a single diagnostic center. These numbers vary greatly by clinician, services, and the size of the diagnostic centers, but families are most often referred to general practitioners, pediatricians, speech and language therapists, psychologists, and psychiatrists [26]. We calculated our statistics based on an estimate that the average diagnostic center is composed of 5 specialists who can attend to about 200 patients each per year. The actual number is likely lower. Two hospitals in Ohio, both of which have 4-5 clinicians on their teams for diagnostics, experience months-long waitlists, with an estimated 40-100 referrals for ASD each month and about 698-1100 patients with ASD who need ongoing care per year [27]. A school district in Texas, staffs 52 diagnosticians and 45 speech language pathologists across about 60 schools, and estimates that each diagnostician has a caseload of 87.5 students per year and each speech language pathologist has a caseload of 53.3 students per year [28], yet the American Speech-Language-Hearing Association estimates that the median monthly caseload of a speech language pathologist working full-time in the United States was 47 cases per month [29]. For analysis, we set the percent of individuals in need of an autism screening set to 0.195% (given that 6.5% of individuals are at an age appropriate for diagnosis from US Census Data, with 3% of those children needing an autism screening).

Resource availability statistics were also computed (Equation 2 in “Results” section) for each county in the United States and used to generate a heatmap. The results from the previous analyses were used for the parameters (z -average distance from an individual with autism) and (RL-resource load).

GapMap

By having individual families submit their diagnostic and demographic information through GapMap [18], we can eventually calculate a more accurate prevalence rate with more granularity. A simple Web portal was created to allow individuals with ASD to enter relevant data and to crowdsource from the social and global databases. Caregivers (aged ≥ 18 years) of persons with a clinical diagnosis of ASD are allowed to submit data, including gender, date of birth, location including city, state, and zip code, specific (clinically defined) diagnosis of an ASD, diagnostic tools, date of diagnosis, any comorbid conditions, email address, and a list of ratable local services used to care for the subject (where local services include medical specialists, therapists, support resources, and “autism-friendly” generic services). We will then validate the disclosed ASD

diagnosis by having participants complete a peer-reviewed autism screening measure such as the Social Communication Questionnaire (SCQ) [30] or the Social Responsiveness Scale (SRS) [31]. IP addresses, date and time of submission, and similarity of data submitted will also be used to detect duplicate or flag anomalies as potentially falsified data. These two measures will ensure higher quality data, as crowdsourced-data have been shown to match the quality of expert-curated data with proper instructions for data submission and reasonable validation on input data [32-34].

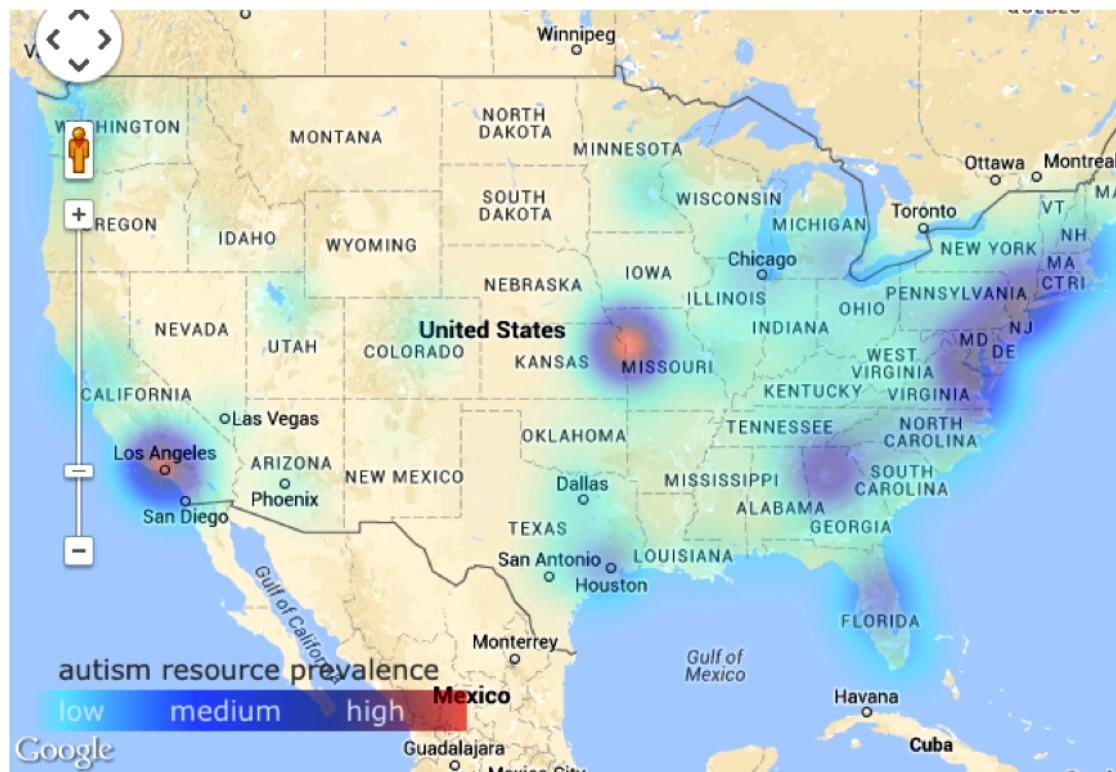
On validation, we will include the family’s data in GapMap; however, if they do not meet ASD criteria, we will not limit the participant from reviewing local resources. Providing families with an accurate list of local resources gives undiagnosed or untreated individuals more chances for earlier diagnosis and an earlier onset of therapy. In addition, we account for the undiagnosed populations within a community by collecting data from participants who select “No diagnosis, but suspicious” under the clinical ASD diagnosis option.

Mainstream social networking platforms, including Facebook, Twitter, colleague networks, consortia, related events and conferences, websites, and fan-based networks, will be used to reach families who have a family member with a clinical diagnosis of autism.

GapMap’s home page also allows users to search and edit its database of autism resources, as shown in Figure 1. Figure 2, which features a Google map with markers and marker clusters for resources and a heatmap overlay, shows resource availability. The metrics resource load, resource gap, and resource availability were defined to aid in this purpose, and estimates were calculated with limited datasets to provide more evidence for the importance of our tool.

Data are encrypted and stored on secure MySQL Databases behind a firewall. GapMap is written in React.js and runs on Amazon Web Services Simple Storage Serve (AWS S3). The backend server runs on AWS app program interface (API) Gateway and AWS Lambda. AWS API Gateway executes specific javascript packages, novel code that interacts with our SQL database, on AWS Lambda. The MySQL relational database is hosted on Amazon Relational Database Service (RDS) and consists of four main databases. Database 1 is unencrypted and stores prevalence rates and resource data; Database 2 is encrypted and stores submitted diagnostic information; Database 3 is encrypted and stores user login information, location, and action-item status; and Database 4 is encrypted and stores the users’ questionnaires. See Figure 3 for an overview of the system architecture.

Figure 1. GapMap features an interactive Google heatmap, comparing resource availability to families with a diagnosed individual with autism. The red coloring on the heatmap shows high autism resource prevalence, while purple coloring shows moderate autism resource prevalence, blue coloring shows low autism resource prevalence, and no coloring shows that, based on our calculations, there are very limited autism-resources available.



Results

Average Distance to the Nearest Diagnostic Center

The distance an individual must travel to the nearest autism diagnostic center is important for several reasons. Studies have shown that individuals are less likely to take advantage of medical services when they are located far away, particularly for low-income families [35,36]. Local care is convenient, and families that must travel long distances to diagnostic facilities may wait to observe more risk factors in their child to schedule an appointment [36]. It may also be the case that areas without any autism resources may lack general knowledge of autism or its characteristic behaviors. These delays may dramatically increase the time to diagnosis, as we know to be the case for other disorders in rural regions [37]. However, given the scarcity of autism diagnostic resources, the typical distance individuals travel for an autism diagnosis may be greater than realized.

Our simulation constructed from US Census data revealed the average distance from an individual in the United States to the nearest diagnostic center to be approximately 182 km (50 miles), with a standard deviation of 235 km (146 miles). The list of

diagnostic centers included hospitals and clinics with dedicated autism diagnosis personnel, but may have missed many solo and small-town diagnosticians. As complete and precise population maps (at the city level) were not available, the simulation also assumed that populations within counties are uniformly distributed, whereas populations are likely to be clustered even within counties. Both of these limitations may slightly overinflate the resulting distances. Even with a fair amount of overestimation, these distances are alarming, and suggest that autism resource allocation is a central issue that needs to be addressed.

The average distance from an individual with ASD to the nearest diagnostic center was also calculated for the United States and the United Kingdom using locational information scraped from the Internet. The dataset for the United States was comprised of 47,622 individuals with autism and 840 developmental medical centers; the dataset for the United Kingdom was comprised of 737 individuals with autism and 135 diagnostic centers. The pool of individuals sampled was restricted to English speakers with Internet access, which can skew toward cities (where autism resources are likely to be more abundant) and younger adults. The results are shown in [Table 1](#).

Figure 2. Example of the mapping interface and home page for GapMap (gapmap.stanford.edu). Participants can electronically consent and participate from any mobile device by clicking on the yellow “Add yourself to the map!” button, as well as toggle between country-level and state-level prevalence of diagnosed autism cases.

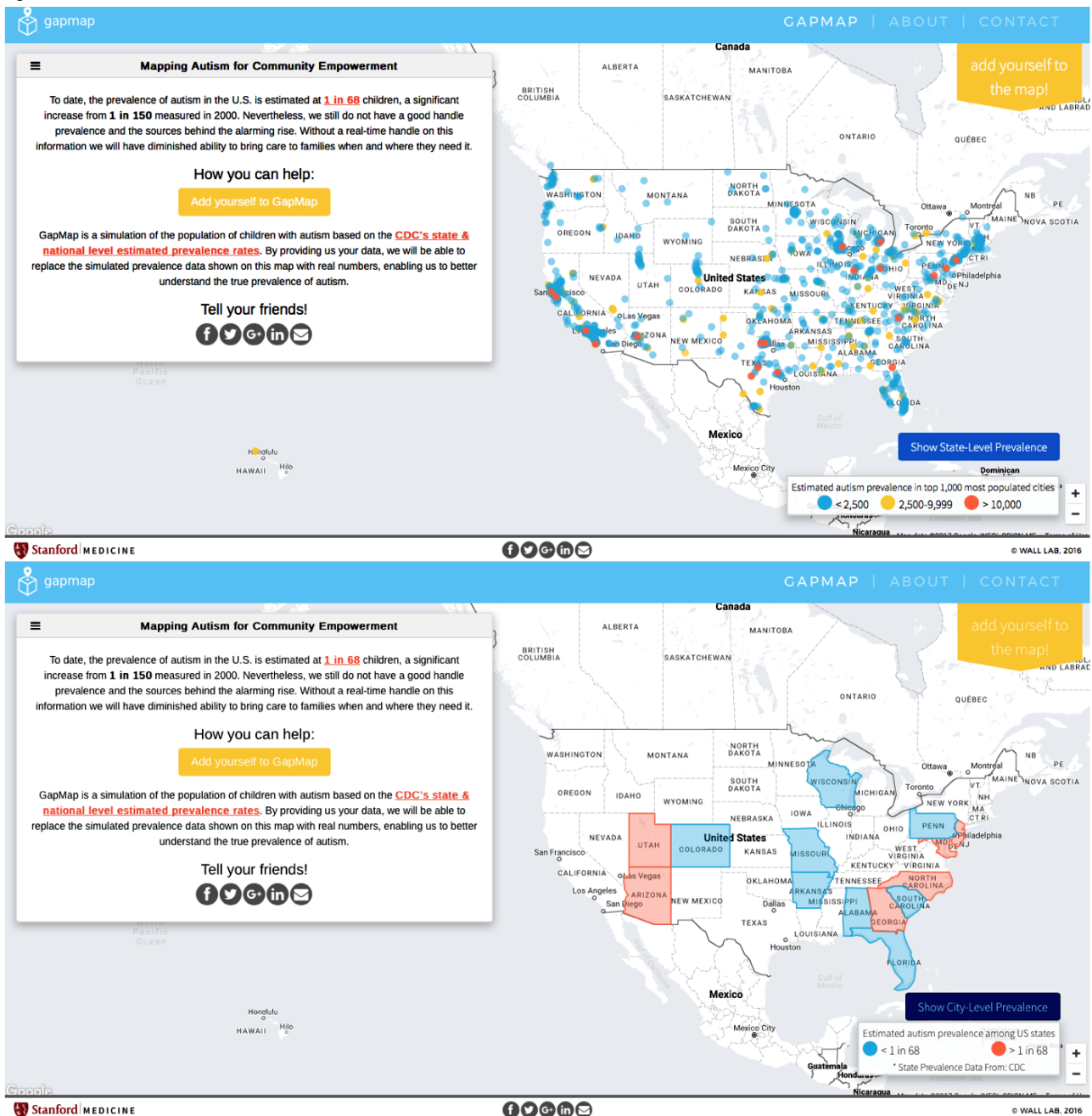


Figure 3. GapMap’s technical architecture. GapMap is hosted on AWS S3, running on a React and Redux front-end framework. The backend framework is comprised of an AWS API Gateway and Lambda Function setup, with secure and scalable end points for retrieving prevalence and resource data, and for submitting participant data. Database 1: unencrypted and stores prevalence rates and resource data; Database 2: encrypted and stores submitted diagnostic information; Database 3: encrypted and stores user login information, location, and action-item status; and Database 4: encrypted and stores the users’ questionnaires.

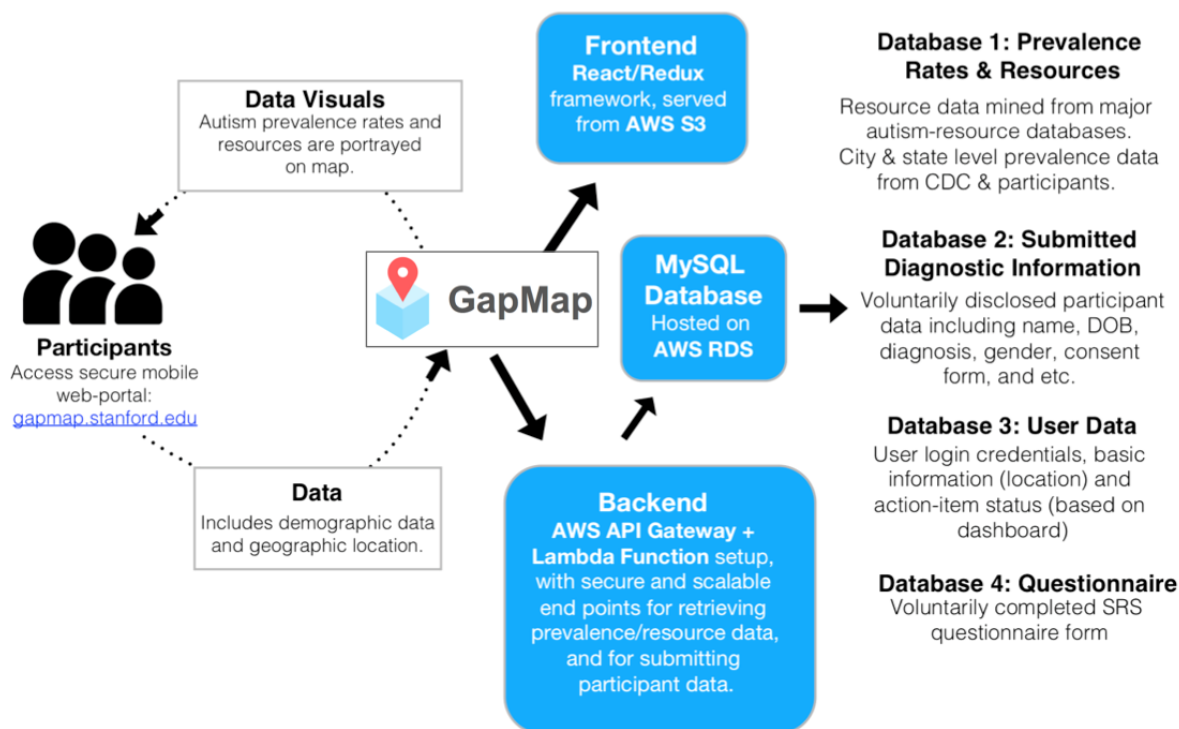


Table 1. Distance between an individual with autism and the nearest diagnostic center in the United States and United Kingdom.

Location	United States (miles)	United Kingdom (miles)
Average distance	32 km (20)	31 km (19)
Median distance	13 km (8)	11 km (7)
Maximum distance	1,819 km (11,130)	519 km (322)
Percent of individuals living within 30 km (19 miles) of a diagnostic center (z)	70%	74%

Notably, the average distances from an individual with autism to the nearest diagnostic center are lesser than the average distance from any individual to the nearest diagnostic center. Although some researchers have hypothesized that this higher rate might be a result of environmental factors in urban areas contributing to this higher incidence [38,39], and distances may be underestimated as a result of biases in sampling methods, these results suggest that individuals who live closer to the diagnostic services are more likely to be diagnosed. This interpretation supports previous research [40], but can also be explained by the incentive for families to move closer to health services if they have concerns for ASD, or other correlational factors such as more advanced parental age (older parents may be more likely to live in urban areas). In order to rule out alternative explanations such as sampling biases or environmental factors and determine whether such an association exists, future work should look at the relationship between age of diagnosis and distance from the nearest diagnostic center. If such a connection does exist, we would expect to see that the average age of diagnosis is lower for individuals closer to a diagnostic provider than for those who must travel extensively.

Although we were not able to collect the date of diagnosis for the individuals in this study, the release of GapMap will soon enable such analyses. GapMap will be an effective tool to support or challenge correlations to certain environmental factors or clusters of high autism incidence. As GapMap collects data increasingly, we will also be able to determine the average distance for individual states in the United States.

Computing Resource Shortages

Although we know autism resources are scarce in many regions, there have not been any attempts to quantify the magnitude or scope of these resource shortages. It will be useful, then, to define several key metrics to determine the existence and magnitude of resource shortages.

We will define resource load (RL) as a measure of how well a center can meet the resource demand placed on it. This is a simple ratio of demand to supply, where demand is the number of individuals who require autism services (Np) and supply is equal to the product of the number of specialists who work at the center (s) and the number of individuals a specialist can attend yearly (o). A resource load of 1 would indicate, then, that

a center is perfectly able to meet the demand. A resource load of 6, however, means that the center is overburdened with 6 times the number of patients it can comfortably handle. We can expect that centers with a resource load greater than one have longer wait times for appointments. See [Figure 4](#) for the equation.

Resource load may be underestimated for some centers and overestimated for others because individuals are often willing to travel greater distances to receive the “best” medical care for their children. In future, we can use ratings and local services usage information collected from GapMap to help explore the extent of this possible phenomenon by checking for correlations between ratings of and distance traveled to each resource.

Although resource load is useful from the perspective of a resource center, it is also important to consider the availability of resources from the perspective of a family looking for ASD resources. We will therefore define resource availability (RA) with respect to a given location. The equation is shown in [Figure 5](#).

In short, resource availability is the sum of the resources available in a given location, where each resource’s contribution is adjusted by its resource load and its distance to the given location. Resources that are very far away and resources that are overburdened are given little weight. Resource availability will be high when resources with reasonable resource load are nearby, and will be at least equal to one if one resource exists within a reasonable distance with a resource load of one. A resource gap, or region in which there exist limited diagnostic

or treatment resources with respect to the demand, is indicated by a resource availability of less than 1. These regions are where we should focus most on improving the accessibility of resources.

The parameter z accounts for how far the average individual will travel to seek out a resource. Note that this distance may vary with the type of resource—for example, individuals may be willing to travel farther for behavioral therapy (once a child has been diagnosed) than for diagnostic resources, when they are acting only on suspicions of developmental delays. For this study, we estimate z for diagnostic centers at 30 km by using the results of our previous analysis, where we estimate that 70% of individuals with autism in the United States live within 30 km of a diagnostic center.

We computed resource load for 840 developmental medical centers in the United States. About 0.5% of the centers had a resource load of less than one, and even after giving some leeway to account for the limited information on diagnosticians and diagnostic centers, only 5% of the centers had a resource load of less than 5. Alarmingly, 18% of centers had a resource load greater than 25, or experienced 25 times the demand they could handle. The average center had a resource load of 18 and a standard deviation of 14. The median center had a resource load of 14, and the maximum resource load experienced by a center was 113. These numbers are distressing, but not particularly surprising, given the difficulty and time involved in obtaining an autism diagnosis. Regardless, they make a convincing case for further research in and reassessment of our investment in diagnostic resources.

Figure 4. Equation for the resource load for a given resource. Where: r = a given resource; RL_r = resource load for a given resource (r); N = the number of individuals nearby; p = the proportion of individuals who are in need of resources; s = the number of specialists; and o = the number of individuals a specialist can attend yearly.

$$RL_r = \frac{Np}{so}$$

Figure 5. Equation for the resource availability for a given location. Where: l = given location; RA_l = resource availability for a given location (l); R = the pool of resource options, where r is one such resource; $d(r,l)$ = the distance between the resource r and the given location (l); RL_r = the resource load for the resource r ; and z = the average distance an individual is willing to travel for a resource.

$$RA_l = \sum_{r \in R} \frac{z}{d(r, l) \cdot RL_r}$$

Discussion

Principal Findings

We are missing the information about diagnostic capabilities in smaller facilities and rural areas. Specialists ranging from general practitioners to developmental pediatricians can (but are not required to) be certified to diagnose ASD, and most use brief, unstandardized assessment instruments to diagnose ASD [15,26]. Not all certified specialists feel comfortable giving a diagnosis and will refer parents to another professional for

diagnosis [26]. Because of this, it is difficult to know which clinicians are also autism diagnosticians. This phenomenon is particularly true in rural areas, where facilities dedicated specifically to ASD are few or nonexistent and generalist clinicians are more prevalent [13,14]. This is problematic because parents who are referred from their pediatrician to a specialist are more likely to receive a diagnosis sooner [14]. In these very rural areas, parents may need to rely on a family doctor or pediatrician to recognize risk factors or early signs of ASD. Soon, we will be able to use GapMap to learn more about these complex dynamics and come to a better understanding of

autism diagnoses, small town diagnosticians, and rural specialists who provide referrals.

Although these analyses and metrics highlight the lack of resources in much of the United States and the overburdening of many centers, they are not enough. We have built GapMap as a tool to collect important information and visually display the results. The collected location information, diagnosis, diagnostic tools, and comorbid conditions will be used to obtain both widespread and highly localizable autism prevalence rates. Date of diagnosis and age will be aggregated and used to obtain localizable average age of diagnosis, a measure that correlates with difficulty obtaining a diagnosis and can be used to help approximate geographic differences in resource accessibility. We will also collect the specific diagnoses of individuals who submit their data, including diagnostic and statistical manual of mental disorders (DSM-IV and DSM-5) diagnoses and an option for “no diagnosis, but suspicious.” This information will allow us to track individuals who remain undiagnosed, providing us with valuable data that is often overlooked in studies of prevalence rates. Ratings and local services will be used to estimate resource usage trends with respect to geography and

resource density. Prevalence rates and local service usage will also be used to calculate resource load and availability for different resource types such as behavioral therapy.

Conclusions

There is a dearth of ASD resources, as well as a lack of understanding of the extent and epidemiology of these resource gaps. Statistics computed from simulations and web scraping suggest that individuals located close to diagnostic centers are more likely to be diagnosed. By quantitatively defining resource load and resource availability, we provide a means to measure and enable the alleviation of increasingly overburdened diagnostic centers and resource-poor areas where parents are unable to diagnose their children as quickly and easily as needed. The release of GapMap will collect crucial information that will provide more accurate data for computing resource loads and availability, uncovering the impact of resource epidemiology on age and likelihood of diagnosis, and gathering localized autism prevalence rates, both from families that have already received a diagnosis and families that haven't received an official diagnosis yet.

Acknowledgments

The authors acknowledge support from Stanford University School of Medicine. They thank Anish Nag, Anika Kumar, and Sylvia Illouz for their work in mining autism-related-resource databases. They also extend thanks to Matthew Hoying, for his advice and contributions with database design and construction. All authors read and approved the final manuscript. The work was supported in part by Stanford's Child Health Research Institute.

Conflicts of Interest

None declared.

References

1. Center for Disease Control and Prevention (CDC). 2016. Prevalence of autism spectrum disorder among children aged 8 years? autism and developmental disabilities monitoring network, 11 sites, United States, 2010 URL: <https://www.cdc.gov/mmwr/preview/mmwrhtml/ss6302a1.htm> [accessed 2016-12-13] [WebCite Cache ID 6mjYzkcZ4]
2. White S. Anxiety in children and adolescents with autism spectrum disorders. *Clin Psychol Rev* 2009;29(3):216-229. [doi: [10.1016/j.cpr.2009.01.003](https://doi.org/10.1016/j.cpr.2009.01.003)]
3. Elsabbagh M, Divan G, Koh Y, Kim YS, Kauchali S, Marcín C, et al. Global prevalence of autism and other pervasive developmental disorders. *Autism Res* 2012 Jun;5(3):160-179 [FREE Full text] [doi: [10.1002/aur.239](https://doi.org/10.1002/aur.239)] [Medline: [22495912](https://pubmed.ncbi.nlm.nih.gov/22495912/)]
4. Maenner MJ, Yeargin-Allsopp M, Van Naarden BK, Christensen DL, Schieve LA. Development of a machine learning algorithm for the surveillance of autism spectrum disorder. *PLoS One* 2016 Dec;11(12):e0168224 [FREE Full text] [doi: [10.1371/journal.pone.0168224](https://doi.org/10.1371/journal.pone.0168224)] [Medline: [28002438](https://pubmed.ncbi.nlm.nih.gov/28002438/)]
5. Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008 Principal Investigators, Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders--Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. *MMWR Surveill Summ* 2012 Mar 30;61(3):1-19 [FREE Full text] [Medline: [22456193](https://pubmed.ncbi.nlm.nih.gov/22456193/)]
6. Developmental, Disabilities Monitoring Network Surveillance Year, and 2010 Principal Investigators. Prevalence of autism spectrum disorder among children aged 8 years--autism and developmental disabilities monitoring network, 11 sites, United States, 2012. *MMWR Surveill Summ* 2016;65(3):1-23. [doi: [10.15585/mmwr.ss6503a1](https://doi.org/10.15585/mmwr.ss6503a1)]
7. Mazurek MO, Handen BL, Wodka EL, Nowinski L, Butter E, Engelhardt CR. Age at first autism spectrum disorder diagnosis: the role of birth cohort, demographic factors, and clinical features. *J Dev Behav Pediatr* 2014;35(9):561-569. [doi: [10.1097/DBP.000000000000097](https://doi.org/10.1097/DBP.000000000000097)] [Medline: [25211371](https://pubmed.ncbi.nlm.nih.gov/25211371/)]
8. Shattuck PT, Durkin M, Maenner M, Newschaffer C, Mandell DS, Wiggins L, et al. Timing of identification among children with an autism spectrum disorder: findings from a population-based surveillance study. *J Am Acad Child Adolesc Psychiatry* 2009 May;48(5):474-483 [FREE Full text] [doi: [10.1097/CHI.0b013e31819b3848](https://doi.org/10.1097/CHI.0b013e31819b3848)] [Medline: [19318992](https://pubmed.ncbi.nlm.nih.gov/19318992/)]

9. Tonge BJ, Bull K, Brereton A, Wilson R. A review of evidence-based early intervention for behavioural problems in children with autism spectrum disorder: the core components of effective programs, child-focused interventions and comprehensive treatment models. *Curr Opin Psychiatry* 2014 Mar;27(2):158-165. [doi: [10.1097/YCO.0000000000000043](https://doi.org/10.1097/YCO.0000000000000043)] [Medline: [24452070](https://pubmed.ncbi.nlm.nih.gov/24452070/)]
10. Russell G, Ford T, Steer C, Golding J. Identification of children with the same level of impairment as children on the autistic spectrum, and analysis of their service use. *J Child Psychol Psychiatry* 2010 Jun;51(6):643-651. [doi: [10.1111/j.1469-7610.2010.02233.x](https://doi.org/10.1111/j.1469-7610.2010.02233.x)] [Medline: [20345841](https://pubmed.ncbi.nlm.nih.gov/20345841/)]
11. Miller L, Reynolds J. Autism and vaccination—the current evidence. *J Spec Pediatr Nurs* 2009 Jul;14(3):166-172. [doi: [10.1111/j.1744-6155.2009.00194.x](https://doi.org/10.1111/j.1744-6155.2009.00194.x)] [Medline: [19614825](https://pubmed.ncbi.nlm.nih.gov/19614825/)]
12. Hutton AM, Caron SL. Experiences of families with children with autism in rural New England. *Focus Autism Other Dev Disabl* 2005 Jan 01;20(3):180-189. [doi: [10.1177/10883576050200030601](https://doi.org/10.1177/10883576050200030601)]
13. Mandell DS, Novak MM, Zubritsky CD. Factors associated with age of diagnosis among children with autism spectrum disorders. *Pediatrics* 2005 Dec;116(6):1480-1486 [FREE Full text] [doi: [10.1542/peds.2005-0185](https://doi.org/10.1542/peds.2005-0185)] [Medline: [16322174](https://pubmed.ncbi.nlm.nih.gov/16322174/)]
14. Wiggins LD, Baio J, Rice C. Examination of the time between first evaluation and first autism spectrum diagnosis in a population-based sample. *J Dev Behav Pediatr* 2006 Apr;27(2 Suppl):S79-S87. [Medline: [16685189](https://pubmed.ncbi.nlm.nih.gov/16685189/)]
15. Bernier R, Mao A, Yen J. Psychopathology, families, and culture: autism. *Child Adolesc Psychiatr Clin N Am* 2010 Oct;19(4):855-867. [doi: [10.1016/j.chc.2010.07.005](https://doi.org/10.1016/j.chc.2010.07.005)] [Medline: [21056350](https://pubmed.ncbi.nlm.nih.gov/21056350/)]
16. Howlin P. *Children with Autism and Asperger's Syndrome: A Guide for Practitioners and Carers*. Chichester, UK: Wiley; 1998.
17. Pinto-Martin JA, Young LM, Mandell DS, Poghosyan L, Giarelli E, Levy SE. Screening strategies for autism spectrum disorders in pediatric primary care. *J Dev Behav Pediatr* 2008 Oct;29(5):345-350. [doi: [10.1097/DBP.0b013e31818914cf](https://doi.org/10.1097/DBP.0b013e31818914cf)] [Medline: [18852608](https://pubmed.ncbi.nlm.nih.gov/18852608/)]
18. GapMap. 2017. The Wall Lab URL: <https://gapmap.stanford.edu/> [accessed 2017-03-28] [WebCite Cache ID 6pJ7eut0E]
19. Jost CC, Mariner JC, Roeder PL, Sawitri E, Macgregor-Skinner GJ. Participatory epidemiology in disease surveillance and research. *Rev Sci Tech* 2007 Dec;26(3):537-549. [Medline: [18293603](https://pubmed.ncbi.nlm.nih.gov/18293603/)]
20. Freifeld CC, Chunara R, Mekaru SR, Chan EH, Kass-Hout T, Ayala IA, et al. Participatory epidemiology: use of mobile phones for community-based health reporting. *PLoS Med* 2010 Dec;7(12):e1000376 [FREE Full text] [doi: [10.1371/journal.pmed.1000376](https://doi.org/10.1371/journal.pmed.1000376)] [Medline: [21151888](https://pubmed.ncbi.nlm.nih.gov/21151888/)]
21. 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/)]
22. U.S. Census Bureau, Population Division. Factfinder. 2016. Annual estimates of the resident population: April 1, 2010 to July 1, 2015 URL: <https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkml> [accessed 2016-11-28] [WebCite Cache ID 6mjYKsr3J]
23. Autism Speaks. 2016. URL: <https://www.autismspeaks.org/> [accessed 2016-11-29] [WebCite Cache ID 6mjYZXDpC]
24. Autism Source. 2004. URL: <http://www.autismsource.org> [accessed 2016-11-13] [WebCite Cache ID 6mjYgryAU]
25. Autism. 2016. The National Autistic Society URL: <http://www.autism.org.uk/> [accessed 2016-12-13] [WebCite Cache ID 6mjYlo6TS]
26. Crane L, Chester J, Goddard L, Henry L, Hill E. Experiences of autism diagnosis: a survey of over 1000 parents in the United Kingdom. *Autism* 2016;20(20):153-162. [doi: [10.1177/1362361315573636](https://doi.org/10.1177/1362361315573636)]
27. Austin J, Manning-Courtney P, Johnson M, Weber R, Johnson H, Murray D, et al. Improving access to care at autism treatment centers: a system analysis approach. *Pediatrics* 2016;137(2):S149-S157 [FREE Full text] [doi: [10.1542/peds.2015-2851M](https://doi.org/10.1542/peds.2015-2851M)]
28. Fort Bend ISD. 2010. Staffing review URL: <http://www.fortbendisd.com/cms/lib09/TX01917858/Centricity/Domain/96/tasb-salay-study.pdf> [accessed 2017-02-24] [WebCite Cache ID 6oWhx6m0o]
29. American Speech-Language-Hearing Association. 2016. ASHA 2014 schools survey: speech language pathologist caseload characteristics report URL: <http://www.asha.org/uploadedFiles/2014-Schools-Survey-SLP-Caseload-Characteristics.pdf> [accessed 2017-02-23] [WebCite Cache ID 6oWhKsVXG]
30. Bölte S, Holtmann M, Poustka F. The social communication questionnaire (SCQ) as a screener for autism spectrum disorders: additional evidence and cross-cultural validity. *J Am Acad Child Adolesc Psychiatry* 2008 Jun;47(6):719-720. [doi: [10.1097/CHI.0b013e31816c42bd](https://doi.org/10.1097/CHI.0b013e31816c42bd)] [Medline: [18496329](https://pubmed.ncbi.nlm.nih.gov/18496329/)]
31. Aldridge FJ, Gibbs VM, Schmidhofer K, Williams M. Investigating the clinical usefulness of the social responsiveness scale (SRS) in a tertiary level, autism spectrum disorder specific assessment clinic. *J Autism Dev Disord* 2012 Feb;42(2):294-300. [doi: [10.1007/s10803-011-1242-9](https://doi.org/10.1007/s10803-011-1242-9)] [Medline: [21516433](https://pubmed.ncbi.nlm.nih.gov/21516433/)]
32. Comber A, Brunson C, See L, Fritz S, McCallum I. Comparing expert and non-expert conceptualisations of the land: an analysis of crowdsourced land cover data. In: Tenbrink T, Stell J, Galton A, Wood Z, editors. *Spatial Information Theory. COSIT 2013. Lecture Notes in Computer Science*, vol 8116. Switzerland: Springer International Publishing; 2013:243-260. [doi: [10.1007/978-3-319-01790-7_14](https://doi.org/10.1007/978-3-319-01790-7_14)]

33. Behrend TS, Sharek DJ, Meade AW, Wiebe EN. The viability of crowdsourcing for survey research. *Behav Res Methods* 2011 Sep;43(3):800-813. [doi: [10.3758/s13428-011-0081-0](https://doi.org/10.3758/s13428-011-0081-0)] [Medline: [21437749](https://pubmed.ncbi.nlm.nih.gov/21437749/)]
34. Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. *J Med Internet Res* 2012;14(2):e46. [doi: [10.2196/jmir.1988](https://doi.org/10.2196/jmir.1988)]
35. Currie J, Reagan P. Distance to hospital and children's use of preventive care: is being closer better, and for whom? *Economic Inquiry* 2003;41(3):378-391. [doi: [10.1093/ei/cbg015](https://doi.org/10.1093/ei/cbg015)]
36. Roberts A, Blunt I, Bardsley M. Focus On: Distance from home to emergency care. London: The Health Foundation and the Nuffield Trust; 2014.
37. Pfeiffer P. Impact of distance facility of initial diagnosis on depression treatment. *Health Serv Res* 2011;46(3):768-786 [FREE Full text] [doi: [10.1111/j.1475-6773.2010.01228.x](https://doi.org/10.1111/j.1475-6773.2010.01228.x)]
38. Rice CE, Baio J, Van Naarden BK, Doernberg N, Meaney FJ, Kirby RS. A public health collaboration for the surveillance of autism spectrum disorders. *Paediatr Perinat Epidemiol* 2007 Mar;21(2):179-190. [doi: [10.1111/j.1365-3016.2007.00801.x](https://doi.org/10.1111/j.1365-3016.2007.00801.x)] [Medline: [17302648](https://pubmed.ncbi.nlm.nih.gov/17302648/)]
39. Nevison CD. A comparison of temporal trends in United States autism prevalence to trends in suspected environmental factors. *Environ Health* 2014 Sep;13(1):73 [FREE Full text] [doi: [10.1186/1476-069X-13-73](https://doi.org/10.1186/1476-069X-13-73)] [Medline: [25189402](https://pubmed.ncbi.nlm.nih.gov/25189402/)]
40. Mazumdar S, Winter A, Liu K, Bearman P. Spatial clusters of autism births and diagnoses point to contextual drivers of increased prevalence. *Soc Sci Med* 2013 Oct;95:87-96 [FREE Full text] [doi: [10.1016/j.socscimed.2012.11.032](https://doi.org/10.1016/j.socscimed.2012.11.032)] [Medline: [23267775](https://pubmed.ncbi.nlm.nih.gov/23267775/)]

Abbreviations

API: application program interface
ASD: Autism spectrum disorders
AWS S3: Amazon Web Services Simple Storage Serve
CDC: The Centers for Disease Control and Prevention
DSM: diagnostic and statistical manual of mental disorders
RDS: Amazon Relational Database Service
SCQ: Social Communication Questionnaire
SRS: Social Responsiveness Scale

Edited by G Eysenbach; submitted 14.12.16; peer-reviewed by K Bowers; comments to author 22.02.17; revised version received 15.03.17; accepted 16.03.17; published 04.05.17.

Please cite as:

Albert N, Daniels J, Schwartz J, Du M, Wall DP
GapMap: Enabling Comprehensive Autism Resource Epidemiology
JMIR Public Health Surveill 2017;3(2):e27
URL: <http://publichealth.jmir.org/2017/2/e27/>
doi: [10.2196/publichealth.7150](https://doi.org/10.2196/publichealth.7150)
PMID: [28473303](https://pubmed.ncbi.nlm.nih.gov/28473303/)

©Nikhila Albert, Jena Daniels, Jessey Schwartz, Michael Du, Dennis P Wall. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 04.05.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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

Making Air Pollution Visible: A Tool for Promoting Environmental Health Literacy

Ekaterina Galkina Cleary¹, MSc, PhD; Allison P Patton², PhD; Hsin-Ching Wu^{3,4}, MA, MPP; Alan Xie⁴, BA; Joseph Stubblefield⁵, BS; William Mass⁵, MPH, PhD; Georges Grinstein^{5,6}, MS, PhD; Susan Koch-Weser⁷, ScM, ScD; Doug Brugge^{7,8,9}, MS, PhD; Carolyn Wong⁴, MS, PhD

¹Center for Integration of Science and Industry, Bentley University, Waltham, MA, United States

²Environmental and Occupational Health Sciences Institute, Rutgers University, Piscataway, NJ, United States

³Department of Public Policy and Public Affairs, University of Massachusetts Boston, Boston, MA, United States

⁴Institute for Asian American Studies, University of Massachusetts Boston, Boston, MA, United States

⁵Computer Science Department, University of Massachusetts Lowell, Lowell, MA, United States

⁶Center for Data Science, University of Massachusetts Amherst, Amherst, MA, United States

⁷Department of Public Health and Community Medicine, Tufts University School of Medicine, Boston, MA, United States

⁸Department of Civil and Environmental Engineering, Tufts University, Medford, MA, United States

⁹Tisch College of Civic Life, Tufts University, Medford, MA, United States

Corresponding Author:

Carolyn Wong, MS, PhD

Institute for Asian American Studies

University of Massachusetts Boston

100 Morrissey Blvd

Boston, MA, 02125

United States

Phone: 1 650 269 4421

Fax: 1 617 287 5656

Email: carolyn.wong@umb.edu

Related Article:

This is a corrected version. See correction statement: <http://publichealth.jmir.org/2017/4/e80/>

Abstract

Background: Digital maps are instrumental in conveying information about environmental hazards geographically. For laypersons, computer-based maps can serve as tools to promote environmental health literacy about invisible traffic-related air pollution and ultrafine particles. Concentrations of these pollutants are higher near major roadways and increasingly linked to adverse health effects. Interactive computer maps provide visualizations that can allow users to build mental models of the spatial distribution of ultrafine particles in a community and learn about the risk of exposure in a geographic context.

Objective: The objective of this work was to develop a new software tool appropriate for educating members of the Boston Chinatown community (Boston, MA, USA) about the nature and potential health risks of traffic-related air pollution. The tool, the Interactive Map of Chinatown Traffic Pollution (“Air Pollution Map” hereafter), is a prototype that can be adapted for the purpose of educating community members across a range of socioeconomic contexts.

Methods: We built the educational visualization tool on the open source Weave software platform. We designed the tool as the centerpiece of a multimodal and intergenerational educational intervention about the health risk of traffic-related air pollution. We used a previously published fine resolution (20 m) hourly land-use regression model of ultrafine particles as the algorithm for predicting pollution levels and applied it to one neighborhood, Boston Chinatown. In designing the map, we consulted community experts to help customize the user interface to communication styles prevalent in the target community.

Results: The product is a map that displays ultrafine particulate concentrations averaged across census blocks using a color gradation from white to dark red. The interactive features allow users to explore and learn how changing meteorological conditions and traffic volume influence ultrafine particle concentrations. Users can also select from multiple map layers, such as a street

map or satellite view. The map legends and labels are available in both Chinese and English, and are thus accessible to immigrants and residents with proficiency in either language. The map can be either Web or desktop based.

Conclusions: The Air Pollution Map incorporates relevant language and landmarks to make complex scientific information about ultrafine particles accessible to members of the Boston Chinatown community. In future work, we will test the map in an educational intervention that features intergenerational colearning and the use of supplementary multimedia presentations.

(*JMIR Public Health Surveill* 2017;3(2):e16) doi:[10.2196/publichealth.7492](https://doi.org/10.2196/publichealth.7492)

KEYWORDS

computer visualization; digital cartography; environmental health literacy; health communication; environmental health; computer-based education; air pollution; ultrafine particles; immigrant education

Introduction

There is a pressing need for new tools to promote environmental health literacy [1]. As research reveals new knowledge about the public health effects of environmental pollutants, it remains challenging to communicate complex scientific information to laypersons. Environmental health literacy will be significantly improved when people not only learn factual knowledge but also understand what tactics can reduce their exposure to risk [1]. Recent advances in the technology of computer mapping hold the potential to increase environmental health literacy by helping users visualize information about environmental hazards in spatial terms and within a geographic area. By forming mental representations of this information, users may be able to make more informed decisions and have more well-defined discussions when considering health-promoting measures [2].

This work aimed to develop a computer visualization tool to promote environmental health literacy about traffic-related pollution in Boston Chinatown, Massachusetts, USA. One component of traffic pollution consists of particulate matter. Particulate matter in ambient air is a leading cause of mortality and morbidity worldwide. It is estimated that ambient particulate matter, which consists of particles 2.5 μm or less in diameter (PM_{2.5}), causes 100,000 to 200,000 deaths in the United States and 3.2 million deaths globally each year [3-5]. Ultrafine particles, a subfraction of PM_{2.5} that are less than 0.1 μm in diameter, tend to be higher locally near heavy traffic. Thus, a community's proximity to highways and major roadways presents a health risk to its inhabitants [6,7]. In Chinatown, where poverty levels are higher than those in most other parts of the city, many residents have low levels of educational attainment and limited English proficiency [8]. For this reason, it is important to improve residents' access to health information.

People vary in their capacity to access, learn, and use information [9]. Efforts to disseminate health information often result in differential learning, with the poor and socially disadvantaged benefiting the least [10,11]. Messages about environmental health need to be tailored to audiences with varying socioeconomic and ethnic backgrounds [1]. Extensive research points to the importance of making computer tools accessible to individuals with low levels of English proficiency and limited computer skills [12]. Development of accessible tools for improving environmental health literacy would help educators cross these linguistic and technological barriers.

Pictures can aid the accessibility of health communication messages [13]. Studies have shown that visual information is better retained than verbal information by users in the cognitive process of building mental models [14-17]. Because of this, visualization technology is particularly well-suited to education about ultrafine particles, as the small size of these particles makes them invisible to the naked eye, and people are often unaware of their presence. Consequently, visualization of such ultrafine particle concentrations on a computer screen provides a means of increasing community awareness.

Visualization tools can also aid community health education. For example, computer maps can offer multiple views of a geographic area and have the potential to improve cognitive performance. The main bottleneck in cognitive processing and visual thinking is often the limited capacity of working memory, which can retain only about 3 to 5 objects [18]. Having access to more than one visualization can allow users to store information in working memory while thinking about, analyzing, and processing that information, and can improve learning and decision making [19]. In a map of pollution, for example, one map layer might show selected features of the built environment, such as roadways and high-rise buildings. Another layer can depict the location and extent of open green space or vegetation. In addition, interactivity of digital maps can further facilitate learning [20]. Users can manipulate the display through visual animations that convey different types of information. This would allow them to explore the effect of temperature and wind severity, and thereby increase the potential for interest and engagement.

Visualization additionally helps overcome the limitations of health literacy promotion reliant on textual representation of information, which often simplifies concepts for readability [21]. Simplification can erase nuances important to comprehension of complex scientific information in general and, in our case, about ultrafine particles. A map showing the spatial distribution of the particles in a community can improve awareness of many factors in the physical environment that affect community residents' daily lives, such as roadways, green space, and residential buildings. As environmental researchers discover how changes in concentrations of ultrafine particles in the air are affected by factors such as temperature and wind, it becomes possible to depict these changing concentrations by means of animated visualizations.

The objective of this work was to develop a visualization tool that would educate members of the Boston Chinatown

community about one form of traffic-related pollution, ultrafine particles. This paper focuses on development of the Interactive Map of Chinatown Traffic Pollution (or Air Pollution Map). Its use in educational interventions will be discussed in a future publication.

Methods

We designed the Air Pollution Map as the first step of a community educational intervention to meet the challenges of disseminating complex scientific knowledge about environmental hazards to laypersons. Users can explore spatial patterns of ultrafine particle concentrations, measured in particle number concentration (PNC) units, and see how concentrations vary with changing conditions, including temperature, wind severity, wind direction, and traffic volume. The map we created provides 4 visual controls as sliders, which allow a user to modify conditions of temperature, wind direction, wind speed, and traffic volume (model explained below). We created bilingual English-Chinese Web-based and desktop-based versions of the visualization.

We created the Air Pollution Map using Weave v1.9.38, a Web-based analysis and visualization environment [22]. Weave was developed at the University of Massachusetts Lowell, Lowell, MA, USA and is now supported as an open source tool. Weave provides the ability to integrate, analyze, and visualize data using a suite of over 20 different interactive tools. Completed visualizations can be disseminated via a website.

Weave [23] is freely available to the public and is simple to install on personal computers. Its accessibility facilitates its use by educators in community settings such as schools, community-based organizations, libraries, and homes. The interactivity of the tool that we developed in Weave is designed to be used as a launching point for group-based conversation about the problems posed by air pollution and tactics that communities can pursue to reduce exposure.

The data used to construct the Air Pollution Map were collected and analyzed by an investigatory team of the Community

Assessment of Freeway Exposure and Health Study (CAFEH). CAFEH measured ultrafine particle concentrations under a variety of weather and traffic conditions in Boston Chinatown [24] and built a statistical model that describes how the ambient conditions were related to the ultrafine particle concentrations [25]. We applied this model as an algorithm and incorporated it into the interactive map in ways that we hoped would allow users to interactively explore hypothetical scenarios that affect ultrafine particle concentrations across the community.

The variables included in our algorithm were wind speed, temperature, traffic volume, and wind direction. These were the primary predictive variables in the ultrafine particles model (Table 1). The model predicts lower ultrafine particle concentrations when winds are stronger, which blow the particles away more effectively. Higher concentrations are expected when there are more cars traveling on Interstate 93, a major north-south highway near Boston Chinatown, resulting in increased emissions. Concentrations are also higher with lower temperatures due to (1) more formation of particles from less-efficient engines, (2) less dispersion of particles in the air, and (3) lower particle removal due to slower evaporation and particles sticking together. The model predicts that when the wind comes from the east, concentrations are highest near Interstate 93, the major highway, and a major train station on the far side of the highway (see Figure 1). Under these conditions, concentrations decrease toward the west side of Chinatown. We developed and ran the model in the R-Project v3.1.1 programming language (R Core Team, R Foundation for Statistical Computing), and we averaged PNC to census blocks for use in our map.

We consulted community experts to assist in developing the user interface for the map. Two community health promoters who had prior experience in educational projects concerning traffic pollution in Chinatown recommended Chinese wording for menu labels and selected community landmarks recognizable to residents.

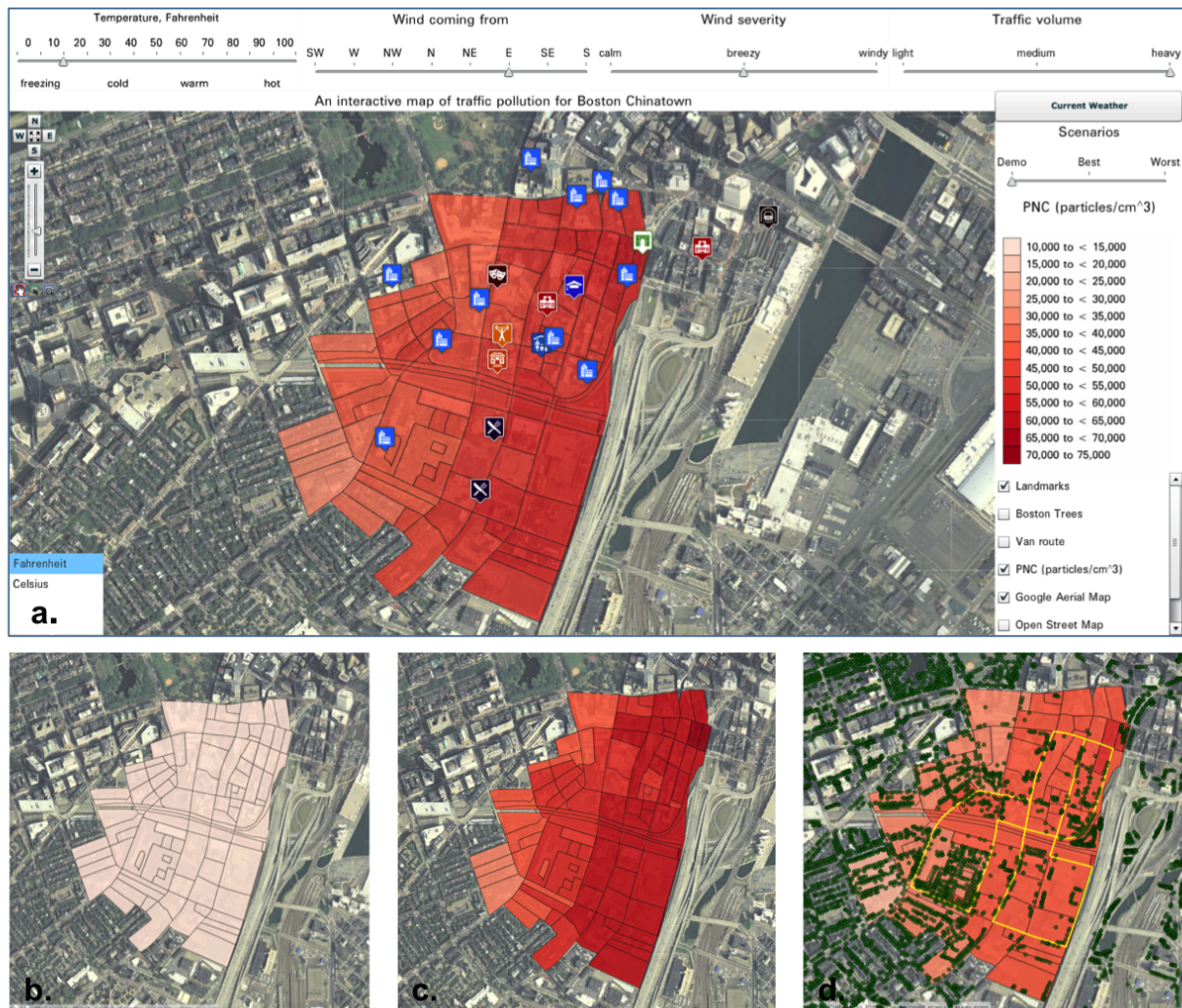
The University of Massachusetts Institutional Review Board approved this study (#2014161).

Table 1. Variables of interest and their effects on particle number concentration (PNC).

Variable	Effect of increase in variable	Selectable values in Weave
Wind speed	Percentage decrease in PNC	Calm, breezy, windy
Temperature	Percentage decrease in PNC	0 to 100°F or -20 to 40°C
Traffic volume	Percentage increase in PNC	Light, medium, heavy
Wind direction	More complicated: highest for east winds, and with larger effect near Interstate 93	N, NE, E, SE, S, SW, W, NW ^a

^aWind direction variables: north (N), northeast (NE), east (E), southeast (SE), south (S), southwest (SW), west (W), northwest (NW).

Figure 1. Several views of the Interactive Map of Chinatown Traffic Pollution: (a) Full aerial Google maps view showing sliders, map layers, scenario slider, and color legend; (b) scenario of the lowest ultrafine particle concentrations; (c) scenario of the highest ultrafine particle concentrations; (d) Community Assessment of Freeway Exposure and Health Study mobile laboratory route (solid line) taken to measure ultrafine particles and locations of surrounding trees (dots). Visualizations created using Weave software v1.9.38. PNC: particle number concentration.



Results

The Air Pollution Map (Figure 1) allows users to see how meteorological and traffic conditions can affect changes in ultrafine particle concentration by moving the 4 sliders for temperature, wind direction, wind severity, and traffic volume (Figure 1 a). PNC at each census block is shown through choropleth mapping, in which brighter color intensity indicates higher pollution levels (key in Figure 1 a). We chose the wording used for the labels on the sliders to conform to popular language conventions among members of the Chinatown community. For example, “calm,” “breezy,” and “windy” are commonly understood terms describing wind severity. Icons for commonly recognized landmarks are used to orient the user within the neighborhood.

There are several layers to the map that can be made visible or hidden. These include a base map showing the street level or a satellite view of the study area with prominent landmarks (Figure 1 a), the route taken by the CAFEH mobile laboratory when measuring ultrafine particles (Figure 1 d), the location of

trees (Figure 1 d), and the ultrafine particle concentration in particles/cm³ (Figure 1 b and 1c). These features allow users to learn experientially by changing values for the variables, such as setting the weather conditions to current values. In doing so, they can discover which combination of conditions leads to the highest (Figure 1 c) and lowest (Figure 1 b) concentrations of ultrafine particles. They can also explore effects of each of the input parameters while holding others constant or in interaction with one another. We included a separate “scenario” slider to quickly navigate to the conditions producing the best and worst levels of pollution (Figure 1 a). Additional exploratory map features include the ability to drill down to Chinatown street levels or zoom out to see adjacent neighborhoods.

Advanced toolbar features of Weave are purposefully hidden, so that buttons not vital to the educational process are not accessible to the novice user. The educational version of the map also hides the scenario slider, so that students may independently arrive at the contributors to the best and worst pollution levels by exploration. Weave’s session state-based architecture allows all significant interactions made with the

system to be recorded (eg, selecting, subsetting, and toggling). By analyzing or even simply replaying the actions each student took while exploring the map, educators may find patterns that are informative about the learning process. For example, teachers can analyze steps students take to identify what combinations of temperature and wind conditions are associated with high levels of pollution.

The bilingual feature and the dual temperature scale (Celsius/Fahrenheit) were necessary to accommodate the needs of the different segments of the community, including longtime residents and newly arrived immigrants. The visualization tool was used in educational workshops targeted to high school youth, English-language adult learners, and seniors in Boston Chinatown. We developed a plan for evaluation of the educational intervention. In future work, we will analyze data collected on the effectiveness of the intervention and produce a report of findings.

A stable version of the Interactive Map of Chinatown Traffic Pollution is available [26]. The bilingual desktop version is available free of charge from the first or last author upon request.

Discussion

Our Air Pollution Map aimed to make complex information about the spatial distribution of ultrafine particulate pollution in a community accessible and usable for residents of Boston Chinatown and to aid them in engaging in health-promoting behaviors. Health educators could, for example, use the visualizations to explain why altering certain choices, such as time and place of outdoor exercise, can reduce exposure [20]. Interactive maps are commonly used by professionals for environmental analysis and planning. However, user interfaces are rarely linguistically tailored to persons with limited English proficiency. One study has shown that limited English proficiency is associated with low health literacy across diverse racial groups. Among a population-based sample of limited English-proficiency adults, a higher percentage of Chinese respondents self-reported lower health literacy (68.3%) compared with their white counterparts (18.8%) [27]. Increasing accessibility of information is an important part of reducing the digital divide in health literacy [28].

We sought to develop the Air Pollution Map with several features not offered in existing mapping tools. Unlike other air pollution maps, which focus on other pollutants, our tool makes ultrafine particles visible to the user. Moreover, concentrations of ultrafine particles are displayed at the census block level, an

appropriate granularity of geographic unit for education about how concentrations of ultrafine particles vary across a neighborhood. We created the user interface for ease of use by individuals with limited English and technical skills. Specifically, the design approach emphasized accessibility by members of the Chinatown community, most of whom are immigrants. Map labels enable users of the tool who live and work in this neighborhood to recognize familiar streets or areas, including highways and local landmarks.

These features of the Air Pollution Map add new capabilities to an evolving technology of mapping air pollutants. Two other tools, Plume Air Report [29] and AirNow [30], display real-time information on pollution and provide recommendations for outdoor activities. The Community LINE Source Model (C-LINE) maps contributions of mobile sources to local air pollution under user-defined traffic emissions and meteorology conditions [31]. Each of these existing tools gives users access to important information about types of air pollution in their community, but not exposure to ultrafine particulate pollution.

In designing an educational intervention using the Air Pollution Map, it is important to consider not only the educational potential of the map but also its limitations. The present version requires users to enter values for temperature, wind severity and direction, and traffic volume by manipulating sliders. In future work, we will add a new capability for the user to automatically import the actual wind and temperature pattern for the current day and time. In addition, the user interface currently uses only English and Chinese. Translations to other languages such as Spanish would be helpful in order to widen accessibility to residents of other neighborhoods. Alterations of the user interface have not yet been implemented for persons who are visually impaired.

The next phase of work entails development and evaluation of an educational intervention using the Air Pollution Map. We have developed multimedia educational presentations, which have been shown to complement visual learning [32,33]. We also devised small-group learning activities and walking tours in the community to aid learning about sources of traffic-related pollution and features of the built environment. These multimodal learning activities are intended to take place in a supportive social environment, which helps motivate learning [34]. In particular, the educational approach emphasizes intergenerational colearning; specifically, teenagers who are familiar with computer technology are first trained to use the interactive map; in turn, the youth teach those who are less familiar with computer technology, such as elders.

Acknowledgments

This project is funded by a grant from the National Library of Medicine, R21LM011919-02. It is part of a study associated with the Community Assessment of Freeway Exposure and Health at Tufts University. The authors would like to thank the following individuals: Fengqing (Tina) Wang and Terry Yin from the Tufts Medical Center, who shared their community expertise and provided English-Chinese translations. Andrew Dufilie of University of Massachusetts Lowell contributed programming support.

Conflicts of Interest

None declared.

References

1. Finn S, O'Fallon L. The emergence of environmental health literacy-from its roots to its future potential. *Environ Health Perspect* 2015 Jun 30 [FREE Full text] [doi: [10.1289/ehp.1409337](https://doi.org/10.1289/ehp.1409337)] [Medline: [26126293](https://pubmed.ncbi.nlm.nih.gov/26126293/)]
2. World Bank. Thinking with mental models. In: *World Development Report 2015: Mind, Society, and Behavior*. Washington, DC: World Bank; 2013:62-75.
3. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012 Dec 15;380(9859):2224-2260 [FREE Full text] [doi: [10.1016/S0140-6736\(12\)61766-8](https://doi.org/10.1016/S0140-6736(12)61766-8)] [Medline: [23245609](https://pubmed.ncbi.nlm.nih.gov/23245609/)]
4. Fann N, Lamson AD, Anenberg SC, Wesson K, Risley D, Hubbell BJ. Estimating the national public health burden associated with exposure to ambient PM_{2.5} and ozone. *Risk Anal* 2012 Jan;32(1):81-95. [doi: [10.1111/j.1539-6924.2011.01630.x](https://doi.org/10.1111/j.1539-6924.2011.01630.x)] [Medline: [21627672](https://pubmed.ncbi.nlm.nih.gov/21627672/)]
5. Pope CA, Dockery DW. Health effects of fine particulate air pollution: lines that connect. *J Air Waste Manag Assoc* 2006 Jun;56(6):709-742. [Medline: [16805397](https://pubmed.ncbi.nlm.nih.gov/16805397/)]
6. Ghosh R, Lurmann F, Perez L, Penfold B, Brandt S, Wilson J, et al. Near-roadway air pollution and coronary heart disease: burden of disease and potential impact of a greenhouse gas reduction strategy in southern California. *Environ Health Perspect* 2016 Feb;124(2):193-200 [FREE Full text] [doi: [10.1289/ehp.1408865](https://doi.org/10.1289/ehp.1408865)] [Medline: [26149207](https://pubmed.ncbi.nlm.nih.gov/26149207/)]
7. HEI Review Panel on Ultrafine Particles. *HEI Perspectives 3: understanding the health effects of ambient ultrafine particles*. Boston, MA: Health Effects Institute; 2013 Jan. URL: <https://www.healtheffects.org/system/files/Perspectives3.pdf> [accessed 2017-03-06] [WebCite Cache ID 6oldkPNh]
8. Boston Public Health Commission. *Health of Boston 2012-2013: a neighborhood focus*. Boston, MA: Boston Public Health Commission Research Office; 2013. URL: http://www.bphc.org/healthdata/health-of-boston-report/Documents/HOB-2012-2013/HOB12-13_FullReport.pdf [accessed 2017-03-06] [WebCite Cache ID 6olmlwoHI]
9. Viswanath K. Public communications and its role in reducing and eliminating health disparities. In: Thomson GE, Mitchell F, Williams MB, editors. *Examining the Health Disparities Research Plan of the National Institutes of Health: Unfinished Business*. Washington, DC: National Academies Press; 2006:215-253.
10. Hornik R. *Public Health Communication: Evidence for Behavior Change*. Mahwah, NJ: L Erlbaum Associates; 2002.
11. Viswanath K, Finnegan JR. The knowledge gap hypothesis: twenty-five years later. *Ann Int Commun Assoc* 1996;19(1):187-228. [doi: [10.1080/23808985.1996.11678931](https://doi.org/10.1080/23808985.1996.11678931)]
12. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, Lopez A, et al. The literacy divide: health literacy and the use of an internet-based patient portal in an integrated health system-results from the diabetes study of northern California (DISTANCE). *J Health Commun* 2010;15 Suppl 2:183-196 [FREE Full text] [doi: [10.1080/10810730.2010.499988](https://doi.org/10.1080/10810730.2010.499988)] [Medline: [20845203](https://pubmed.ncbi.nlm.nih.gov/20845203/)]
13. Houts PS, Doak CC, Doak LG, Loscalzo MJ. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. *Patient Educ Couns* 2006 May;61(2):173-190. [doi: [10.1016/j.pec.2005.05.004](https://doi.org/10.1016/j.pec.2005.05.004)] [Medline: [16122896](https://pubmed.ncbi.nlm.nih.gov/16122896/)]
14. Standing L, Conezio J, Haber RN. Perception and memory for pictures: single-trial learning of 2500 visual stimuli. *Psychon Sci* 1970;19(2):73-74. [doi: [10.3758/bf03337426](https://doi.org/10.3758/bf03337426)]
15. Bower GH, Karlin MB, Dueck A. Comprehension and memory for pictures. *Mem Cognit* 1975 Mar;3(2):216-220. [doi: [10.3758/BF03212900](https://doi.org/10.3758/BF03212900)] [Medline: [21287062](https://pubmed.ncbi.nlm.nih.gov/21287062/)]
16. Haring MJ, Fry MA. Effect of pictures on children's comprehension of written text. *Educ Commun Technol J* 1979;27(3):185-190 [FREE Full text] [doi: [10.1007/BF02765450](https://doi.org/10.1007/BF02765450)]
17. Bartram DJ. Comprehending spatial information: the relative efficiency of different methods of presenting information about bus routes. *J Appl Psychol* 1980 Feb;65(1):103-110. [Medline: [7364703](https://pubmed.ncbi.nlm.nih.gov/7364703/)]
18. Vogel EK, Woodman GF, Luck SJ. Storage of features, conjunctions, and objects in visual working memory. *J Exp Psychol* 2001;27(1):92-114. [doi: [10.1037//0096-1523.27.1.92](https://doi.org/10.1037//0096-1523.27.1.92)]
19. Ware C. *Information Visualization: Perception for Design (Interactive Technologies)*. 3rd edition. Waltham, MA: Morgan Kaufmann; 2012.
20. Domagk S, Schwartz RN, Plass JL. Interactivity in multimedia learning: an integrated model. *Comput Hum Behav* 2010 Sep;26(5):1024-1033. [doi: [10.1016/j.chb.2010.03.003](https://doi.org/10.1016/j.chb.2010.03.003)]
21. Zarcadoolas C. The simplicity complex: exploring simplified health messages in a complex world. *Health Promot Int* 2011 Sep;26(3):338-350. [doi: [10.1093/heapro/daq075](https://doi.org/10.1093/heapro/daq075)] [Medline: [21149317](https://pubmed.ncbi.nlm.nih.gov/21149317/)]
22. Weave. *Visualize your data*. Lowell, MA: Weave Visual Analytics URL: <http://iweave.com> [accessed 2017-03-07] [WebCite Cache ID 6on2NC8uu]
23. Duffile A, Fallon J, Stickney P, Grinstein G. Weave: a web-based architecture supporting asynchronous and real-time collaboration. 2012 Presented at: *AVI 2012 Advanced Visual Interfaces International Working Conference*; May 22-26, 2012; Capri, Italy URL: <https://pdfs.semanticscholar.org/38c4/115e0d11c5530c6128cfa2d87ad7a98479f7.pdf>

24. Patton AP, Perkins J, Zamore W, Levy JI, Brugge D, Durant JL. Spatial and temporal differences in traffic-related air pollution in three urban neighborhoods near an interstate highway. *Atmos Environ* 2014 Dec 01;99:309-321 [FREE Full text] [doi: [10.1016/j.atmosenv.2014.09.072](https://doi.org/10.1016/j.atmosenv.2014.09.072)] [Medline: [25364295](https://pubmed.ncbi.nlm.nih.gov/25364295/)]
25. Patton AP, Zamore W, Naumova EN, Levy JI, Brugge D, Durant JL. Transferability and generalizability of regression models of ultrafine particles in urban neighborhoods in the Boston area. *Environ Sci Technol* 2015 May 19;49(10):6051-6060 [FREE Full text] [doi: [10.1021/es5061676](https://doi.org/10.1021/es5061676)] [Medline: [25867675](https://pubmed.ncbi.nlm.nih.gov/25867675/)]
26. An interactive map of traffic pollution for Boston Chinatown. URL: <http://chinatown.oicweave.com> [accessed 2017-08-14] [WebCite Cache ID 6siTaGLNk]
27. Sentell T, Braun KL. Low health literacy, limited English proficiency, and health status in Asians, Latinos, and other racial/ethnic groups in California. *J Health Commun* 2012;17(suppl 3):82-99 [FREE Full text] [doi: [10.1080/10810730.2012.712621](https://doi.org/10.1080/10810730.2012.712621)] [Medline: [23030563](https://pubmed.ncbi.nlm.nih.gov/23030563/)]
28. Mackert M, Mabry-Flynn A, Champlin S, Donovan EE, Pounders K. Health literacy and health information technology adoption: the potential for a new digital divide. *J Med Internet Res* 2016 Oct 04;18(10):e264 [FREE Full text] [doi: [10.2196/jmir.6349](https://doi.org/10.2196/jmir.6349)] [Medline: [27702738](https://pubmed.ncbi.nlm.nih.gov/27702738/)]
29. Plume Labs. Plume air report: the air pollution forecast app in your city. Paris, France: Plume Labs URL: <https://plumelabs.com/en/products/air-report> [accessed 2017-03-21] [WebCite Cache ID 6p8QT52I0]
30. U.S. Environmental Protection Agency, Office of Air Quality Planning Standards. AirNow. Research Triangle Park, NC: U.S. Environmental Protection Agency, National Service Center for Environmental Publications; 2017. URL: <https://www.airnow.gov/> [accessed 2017-03-20] [WebCite Cache ID 6p6y1plY4]
31. Barzyk TM, Isakov V, Arunachalam S, Venkatram A, Cook R, Naess B. A near-road modeling system for community-scale assessments of traffic-related air pollution in the United States. *Environ Modelling Software* 2015 Apr;66:46-56. [doi: [10.1016/j.envsoft.2014.12.004](https://doi.org/10.1016/j.envsoft.2014.12.004)]
32. Shiffer MJ. Spatial multimedia representations to support community participation. In: Craig WJ, Harris TM, Weiner D, editors. *Community Participation and Geographic Information Systems*. London, UK: Taylor and Francis; 2002:309-319.
33. Meppelink CS, van Weert JC, Haven CJ, Smit EG. The effectiveness of health animations in audiences with different health literacy levels: an experimental study. *J Med Internet Res* 2015;17(1):e11 [FREE Full text] [doi: [10.2196/jmir.3979](https://doi.org/10.2196/jmir.3979)] [Medline: [25586711](https://pubmed.ncbi.nlm.nih.gov/25586711/)]
34. Boratto L, Carta S, Fenu G, Manca M, Mulas F, Pilloni P. The role of social interaction on users motivation to exercise: a persuasive web framework to enhance the self-management of a healthy lifestyle. *Pervasive Mobile Comput* 2017;35:98-114. [doi: [10.1016/j.pmcj.2016.08.009](https://doi.org/10.1016/j.pmcj.2016.08.009)]

Abbreviations

CAFEH: Community Assessment of Freeway Exposure and Health Study

PM2.5: particulate matter $\leq 2.5 \mu\text{m}$ in diameter

PNC: particle number concentration

Edited by G Eysenbach; submitted 12.02.17; peer-reviewed by S Finn, FA Shafie; comments to author 28.02.17; revised version received 07.03.17; accepted 07.03.17; published 12.04.17.

Please cite as:

Cleary EG, Patton AP, Wu HC, Xie A, Stubblefield J, Mass W, Grinstein G, Koch-Weser S, Brugge D, Wong C

Making Air Pollution Visible: A Tool for Promoting Environmental Health Literacy

JMIR Public Health Surveill 2017;3(2):e16

URL: <http://publichealth.jmir.org/2017/2/e16/>

doi: [10.2196/publichealth.7492](https://doi.org/10.2196/publichealth.7492)

PMID: [28404541](https://pubmed.ncbi.nlm.nih.gov/28404541/)

©Ekaterina Galkina Cleary, Allison P Patton, Hsin-Ching Wu, Alan Xie, Joseph Stubblefield, William Mass, Georges Grinstein, Susan Koch-Weser, Doug Brugge, Carolyn Wong. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 12.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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

Determinants of Participants' Follow-Up and Characterization of Representativeness in Flu Near You, A Participatory Disease Surveillance System

Kristin Baltrusaitis^{1,2}, MA; Mauricio Santillana^{1,3,4}, MS, PhD; Adam W Crawley⁵, MPH; Rumi Chunara^{6,7}, MS, PhD; Mark Smolinski⁵, MPH, MD; John S Brownstein^{1,3}, MPH, PhD

¹Computational Health Informatics Program, Boston Children's Hospital, Boston, MA, United States

²Department of Biostatistics, Boston University School of Public Health, Boston, MA, United States

³Harvard Medical School, Boston, MA, United States

⁴Harvard School of Engineering and Applied Sciences, Cambridge, MA, United States

⁵Skoll Global Threats Fund, San Francisco, CA, United States

⁶The Global Institute of Public Health, New York University, New York, NY, United States

⁷Computer Science & Engineering, New York University, New York, NY, United States

Corresponding Author:

Kristin Baltrusaitis, MA

Department of Biostatistics

Boston University School of Public Health

Crosstown Building

801 Massachusetts Avenue 3rd Floor

Boston, MA, 02118

United States

Phone: 1 617 638 5172

Fax: 1 617 638 6484

Email: kbalt2@bu.edu

Abstract

Background: Flu Near You (FNY) is an Internet-based participatory surveillance system in the United States and Canada that allows volunteers to report influenza-like symptoms using a brief weekly symptom report.

Objective: Our objective was to evaluate the representativeness of the FNY population compared with the general population of the United States, explore the demographic and behavioral characteristics associated with FNY's high-participation users, and summarize results from a user survey of a cohort of FNY participants.

Methods: We compared (1) the representativeness of sex and age groups of FNY participants during the 2014-2015 flu season versus the general US population and (2) the distribution of Human Development Index (HDI) scores of FNY participants versus that of the general US population. We analyzed associations between demographic and behavioral factors and the level of participant follow-up (ie, high vs low). Finally, descriptive statistics of responses from FNY's 2015 and 2016 end-of-season user surveys were calculated.

Results: During the 2014-2015 influenza season, 47,234 unique participants had at least one FNY symptom report that was either self-reported (users) or submitted on their behalf (household members). The proportion of female FNY participants was significantly higher than that of the general US population (n=28,906, 61.2% vs 51.1%, $P<.001$). Although each age group was represented in the FNY population, the age distribution was significantly different from that of the US population ($P<.001$). Compared with the US population, FNY had a greater proportion of individuals with HDI >5.0, signaling that the FNY user distribution was more affluent and educated than the US population baseline. We found that high-participation use (ie, higher participation in follow-up symptom reports) was associated with sex (females were 25% less likely than men to be high-participation users), higher HDI, not reporting an influenza-like illness at the first symptom report, older age, and reporting for household members (all differences between high- and low-participation users $P<.001$). Approximately 10% of FNY users completed an additional survey at the end of the flu season that assessed detailed user characteristics (3217/33,324 in 2015; 4850/44,313 in 2016). Of these users, most identified as being either retired or employed in the health, education, and social services sectors and indicated that they achieved a bachelor's degree or higher.

Conclusions: The representativeness of the FNY population and characteristics of its high-participation users are consistent with what has been observed in other Internet-based influenza surveillance systems. With targeted recruitment of underrepresented populations, FNY may improve as a complementary system to timely tracking of flu activity, especially in populations that do not seek medical attention and in areas with poor official surveillance data.

(*JMIR Public Health Surveill* 2017;3(2):e18) doi:[10.2196/publichealth.7304](https://doi.org/10.2196/publichealth.7304)

KEYWORDS

public health surveillance; influenza, human; community-based participatory research; crowdsourcing; public health informatics; digital disease detection

Introduction

Influenza infections are associated with thousands of deaths each year in the United States [1]. Reliable real-time estimates of the temporal and geographical trends of the spread of influenza in the population are crucial to prepare for unusual influenza epidemics, for clinical resource allocation, and to assess vaccine effectiveness. In the United States, the Centers for Disease Control and Prevention (CDC) collects, compiles, and analyzes data from laboratories, outpatient health care offices, mortality surveillance systems, hospitals, and state health departments, and summarizes these datasets in weekly influenza surveillance reports. However, there is typically a 2-week lag between time of illnesses and publication of these reports. Furthermore, because these data sources report only on those individuals who seek medical care, an unknown proportion of individuals who do not visit health care facilities are unrepresented.

Over the past decade, Internet-based biosurveillance systems have developed as a way to provide an informal, complementary approach to traditional (syndromic) surveillance methods. These systems have the potential to reach a wider population and provide real-time access to users' symptom reports because they leverage alternative data sources, such as Google [2], Yahoo [3], and Baidu [4] Internet searches; Twitter posts [5-7]; Wikipedia article views [8,9]; and clinicians' database queries [10]. Internet search-based efforts in particular (ie, methods that use patterns of flu-related Internet searches to track flu) have led to very accurate Internet flu tracking systems [11-13]. Crowdsourced Internet-based participatory syndromic surveillance programs, such as Influenzanet in Europe, FluTracking in Australia, and Flu Near You (FNY) in the United States and Canada, have also been developed to track community influenza activity. These systems correlate well with traditional, clinical-based influenza-like illness (ILI) activity surveillance tools [14-17], and other platforms, such as GoViral, have validated the use of participatory information for disease surveillance by comparing volunteers' self-reported symptoms with specimens [18]. Although participatory surveillance systems track flu activity in a timely fashion, a large, diverse cohort of users who participate regularly and are representative of the population is essential for these systems to work effectively. Current surveillance systems have made efforts to quantify the biases between participant and general populations and investigate factors that influence follow-up participation [19-22].

We focused on FNY, which is administered by HealthMap of Boston Children's Hospital in partnership with the Skoll Global Threats Fund [23]. Specifically, we evaluated the representativeness of the FNY participant population compared with the general population of the United States, explored the demographic and behavioral characteristics that are associated with FNY's high-participation users, and summarized the results from a survey of a cohort of FNY users.

Methods

Representativeness of Participant Population

FNY Participants

Any resident of the United States or Canada can register as a user through the FNY website [24], mobile app, or Facebook. Upon registration, users provide information on their sex, date of birth, residential zip code, and email address. Although individuals must be at least 13 years of age to register, users can also add household members of any age and submit reports on their behalf. In exchange for participating in FNY, users can visualize local flu activity on maps, connect with local public health organizations, and find nearby locations offering flu vaccines. In this paper we define users as the population of individuals who registered with FNY and participants as the combined population of users and household members.

FNY Data Collection

Following registration, FNY users are asked to submit brief weekly reports where they can report any symptoms that they or any registered household members had during the previous week (Monday through Sunday). The symptoms in the report include fever, cough, sore throat, shortness of breath, chills or night sweats, fatigue, nausea or vomiting, diarrhea, headache, and body ache. If a user did not have any of these symptoms, he or she can also choose "I did not have any of the listed symptoms." However, if a user reports any of these symptoms, he or she is asked to provide the date of symptom onset. In addition, users are asked if they have received an influenza vaccination for the current flu season. Users are sent a reminder to complete the symptom report every Monday through either an email with a survey link or a push notification on their mobile phone. Although data are collected throughout the year, users have the option to suspend symptom reporting during the summer.

Census and Social Data Collection

We obtained national estimates of sex and age from the United States Census Bureau's 2014 annual estimates of resident

population [25]. Socioeconomic status was estimated at the county level using the Human Development Index (HDI) as a proxy [26]. The HDI is the county-level average of the educational index and income index and is measured on scale from 0-10, where 0 is the lowest HDI score and 10 is the highest. The educational index is a weighted average of the educational attainment index (ie, the measure of the overall level of educational attainment achieved by the adult population) and the enrollment index (ie, total number of students enrolled in school, of any age at any level, divided by the total school-aged population of 3- to 24-year-olds, inclusive). We calculated the income index from county-level median income [26]. The use of county-level HDI as a proxy for socioeconomic status was further assessed by estimating user-specific HDI from the cohort survey results (see below) and comparing these estimates with the corresponding county-level HDI estimate. Consistent with the method established by the Measure of America, the income index of the user-specific HDI was estimated by dividing the difference between the log of the zip code-level median income of the user and the log of the minimum US median income by the difference between the log of the maximum and the log of the minimum US median incomes, and then multiplying this ratio by 10 to scale the index between 0 and 10 [27]. We estimated the educational index from the level of education response of the user survey, where we assigned smaller values to lower educational attainment, such as “did not graduate high school,” than to higher educational attainment, such as a doctoral, law, or medical degree.

Comparison of FNY Participant Population With the US Population

Although FNY includes participants from the United States, Canada, and US territories, such as Puerto Rico, the majority of users reside within the 50 US states. As a result, the participant population used in this analysis included all registered users and household members residing within the 50 US states with complete sex, date of birth, and zip code information who submitted at least one symptom report during the 2014-2015 flu season, as defined by CDC weeks 40 (week ending October 4, 2014) through 20 (week ending May 23, 2015). We addressed the representativeness of sex (male and female) and age groups (<5, 5-14, 15-29, 30-39, 40-49, 50-59, 60-69, 70-79, ≥80 years) of FNY participants compared with the general US population using a 2-sided chi-square goodness-of-fit test. We compared the county-level distribution of HDI scores of FNY participants with that of the general US population using a 2-sample Kolmogorov-Smirnov test. For sensitivity purposes, we also performed these calculations using FNY data from the 2012-2013 and 2013-2014 flu seasons.

Characteristics of High-Participation Users

For this analysis, we considered only users who reported their own information, completed at least one symptom report during

the 2014-2015 flu season during or before CDC week 17, and provided sex information at registration. In addition, we chose only residents of the United States between ages 13 and 80 years at their registration date because users must be at least 13 years of age to register. A limit of 80 years of age was used to account for possible errors in date of birth input at user registration. Users who met these criteria were classified as either a high-participation user or a low-participation user based on the number of symptom reports they submitted during the 2014-2015 flu season. Users who completed more than 3 symptom reports during the 2014-2015 flu season were identified as high-participation users.

The demographic factors used in this analysis were sex (male or female), age group (13-29, 30-39, 40-49, 50-59, 60-69, and 70-79 years), and HDI as a continuous variable (see above for a description of methods used to calculate HDI). In addition, we included whether or not an ILI, as defined by the CDC, was reported at first entry. Although we did not examine information from individual household members in the analysis, we also included whether or not primary participants reported on behalf of other household members.

We analyzed associations between these demographic and behavioral factors and the level of participant follow-up using multivariable logistic regression. For odds ratio (OR) comparisons among age groups, we used 50-59 years as the reference group because it had the largest number of users. The demographic and behavioral factors were treated as independent variables, while level of follow-up was a dichotomous outcome (high-participation user vs low-participation user). We dichotomized the outcome because the distribution of number of reports was not normally distributed, and we determined the cutoff value of 3 empirically by assessing the histogram of number of reports. Sensitivity analyses confirmed the robustness of our findings. These additional analyses were conducted using more- and less-stringent definitions of high-participation users—specifically, more than 10 entries and more than 1 entry, respectively—for the 2012-2013, 2013-2014, and 2014-2015 flu seasons. Data were analyzed using R for Mac OS X version 3.1.1 (R Foundation).

Cohort Survey

To supplement data from user registrations, the FNY team conducted end-of-season user surveys in June 2015 and May 2016. The surveys were administered through SurveyGizmo online survey software (Widgix LLC), with survey invitations sent via email to all active FNY users. A completed survey entered the user into a raffle for incentives that included an iPad and US \$100 gift cards. Users were asked a variety of questions designed to better understand their interest in and motivations for reporting to FNY. Here we report on responses to a subset of questions (Textbox 1).

Textbox 1. Subset of survey questions administered to users of Flu Near You.

- How did you first hear about Flu Near You?
- Why did you sign up for Flu Near You?
- What is your primary motivation for continuing to use Flu Near You?
- Do you consider yourself more or less likely to report to Flu Near You when you have symptoms of illness to report?
- When you report to Flu Near You, do you report on any symptoms you've experienced during the week, or only symptoms that you think may be linked to the flu?
- What symptoms do you think are the primary symptoms of a flu infection?
- What is your industry/occupation?
- What is your level of education?

We exported survey responses to Excel 2016 (Microsoft Corporation) and Stata version 13 (StataCorp LP) for analysis and tabulation of descriptive statistics.

Results

Representativeness of Participant Population

Among states, California had the largest number of participants for the 2014-2015 flu season ($n=6595$), while Wyoming had the fewest ($n=89$) (see [Figure 1](#)). When we adjusted for state population size, Rhode Island had the greatest per capita representation (0.04%), and Mississippi had the lowest (0.008%) ([Figure 1](#)). The 2012-2013 and 2013-2014 flu seasons displayed a similar geographic distribution.

During the 2014-2015 influenza season, 47,234 unique participants had at least one symptom report that was either self-reported or submitted on their behalf. Of these participants,

28,906 (61.20%) were female and 18,328 (38.80%) were male. The proportion of female FNY participants was significantly overrepresented when compared with the general US population (51.1% female, $P<.001$) ([Figure 2 A](#)). Although each age group was represented in the FNY population, the distribution of age was significantly different from that of the US population ($P<.001$; data not shown). Overall, adult populations were overrepresented (ages 40-79 years), and both younger populations (ages <30 years) and older populations (ages ≥ 80 years) were underrepresented ([Figure 2 B](#)). The HDI range in the FNY population was 0-9.54 with a median of 5.03. As [Figure 2 C](#) and [Figure 2 D](#) show, the distribution of HDI scores was significantly different between the FNY population and the US population ($P<.001$). In general, FNY had a greater proportion of individuals with HDI scores greater than 5.0 ([Figure 2 E](#)). These descriptive statistics were similar across all 3 flu seasons (2012-2013, 2013-2014, and 2014-2015).

Figure 1. Distribution of participants for 2014- 2015 flu season by state (A) unadjusted distribution (B) adjusted distribution by state population.

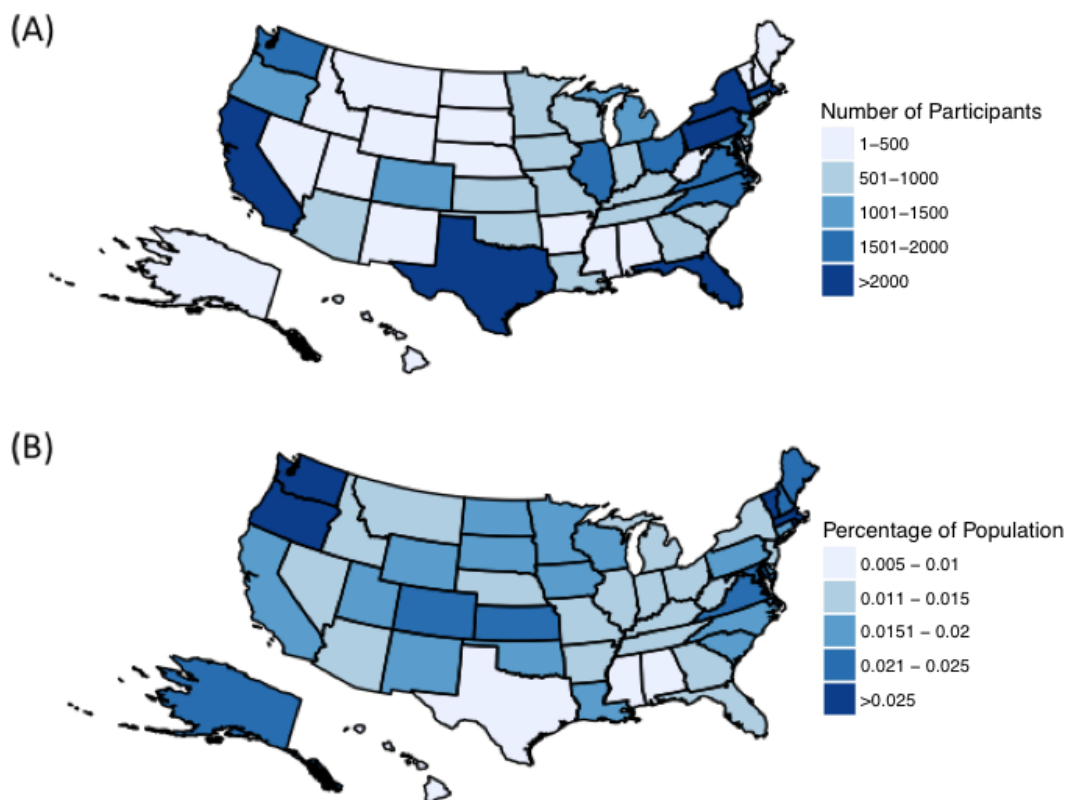
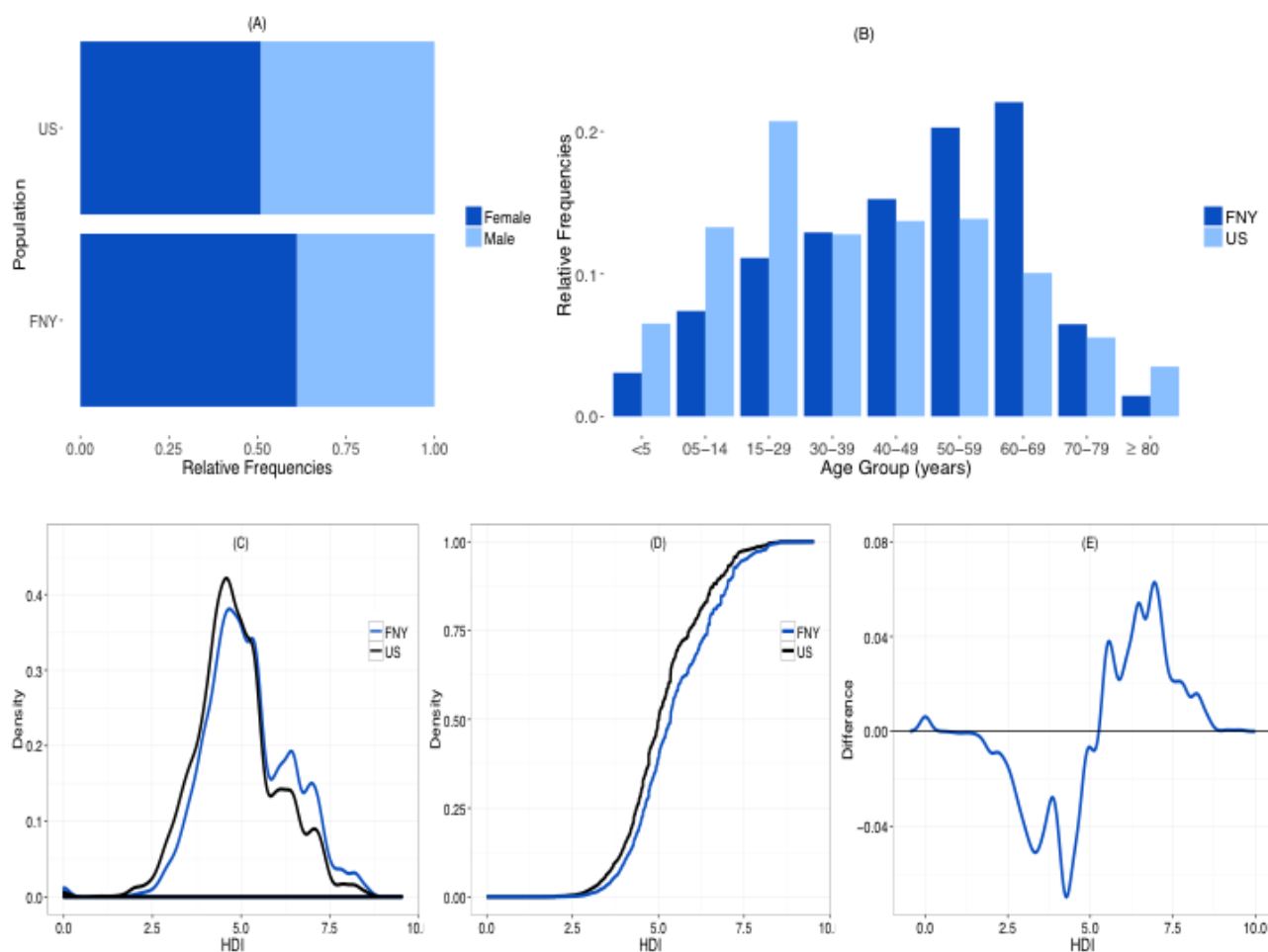


Figure 2. (A) Sex and (B) age profiles of Flu Near You (FNY) participants and comparison with the general population. (C) Distribution and (D) cumulative distribution of Human Development Index (HDI) scores by county for FNY and US populations and (E) difference between US and FNY distributions.



Characteristics of High-Participation Users

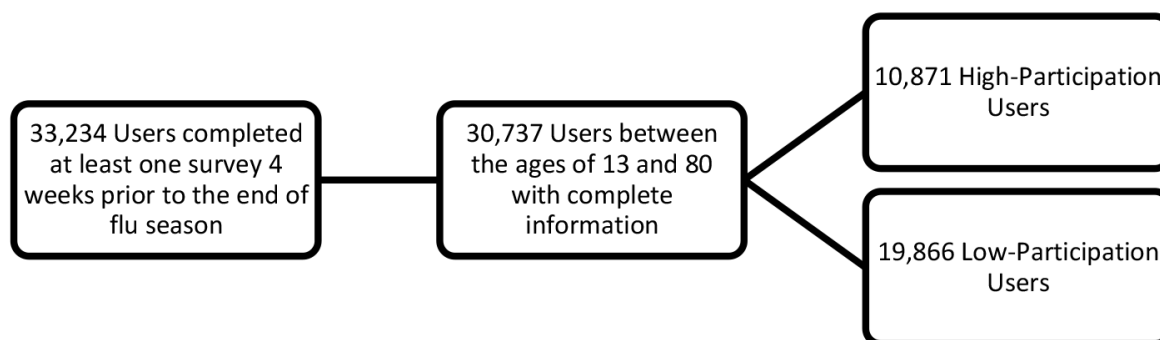
Figure 3 shows a flowchart of FNY user enrollment, including the number of users classified as either a high-participation user or a low-participation user. Table 1 summarizes results (adjusted ORs). Overall, females were 25% less likely than males to be high-participation users ($P < .001$). Users who reported for additional household members had 3.29 times the odds of being high-participation users compared with users who did not report

for additional household members ($P < .001$). Each unit increase in HDI was also associated with an increase in the odds of being a high-participation user (OR 1.12, $P < .001$). Users who reported symptoms meeting the definition of ILI at the first entry were 78% less likely than those who did not to be high-participation users ($P < .001$). There was a significant difference in participation among age groups ($P < .001$). In general, the odds of being a high-participation user increased with age.

Table 1. Summary of adjusted odds ratios (ORs) of being a high-participation user of Flu Near You.

Variable	Reference group	Adjusted OR	95% CI	P value
Sex: female	Male	0.75	0.71-0.79	<.001
Household members: yes	No	3.29	3.12-3.35	<.001
Human Development Index	N/A ^a	1.12	1.09-1.14	<.001
Influenza-like illness status at first survey: yes	No	0.22	0.19-0.25	<.001
Age group (years)				
13-29	50-59	0.67	0.61-0.74	<.001
30-39	50-59	0.54	0.49-0.58	<.001
40-49	50-59	0.70	0.64-0.75	<.001
60-69	50-59	1.14	1.07-1.23	<.001
70-79	50-59	1.23	1.11-1.36	<.001

^aN/A: not applicable.

Figure 3. Flowchart of user enrollment.

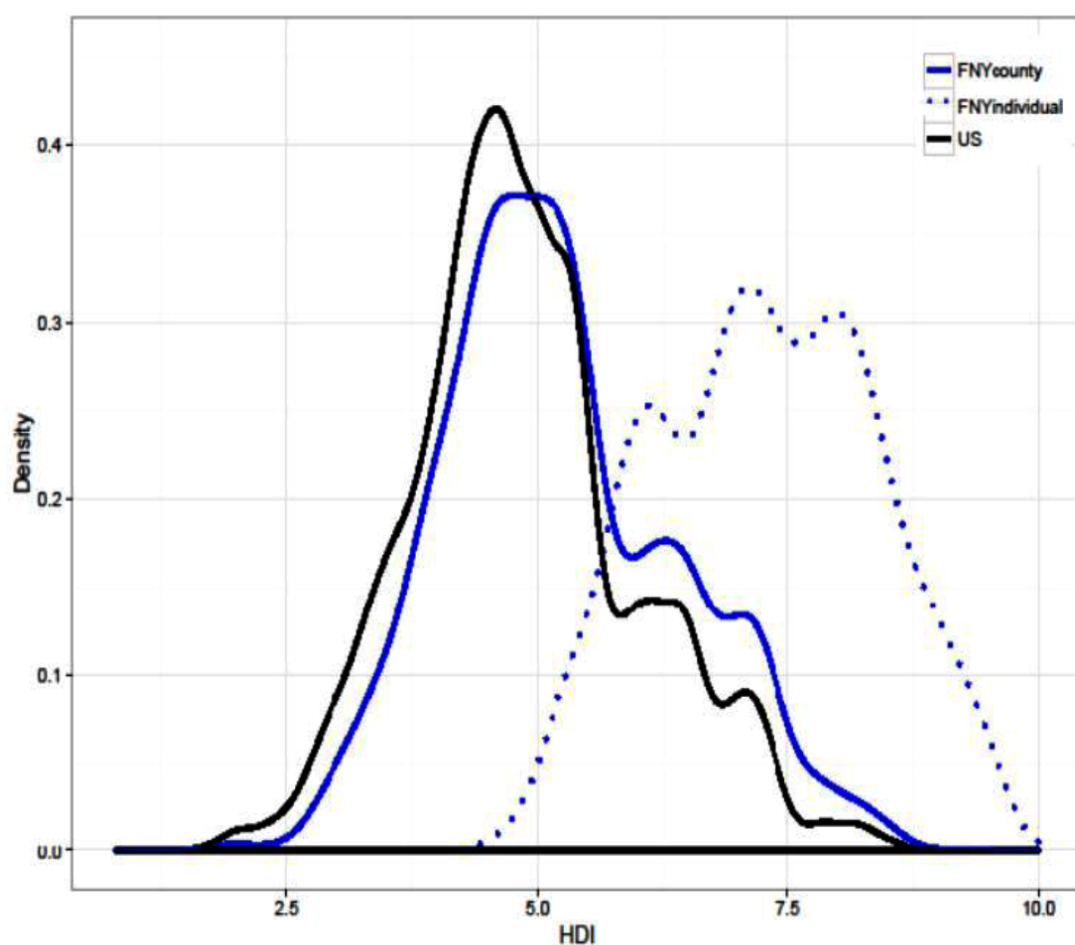
Cohort Survey

In 2015 the FNY user survey received 3217 responses from users, and the total numbers of users responding to the survey increased to 4850 in 2016. In both years the largest proportion of users identified as being retired (878/3217, 27.29%, in 2015; 1651/4850, 34.04%, in 2016), followed by users employed in the fields of health care and social assistance (620/3217, 19.27%, in 2015; 902/4850, 18.59%, in 2016), professional, scientific, and technical services (370/3217, 11.50%, in 2015; 453/4850, 9.34%, in 2016), and educational services (309/3217, 9.61%, in 2015; 409/4850, 8.43%, in 2016). These 4 categories accounted for 67.67% (2177/3217) of survey respondents in 2015 and 70.41% (3415/4850) in 2016. These results suggest

that FNY relies heavily on retirees and those employed in the health, education, and social services sectors.

In addition, our user surveys indicated that the majority of respondents had achieved a bachelor's degree or higher (2322/3217, 72.18%, in 2015; 3315/4850, 68.35%, in 2016), while less than 1% had not graduated high school. Approximately one-quarter of respondents had attained a master's degree in each survey year (845/3217, 26.27%, in 2015; 1181/4850, 24.35%, in 2016), while some respondents held doctoral or other advanced degrees (435/3217, 13.52%, in 2015; 602/4850, 12.41%, in 2016). These results support trends seen in participants' HDI scores, suggesting that the FNY participant population may have a higher educational attainment than the general US population (Figure 4).

Figure 4. Distributions of county-level Human Development Index (HDI) scores and individual user-specific HDI scores in Flu Near You (FNY) survey participants.



Discussion

Our results, in combination with previous work [15], show that FNY has the potential to achieve its goals, which include collecting and sharing participant-reported symptoms in order to increase awareness of ILI activity, generate early signals of ILI occurrence, and track ILI symptoms across the United States. Here, we show that participation varied across geographic location, sex, age, and HDI. Although all 50 states were represented during the 2014-2015 flu season, a few states had fewer than 500 participants, and the geographic distribution shows large gaps of information, especially in areas in the Midwest and South. FNY participants tended to cluster around large urban areas, with especially large user bases in the greater metropolitan areas surrounding Boston, New York City, and San Francisco.

Overall, females were overrepresented in our participant population (61.20%). This overrepresentation is consistent with findings from other participatory surveillance systems. During the 2011-2012 flu season, Influenzanet participants were more likely to be female than in the general population (56.8% vs 50.9%, $P < .001$), and among FluTracking participants who completed at least one survey each year, 66% in 2011 and 64% in 2012 were female [19,28]. This overrepresentation of female

participants is reflective of other studies showing that women are more likely than men to seek online health information [29,30].

FNY participants also had a higher HDI than that of the US population. This finding aligns with the results of our user survey conducted in 2015 and 2016. When comparing county-level HDI estimates with the user-specific HDI estimates within the population of FNY users who completed the 2016 survey, we found that for most survey participants, the county-level HDI underestimated the user-specific HDI, which further supports our initial findings that FNY participants have a higher HDI than the US population. These relatively high levels of HDI in the FNY population can be in part due to patterns in Internet penetration. Studies from the Pew Research Center have shown that Americans with high education levels and those in relatively affluent households have high Internet penetration [31].

Nor was the FNY population representative of the general US population in terms of age. Both younger populations (ages <30 years) and older populations (ages ≥ 80 years) were underrepresented, while the age groups between 40 and 80 years were overrepresented. As with sex, patterns of age representations were similar for both Influenzanet and FluTracking participants [19,28]. All of these differences in

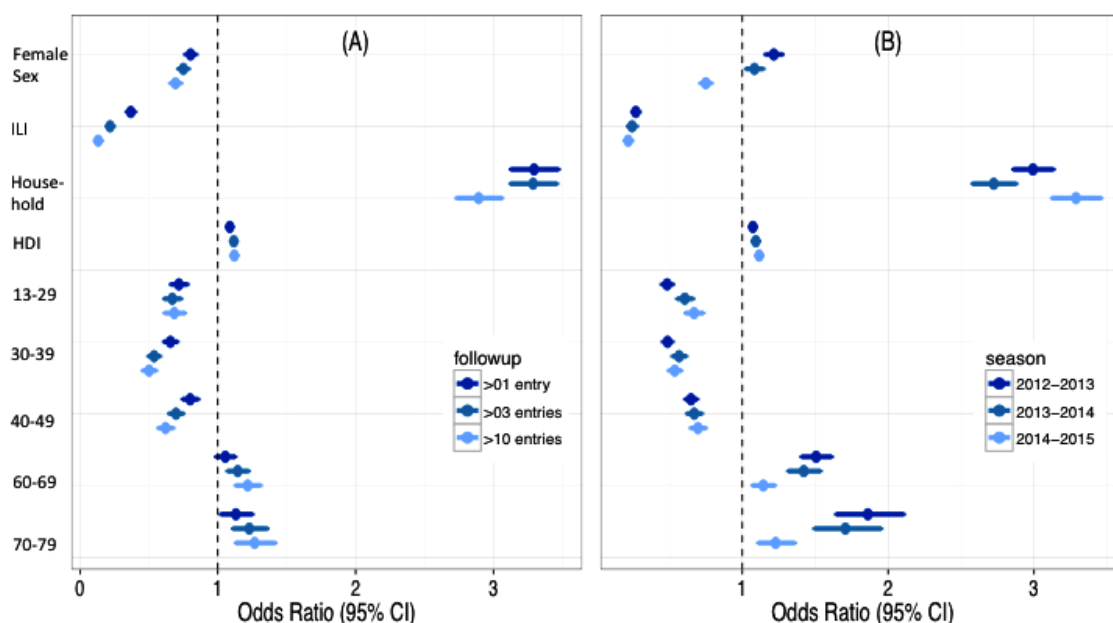
population characteristics (ie, sex, HDI, and age) were consistent across all 3 flu seasons we assessed.

We found that higher participation in follow-up symptom reports was associated with sex (male), reporting for household members, higher HDI score, not reporting an ILI at the first survey, and older age. These findings were consistent using both more-stringent (>10 entries) and less-stringent (>1 entry) definitions of good users (Figure 5 A). The results were consistent across all 3 seasons (Figure 5 B), except for sex. While females were less likely to be better users during the 2014-2015 season, this was not the case during the 2012-2013 and 2013-2014 seasons. Given the differences in reporting patterns by sex across years, an underlying factor, such as method of member recruitment, may be a confounder of this association. In addition, the ORs comparing participation habits between males and females were close to 1 (Figure 5 B), and a previously published study from Influenzanet found that there

were no significant differences between reporting for males and females [21].

The biases intrinsic within the FNY population are consistent with biases found in other Internet-based influenza surveillance systems. Despite these biases in the sociodemographic characteristics of the population, previous studies have shown that FNY and traditional disease surveillance systems capture similar trends of ILI rates at the national level [12,15]. The CDC has established a flu surveillance system that is robust and well accepted, by tracking individuals with ILI symptoms who seek medical care. Because only 35% of FNY users who report ILI symptoms seek medical attention, FNY captures flu activity in populations not accounted for by official surveillance data. As a result, with targeted recruitment, FNY may become a robust and complementary surveillance system that will benefit public health officials and the general population.

Figure 5. Odds ratios of high-participation users, by sex, influenza-like illness (ILI), Human Development Index (HDI), and age range, for (A) different levels of follow-up during the 2014-2015 flu season and (B) the 2012-2013, 2013-2014, and 2014-2015 flu seasons using more than 3 entries to define high-participation users.



Acknowledgments

KB, MSa, AWC, RC, MSm, and JSB conceived the research. KB and AWC conducted the statistical analysis. KB, MSa, AWC, and RC drafted the manuscript. All authors contributed to the final version of the paper. Research supported by the NIH/GMS: T32GM074905 (to KB).

Conflicts of Interest

None declared.

References

1. U.S. Centers for Disease Control and Prevention. Estimates of deaths associated with seasonal influenza --- United States, 1976-2007. *MMWR Morb Mortal Wkly Rep* 2010 Aug 27;59(33):1057-1062 [[FREE Full text](#)] [Medline: [20798667](#)]

2. 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/)]
3. Polgreen PM, Chen Y, Pennock DM, Nelson FD. Using internet searches for influenza surveillance. *Clin Infect Dis* 2008 Dec 01;47(11):1443-1448. [doi: [10.1086/593098](https://doi.org/10.1086/593098)] [Medline: [18954267](https://pubmed.ncbi.nlm.nih.gov/18954267/)]
4. Yuan Q, Nsoesie EO, Lv B, Peng G, Chunara R, Brownstein JS. Monitoring influenza epidemics in china with search query from baidu. *PLoS One* 2013;8(5):e64323 [FREE Full text] [doi: [10.1371/journal.pone.0064323](https://doi.org/10.1371/journal.pone.0064323)] [Medline: [23750192](https://pubmed.ncbi.nlm.nih.gov/23750192/)]
5. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS One* 2011 May 04;6(5):e19467 [FREE Full text] [doi: [10.1371/journal.pone.0019467](https://doi.org/10.1371/journal.pone.0019467)] [Medline: [21573238](https://pubmed.ncbi.nlm.nih.gov/21573238/)]
6. Paul MJ, Dredze M, Broniatowski D. Twitter improves influenza forecasting. *PLoS Curr* 2014 Oct 28;6 [FREE Full text] [doi: [10.1371/currents.outbreaks.90b9ed0f59bae4ccaa683a39865d9117](https://doi.org/10.1371/currents.outbreaks.90b9ed0f59bae4ccaa683a39865d9117)] [Medline: [25642377](https://pubmed.ncbi.nlm.nih.gov/25642377/)]
7. Chen L, Tozammel Hossain KSM, Butler P, Ramakrishnan N, Prakash BA. Syndromic surveillance of Flu on Twitter using weakly supervised temporal topic models. *Data Min Knowl Disc* 2015 Sep 8;30(3):681-710. [doi: [10.1007/s10618-015-0434-x](https://doi.org/10.1007/s10618-015-0434-x)]
8. McIver DJ, Brownstein JS. Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time. *PLoS Comput Biol* 2014 Apr;10(4):e1003581 [FREE Full text] [doi: [10.1371/journal.pcbi.1003581](https://doi.org/10.1371/journal.pcbi.1003581)] [Medline: [24743682](https://pubmed.ncbi.nlm.nih.gov/24743682/)]
9. 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/)]
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. 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/)]
12. 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/)]
13. 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/)]
14. van Noort SP, Codeço CT, Koppeschaar CE, van Ransst M, Paolotti D, Gomes MG. Ten-year performance of Influenzanet: ILI time series, risks, vaccine effects, and care-seeking behaviour. *Epidemics* 2015 Dec;13:28-36 [FREE Full text] [doi: [10.1016/j.epidem.2015.05.001](https://doi.org/10.1016/j.epidem.2015.05.001)] [Medline: [26616039](https://pubmed.ncbi.nlm.nih.gov/26616039/)]
15. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wójcik O, et al. Flu Near You: crowdsourced symptom reporting spanning 2 influenza seasons. *Am J Public Health* 2015 Oct;105(10):2124-2130. [doi: [10.2105/AJPH.2015.302696](https://doi.org/10.2105/AJPH.2015.302696)] [Medline: [26270299](https://pubmed.ncbi.nlm.nih.gov/26270299/)]
16. Dalton CB, Carlson SJ, Butler MT, Elvidge E, Durrheim DN. Building influenza surveillance pyramids in near real time, Australia. *Emerg Infect Dis* 2013 Nov;19(11):1863-1865 [FREE Full text] [doi: [10.3201/eid1911.121878](https://doi.org/10.3201/eid1911.121878)] [Medline: [24207165](https://pubmed.ncbi.nlm.nih.gov/24207165/)]
17. 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 02;5:9540 [FREE Full text] [doi: [10.1038/srep09540](https://doi.org/10.1038/srep09540)] [Medline: [25835538](https://pubmed.ncbi.nlm.nih.gov/25835538/)]
18. Goff J, Rowe A, Brownstein JS, Chunara R. Surveillance of acute respiratory infections using community-submitted symptoms and specimens for molecular diagnostic testing. *PLoS Curr* 2015 May 27;7 [FREE Full text] [doi: [10.1371/currents.outbreaks.0371243baa7f3810ba1279e30b96d3b6](https://doi.org/10.1371/currents.outbreaks.0371243baa7f3810ba1279e30b96d3b6)] [Medline: [26075141](https://pubmed.ncbi.nlm.nih.gov/26075141/)]
19. Cantarelli P, Debin M, Turbelin C, Poletto C, Blanchon T, Falchi A, et al. The representativeness of a European multi-center network for influenza-like-illness participatory surveillance. *BMC Public Health* 2014 Sep 20;14:984 [FREE Full text] [doi: [10.1186/1471-2458-14-984](https://doi.org/10.1186/1471-2458-14-984)] [Medline: [25240865](https://pubmed.ncbi.nlm.nih.gov/25240865/)]
20. Debin M, Turbelin C, Blanchon T, Bonmarin I, Falchi A, Hanslik T, et al. Evaluating the feasibility and participants' representativeness of an online nationwide surveillance system for influenza in France. *PLoS One* 2013;8(9):e73675 [FREE Full text] [doi: [10.1371/journal.pone.0073675](https://doi.org/10.1371/journal.pone.0073675)] [Medline: [24040020](https://pubmed.ncbi.nlm.nih.gov/24040020/)]
21. Bajardi P, Vespignani A, Funk S, Eames KT, Edmunds WJ, Turbelin C, et al. Determinants of follow-up participation in the Internet-based European influenza surveillance platform Influenzanet. *J Med Internet Res* 2014 Mar 10;16(3):e78 [FREE Full text] [doi: [10.2196/jmir.3010](https://doi.org/10.2196/jmir.3010)] [Medline: [24613818](https://pubmed.ncbi.nlm.nih.gov/24613818/)]
22. Bajardi P, Paolotti D, Vespignani A, Eames K, Funk S, Edmunds WJ, et al. Association between recruitment methods and attrition in Internet-based studies. *PLoS One* 2014;9(12):e114925 [FREE Full text] [doi: [10.1371/journal.pone.0114925](https://doi.org/10.1371/journal.pone.0114925)] [Medline: [25490045](https://pubmed.ncbi.nlm.nih.gov/25490045/)]
23. Chunara R, Aman S, Smolinski M, Brownstein J. Flu Near You: an online self-reported influenza surveillance system in the USA. *Online J Public Health Inform* 2013;5(1). [doi: [10.5210/ojphi.v5i1.4456](https://doi.org/10.5210/ojphi.v5i1.4456)]
24. HealthMap, Skoll Global Threats Fund, Boston Children's Hospital. Flu Near You. Help track the flu. Save lives. 2016. URL: <https://flunearyou.org/> [accessed 2017-03-23] [WebCite Cache ID 6pBRBW6fP]

25. United States Census Bureau, Population Division. Annual estimates of the resident population for selected age groups by sex for the United States, states, counties, and Puerto Rico Commonwealth and Municipios: April 1, 2010 to July 1, 2014. 2015 Jun. URL: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?_afpt=table [accessed 2017-04-04] [[WebCite Cache ID 6pTRtFulb](#)]
26. Lewis K, Burd-Sharps S. The Measure of America 2013-2014 URL: http://www.measureofamerica.org/Measure_of_America2013-2014MethodNote.pdf [accessed 2017-04-04] [[WebCite Cache ID 6pTPGkaxd](#)]
27. Michigan Population Studies Center, Institute for Social Research. 2016. Zip code characteristics: mean and median household income URL: <http://www.psc.isr.umich.edu/dis/census/Features/tract2zip/>; [accessed 2016-09-22] [[WebCite Cache ID 6kiQWUZud](#)]
28. Carlson SJ, Dalton CB, Butler MT, Fejsa J, Elvidge E, Durrheim DN. Flutracking weekly online community survey of influenza-like illness annual report 2011 and 2012. *Commun Dis Intell Q Rep* 2013 Dec 31;37(4):E398-E406 [[FREE Full text](#)] [Medline: [24882237](#)]
29. Nölke L, Mensing M, Krämer A, Hornberg C. Sociodemographic and health-(care-)related characteristics of online health information seekers: a cross-sectional German study. *BMC Public Health* 2015 Jan 29;15:31 [[FREE Full text](#)] [doi: [10.1186/s12889-015-1423-0](https://doi.org/10.1186/s12889-015-1423-0)] [Medline: [25631456](#)]
30. Fox S, Rainie L. The online health care revolution: how the Web helps Americans take better care of themselves. Washington, DC: Pew Internet & American Life Project; 2000 Nov 26. URL: http://www.pewinternet.org/files/old-media/Files/Reports/2000/PIP_Health_Report.pdf.pdf [accessed 2017-03-23] [[WebCite Cache ID 6pBHZLT6d](#)]
31. Perrin A, Duggan M. Americans' internet access:2000-2015. Washington, DC: Pew Research Center; 2015 Jun 26. URL: <http://www.pewinternet.org/2015/06/26/americans-internet-access-2000-2015/> [accessed 2017-03-23] [[WebCite Cache ID 6pBIA2Ky3](#)]

Abbreviations

CDC: Centers for Disease Control and Prevention
FNY: Flu Near You
HDI: Human Development Index
ILI: influenza-like illness
OR: odds ratio

Edited by T Sanchez; submitted 11.01.17; peer-reviewed by K Saha, C Dalton, P Bajardi; comments to author 16.02.17; revised version received 03.03.17; accepted 16.03.17; published 07.04.17.

Please cite as:

Baltrusaitis K, Santillana M, Crawley AW, Chunara R, Smolinski M, Brownstein JS
Determinants of Participants' Follow-Up and Characterization of Representativeness in Flu Near You, A Participatory Disease Surveillance System
JMIR Public Health Surveill 2017;3(2):e18
URL: <http://publichealth.jmir.org/2017/2/e18/>
doi: [10.2196/publichealth.7304](https://doi.org/10.2196/publichealth.7304)
PMID: [28389417](https://pubmed.ncbi.nlm.nih.gov/28389417/)

©Kristin Baltrusaitis, Mauricio Santillana, Adam W Crawley, Rumi Chunara, Mark Smolinski, John S Brownstein. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 07.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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

Saúde na Copa: The World's First Application of Participatory Surveillance for a Mass Gathering at FIFA World Cup 2014, Brazil

Onicio Leal Neto^{1,2,3}, MPH; George Santiago Dimech⁴, MPH; Marlo Libel⁵, MD; Wayner Vieira de Souza⁶, PhD; Eduarda Cesse⁶, PhD; Mark Smolinski⁵, MPH; Wanderson Oliveira⁷, MPH; Jones Albuquerque^{1,8}, PhD

¹Epitrack, Recife, Brazil

²SingularityU Recife Chapter, Recife, Brazil

³Aggeu Magalhães Research Center, Department of Health Collective, Recife, Brazil

⁴Pernambuco Health Department, Recife, Brazil

⁵Skoll Global Threats Fund, Pandemics Team, San Francisco, CA, United States

⁶Aggeu Magalhães Research Center, Department of Public Health, Recife, Brazil

⁷Brazil's Ministry of Health, General Coordination of Public Health Emergencies Response, Brasilia, Brazil

⁸Federal Rural University of Pernambuco, Informatics Department, Recife, Brazil

Corresponding Author:

Onicio Leal Neto, MPH

Epitrack

Republica do Libano, 251. 2508 A

Pina

Recife, 51110-160

Brazil

Phone: 55 81991849209

Fax: 55 8137218600

Email: onicio@epitrack.tech

Abstract

Background: The 2005 International Health Regulations (IHRs) established parameters for event assessments and notifications that may constitute public health emergencies of international concern. These requirements and parameters opened up space for the use of nonofficial mechanisms (such as websites, blogs, and social networks) and technological improvements of communication that can streamline the detection, monitoring, and response to health problems, and thus reduce damage caused by these problems. Specifically, the revised IHR created space for participatory surveillance to function, in addition to the traditional surveillance mechanisms of detection, monitoring, and response. Participatory surveillance is based on crowdsourcing methods that collect information from society and then return the collective knowledge gained from that information back to society. The spread of digital social networks and wiki-style knowledge platforms has created a very favorable environment for this model of production and social control of information.

Objective: The aim of this study was to describe the use of a participatory surveillance app, Healthy Cup, for the early detection of acute disease outbreaks during the Fédération Internationale de Football Association (FIFA) World Cup 2014. Our focus was on three specific syndromes (respiratory, diarrheal, and rash) related to six diseases that were considered important in a mass gathering context (influenza, measles, rubella, cholera, acute diarrhea, and dengue fever).

Methods: From May 12 to July 13, 2014, users from anywhere in the world were able to download the Healthy Cup app and record their health condition, reporting whether they were good, very good, ill, or very ill. For users that reported being ill or very ill, a screen with a list of 10 symptoms was displayed. Participatory surveillance allows for the real-time identification of aggregates of symptoms that indicate possible cases of infectious diseases.

Results: From May 12 through July 13, 2014, there were 9434 downloads of the Healthy Cup app and 7155 (75.84%) registered users. Among the registered users, 4706 (4706/7155, 65.77%) were active users who posted a total of 47,879 times during the study period. The maximum number of users that signed up in one day occurred on May 30, 2014, the day that the app was officially launched by the Minister of Health during a press conference. During this event, the Minister of Health announced the special government program Health in the World Cup on national television media. On that date, 3633 logins were recorded, which accounted for more than half of all sign-ups across the entire duration of the study (50.78%, 3633/7155).

Conclusions: Participatory surveillance through community engagement is an innovative way to conduct epidemiological surveillance. Compared to traditional epidemiological surveillance, advantages include lower costs of data acquisition, timeliness of information collected and shared, platform scalability, and capacity for integration between the population being served and public health services.

(*JMIR Public Health Surveill* 2017;3(2):e26) doi:[10.2196/publichealth.7313](https://doi.org/10.2196/publichealth.7313)

KEYWORDS

mass gatherings; participatory surveillance; public health; epidemiology

Introduction

Mass gatherings (MGs) are situations that involve large populations that come together for specific causes related to leisure (eg, sports events, carnivals, and concerts), religion (eg, Hajj, World Youth Day), politics (eg, marches, protests, presidential inaugurations), or similar purposes [1-4]. By changing business, media, and public health environments, MGs create both new opportunities and novel risks [3-5].

During any MG, two changes in particular have the potential to significantly increase pressure on the local health care system. The first is the increase in simple contact between travelers (ie, tourists, athletes, workers, volunteers, press staff, authorities) and individuals in the local population [1-3]. This contact can promote both the introduction of diseases into the local population by visitors and the transmission of diseases to visitors through either visitor-to-visitor or native-to-visitor contact. The potential for transmission among visitors and between natives and visitors (both native-to-visitor and visitor-to-native) is further influenced by the increased mobility of people and goods made possible by the forces of globalization [3]. This mobility increases contact among individuals and goods from different epidemiological settings, promotes the exchange of potential threats to public health, and generates new risks [3,4,6,7]. Recent experiences with the persistence of dengue fever and increased circulation of Zika virus [8-10] and Chikungunya fever have made the threat of infectious disease transmission a tangible reality that requires rapid detection and response preparation (ie, including case findings, vaccine availability, diagnostic procedures, medical services, epidemiological research, medicines, and guidance for the public and professionals) [3,4,7]. A second MG-associated change that has the potential to significantly increase pressure on the local health care system is the size of the MG itself, which can impact routine demand not only for health, but for all public services, especially security and transportation [3,4,6].

The revised International Health Regulations (IHRs) 2005 required states to develop, strengthen, and maintain the capabilities to detect, assess, notify, and report risk events to international public health authorities, including situations resulting from MGs [11]. Additionally, the 2005 IHRs established parameters for event assessments and notifications that may constitute public health emergencies of international concern [12-14]. These requirements and parameters opened up space for the use of nonofficial mechanisms (eg, websites, blogs, and social networks) and technological improvements of communication that can streamline the detection, monitoring,

and response to health problems, and thus reduce damage caused by these problems [15-19]. Specifically, the revised IHRs created space for participatory surveillance to function, in addition to the traditional surveillance mechanisms of detection, monitoring, and response [20-27]. Participatory surveillance is based on crowdsourcing methods that collect information from society and then return the collective knowledge gained from that information back to society. The spread of digital social networks and wiki-style knowledge platforms has created a very favorable environment for this model of production and social control of information [18,19,28-31].

The aim of this study was to describe the use of a participatory surveillance app, Healthy Cup, for the early detection of acute disease outbreaks during the Fédération Internationale de Football Association (FIFA) World Cup 2014. The main health outcomes (ie, outcomes related to infectious diseases) that have been associated with MG sporting events like this are respiratory, cardiovascular, and gastrointestinal (ie, diarrhea) [32-36]. Respiratory outcomes are mainly associated with viruses that can spread easily between individuals, cardiovascular outcomes are usually associated with the emotional stress that fans experience during sports events, and gastrointestinal outcomes are often associated with the expansion of the informal food trade that typically occurs during MG events (eg, foods being sold on the street) and the desire of tourists to sample local cuisine, which may not be well tolerated by travelers' bodies. Our focus was on these three main infectious disease outcomes, specifically respiratory, diarrheal, and rash syndromes related to diseases that are considered important in an MG context (ie, influenza, measles, rubella, cholera, acute diarrhea, and dengue fever).

Methods

From May 12 to July 13, 2014, users from anywhere in the world were able to download the Healthy Cup app and record their health condition, reporting whether they were *good*, *very good*, *ill*, or *very ill*. For users that reported being *ill* or *very ill*, a screen with a list of 10 symptoms was displayed. Participatory surveillance allows for the real-time identification of aggregates of symptoms that indicate possible cases of infectious diseases. Table 1 lists the symptoms, and syndromes and diseases associated with these symptoms, that were included in the Healthy Cup app. In addition to these 10 symptoms, the app also had 2 additional queries regarding contact chain (eg, "I got in touch or know someone with some of these symptoms in the last 7 days") and severity of the symptoms ("I looked for a health care service").

Table 1. Symptoms, syndromes, and diseases searched using Healthy Cup app.

Symptoms	Syndromes			Diseases					
	Respiratory	Diarrheal	Rash	Influenza	Measles	Rubella	Cholera	Acute diarrhoea	Dengue
Fever	X	X	X	X	X	X	-	X	X
Cough	X	-	X	X	X	X	-	-	-
Sore throat	X	-	-	X	-	-	-	-	-
Nausea	-	X	-	-	-	-	-	X	-
Joint pain	-	-	X	-	-	-	-	-	X
Headache	-	-	X	-	-	-	-	-	X
Diarrhea	-	X	-	-	-	-	X	X	-
Rash	-	-	X	-	X	X	-	-	X
Bleeding	-	-	X	-	-	-	-	-	X
Shortness of breath	X	-	-	X	-	-	-	-	-

Healthy Cup was developed by a partnership between the Secretariat of Health Surveillance (Brazil's Ministry of Health), Skoll Global Threats Fund, and EpiTrack eHealth. The app was designed on an open source platform for use on mobile devices. The platform was developed as a hybrid app for both iOS and Android operating systems that could be accessed by anyone using an iOS or Android smartphone, or as a Web app in any Internet browser. The iOS native and Web apps were developed using PhoneGap (built with JavaScript, HTML5 and CSS) [37]; the Android app was developed in native language. The mobile and Web apps use external interface capabilities to support Application Program Interface (API) Google Places (for location of nearby hospitals and pharmacies), Google Maps API (for user navigation to points of interest and viewpoints on the dashboard), and Twitter API (for streaming of the Ministry of Health profile) [38,39]. We used a MySQL-type server developed in PHP language, and we managed the database using phpMyAdmin.

The Healthy Cup project was hosted throughout the study period by Dreamhost [40], and its code and versioning used GitHub. To ensure safety of the platform, we also set up a virtual private

server with 60 gigabytes of storage, 2 gigabytes of random-access memory, and unlimited bandwidth. Both reports and registers had geolocation features, in which the system captured the coordinates related to each use. This function was implemented following these standards to acquire geolocation data: (1) asking the permission of the users; (2) starting up a cron job to get the coordinates; and (3) inserting this coordinate for each respective event, whether it was a report or register.

Results

From May 12 through July 13, 2014, there were 9434 downloads of the Healthy Cup app and 7155 (75.84%) registered users. Among the registered users, 4706 (4706/7155, 65.77%) were active users who posted a total of 47,879 times during the study period. Of these posts, 89.43% (42,818/47,879) reported no symptoms. One symptom was reported in 3173 posts (3173/47,879, 6.63%); one or more symptoms were reported in 5329 posts (5329/47,879, 11.13%; with an average of 1.8 per post); five or more symptoms were reported in 220 posts (220/47,879, 0.46%); and all 10 symptoms were reported in 99 posts (99/47,879, 0.21%; see Table 1 and Table 2).

Table 2. Distribution of the posts, symptoms profile, and syndromes by host city. Saúde na Copa 2014 percentages for posts are in relation to total posts among all host cities; percentages for symptoms are in relation to host city posts.

Host city	Posts, n (%)	With symptoms, n (%)	Diarrhea syndrome, n (%)	Respiratory syndrome, n (%)	Rash Syndrome, n (%)
Belo Horizonte	1133 (4.32)	128 (11.30)	7 (5.5)	12 (9.4)	5 (3.9)
Brasilia	7951 (30.33)	573 (7.20)	14 (2.4)	37 (6.5)	7 (1.2)
Cuiabá	1109 (4.23)	173 (15.60)	1 (0.6)	15 (8.7)	0 (0.0)
Curitiba	824 (3.14)	92 (11.2)	6 (6.5)	9 (9.8)	2 (2.2)
Fortaleza	1519 (5.80)	174 (11.45)	6 (3.4)	9 (5.2)	3 (1.7)
Manaus	985 (3.76)	136 (13.8)	4 (2.9)	12 (8.8)	2 (1.5)
Natal	938 (3.58)	77 (8.2)	2 (2.6)	4 (5.2)	0 (0.0)
Porto Alegre	925 (3.53)	113 (12.2)	6 (5.3)	7 (6.2)	3 (2.7)
Recife	4316 (16.47)	282 (6.53)	5 (1.8)	9 (3.2)	1 (0.4)
Rio de Janeiro	3069 (11.70)	348 (11.34)	15 (4.3)	28 (8.0)	13 (3.7)
Salvador	1324 (5.04)	230 (17.38)	5 (2.2)	19 (8.3)	2 (0.9)
São Paulo	2125 (8.10)	358 (16.84)	16 (4.5)	37 (10.3)	5 (1.4)
Total	26,218 (100.00)	2684 (10.24)	87 (3.2)	198 (7.4)	43 (1.6)

The maximum number of users that signed up in one day occurred on May 30, 2014, the day that the app was officially launched by the Minister of Health during a press conference. During this event, the Minister of Health announced the special government program *Health in the World Cup* on national television media (Figure 1). On that date, 3633 logins were recorded, which accounted for more than half of all sign-ups across the entire duration of the study (50.78%, 3633/7155). Based on cost reduction efforts by the Brazilian government, this was the only advertising action that was undertaken for the app.

Most of the active users (3526/4706, 74.95%) installed the app on Android mobile devices, 1167 (1167/4706, 24.80%) were on Apple iOS mobile devices, and 13 (13/4706, 0.28%) were

on desktop computers. Slightly more than half of all users were male (2478/4706, 52.66%). Users ranged in age from 13 to 77 years with a median of 32 years (only individuals aged 13 years and older were allowed to use the app; see Figure 2). Due to privacy rules, we were not allowed to collect nationality data. A total of 4661 users (4661/4706, 99.04%) preferred the app in Portuguese, 34 (34/4706, 0.72%) in Spanish, and 12 (12/4706, 0.25%) in English.

The intensity of participation in relation to the match dates of the Brazilian team and the World Cup calendar are illustrated in Figure 3 and Table 3, respectively. It should be emphasized that on the days the Brazilian team played (Figure 1) and in the second phase of the World Cup matches (Table 3), push notifications were sent to users through the app.

Table 3. Distribution of reports according to World Cup Saúde na Copa, 2014.

World Cup timeframe	All posts (reports)	% (of total posts)	Posts with symptoms	% (of all posts during timeframe)
1. Pre-World Cup (May 12 - June 11)	19,737	41.22	3490	17.68
2. Group phase (June 12 - 27)	16,868	35.23	1241	7.36
3. Eighth finals (June 28 - July 3)	3762	7.86	217	5.77
4. Fourth finals (Jul 4 - 7)	2438	5.09	103	4.22
5. Semifinals (Jul 9 - 11)	1910	3.99	128	6.70
6. Finals (Jul 12 and 13)	1533	3.20	98	6.39
7. Post-World Cup (Jul 14 - 23)	1631	3.41	82	5.03
Total	47,879	100.00	5359	11.19

Upon user authorization and if the device had an active global positioning system, the app automatically recorded the location of each post. Only 6.00% (2,824/47,879) of posts provided no location information. More than half of all posts were recorded in World Cup host cities (26,218/47,879, 54.76%; Table 2), 37.90% (18,146/47,879) were elsewhere in Brazil, and 1.40%

(670/47,879) were in other countries. Table 2 also displays the locations with the highest number of known posts: Brasília (30.33%, 7951/26,218), Recife (16.46%, 4316/26,218), and Rio de Janeiro (11.71%, 3069/26,218), which together accounted for more than half (54.76%, 26,218/47,879) of all posts located in World Cup host cities. Among the 26,218 posts in World

Cup host cities, the greatest number of symptoms were recorded in Brasilia (573 posts with symptoms), São Paulo (358 posts with symptoms), and Rio de Janeiro (348 posts with symptoms). Also among the 26,218 host city posts, the greatest frequency of symptoms proportionally were recorded in Salvador (17.38%, 230/1324), São Paulo (16.84%, 358/2125), and Cuiabá (15.60%, 173/1109).

Of the three syndromes detected by Healthy Cup (ie, respiratory, diarrheal, and rash), respiratory syndrome occurred with the greatest frequency based on reported symptoms. The greatest number of posts of respiratory syndromes were reported in São

Paulo (n=37), Brasilia (n=37), and Rio de Janeiro (n=28). The highest frequencies of rash syndrome (ie, percentage of posts signaling rashes) were reported in Belo Horizonte (3.9%, 5/128), Rio de Janeiro (3.7%, 13/348), and Porto Alegre (2.7%, 3/113). The greatest number of diarrheal syndromes were reported in São Paulo (n=16), Rio de Janeiro (n=15), and Brasilia (n=14). The highest frequencies of diarrheal syndrome were reported in Curitiba (6.5%, 6/92), Belo Horizonte (5.5%, 7/128), Minas Gerais (5.5%, 7/128), and Porto Alegre (5.3%, 6/113). Finally, the greatest number of rash syndromes were reported in Rio de Janeiro (n=13), Brasilia (n=7), São Paulo (n=5), and Belo Horizonte (n=5).

Figure 1. Temporal distribution of posts by dates, Sau´de na Copa, 2014.

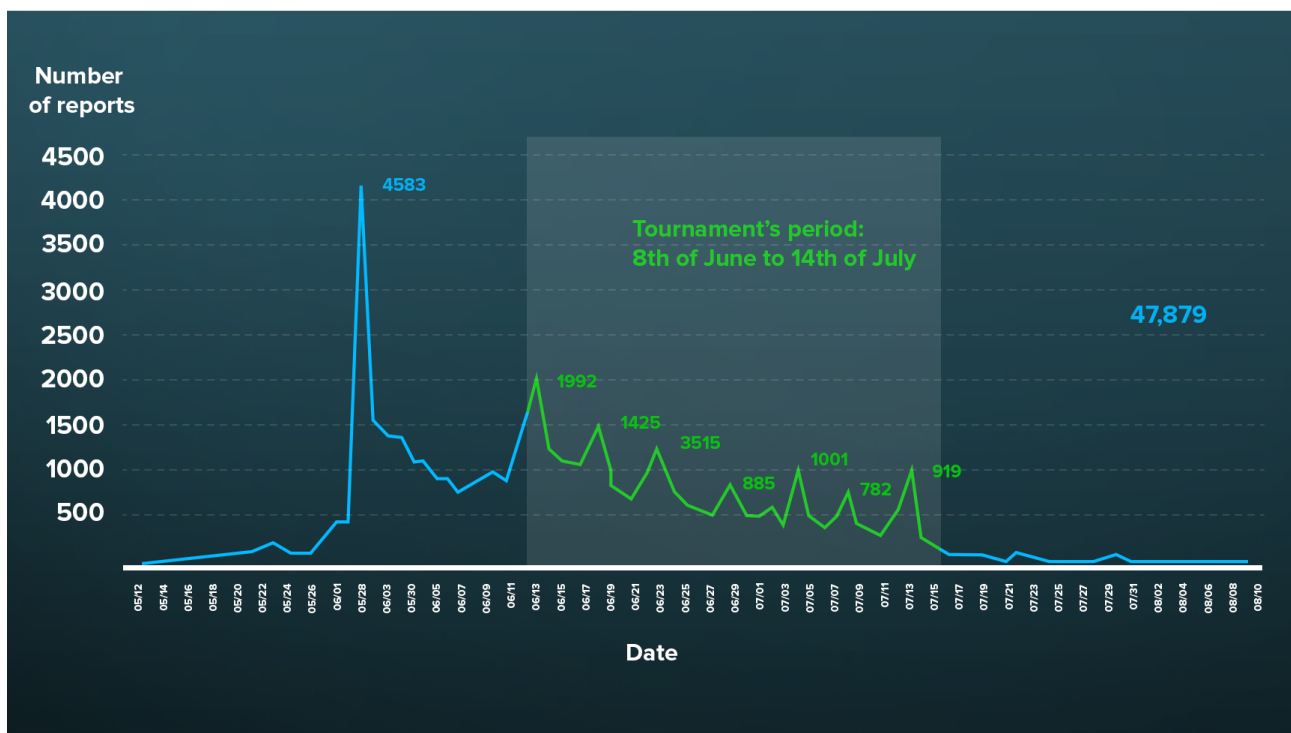


Figure 2. Distribution of users according to sex and age, Saú de na Copa, 2014.

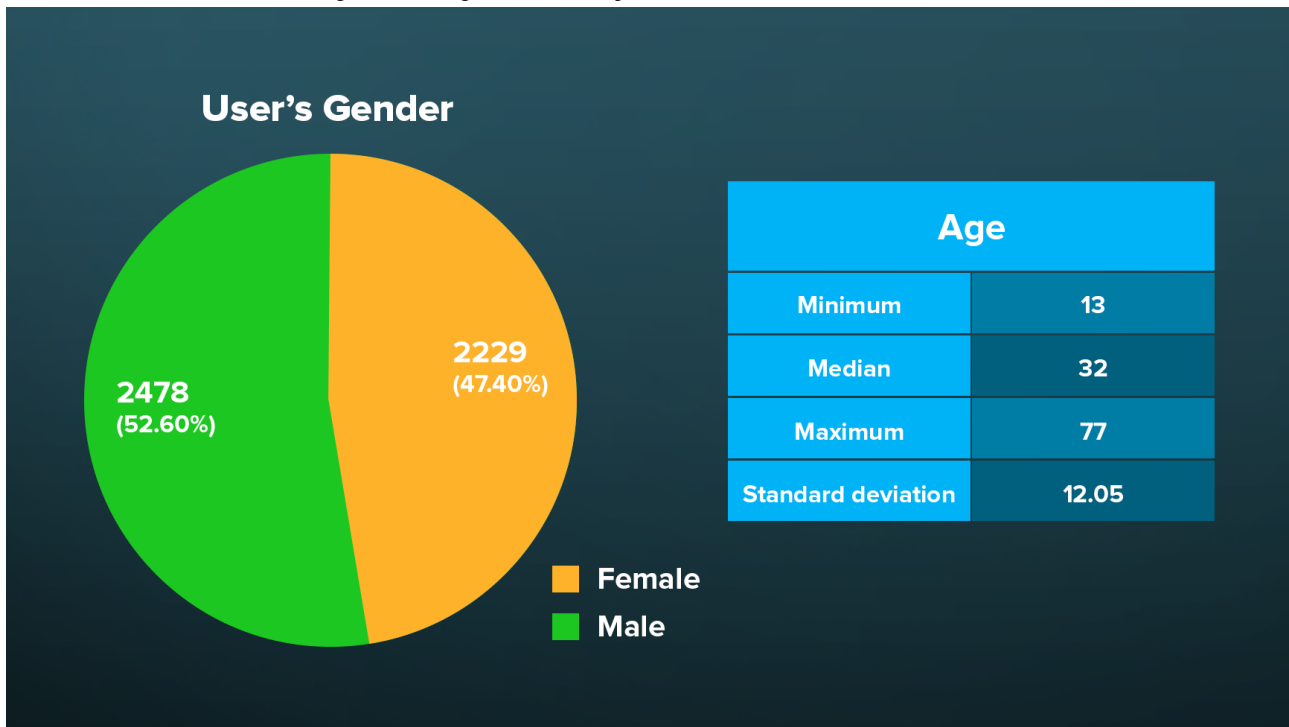
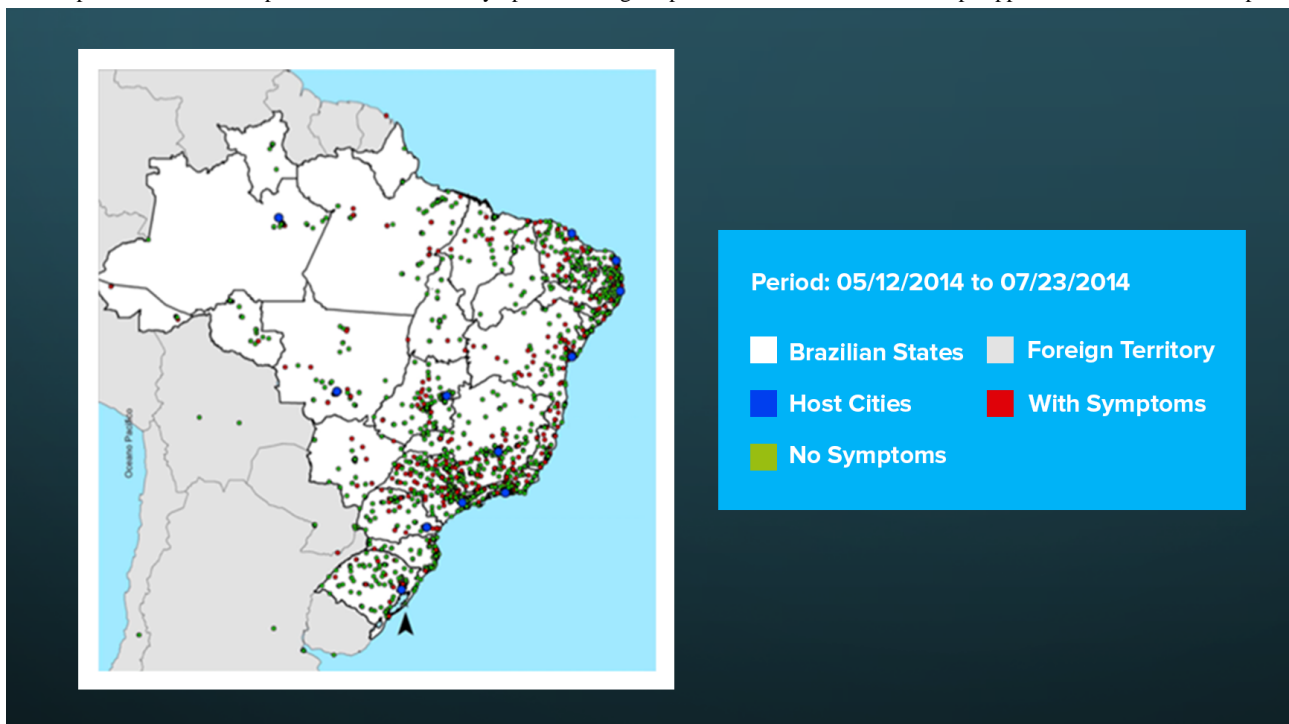


Figure 3. Spatial distribution of posts with and without symptoms during the period of launch of Saú de na Copa app until the end of the championship.



Discussion

Although the initial impact of the media was remarkable, with the greatest number of users signing up for Healthy Cup on the day the app was officially launched and first advertised on national television (Figure 1), this impact was not sustained even after 7 days of intense media (eg, television and websites) news about the app. The fact that the median age among those who signed up was 30 years suggests that this age group was more exposed to Healthy Cup promotion through television and

website news [21,24,41,42]. The limited participation of foreigners (less than 1%) may reflect the limited language options of the app (English and Spanish) [24].

The increased number of posts during the Brazilian matches may be related to the sending of push notifications before those matches, which reminded participants to use the app [24,29]. The host city with the highest concentration of posts was Brasilia (30.33%, 7951/26,218), which may be related to the fact that

Brasilia had the highest number of local news reports about the app.

More than half of all posts were recorded in World Cup host cities (54.76%, 26,218/47,879), suggesting that the app can be used as a tool to identify potential alerts for outbreaks associated with MGs [31-36]. The remaining 45.24% of posts (21,661/47,879) were recorded in cities elsewhere in Brazil (ie, other than the host cities of the World Cup), illustrating rapid penetration of the app across the country, even in states not hosting the FIFA World Cup 2014 [24].

Posts with six or more reported symptoms that were incompatible with the sought-after syndromes were considered *spam reports*. The higher prevalence of posts reported during the first two weeks (Figure 1) may be related to curiosity about this new type of tool being used in public health, with users wanting to record their health situation even if they were not showing symptoms [24].

Respiratory syndromes were reported more often than any other (Table 2), suggesting that the tool may have the capacity for early detection of epidemiological changes associated with influenza [41]. It is noteworthy that reports of bleeding and rash showed a high frequency of demand for health services [43,44].

The criteria used to classify syndromes (ie, based on parameters used by the Ministry of Health's Secretariat of Health Surveillance) may have underestimated the number of users with any of the three syndromes (ie, respiratory, diarrheal, or rash) [45]. However, during this same time period, the official Health Surveillance System (at the Ministry of Health) routinely used by the Integrated Health Center of Joint Operations did not identify the occurrence of any public health emergency events (ie, syndromic clusters) warranting intervention. Thus, data and findings from the Healthy Cup platform were validated by traditional sources.

Conclusions

Participatory surveillance through community engagement is an innovative way to conduct epidemiological surveillance. Compared to traditional epidemiological surveillance, the advantages of participatory surveillance include its lower cost for data acquisition, timeliness of information collected and

shared, platform scalability, and capacity for integration between the population being served and public health services.

The pilot of the Healthy Cup app during the FIFA World Cup 2014 allowed us to evaluate the potential for participatory surveillance in Brazil. Based on our results, participatory surveillance appears to have the potential to become a routine component of national health surveillance and to help improve the early detection of outbreaks and epidemics, timely intervention, and risk minimization. The Healthy Cup platform in particular appears to be sensitive to multiple symptoms and syndromes associated with a range of potential threats.

We also learned several lessons from our piloting experience, including the idea that investment in communication, marketing, and advertising is necessary to penetrate multiple social strata (eg, different age groups) and to reach as many users as possible. Relying on spontaneous media (ie, news and nonpaid ads) and press conferences alone limited use of the app to consumer groups that seek out this specific type of information. Investment in digital media could create a great opportunity to not only boost the number of users, but also enhance the engagement of users.

Another lesson learned is the need for reciprocity. Citizens may feel more motivated to participate if they receive something in return, such as information about diseases that are being reported in their area. This information could be sent back to users on maps or via specific screens within the app. Providing population-level information in return for individual participation could ensure continuous engagement with the app, which would improve data quality. The value of reciprocity should be tested in future MG participatory surveillance scenarios.

The Healthy Cup app served as the basis for the Guardians of Health (*Guardiões da Saúde*) participatory surveillance app that was used during the Olympic and Paralympic games in Brazil. An additional lesson learned through use of Guardians of Health is the importance of the need for clear expectations of a government's role in participatory surveillance. Some governments may not have dedicated teams to examine and interpret data generated through participatory surveillance. This issue underscores the importance of designing an intuitive platform that generates easily-visualized data.

Acknowledgments

This work was funded by a grant from Skoll Global Threats Fund, mediated by TEPHINET. We would like to especially thank Rodrigo Carneiro, Dennis Calazans, Lucas Medeiros, Gustavo Almeida, Thulio Philipe, Tulio Assis, Geraldo Augusto, and Cesar Albuquerque.

Authors' Contributions

OBLN organized and conceived of the paper, and analyzed and wrote the first and final versions of the paper. GSD conceived of analysis plans and analyzed the data. ML critically reviewed the paper. WKO conceptualized the aims. WS and EC conceptualized the aims, and reviewed the paper and intellectual content. MS and JA reviewed the paper and intellectual content. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Memish Z, Stephens G, Al Rabeeah A. Mass gatherings health. *Lancet Infect Dis* 2012;3099(11):10. [doi: [10.1016/S1473-3099\(11\)70287-0](https://doi.org/10.1016/S1473-3099(11)70287-0)]
2. Johansson A, Batty M, Hayashi K, Al Bar O, Marcozzi D, Memish ZA. Crowd and environmental management during mass gatherings. *Lancet Infect Dis* 2012 Feb;12(2):150-156 [FREE Full text] [doi: [10.1016/S1473-3099\(11\)70287-0](https://doi.org/10.1016/S1473-3099(11)70287-0)] [Medline: [22252150](https://pubmed.ncbi.nlm.nih.gov/22252150/)]
3. Abubakar I, Gautret P, Brunette GW, Blumberg L, Johnson D, Pomeroy G, et al. Global perspectives for prevention of infectious diseases associated with mass gatherings. *Lancet Infect Dis* 2012 Jan;12(1):66-74. [doi: [10.1016/S1473-3099\(11\)70246-8](https://doi.org/10.1016/S1473-3099(11)70246-8)] [Medline: [22192131](https://pubmed.ncbi.nlm.nih.gov/22192131/)]
4. Memish ZA, Stephens GM, Steffen R, Ahmed QA. Emergence of medicine for mass gatherings: lessons from the Hajj. *Lancet Infect Dis* 2012 Jan;12(1):56-65. [doi: [10.1016/S1473-3099\(11\)70337-1](https://doi.org/10.1016/S1473-3099(11)70337-1)] [Medline: [22192130](https://pubmed.ncbi.nlm.nih.gov/22192130/)]
5. Steffen R, Bouchama A, Johansson A, Dvorak J, Isla N, Smallwood C, et al. Non-communicable health risks during mass gatherings. *Lancet Infect Dis* 2012 Feb;12(2):142-149. [doi: [10.1016/S1473-3099\(11\)70293-6](https://doi.org/10.1016/S1473-3099(11)70293-6)] [Medline: [22252147](https://pubmed.ncbi.nlm.nih.gov/22252147/)]
6. Deris ZZ, Hasan H, Sulaiman SA, Wahab MS, Naing NN, Othman NH. The prevalence of acute respiratory symptoms and role of protective measures among Malaysian hajj pilgrims. *J Travel Med* 2010;17(2):82-88 [FREE Full text] [doi: [10.1111/j.1708-8305.2009.00384.x](https://doi.org/10.1111/j.1708-8305.2009.00384.x)] [Medline: [20412173](https://pubmed.ncbi.nlm.nih.gov/20412173/)]
7. Khan K, Freifeld CC, Wang J, Mekar SR, Kossowsky D, Sonricker AL, et al. Preparing for infectious disease threats at mass gatherings: the case of the Vancouver 2010 Olympic Winter Games. *CMAJ* 2010 Apr 06;182(6):579-583 [FREE Full text] [doi: [10.1503/cmaj.100093](https://doi.org/10.1503/cmaj.100093)] [Medline: [20181726](https://pubmed.ncbi.nlm.nih.gov/20181726/)]
8. Noël H, Rizzo C. Spread of chikungunya from the Caribbean to mainland Central and South America: a greater risk of spillover in Europe? *Euro Surveill* 2014 Jul 17;19(28):20855. [Medline: [25060570](https://pubmed.ncbi.nlm.nih.gov/25060570/)]
9. Donalisio MR, Freitas AR. Chikungunya in Brazil: an emerging challenge. *Rev Bras Epidemiol* 2015 Mar;18(1):283-285 [FREE Full text] [doi: [10.1590/1980-5497201500010022](https://doi.org/10.1590/1980-5497201500010022)] [Medline: [25651028](https://pubmed.ncbi.nlm.nih.gov/25651028/)]
10. Dyer O. Zika virus spreads across Americas as concerns mount over birth defects. *BMJ* 2015 Dec 23;351:h6983. [Medline: [26698165](https://pubmed.ncbi.nlm.nih.gov/26698165/)]
11. World Health Organization. International Health Regulations 2005 2nd Edition. Geneva: World Health Organization; 2005.
12. Bogich TL, Chunara R, Scales D, Chan E, Pinheiro LC, Chmura AA, et al. Preventing pandemics via international development: a systems approach. *PLoS Med* 2012;9(12):e1001354 [FREE Full text] [doi: [10.1371/journal.pmed.1001354](https://doi.org/10.1371/journal.pmed.1001354)] [Medline: [23239944](https://pubmed.ncbi.nlm.nih.gov/23239944/)]
13. Christakis NA, Fowler JH. Social network sensors for early detection of contagious outbreaks. *PLoS One* 2010;5(9):e12948 [FREE Full text] [doi: [10.1371/journal.pone.0012948](https://doi.org/10.1371/journal.pone.0012948)] [Medline: [20856792](https://pubmed.ncbi.nlm.nih.gov/20856792/)]
14. Chan EH, Brewer TF, Madoff LC, Pollack MP, Sonricker AL, Keller M, et al. Global capacity for emerging infectious disease detection. *Proc Natl Acad Sci U S A* 2010 Dec 14;107(50):21701-21706 [FREE Full text] [doi: [10.1073/pnas.1006219107](https://doi.org/10.1073/pnas.1006219107)] [Medline: [21115835](https://pubmed.ncbi.nlm.nih.gov/21115835/)]
15. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection--harnessing the Web for public health surveillance. *N Engl J Med* 2009 May 21;360(21):2153-5, 2157 [FREE Full text] [doi: [10.1056/NEJMp0900702](https://doi.org/10.1056/NEJMp0900702)] [Medline: [19423867](https://pubmed.ncbi.nlm.nih.gov/19423867/)]
16. Brownstein JS, Freifeld CC. HealthMap: the development of automated real-time Internet surveillance for epidemic intelligence. *Euro Surveill* 2007 Nov 29;12(11):E071129.5 [FREE Full text] [Medline: [18053570](https://pubmed.ncbi.nlm.nih.gov/18053570/)]
17. Madoff LC, Woodall JP. The Internet and the global monitoring of emerging diseases: lessons from the first 10 years of ProMED-mail. *Arch Med Res* 2005;36(6):724-730. [doi: [10.1016/j.arcmed.2005.06.005](https://doi.org/10.1016/j.arcmed.2005.06.005)] [Medline: [16216654](https://pubmed.ncbi.nlm.nih.gov/16216654/)]
18. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PLoS One* 2010;5(11):e14118 [FREE Full text] [doi: [10.1371/journal.pone.0014118](https://doi.org/10.1371/journal.pone.0014118)] [Medline: [21124761](https://pubmed.ncbi.nlm.nih.gov/21124761/)]
19. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS One* 2011;6(5):e19467 [FREE Full text] [doi: [10.1371/journal.pone.0019467](https://doi.org/10.1371/journal.pone.0019467)] [Medline: [21573238](https://pubmed.ncbi.nlm.nih.gov/21573238/)]
20. Salathé M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, Buckee C, et al. Digital epidemiology. *PLoS Comput Biol* 2012;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/)]
21. Freifeld CC, Chunara R, Mekar SR, Chan EH, Kass-Hout T, Ayala IA, et al. Participatory epidemiology: use of mobile phones for community-based health reporting. *PLoS Med* 2010 Dec 07;7(12):e1000376 [FREE Full text] [doi: [10.1371/journal.pmed.1000376](https://doi.org/10.1371/journal.pmed.1000376)] [Medline: [21151888](https://pubmed.ncbi.nlm.nih.gov/21151888/)]
22. Keller M, Blench M, Tolentino H, Freifeld CC, Mandl KD, Mawudeku A, et al. Use of unstructured event-based reports for global infectious disease surveillance. *Emerg Infect Dis* 2009 May;15(5):689-695 [FREE Full text] [doi: [10.3201/eid1505.081114](https://doi.org/10.3201/eid1505.081114)] [Medline: [19402953](https://pubmed.ncbi.nlm.nih.gov/19402953/)]
23. 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/)]
24. Oliveira W, Dimech GS, Leal-Neto OB, Libel M, Smolinski M. The world's first application of participatory surveillance at a mass gathering: FIFA World Cup 2014, Brazil. 2014 Presented at: International Meeting on Emerging Diseases and Surveillance 2014; Nov 1 2014; Vienna, Austria p. 44.

25. Observatório da Dengue. 2017. URL: <http://observatorio.inweb.org.br/dengue/conteudo/inicial%3E> [accessed 2017-04-04] [WebCite Cache ID 6pTmjyZSn]
26. 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/)]
27. Dengue na Web. 2017. URL: <http://www.denguenaweb.org/projeto%3E> [accessed 2017-04-04] [WebCite Cache ID 6pTmmp9lh]
28. Wójcik OP, Brownstein JS, Chunara R, Johansson MA. Public health for the people: participatory infectious disease surveillance in the digital age. *Emerg Themes Epidemiol* 2014;11:7 [FREE Full text] [doi: [10.1186/1742-7622-11-7](https://doi.org/10.1186/1742-7622-11-7)] [Medline: [24991229](https://pubmed.ncbi.nlm.nih.gov/24991229/)]
29. 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](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](https://pubmed.ncbi.nlm.nih.gov/19329408/)]
30. Chunara R, Smolinski MS, Brownstein JS. Why we need crowdsourced data in infectious disease surveillance. *Curr Infect Dis Rep* 2013 Aug;15(4):316-319 [FREE Full text] [doi: [10.1007/s11908-013-0341-5](https://doi.org/10.1007/s11908-013-0341-5)] [Medline: [23689991](https://pubmed.ncbi.nlm.nih.gov/23689991/)]
31. Jost CC, Mariner JC, Roeder PL, Sawitri E, Macgregor-Skinner GJ. Participatory epidemiology in disease surveillance and research. *Rev Sci Tech* 2007 Dec;26(3):537-549. [Medline: [18293603](https://pubmed.ncbi.nlm.nih.gov/18293603/)]
32. Brockmann S, Piechotowski I, Bock-Hensley O, Winter C, Oehme R, Zimmermann S, et al. Outbreak of leptospirosis among triathlon participants in Germany, 2006. *BMC Infect Dis* 2010 Apr 10;10:91 [FREE Full text] [doi: [10.1186/1471-2334-10-91](https://doi.org/10.1186/1471-2334-10-91)] [Medline: [20380736](https://pubmed.ncbi.nlm.nih.gov/20380736/)]
33. Pfaff G, Lohr D, Santibanez S, Mankertz A, van Treeck U, Schonberger K, et al. Spotlight on measles 2010: Measles outbreak among travellers returning from a mass gathering, Germany, September to October 2010. *Euro Surveill* 2010 Dec 16;15(50) [FREE Full text] [Medline: [21172175](https://pubmed.ncbi.nlm.nih.gov/21172175/)]
34. Schmid D, Holzmann H, Alfery C, Wallenko H, Popow-Kraupp TH, Allerberger F. Mumps outbreak in young adults following a festival in Austria, 2006. *Euro Surveill* 2008 Feb 14;13(7) [FREE Full text] [Medline: [18445415](https://pubmed.ncbi.nlm.nih.gov/18445415/)]
35. Centers for Disease Control and Prevention (CDC). Update: outbreak of acute febrile illness among athletes participating in Eco-Challenge-Sabah 2000--Borneo, Malaysia, 2000. *MMWR Morb Mortal Wkly Rep* 2001 Jan 19;50(2):21-24 [FREE Full text] [Medline: [11215718](https://pubmed.ncbi.nlm.nih.gov/11215718/)]
36. Wilder-Smith A, Earnest A, Ravindran S, Paton NI. High incidence of pertussis among Hajj pilgrims. *Clin Infect Dis* 2003 Nov 01;37(9):1270-1272 [FREE Full text] [doi: [10.1086/378748](https://doi.org/10.1086/378748)] [Medline: [14557975](https://pubmed.ncbi.nlm.nih.gov/14557975/)]
37. PhoneGap. Adobe Systems Inc. 2016. URL: <http://phonegap.com/> [accessed 2017-04-04] [WebCite Cache ID 6pTmxv6Vh]
38. Google. Google Developers Consoles. 2017. URL: <https://accounts.google.com/ServiceLogin?service=cloudconsole<mpl=api&osid=1&passive=true&continue=https://console.developers.google.com/> [accessed 2017-04-04] [WebCite Cache ID 6pTn0GIwI]
39. Twitter APIs. Twitter Developers. Disponível em URL: <https://dev.twitter.com/> [WebCite Cache ID 6pTmxv6Vh]
40. DreamHost. DreamHost. 2017. URL: <http://www.dreamhost.com/> [accessed 2017-04-16] [WebCite Cache ID 6pmAVhCZK]
41. Chunara R, Freifeld CC, Brownstein JS. New technologies for reporting real-time emergent infections. *Parasitology* 2012 Dec;139(14):1843-1851 [FREE Full text] [doi: [10.1017/S0031182012000923](https://doi.org/10.1017/S0031182012000923)] [Medline: [22894823](https://pubmed.ncbi.nlm.nih.gov/22894823/)]
42. Debin M, Turbelin C, Blanchon T, Bonmarin I, Falchi A, Hanslik T, et al. Evaluating the feasibility and participants' representativeness of an online nationwide surveillance system for influenza in France. *PLoS One* 2013;8(9):e73675 [FREE Full text] [doi: [10.1371/journal.pone.0073675](https://doi.org/10.1371/journal.pone.0073675)] [Medline: [24040020](https://pubmed.ncbi.nlm.nih.gov/24040020/)]
43. Eysenbach G. SARS and population health technology. *J Med Internet Res* 2003;5(2):e14 [FREE Full text] [doi: [10.2196/jmir.5.2.e14](https://doi.org/10.2196/jmir.5.2.e14)] [Medline: [12857670](https://pubmed.ncbi.nlm.nih.gov/12857670/)]
44. Collier N, Son NT, Nguyen NM. OMG U got flu? Analysis of shared health messages for bio-surveillance. *J Biomed Semantics* 2011;2 Suppl 5:S9 [FREE Full text] [doi: [10.1186/2041-1480-2-S5-S9](https://doi.org/10.1186/2041-1480-2-S5-S9)] [Medline: [22166368](https://pubmed.ncbi.nlm.nih.gov/22166368/)]
45. Patterson-Lomba O, Van Noort NS, Cowling BJ, Wallinga J, Gomes MG, Lipsitch M, et al. Utilizing syndromic surveillance data for estimating levels of influenza circulation. *Am J Epidemiol* 2014 Jun 01;179(11):1394-1401 [FREE Full text] [doi: [10.1093/aje/kwu061](https://doi.org/10.1093/aje/kwu061)] [Medline: [24748609](https://pubmed.ncbi.nlm.nih.gov/24748609/)]

Abbreviations

- API:** Application Program Interface
FIFA: Fédération Internationale de Football Association
IHR: International Health Regulation
MG: mass gathering
-

Edited by T Sanchez; submitted 12.01.17; peer-reviewed by E Yom-Tov, D Bateman; comments to author 16.02.17; revised version received 23.02.17; accepted 20.03.17; published 04.05.17.

Please cite as:

*Leal Neto O, Dimech GS, Libel M, de Souza WV, Cesse E, Smolinski M, Oliveira W, Albuquerque J
Saúde na Copa: The World's First Application of Participatory Surveillance for a Mass Gathering at FIFA World Cup 2014, Brazil
JMIR Public Health Surveill 2017;3(2):e26*

URL: <http://publichealth.jmir.org/2017/2/e26/>

doi: [10.2196/publichealth.7313](https://doi.org/10.2196/publichealth.7313)

PMID: [28473308](https://pubmed.ncbi.nlm.nih.gov/28473308/)

©Onicio Leal Neto, George Santiago Dimech, Marlo Libel, Wayner Vieira de Souza, Eduarda Cesse, Mark Smolinski, Wanderson Oliveira, Jones Albuquerque. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 04.05.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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

Understanding Environmental and Contextual Influences of Physical Activity During First-Year University: The Feasibility of Using Ecological Momentary Assessment in the MovingU Study

Chloe Bedard^{1,2}, BHSc (Hons); Sara King-Dowling^{2,3}, BSc Kin (Hons); Madeline McDonald^{2,3}, BSc Kin (Hons); Genevieve Dunton⁴, MPH, PhD; John Cairney^{2,5}, MA, PhD; Matthew Kwan^{2,3}, MSc, PhD

¹Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

²Infant and Child Health (INCH) Lab, Department of Family Medicine, McMaster University, Hamilton, ON, Canada

³Department of Kinesiology, McMaster University, Hamilton, ON, Canada

⁴Institute for Health Promotion & Disease Prevention Research, Department of Preventive Medicine, University of Southern California, Los Angeles, CA, United States

⁵Faculty of Kinesiology and Physical Education, University of Toronto, Toronto, ON, Canada

Corresponding Author:

Matthew Kwan, MSc, PhD

Infant and Child Health (INCH) Lab, Department of Family Medicine

McMaster University

DBHSC, 5th Floor

1280 Main Street West

Hamilton, ON, L8S 4L8

Canada

Phone: 1 9055259140 ext 20303

Fax: 1 9055274440

Email: kwanmy@mcmaster.ca

Abstract

Background: It is well established that drastic declines in physical activity (PA) occur during young adults' transition into university; however, our understanding of contextual and environmental factors as it relates to young adults' PA is limited.

Objective: The purpose of our study was to examine the feasibility of using wrist-worn accelerometers and the use of ecological momentary assessment (EMA) to assess the context and momentary correlates of PA on multiple occasions each day during first-year university.

Methods: First-year university students were asked to participate in the study. The participants completed a brief questionnaire and were subsequently asked to wear an ActiGraph GT9X-Link accelerometer and respond to a series of EMA prompts (7/day) via their phones for 5 consecutive days.

Results: A total of 96 first-year university students with smartphones agreed to participate in the study (mean age 18.3 [SD 0.51]; n=45 females). Overall, there was good compliance for wearing the accelerometers, with 91% (78/86) of the participants having ≥ 2 days of ≥ 10 hours of wear time (mean=3.53 valid days). Students were generally active, averaging 10,895 steps/day (SD 3413) or 1123.23 activity counts/min (SD 356.10). Compliance to EMA prompts was less desirable, with 64% (55/86) of the participants having usable EMA data (responding to a minimum of ≥ 3 days of 3 prompts/day or ≥ 4 days of 2 prompts/day), and only 47% (26/55) of these participants were considered to have excellent EMA compliance (responding to ≥ 5 days of 4 prompts/day or ≥ 4 days of 5 prompts/day).

Conclusions: This study represents one of the first studies to use an intensive real-time data capture strategy to examine time-varying correlates of PA among first-year university students. These data will aim to describe the physical and social contexts in which PA occurs and examine the relationships between momentary correlates of PA among the first-year university students. Overall, current results suggest that wrist-worn accelerometers and EMA are feasible methods for data collection among the young adult population; however, more work is needed to understand how to improve upon compliance to a real-time data capture method such as EMA.

(JMIR Public Health Surveill 2017;3(2):e32) doi:[10.2196/publichealth.7010](https://doi.org/10.2196/publichealth.7010)

KEYWORDS

exercise; compliance; feasibility studies; young adult; students

Introduction

Background

There is mounting evidence that engagement in habitual moderate-to-vigorous aerobic physical activity (MVPA) provides many physiological and psychosocial benefits as well as attenuates the risk of over 25 chronic health conditions [1-3]. In Canada, it is recommended that adults over the age of 18 years engage in 150 minutes of MVPA per week [4]. Despite the known benefits of regular physical activity (PA), accelerometry data collected in the 2012-2013 Canadian Health Measures Survey suggest that approximately 20% of Canadian adults engage in the recommended levels of PA [5]. There is a documented widespread pandemic of physical inactivity across developed nations [6-10], and this is particularly salient among the emerging adult population (ie, ages 18 to 25 years) [11-16]. Specifically, the transition out of high school has been found to be a period of time marked by drastic declines in leisure-time PA participation [17-19]. These declines in PA during this life stage are particularly concerning because behavior patterns exhibited during emerging adulthood track through to adulthood [20]. Therefore, it is critical that we understand how to best attenuate these drastic declines in PA, thereby developing strategies to help facilitate PA behaviors during early adulthood.

Before effective interventions can be designed and applied, salient determinants of PA and the specific contexts in which PA occurs must be well understood. Current evidence suggests that self-efficacy, behavioral intentions, past behavior, time constraints, and changing academic pressures are important factors related to PA behaviors in young adults transitioning to university or college [14,21-25]. To date, few studies have investigated the environmental and contextual influences of PA as these young adults move away from home for the first time. For example, little is known about the specific times, places, and settings in which PA tends to take place, as well as how momentary changes in affective or mood states, state self-control, or PA motivations impact PA behaviors. Furthermore, extant literature is often limited due to cross-sectional or prospective designs with single, self-reported retrospective assessments of PA and its correlates [14,21-23]. The transition out of high school is considered to be the first major life transition that an individual faces during the life course, often resulting in corresponding changes in priorities and actions [26]. This reinforces the need for more research to understand how this transitory period influences PA cognitions and, potentially, its variations during this major transition period.

Measurement Issues

Traditional methods examining salient determinants of PA use self-report surveys, requiring respondents to retrospectively report their perceptions, thoughts, and feelings related to general PA over a predetermined period of time (eg, in the last 2 months, over the past week). A major limitation of these self-reported measures is the presence of social desirability and recall biases [27,28]. Self-reported PA frequencies and durations have been

correlated to social desirability and social approval personality traits, leading to an overestimation of activity levels [29]. Particularly within the PA literature, it has been noted that our ability to recall over the past week is limited and highly subject to over-reporting errors [30-32]. Even in an honest attempt to recollect PA behavior in the time between PA occurrence and the time at which the respondent is asked about it, many details become distorted, thus resulting in invalid answers reported on the survey [27,31]. Furthermore, psychosocial assessments will ask respondents to reflect back over time to generalize their thoughts, raising questions of ecological validity for these measures [33]. More recently, however, emerging literature is attempting to minimize the time between the events and reporting, using real-time data capture methods [34].

New Measurement Tools

Ecological momentary assessments (EMA) have become increasingly popular as a real-time data capture method in PA research. Although the method itself is not entirely new, as there are many examples of researchers using diaries and logs to record momentary feeling states and behaviors [34,35], the proliferation of smartphones has created new opportunities to conduct EMA studies, while limiting the burden of constant recording of occurrences using the traditional paper-and-pencil method. Smartphones have become ubiquitous in the young adult population [36] and therefore represent an intuitive way to implement data-intensive recruitment strategies by collecting multiple responses from participants via electronic surveys sent multiple times throughout the day [34]. These data would also be advantageous as they represent more ecologically valid assessments that are representative of everyday life [34].

The emergence of objective PA measurement using devices such as accelerometers has also become increasingly popular due to their ability to capture the occurrence, duration, and intensity of PA [37]. More recently, validated wrist-worn accelerometers have become available for researchers, providing an alternative to the traditional waist-worn accelerometers, which can be more aesthetically appealing and comfortable to wear—both of which have been cited as reasons for noncompliance to waist-worn [38]. According to Troiano et al [39], the compliance rate for wrist-worn accelerometers tends to be higher than the compliance rate for waist-worn accelerometers, thus supporting the shift in research from waist-worn to wrist-worn accelerometers. In addition, results from the National Health and Nutrition Examination Survey (NHANES) study found that the compliance rate was 70% to 80% for participants providing valid data with wrist-worn accelerometers as compared with only 40% to 70% for those who provided valid data with waist-worn accelerometers [39]. Despite the emergence of these tools, few studies have investigated the utility of using these methods in PA research among young adults who recently transitioned out of high school.

Study Purpose

The purpose of this study was to explore the feasibility of using wrist-worn accelerometers and using EMA as a data-capture tool to better understand the salient correlates of PA for young adults who recently transitioned from high school into university.

Methods

Participants and Recruitment

The sample used in this study consisted of 86 first-year university students living in university residences as a part of the MovingU study. The mean age of the sample was 18.31 years (SD 0.51), primarily white (38/85, 45%) or Asian students (26/85, 31%), with a relatively equal proportion of males (40/85, 47%) and females (45/85, 53%) (see [Table 1](#)). To be eligible, participants must have been in their first year of studies, living on campus, and have a smartphone capable of downloading the EMA app (iOS or Android operating systems). Participants were recruited over the course of a 2-week period at two university residence buildings. Recruitment materials (ie, social media and flyers) briefly outlined the study's purpose and were advertised to all the students living in those residence buildings. All students interested in participating in the study were asked to meet with trained research assistants in a common area on a Tuesday evening.

Procedures

All interested students were first provided with a detailed description of the study purposes and requirements. Participants

who provided written consent were then asked to complete a Web-based questionnaire using an Android tablet. Once completed, participants were asked to download on their smartphone (iPhone or Android) the mobile EMA (mEMA) app designed and developed by illumavu Inc [40]. Each participant was given a unique personalized code to enter into the app, with EMA prompts scheduled to begin the following morning. A random sampling schedule was used, whereby participants were prompted at random times within predetermined time frames throughout the day (ie, every 2 hours) between 9:00 AM and 11:00 PM. Participants were instructed to complete all or as many of the EMA prompts over the 5-day sampling period and were also given a wrist-worn accelerometer (ActiGraph GT9X Link; Actigraph, Pensacola, FL, USA) to wear for each of those days. This 5-day sample is generally consistent with previous EMA research, which typically spans between 4 and 14 days [41]. With the exception of showering and participation in water-based activities, participants were instructed to wear the accelerometer for as long as possible (including sleep). Participants were provided a sleep log to indicate the times they went to bed each night and woke up each morning. They were also instructed to contact the study team if they encountered any issues with the EMA app. To compensate the students for their time and efforts, participants were given a Can \$10 Starbucks gift card for completion of the questionnaire and agreeing to wear the accelerometer. For the additional burden of completing EMA prompts, they were compensated another \$1 (in Starbucks gift cards) for each prompt they completed to a maximum of Can \$5 per day (or Can \$25 over the 5-day study period). All study procedures were approved by the Hamilton Integrated Research Ethics Board.

Table 1. Demographic characteristics of participants (N=85).

Characteristics	n (%) or mean (SD)
Age in years, mean (SD) ^a	18.3 (0.51)
Gender, n (%)	
Male	40 (47)
Female	45 (53)
Ethnicity, n (%)	
White	38 (45)
Asian	26 (30)
South Asian	6 (7)
Middle Eastern	5 (6)
Black	3 (4)
Indigenous	2 (2)
Mixed race	3 (4)
Prefer not to answer	2 (2)
12-point GPA ^b , mean (SD)	8.1 (2.3)

^aSD: standard deviation.

^bGPA: grade point average.

Table 2. Compliance rates for accelerometer wear time (N=86).

Category	Definition	n (%)
Excellent	≥10 hours of wear time for ≥5 days	20 (23)
Good	≥10 hours of wear time for 4 days	34 (40)
Fair	≥10 hours of wear time for 3 days	15 (17)
Minimum	≥10 hours of wear time for 2 days	9 (11)
Noncompliant	Did not have ≥10 hours of wear time for 2 days	8 (9)

Feasibility Indicators

Accelerometer Compliance

Accelerometry is an objective method of assessing free-living PA and an established method for measuring activity [42]. Participants were instructed to wear the wrist-worn ActiGraph GT9X Link accelerometer on their nondominant wrist for 5 consecutive days. Given that the ActiGraph Link is relatively new, without validated cut-points relative to intensity of PA in this population, total counts of PA will be used and analyzed in 60s epochs. Compliance to accelerometer use was categorized into four domains distinguished by the number of days of valid wear time of 10 or more hours per day: minimal, fair, good, and excellent. The minimum compliance wear time of 10 hours per day is the most often used cut-off for accelerometry data in the United States and Canada [43,44]. Given that the purpose of this study was to describe the feasibility of using wrist-worn accelerometers, we categorized the data to reflect fair, good, and excellent compliance to accelerometer wear (see Table 2). Wear time was calculated using ActiLife based on the Troiano [43] wear time validation parameters with non-wear time defined as ≥60 minutes of consecutive zero counts.

EMA Compliance

Participants received 7 EMA prompts per day over the 5-day study period, resulting in 35 prompts in total. Using a random-sampling signal-contingent schedule, participants were asked to respond to prompts whenever a notification was sent to their device (taking approximately 1 to 2 minutes to complete). Consistent with several previous EMA studies [45,46], each prompt includes a very brief questionnaire, assessing contextual information on current activity, physical location, type of social company, current affective and feeling states, as well as state motivation to be active and self-control. Specifically, it included questions to obtain contextual

information on three domains: “What are you currently doing?” “Where are you right now?” and “Who are you currently with?” Each question had a range of response options outlining expected responses and an “other” response option with a textbox for options not listed. Each prompt also included measures of acute outcome expectancy (eg, “doing 10+ min of physical activity in the next few hours would help me feel less stressed”), barrier self-efficacy (eg, “Can you do 10+ min of physical activity sometime within the next few hours even if you get busy?”), and intentions (eg, “I intend to be physically active for 10+ min sometime within the next few hours”), and was evaluated on a 5-point Likert scale. These items have been used in previous research and have been found to be valid measures [41,46-48]. Additionally, 5 items to assess affective states were included (eg, “How happy do you feel right now?” and “How tense or anxious do you feel right now?”), with response options ranging from “not at all” to “extremely,” which were based on the Positive and Negative Affect Schedule for Children [49]. Finally, we included 2 items from the State Self-Control Capacity Scale (eg, “If I were tempted by something right now, it would be very difficult to resist”), assessed on a 7-point Likert scale as used in Schöndube et al [50]. Further details on the MovingU study design, including the specific questions asked at each prompt, can be found in a previous publication [45]. Compliance to EMA prompts were categorized into 4 groups distinguished by the number of days prompts were answered and the number of prompts answered per day: noncompliant, minimum, fair, and excellent. The noncompliant to minimum compliance threshold identified was based on the minimum responses we estimate are required for using statistical techniques such as mixed-effects modeling (ie, at least 3 days of 3 prompts/day OR ≥ 4 days of 2 prompts/day). The categorizations for fair and excellent compliance were to further describe how compliant the participants were to the EMA prompts (see Table 3).

Table 3. Compliance rates for EMA responses (N=86).

Category	Definition	n (%)
Excellent	≥5 days of 4 prompts/day or ≥4 days of 5 prompts/day; response to ≥57% of total prompts.	26 (30)
Fair	≥4 days of 3 prompts/day or at least three days (including 1 weekend day) of 4 prompts/day; response to ≥34% of total prompts.	18 (21)
Minimum	≥3 days of 3 prompts/day or ≥4 days of 2 prompts/day; response to ≥23% of total prompts.	11 (13)
Noncompliant	≤4 days of fewer than 2 prompts/day.	31 (36)

Process Evaluation Survey

Following the 5-day study period, all participants were invited to voluntarily complete an anonymous Web-based questionnaire assessing the acceptability and receptivity of our EMA study. An invitation to participate was emailed to all participating students, including a link to the questionnaire on the Web. A total of 47 of the 86 study participants (55% response rate) completed the brief 10-item questionnaire related to *compliance*, *perceived burden*, and *compensation*.

Compliance

Three questions were asked related to participants' general compliance to wearing the accelerometer or completing the EMA prompts. For example, we asked, "How often were you able to wear the accelerometer watch?" with responses ranging from 0=I didn't wear it to 4=I wore it all the time; and "On average, how many prompts per day did you respond to?" with responses ranging from 0=none to 4=5 or more.

Perceived Burden

Two questions sought feedback regarding the perceived study burden, asking, "What did you think about the number of prompts that you received each day?" with the response options of 0=far too few, 1=somewhat too few, 2=just right, 3=somewhat too many, and 4=far too many; and "What did you think about the 5 days of being prompted?" with the response of 0=5 days was appropriate or 1=5 days was too long.

Compensation

There were 5 items that asked about compensation and participants' overall motivation to participate in the study. For example, there was a question that asked, "What did you think of the study compensation or gift card amount?" Response options included 0=wasn't worth my time, 1=too little for my time, 2=fair, 3=good for the time required, and 4= very good , generous for the time required.

Analyses

Descriptive statistics were used to summarize the demographic characteristics of the study sample, compliance rates of both accelerometer wear and EMA prompt responses, as well as responses to the process evaluation survey. Statistical analyses were conducted using SPSS version 22.

Results

Sample Characteristics

A total of 98 students initially provided written consent to participate in the study. There were 2 students who provided formal withdrawal, 1 student due to smartphone incapability with the mEMA app and 1 student because of expressed participant burden. There were 10 students who expressed interest and provided written consent, but for one reason or another did not take part in any of the data collection activities (ie, survey, accelerometer, EMA responses). Details of the sample characteristics for all participants are shown in [Table 1](#), with one exception. A single participant provided accelerometer data and responded to EMA prompts, but failed to complete the baseline survey. This participant was included in the compliance

results but not in the demographic characteristic table. Sample characteristics of the 85 participants are presented in [Table 1](#).

Accelerometer Compliance

Of the 86 participants, a total of 91% (78/86) wore the activity monitor for at least two days for 10 or more hours on each those days (see [Table 2](#)). The majority of our sample (54/86, 63%) met the requirements for good or excellent compliance, meaning they had valid accelerometry data (ie, ≥ 10 hours of wear time) on at least 80% of the study period. Only 9% (8/86) of the sample was considered noncompliant, failing to wear the accelerometer for the minimum of 2 days for 10 or more hours.

EMA Compliance

Of the 86 participants in our sample, only 55 (64%) provided what we estimated to be usable data for analyses (ie, ≥ 3 days of 3 prompts/day or ≥ 4 days of 2 prompts/day). Among those with usable data based on compliance, nearly half of the participants (26/55, 47%) had excellent EMA compliance (defined as having 5 days of 4 or more prompts each day or having 4 days of 5 or more prompts each day), approximately one-third (18/55, 33%) had fair compliance (defined as responding to 4 days of 3 or more prompts per day or at least 3 days including 1 weekend day of 4 or more prompts per day), and 20% (11/55) have only met the minimum compliance requirements. The average number of prompts that participants with minimum compliance responded to over the 5 days was 19.6 (19.6/35, 56%). A complete breakdown of compliance rates for EMA responses are shown in [Table 3](#).

Combined Accelerometer and EMA Compliance

Of the 78 participants with valid accelerometry data, 64% (50/78) met the minimum compliance threshold of EMA responses. There were 5 participants who had compliant EMA data but did not have compliant accelerometer data, and a total of 3 participants who were noncompliant for both EMA responses and accelerometer wear.

Process Evaluation Survey

Among the 47 participants who completed the process evaluation survey, 91% (43/47) self-reported wearing the accelerometer most or all of the time and 89% (40/45; 2 participants skipped this question) reported answering 3 or more EMA prompts a day. In terms of the compensation, only 2% (1/47) reported that the gift card amount was either too little for the time required in the study or was not worth their time. However, at least 75% (35/47) reported that they would have responded to more EMA prompts if they were provided with more compensation for their time. The majority of participants, 51% (24/47) and 96% (45/47), reported that the number of prompts they received each day and the number of days of prompting, respectively, were appropriate. Only 4% (2/47) reported that 5 days of prompting was too long, but 34% (16/47) reported that there were either somewhat or far too many prompts.

Discussion

Principal Findings

Given the drastic declines in PA during the emerging adulthood period, it is becoming increasingly important to intervene before individuals form stable unhealthy behavioral patterns. Considering the ubiquitous nature of smartphones combined with the accessibility of accelerometry, there is a unique opportunity to collect precise estimates of PA and greater ecologically valid measurements of the salient PA influences to inform the design of interventions. This study demonstrated that it is feasible to combine methods of wrist-worn accelerometers and a mobile EMA app among first-year university students; however, more can be done to improve compliance.

The compliance rate for the accelerometers was good to excellent, with 91% of the sample providing valid accelerometry data. The proportion of our sample providing valid accelerometry data is comparable with that of waist-worn accelerometers used in the 2007-2009 Canadian Health Measures Survey. Additionally, the average number of steps per day, 10,895, is similar to the population average in Canada [51]. The similarity of both of our accelerometer compliance rate and average step count suggests that our sample was similar to the broader Canadian population [51].

The compliance rate for accelerometer wear, however, was considerably better than the compliance rate for EMA responses. Among the 86 participants in the study, only 55 responded for 3 or more days of 3 or more prompts per day or for 4 days of at least 2 prompts each day. This compliance rate somewhat aligns with the information obtained on the process evaluation survey, indicating that one-third of respondents thought there were too many prompts; however, critically, it is considerably lower than the compliance from a late adolescent sample from Southern California [52]. This difference in compliance rates may be in part due to differences in sample age, EMA sampling schedule, and compensation. In Dunton et al [52], the sample included high school students. As a result, the students may be more likely to appease research staff because they are not yet at the stage where they feel fully autonomous and independent, or they may be prompted by their parents to acquiesce with the research study requirements. Alternatively, there may be more novelty for high school students to participate in research, compared with the ambivalence that university students may be demonstrating as they often get asked to participate in a variety of studies on campus. This hypothesis is somewhat supported because 53% of our sample reported in the process evaluation survey that the "study was somewhat interesting, but (they) had other things to do," suggesting that our sample was prioritizing other activities over participation in research.

Another difference between our study design and the Dunton et al [52] study was the prompting schedule. Participants in Dunton et al [52] were not prompted throughout school hours and, thus, were prompted less frequently overall. Given that more than one-third of the postevaluation respondents indicated that they received too many prompts, this may have resulted in the greater rates of noncompliance in this study. Future research

aimed at investigating the specific factors predicting EMA responses is required. This could determine whether there may be specific times within a day or days within the study period that result in greater noncompliance.

Finally, the role of compensation may be something that can help to improve compliance rates. Although the vast majority of participants indicated their satisfaction with the compensation provided, over 75% reported that their compliance could be improved with increasing value of compensation per prompt. Interestingly, fewer respondents indicated they would improve their response rate if offered larger and guaranteed compensation regardless of the number of prompts they answered. The study by Dunton et al [52] compensated \$100 for 14 days of responses to EMA, which was significantly higher than our maximum of \$35 in gift cards that could be earned. While more compensation may be required more generally, future research may want to investigate the differences in compliance based on compensation per prompt or for a guaranteed compensation. The findings from our postevaluation survey would suggest that participants may have been more motivated to answer prompts based on the value attached to each prompt rather than the guarantee of a large gift card after study completion. The caveat, however, is that 95.7% of respondents indicated that the length of the study period was appropriate, yet the compliance was relatively modest. More research to help identify optimal compensation methods and study duration length is needed.

Limitations

There are a number of limitations associated with this study. First, the issue of generalizability of the sample must be acknowledged. The sample in this study included a relatively small and homogeneous group of students living on campus, and relatively passive recruitment methods were used. In addition to a response bias, our findings may not be generalizable to the entire first-year university student population, particularly for those who do not live on campus. Second, as this study took place at a university, the lower compliance rate may be due to the students' inability to answer prompts being given during class times. More research aimed at understanding factors related to overall compliance and compliance during certain times of the day is needed. Third, EMA compliance rates may be partly attributed to a lack of familiarity with the app, as the push notification that sent the EMA survey to the participants required participants to navigate into the app when prompted. This may not have been entirely intuitive for participants. Fourth, it may be possible that our single, brief 5-day data collection period is not representative of the student's typical routine. It may have been a particularly busy week for students that may be impacting compliance, thus further studies using multiple EMA collection periods are needed. Finally, we acknowledge that the response rate for the process evaluation survey was low, with less than half the sample providing feedback related to the acceptability and receptivity of using EMA in university students. The low response rate may be in part due to the anonymity of the survey and no compensation being given to respond. Given that they were completed anonymously, individual responses cannot be matched to their objectively measured compliance rate, although the vast majority of respondents reported being largely

compliant to both the accelerometer wear and EMA responses. This is indicative of a response bias, and warrants caution when interpreting the process evaluation results.

Conclusions

Overall, the MovingU study represents one of the first studies to use an intensive data recruitment strategy through the use of EMA, aimed at understanding the factors related to PA during students' transition from high school into university. Current results suggest that wrist-worn accelerometry is a feasible method to assess objective PA among the young adult population, but that more work is needed to understand how to increase compliance to real-time data capture methods such as

EMA. A greater understanding of the predictors of compliance to the EMA and accelerometer protocol in a population of emerging adults will inform the design of the next large-scale EMA study. Future work will then explore some of the time-varying and time-invariant predictors of EMA compliance, as well as begin to examine how PA correlates measured multiple times each day across the 5-day study period relate to objectively assessed PA in these first-year university students. With enhanced knowledge regarding the salient predictors of PA in an emerging adult population, we can move toward designing interventions that target these predictors to have a larger impact on PA behavior change during this major life transition.

Acknowledgments

This research was supported by an Insight Development grant from the Social Sciences and Humanities Research Council of Canada (430-2014-00631). We would also like to thank McMaster University Residence Life in facilitating the data collection process, and Sarah Wellman and other members of the Infant and Child Health (INCH) lab for helping with recruitment for the study.

Conflicts of Interest

None declared.

References

1. Warburton DE, Katzmarzyk PT, Rhodes RE, Shephard RJ. Evidence-informed physical activity guidelines for Canadian adults. *Can J Public Health* 2007;98(Suppl 2):S16-S68. [Medline: [18213940](#)]
2. Warburton DE, Charlesworth S, Ivey A, Nettlefold L, Bredin SS. A systematic review of the evidence for Canada's physical activity guidelines for adults. *Int J Behav Nutr Phys Act* 2010 May 11;7:39 [FREE Full text] [doi: [10.1186/1479-5868-7-39](#)] [Medline: [20459783](#)]
3. Bouchard C, Shephard R, Stephens T. ULB.: Human Kinetics Publishers; 1993. Physical activity, fitness, and health International Proceedings and Consensus Statement URL: <http://tocs.ulb.tu-darmstadt.de/34091963.pdf> [accessed 2017-05-17] [WebCite Cache ID 6qWtsu6sW]
4. Tremblay MS, Warburton DE, Janssen I, Paterson DH, Latimer AE, Rhodes RE, et al. New Canadian physical activity guidelines. *Appl Physiol Nutr Metab* 2011 Feb;36(1):36-46 [FREE Full text] [doi: [10.1139/H11-009](#)] [Medline: [21326376](#)]
5. Government of Canada Statistics Canada. 2015 Feb 18. Directly measured physical activity of adults, 2012 and 2013 URL: <http://www.statcan.gc.ca/pub/82-625-x/2015001/article/14135-eng.htm> [accessed 2017-03-07] [WebCite Cache ID 6omxg15T2]
6. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U. Global physical activity levels: surveillance progress, pitfalls, and prospects. *The Lancet* 2012 Jul 21;380(9838):247-257. [doi: [10.1016/S0140-6736\(12\)60646-1](#)] [Medline: [22818937](#)]
7. Bauman AE, Reis RS, Sallis JF, Wells JC, Loos RJF, Martin BW. Correlates of physical activity: why are some people physically active and others not? *The Lancet* 2012 Jul 21;380(9838):258-271. [doi: [10.1016/S0140-6736\(12\)60735-1](#)] [Medline: [22818938](#)]
8. Heath GW, Parra DC, Sarmiento OL, Andersen LB, Owen N, Goenka S, et al. Evidence-based intervention in physical activity: lessons from around the world. *The Lancet* 2012 Jul 21;380(9838):272-281. [doi: [10.1016/S0140-6736\(12\)60816-2](#)] [Medline: [22818939](#)]
9. Pratt M, Sarmiento OL, Montes F, Ogilvie D, Marcus BH, Perez LG, et al. The implications of megatrends in information and communication technology and transportation for changes in global physical activity. *The Lancet* 2012 Jul 21;380(9838):282-293. [doi: [10.1016/S0140-6736\(12\)60736-3](#)] [Medline: [22818940](#)]
10. Kohl HW, Craig CL, Lambert EV, Inoue S, Alkandari JR, Leetongin G, et al. The pandemic of physical inactivity: global action for public health. *The Lancet* 2012 Jul 21;380(9838):294-305. [doi: [10.1016/S0140-6736\(12\)60898-8](#)] [Medline: [22818941](#)]
11. Jung ME, Bray SR, Ginis KAM. Behavior change and the freshman 15: tracking physical activity and dietary patterns in 1st-year university women. *J Am Coll Health* 2008;56(5):523-530. [doi: [10.3200/JACH.56.5.523-530](#)] [Medline: [18400664](#)]
12. Bray SR, Born HA. Transition to university and vigorous physical activity: implications for health and psychological well-being. *J Am Coll Health* 2004;52(4):181-188. [doi: [10.3200/JACH.52.4.181-188](#)] [Medline: [15018429](#)]

13. Van Dyck D, De Bourdeaudhuij I, Deliens T, Deforche B. Can changes in psychosocial factors and residency explain the decrease in physical activity during the transition from high school to college or university? *Int J Behav Med* 2015 Apr;22(2):178-186. [doi: [10.1007/s12529-014-9424-4](https://doi.org/10.1007/s12529-014-9424-4)] [Medline: [25031186](#)]
14. Kwan MY, Bray SR, Ginis KAM. Predicting physical activity of first-year university students: an application of the theory of planned behavior. *J Am Coll Health* 2009;58(1):45-52. [doi: [10.3200/JACH.58.1.45-55](https://doi.org/10.3200/JACH.58.1.45-55)] [Medline: [19592353](#)]
15. Wengreen HJ, Moncur C. Change in diet, physical activity, and body weight among young-adults during the transition from high school to college. *Nutr J* 2009 Jul 22;8:32 [FREE Full text] [doi: [10.1186/1475-2891-8-32](https://doi.org/10.1186/1475-2891-8-32)] [Medline: [19624820](#)]
16. Arnett JJ. Emerging adulthood. A theory of development from the late teens through the twenties. *Am Psychol* 2000 May;55(5):469-480. [Medline: [10842426](#)]
17. Kwan MY, Cairney J, Faulkner GE, Pullenayegum EE. Physical activity and other health-risk behaviors during the transition into early adulthood: a longitudinal cohort study. *Am J Prev Med* 2012 Jan;42(1):14-20. [doi: [10.1016/j.amepre.2011.08.026](https://doi.org/10.1016/j.amepre.2011.08.026)] [Medline: [22176841](#)]
18. Dumith SC, Gigante DP, Domingues MR, Kohl HW. Physical activity change during adolescence: a systematic review and a pooled analysis. *Int J Epidemiol* 2011 Jun;40(3):685-698 [FREE Full text] [doi: [10.1093/ije/dyq272](https://doi.org/10.1093/ije/dyq272)] [Medline: [21245072](#)]
19. Gordon-Larsen P, Nelson MC, Popkin BM. Longitudinal physical activity and sedentary behavior trends: adolescence to adulthood. *Am J Prev Med* 2004 Nov;27(4):277-283. [doi: [10.1016/j.amepre.2004.07.006](https://doi.org/10.1016/j.amepre.2004.07.006)] [Medline: [15488356](#)]
20. Azevedo MR, Araújo CL, Cozzensa da Silva M, Hallal PC. Tracking of physical activity from adolescence to adulthood: a population-based study. *Rev Saude Publica* 2007 Feb;41(1):69-75 [FREE Full text] [Medline: [17273636](#)]
21. Kwan MY, Faulkner G. Perceptions and barriers to physical activity during the transition to university. *Am J Health Stud* 2011;26(2):87-96.
22. Gyuresik NC, Spink KS, Bray SR, Chad K, Kwan M. An ecologically based examination of barriers to physical activity in students from grade seven through first-year university. *J Adolesc Health* 2006 Jun;38(6):704-711. [doi: [10.1016/j.jadohealth.2005.06.007](https://doi.org/10.1016/j.jadohealth.2005.06.007)] [Medline: [16730599](#)]
23. Gómez-López M, Gallegos AG, Extremera AB. Perceived barriers by university students in the practice of physical activities. *J Sports Sci Med* 2010;9(3):374-381 [FREE Full text] [Medline: [24149629](#)]
24. Keating XD, Guan J, Piñero JC, Bridges DM. A meta-analysis of college students' physical activity behaviors. *J Am Coll Health* 2005;54(2):116-125. [doi: [10.3200/JACH.54.2.116-126](https://doi.org/10.3200/JACH.54.2.116-126)] [Medline: [16255324](#)]
25. Bray SR. Self-efficacy for coping with barriers helps students stay physically active during transition to their first year at a university. *Res Q Exerc Sport* 2007 Mar;78(2):61-70. [doi: [10.1080/02701367.2007.10599404](https://doi.org/10.1080/02701367.2007.10599404)] [Medline: [17479575](#)]
26. Gottlieb BH, Still E, Newby-Clark IR. Types and precipitants of growth and decline in emerging adulthood. *J Adolesc Res* 2007 Mar 01;22(2):132-155. [doi: [10.1177/0743558406298201](https://doi.org/10.1177/0743558406298201)]
27. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use, 5th edition. Oxford: Oxford University Press; 2014.
28. Van de Mortel TF. Faking it: social desirability response bias in self-report research. *Aust J Adv Nurs* 2008;25(4):40.
29. Adams SA, Matthews CE, Ebbeling CB, Moore CG, Cunningham JE, Fulton J, et al. The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol* 2005 Feb 15;161(4):389-398 [FREE Full text] [doi: [10.1093/aje/kwi054](https://doi.org/10.1093/aje/kwi054)] [Medline: [15692083](#)]
30. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc* 2000 Sep;32(9 Suppl):S498-S504. [Medline: [10993420](#)]
31. Dishman RK, Washburn RA, Schoeller DA. Measurement of physical activity. *Quest* 2001 Aug;53(3):295-309. [doi: [10.1080/00336297.2001.10491746](https://doi.org/10.1080/00336297.2001.10491746)]
32. Rzewnicki R, Auweele YV, De Bourdeaudhuij I. Addressing overreporting on the International physical activity questionnaire (IPAQ) telephone survey with a population sample. *Public Health Nutr* 2003 May;6(3):299-305. [doi: [10.1079/PHN2002427](https://doi.org/10.1079/PHN2002427)] [Medline: [12740079](#)]
33. Cicourel A. Interviews, surveys, and the problem of ecological validity. *The American Sociologist* 1982;17(1):11-20 [FREE Full text]
34. Stone A, Shiffman S, Atienza A, Nebeling L. The Science of Real-Time Data Capture: Self-Reports in Health Research. New York, NY: Oxford University Press; 2007.
35. Welk G. Physical activity assessments for health-related research. Champaign, IL: Human Kinetics; 2002.
36. Catalyst. Smartphone behaviour in Canada and the implications for marketers in 2016 URL: <http://catalyst.ca/2016-canadian-smartphone-behaviour/> [accessed 2016-09-07] [WebCite Cache ID 6kLwMiaFX]
37. Hallal PC, Reichert FF, Clark VL, Cordeira KL, Menezes AM, Eaton S, et al. Energy expenditure compared to physical activity measured by accelerometry and self-report in adolescents: a validation study. *PLoS One* 2013;8(11):e77036 [FREE Full text] [doi: [10.1371/journal.pone.0077036](https://doi.org/10.1371/journal.pone.0077036)] [Medline: [24223707](#)]
38. Perry MA, Hendrick PA, Hale L, Baxter GD, Milosavljevic S, Dean SG, et al. Utility of the RT3 triaxial accelerometer in free living: an investigation of adherence and data loss. *Appl Ergon* 2010 May;41(3):469-476. [doi: [10.1016/j.apergo.2009.10.001](https://doi.org/10.1016/j.apergo.2009.10.001)] [Medline: [19875099](#)]
39. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. *Br J Sports Med* 2014 Jul;48(13):1019-1023 [FREE Full text] [doi: [10.1136/bjsports-2014-093546](https://doi.org/10.1136/bjsports-2014-093546)] [Medline: [24782483](#)]

40. ilumivu. mEMA URL: <https://ilumivu.wordpress.com/> [accessed 2016-09-15] [WebCite Cache ID 6kXztYNz5]
41. Liao Y, Skelton K, Dunton G, Bruening M. A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: an adapted STROBE checklist for reporting EMA studies (CREMAS). *J Med Internet Res* 2016 Jun 21;18(6):e151 [FREE Full text] [doi: [10.2196/jmir.4954](https://doi.org/10.2196/jmir.4954)] [Medline: [27328833](https://pubmed.ncbi.nlm.nih.gov/27328833/)]
42. Bouten CV, Westerterp KR, Verduin M, Janssen JD. Assessment of energy expenditure for physical activity using a triaxial accelerometer. *Med Sci Sports Exerc* 1994 Dec;26(12):1516-1523. [Medline: [7869887](https://pubmed.ncbi.nlm.nih.gov/7869887/)]
43. Troiano R, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc* 2008 Jan;40(1):181-188. [Medline: [18091006](https://pubmed.ncbi.nlm.nih.gov/18091006/)]
44. Colley RC, Garriguet D, Janssen I, Craig CL, Clarke J, Tremblay MS. Physical activity of Canadian adults: accelerometer results from the 2007 to 2009 Canadian health measures survey. *Health Rep* 2011 Mar;22(1):7-14 [FREE Full text] [Medline: [21510585](https://pubmed.ncbi.nlm.nih.gov/21510585/)]
45. Dunton GF, Atienza AA, Castro CM, King AC. Using ecological momentary assessment to examine antecedents and correlates of physical activity bouts in adults age 50+ years: a pilot study. *Ann Behav Med* 2009 Dec;38(3):249-255. [doi: [10.1007/s12160-009-9141-4](https://doi.org/10.1007/s12160-009-9141-4)] [Medline: [20052568](https://pubmed.ncbi.nlm.nih.gov/20052568/)]
46. Dunton GF, Intille SS, Wolch J, Pentz MA. Children's perceptions of physical activity environments captured through ecological momentary assessment: a validation study. *Prev Med* 2012 Aug;55(2):119-121. [doi: [10.1016/j.ypmed.2012.05.015](https://doi.org/10.1016/j.ypmed.2012.05.015)] [Medline: [22659225](https://pubmed.ncbi.nlm.nih.gov/22659225/)]
47. Dunton GF, Huh J, Leventhal AM, Riggs N, Hedeker D, Spruijt-Metz D, et al. Momentary assessment of affect, physical feeling states, and physical activity in children. *Health Psychol* 2014 Mar;33(3):255-263 [FREE Full text] [doi: [10.1037/a0032640](https://doi.org/10.1037/a0032640)] [Medline: [23668846](https://pubmed.ncbi.nlm.nih.gov/23668846/)]
48. Dunton GF, Liao Y, Intille SS, Spruijt-Metz D, Pentz M. Investigating children's physical activity and sedentary behavior using ecological momentary assessment with mobile phones. *Obesity (Silver Spring)* 2011 Jun;19(6):1205-1212. [doi: [10.1038/oby.2010.302](https://doi.org/10.1038/oby.2010.302)] [Medline: [21164502](https://pubmed.ncbi.nlm.nih.gov/21164502/)]
49. Ebesutani C, Regan J, Smith A, Reise S, Higa-McMillan C, Chorpita BF. The 10-item positive and negative affect schedule for children, child and parent shortened versions: application of item response theory for more efficient assessment. *J Psychopathol Behav Assess* 2012 Jan 4;34(2):191-203. [doi: [10.1007/s10862-011-9273-2](https://doi.org/10.1007/s10862-011-9273-2)]
50. Schöndube A, Bertrams A, Sudeck G, Fuchs R. Self-control strength and physical exercise: an ecological momentary assessment study. *Psychol Sport Exerc* 2017 Mar;29:19-26. [doi: [10.1016/j.psychsport.2016.11.006](https://doi.org/10.1016/j.psychsport.2016.11.006)]
51. Colley RC, Garriguet D, Janssen I, Craig CL, Clarke J, Tremblay MS. Physical activity of Canadian children and youth: accelerometer results from the 2007 to 2009 Canadian health measures survey. *Health Rep* 2011 Mar;22(1):15-23 [FREE Full text] [Medline: [21510586](https://pubmed.ncbi.nlm.nih.gov/21510586/)]
52. Dunton GF, Dzibur E, Intille S. Feasibility and performance test of a real-time sensor-informed context-sensitive ecological momentary assessment to capture physical activity. *J Med Internet Res* 2016 Jun 01;18(6):e106 [FREE Full text] [doi: [10.2196/jmir.5398](https://doi.org/10.2196/jmir.5398)] [Medline: [27251313](https://pubmed.ncbi.nlm.nih.gov/27251313/)]

Abbreviations

- EMA:** ecological momentary assessment
- mEMA:** mobile ecological momentary assessment
- MVPA:** moderate-to-vigorous aerobic physical activity
- NHANES:** National Health and Nutrition Examination Survey
- PA:** physical activity

Edited by G Eysenbach; submitted 17.11.16; peer-reviewed by G Heath, A Leblanc, S Smiley; comments to author 26.01.17; revised version received 08.03.17; accepted 30.03.17; published 31.05.17.

Please cite as:

Bedard C, King-Dowling S, McDonald M, Dunton G, Cairney J, Kwan M
Understanding Environmental and Contextual Influences of Physical Activity During First-Year University: The Feasibility of Using Ecological Momentary Assessment in the MovingU Study
JMIR Public Health Surveill 2017;3(2):e32
URL: <http://publichealth.jmir.org/2017/2/e32/>
doi: [10.2196/publichealth.7010](https://doi.org/10.2196/publichealth.7010)
PMID: [28566264](https://pubmed.ncbi.nlm.nih.gov/28566264/)

©Chloe Bedard, Sara King-Dowling, Madeline McDonald, Genevieve Dunton, John Cairney, Matthew Kwan. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 31.05.2017. This is an open-access article distributed under

the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.

Viewpoint

Ensuring Inclusion of Adolescent Key Populations at Higher Risk of HIV Exposure: Recommendations for Conducting Biological Behavioral Surveillance Surveys

Lisa Grazina Johnston^{1*}, MA, MPH, PhD; Justine Sass^{2*}, MA; Jeffry Acaba^{3*}, BA; Wing-Sie Cheng⁴, MA; Shirley Mark Prabhu^{4*}, MSc

¹Independent Consultant, UNICEF, East Asia and Pacific Regional Office, Valencia, Spain

²UNESCO, Section of Education for Inclusion and Gender Equality, Paris, France

³Youth Lead, Bangkok, Thailand

⁴UNICEF, East Asia and Pacific Regional Office, Bangkok, Thailand

*these authors contributed equally

Corresponding Author:

Lisa Grazina Johnston, MA, MPH, PhD

Independent Consultant, UNICEF, East Asia and Pacific Regional Office

14 Arzobispo Mayoral Puerta 5

Valencia, 46002

Spain

Phone: 34 634913918

Fax: 34 634913918

Email: lsjohnston.global@gmail.com

Abstract

Ending acquired immune deficiency syndrome (AIDS) depends on greater efforts to reduce new human immunodeficiency virus (HIV) infections and prevent AIDS-related deaths among key populations at highest HIV risk, including males who have sex with males, sex workers, and people who inject drugs. Although adolescent key populations (AKP) are disproportionately affected by HIV, they have been largely ignored in HIV biological behavioral surveillance survey (BBSS) activities to date. This paper reviews current ethical and sampling challenges and provides suggestions to ensure AKP are included in surveillance activities, with the aim being to enhance evidence-informed, strategic, and targeted funding allocations and programs toward ending AIDS among AKP. HIV BBSS, conducted every few years worldwide among adult key populations, provide information on HIV and other infections' prevalence, HIV testing, risk behaviors, program coverage, and when at least three of these surveys are conducted, trend data with which to evaluate progress. We provide suggestions and recommendations on how to make the case to ethical review boards to involve AKP in surveillance while assuring that AKP are properly protected. We also describe two widely used probability sampling methods, time location sampling and respondent driven sampling, and offer considerations of feature modifications when sampling AKP. Effectively responding to AKP's HIV and sexual risks requires the inclusion of AKP in HIV BBSS activities. The implementation of strategies to overcome barriers to including AKP in HIV BBSS will result in more effective and targeted prevention and intervention programs directly suited to the needs of AKP.

(*JMIR Public Health Surveill* 2017;3(2):e40) doi:[10.2196/publichealth.7459](https://doi.org/10.2196/publichealth.7459)

KEYWORDS

adolescents; sex workers; sexual minorities; drug users; ethics; sampling

Introduction

Ending human immunodeficiency virus (HIV) or acquired immune deficiency syndrome (AIDS) relies on greater efforts to reduce new HIV infections and prevent AIDS-related deaths among key populations at highest risk. Whereas data are limited, studies from low and concentrated epidemic countries suggest

that HIV prevalence is disproportionately high among adolescents, aged 10-19 years, who sell sex, engage in same-sex relationships, and inject drugs [1]. Ignoring that these behaviors occur among adolescents places them at even higher risk of HIV infection and creates barriers to HIV testing, and, when needed, essential care and treatment [2]. Adolescents under the age of 18 years who sell sex and are likely to be defined as

sexually exploited under human rights law [3,4], tend to be the most ignored and least protected group due to controversy and lack of clarity on how to meet their needs [5-7]. In addition, adolescent key populations (AKP) including adolescent males who have sex with males (MSM); male, female, and transgender adolescents who sell sex; and adolescents who inject drugs face numerous vulnerabilities including low education; low access to health care; and high levels of stigma, discrimination, and violence [2]. A punishing legal environment and severe taboos around same-sex relations, selling sex, and injecting drug use tend to drive these behaviors underground, reinforcing the exclusion of adolescents and perpetuating the infection cycle [2,8].

Since early 2000, HIV biological behavioral surveillance surveys (BBSS) have been a key component of national responses to HIV [9,10]. BBSS, conducted every few years worldwide among adult key populations, is the routine collection of strategic information to measure HIV and other infections and biological and behavioral trends; to model the HIV epidemic; to determine allocation of limited resources and funding; to enhance efforts to more effectively respond to local, national, and global HIV prevention strategies; and to measure program coverage. For countries that have conducted at least three BBSS, trends provide information about different countries' progress in responding to HIV. Many questions in the BBSS are sensitive, including those related to sexual behaviors; drug use; HIV testing history; visits to sexually transmitted infections services; and personal experiences of being arrested, raped, stigmatized, and discriminated against. Unfortunately, many of these surveillance activities do not collect data from persons under the age of 18 years. Some surveys have managed to collect data from participants as young as 15 years [11,12]; however, these data are not disaggregated (by age and sex), and the sample sizes comprising 15 to 19 year olds are often not sufficient for meaningful analysis [13,14]. To fill this gap, some countries have conducted separate studies on younger cohorts; however, many studies use non-probability sampling methods from which inferences about the population cannot be made [15,16].

The lack of continuous, systematic collection, analysis, and interpretation of data on AKP has resulted in a dearth of strategic

information needed for the planning, implementation, and evaluation of essential HIV programs. This paper discusses the ethical barriers related to sampling AKP, and suggestions on how to overcome them, and presents recommendations on how to include AKP in BBSS using time location sampling (TLS) and respondent driven sampling (RDS).

Methods

Ethical Barriers to Sampling Adolescent Key Populations (AKP)

One of the barriers to AKP inclusion in BBSS is ethical constraints. In almost all countries, adolescents under 18 years are considered children and entitled to the protection of their rights under the Convention on Rights of the Child [17]. Compared with adults, children are more vulnerable to exploitation, abuse, violence, and other harmful outcomes of research, and therefore, require additional safeguards [18-20]. In order for children to participate in research, informed consent is usually required from both the child and the child's parent or guardian. The failure of institutional review boards to approve self-consent and waive guardian permission in conducting HIV surveys is a significant barrier to AKP participation. AKP may fear being stigmatized, punished, or in some cases victimized by their families if guardian permission results in the disclosure of their behaviors, sexual orientation or gender identity, or HIV status [21]. As a result, surveys may end up with smaller samples and unrepresentative findings. Convincing governments to approve self or proxy (eg, trained social worker serving as a proxy guardian) consent may require building consensus of the value of AKP inclusion and promoting their right to participate, whereas at the same time ensuring their protection from harm. Key ethical parameters of data collection involving key populations under the age of 18 years include (1) informed consent; (2) domestic laws governing child protection; (3) identification of, and referral to, services for AKP; and (4) biological testing. Within each of these parameters, considerations and suggestions for supporting the inclusion of AKP in BBSS are presented in [Table 1](#).

Table 1. Ethical considerations and suggestions for conducting biological behavioral surveillance surveys (BBSS) among adolescent key populations (AKP) under the age of 18 years.

Topic	Considerations and suggestions
Obtaining informed consent for studies involving adolescents	<p>Meet with and receive input from all relevant groups and agencies when designing protocols and consent forms, and when considering who will need consent from parents or guardians. For instance, in some countries (1) the age of majority may be younger than 18 years, and (2) children who are married, living independently, or in child-headed households may be considered “liberated minors” and thus, may not require additional consent from a parent or guardian.</p> <p>When requesting consent for surveys on sensitive issues from a parent or guardian on behalf of AKP^a, it may be necessary to keep the nature of the survey vague (ie, refer to a survey on sexual risk or drug use as a “health” survey) and list sex or drugs as one of many health issues being assessed.</p>
Domestic laws governing adolescent protection	<p>Review domestic laws for mandatory reporting of disclosures of adolescent abuse, neglect, violence, or exploitation; consider how mandatory reporting would affect the final research outcome (eg, Will AKP refuse to participate if their information is not confidential? and Will AKP refuse to report certain types of information?) and decide whether a waiver is needed.</p> <p>Discuss issues of including or waiving mandatory reporting with adolescent protection officials, social workers, rights advocates, partner agencies, etc.</p> <p>Approval of such waivers may only be possible through the national ethical review board or a senior-level adolescent protection authority.</p>
Identification of, and referral to, services for adolescents	<p>Identify and develop a comprehensive list of potential services for AKP (ie, services able to address issues of violence, abuse, neglect or exploitation, and other services such as drop-in centers, shelters, helplines, government or non-government social welfare, vocational training, school programs, health care, legal aid) to respond to AKP special situations.</p> <p>Provide referral lists to AKP who participate in a study, whether or not they disclose harm or high-risk behavior.</p> <p>Be able to identify which specific service an adolescent may need to access (ie, do not have a list only containing referrals for those who need sexual abuse counseling or detoxification from drugs).</p> <p>Ensure that research protocols include clear procedures for making discreet, direct referrals to service providers, should AKP request such assistance. Unless there are mandatory reporting requirements, confidentiality must be respected.</p> <p>Work with qualified adolescent protection or health professionals to establish guidelines and meet with potential service providers to assess the capacity, expertise, and resources to respond to direct referrals, including how to respond to urgent or acute abuse, neglect, violence, or exploitation cases.</p> <p>If no services exist in the survey area or staff is not equipped to respond to referrals from AKP participants, consider budgeting for and establishing a temporary team of trained service providers to whom AKP can be referred during and shortly following data collection.</p> <p>Consider providing transportation arrangements so that service providers in nearby areas can meet with AKP who request support. Plan these arrangements in cooperation with protection and health professionals.</p> <p>Develop plans to determine how to accommodate AKP needing long-term or specialized support beyond the research.</p>
Eligibility	<p>Many labels used by researchers to describe key populations (ie, people who inject drugs, male or female sex workers) may not be recognized by AKP engaging in the same behaviors. Refer to the behavior guiding the eligibility, rather than the population group.</p>
Support during the survey	<p>Have available a professionally trained social worker or advocate (a person or service provider with qualifications to provide information and support to an adolescent in distress).</p>
Biological testing	<p>When considering inclusion of a biological component, find out (1) if the country’s ethical standards allow testing on adolescents. If so, what are the laws governing HIV^b or other infections testing of adolescents? (2) How is pre- and post-test counseling conducted for and test results provided to AKP and/or parents and guardians? and (3) Whether there are available and appropriate referrals for care, management, and treatment for AKP with positive test results.</p>

^aAKP: adolescent key populations.

^bHIV: human immunodeficiency virus.

Informed consent for AKP under the age of 18 years should include considerations beyond general ethical assurances included in any protocol. Extra effort may be needed to ensure that adolescents understand all of the elements in a consent process, including the purpose of the research, the kinds of information to be collected, how confidentiality will be

maintained, the interview procedure (in particular that the participant does not have to answer questions with which she or he does not feel comfortable, and that the interview can be stopped at any time), and a contact number for more information about the study or to make a complaint [22]. For instance, in the Philippines, AKP were required to understand and repeat

back 4 key items of consent in order to participate in a BBSS using RDS and TLS: (1) participation is voluntary, (2) information is confidential (no one will know what you tell me), (3) participation involves an interview and HIV counseling and testing, and (4) participation will help improve services for adolescents [23]. Another safeguard has been to have on-site social or health workers as “parental proxies” to provide consent on behalf of or in addition to AKP under the age of 18 years [24].

Some countries have specific regulations that disclosures of violence, abuse, neglect, or exploitation of a child override confidentiality and must be reported to relevant authorities. Government employees or particular professions (eg, social workers, health workers, and teachers) or any person aware of an incident must report it. If there are no exemptions for mandatory reporting for the research, then a waiver from an appropriate authority is needed so that interviewers are not required to report abuse disclosures without the adolescent’s approval. It is essential for survey planners to (1) review domestic mandatory reporting laws of disclosures of child abuse, neglect, violence, or exploitation and consider how reporting would affect the final research outcome; and (2) discuss options to waive mandatory reporting by adolescent protection officials, social workers, rights advocates, and partner agencies.

Collecting biological specimens from AKP under the age of 18 years will normally require informed consent from parents and guardians. In many situations, parents and guardians will have access to the test results [25], which can be a strong impediment to AKP getting tested, especially given the sensitivity of an HIV result and the implications of sexual activity or injecting drug use [2]. In light of the increased risk of HIV infection faced by adolescents, some countries are reevaluating and adapting their current legal and policy constraints requiring parental consent for adolescents wanting confidential HIV counseling and testing [21,26,27]. Twelve countries in Asia and the Pacific now have HIV testing laws in place that allow people under 18 years to consent independently to HIV testing [2].

Methods to Sample Adolescent Key Population (AKP)

Much of our knowledge about adolescent health comes from household- and school-based surveillance, both of which rely on populations that have sampling frames [28]. However, these surveys typically miss populations at higher risk for HIV exposure, many of whom have unstable living environments and housing and prefer to remain “hidden” from law enforcement and other government authorities. Furthermore, these studies fail to capture strategic information on HIV, including HIV and other infections prevalence, program coverage, or specific risk behaviors. There are currently 2 probability-based sampling methods that have been used successfully to sample adult key populations without sampling frames for BBSS—TLS (also known as venue-day-time sampling) and RDS [10,29,30]. Both of these methods allow participants to remain anonymous (ie, no need to collect name, address, or other personal details that can be used to contact a participant) and are therefore, effective at sampling populations that practice risky behaviors and/or face stigma and discrimination. Knowledge about the target population is needed

before deciding on which methodology to use. For instance, TLS requires that AKP congregate in identifiable and accessible locations such as street corners, markets, transportation centers, clubs, bars, or other areas [9,31,32], and RDS requires that the population be socially networked so that AKP can identify and recruit other eligible AKP.

TLS was first used to estimate HIV seroprevalence among young MSM (15-22 years) in the United States [33]. The method entails identifying days and times when populations congregate at specific locations (ie, brothels, city blocks, bars, and so on), constructing a sampling frame of time and location units, randomly selecting and visiting time and location units (the primary sampling units), and systematically intercepting and collecting information from consenting individuals. The number of persons at each location provides a sampling weight that can be used a priori, to draw a self-weighting sample, or post priori, in analysis. Data collection takes place at the venue, if space (or venue owner) permits, or in a mobile site near the location, such as a van, or by making appointments to come to a designated study site. The major contribution of TLS over other cluster sampling methods is the ability to account for the fact that key populations are not statically associated with a particular location and often move between multiple locations during a single day. As such, TLS allows researchers to construct a sample with known properties, make statistical inference to the larger population of location visitors, and theorize about the introduction of biases that may limit generalization of results to the target population.

RDS was first used to sample people who inject drugs in the United States [34]. Briefly, RDS begins with a handful of non-randomly selected participants (referred to as seeds) who enroll in the survey and upon completion, receive recruitment coupons that they use to recruit their peers [34-37]. Participants recruited by seeds make up the first wave of participants, who in turn recruit the second wave of participants. This recruitment process continues until the sample comprises multiple recruitment waves and ends once the desired sample size is reached. RDS provides a primary incentive for completing an interview and for recruiting peers. The use of uniquely numbered coupons identifies who recruited whom, while maintaining anonymity. Someone receiving a coupon can decide whether to enroll or not. The major contribution of RDS over other chain referral sampling methods is the ability to harness information about people’s social networks to make inferences about the population. Data on who recruited whom are used to adjust for differential recruitment, and the measurement of each participant’s social network size (the number of peers they know who also know them) is used to weigh the data to control for the overestimation of those with larger social networks and the underestimation of those with smaller social networks.

Both methods can, and have been, used for HIV and non-HIV related surveys of AKP (10-19 years) [38,39]. For example, in Asia, TLS was used in a survey conducted among MSM, male sex workers, and transgender persons (15-24 years; 30.6% of which were between the ages of 15-19 years) in Chiang Mai and Phuket, Thailand (n=827) [40]. In this study, using oral testing for HIV, 13.1% of AKP were HIV seropositive [40]. In Cambodia, TLS was used to sample 1204 males and 1166

females (10-24 years; 52.4% of which were boys and 53.2% of which were girls between the ages of 15-19 years), who were unmarried, and considered high risk based on their visibility at high risk venues. The sample comprised 252 hotspots (ie, bars, night clubs, karaoke parlors, massage parlors, street corners, football fields, snooker clubs, and computer game shops) in the capital city and 7 other provinces [41]. Although information about HIV-related risk behaviors was captured, no biological testing was performed, and there is no indication of adjustments in the analysis for sizes of venues and frequency of visits. Another survey using TLS that directly targeted young key populations living on the streets in the Russian Federation and Ukraine, mapped metro and train stations, street markets, feeding centers or fast-food sites, parks, and computer clubs [42]. This survey used 2 mobile teams in vans in which participants consented to participate, were interviewed, and underwent pre- and post-test counseling and a rapid HIV test.

RDS was used in several provinces of Thailand among young (15-24 years) females who exchange sex for money and goods, MSM, young non-Thai migrants, and transgender youth [43]. Although RDS recruitment worked well, the sample sizes were not sufficient enough to capture many adolescents, and although information about HIV-related risk behaviors and testing was captured, no biological testing was performed. These limitations were also found in RDS surveys conducted among young MSM in Monywa and Yangon, Myanmar [44].

Web-based RDS has been developed to sample hard to reach populations through messaging and emails [45-47]. Although this method does not allow for the collection of biological information since there is no face-to-face contact between research staff and participants, it can be useful and efficient for collecting behavioral information from AKP. For instance, in China, an RDS survey is being planned for adolescent males who have sex with males using a Web-based application. A working group of young MSM and representatives from

community-based networks are involved in the questionnaire and survey design and will help in the selection of seeds. Although the inclusion of young key populations, most of whom do not have research skills, involves additional time, their input into the questionnaire has resulted in the prioritization of interview questions, the inclusion of adolescent appropriate language, and important local terms.

When using TLS and RDS to conduct research on AKP, it is important to consider how the methodology should be modified in order for it to be accepted by, and appropriate for, adolescents. When selecting time periods for sampling AKP, one TLS survey of adolescents who use drugs adjusted their data collection activities to avoid hours when activities such as school, work, or chores were most likely to occur. When selecting a data collection site, an RDS survey of street youth in Albania anticipated that waiting space would be needed for older siblings, parents, or other caretakers accompanying an adolescent participant or for younger siblings who were in the care of an adolescent participant [39]. An extra staff person was hired to engage younger siblings with toys, puzzles, games, and other activities as they waited for their adolescent caretaker to finish the survey process. In place of cash incentives, surveys of adolescents have used food items, soap, clothing, games, and other items [39]. Surveys of AKP conducted in Thailand and Myanmar used phone credit as incentives [43,44]. When asking AKP about social network sizes, some RDS surveys have used pictures and counting techniques to help adolescents think about size differentials. For instance, one survey used several different pictures of circles of increasing sizes to help adolescents visualize their most accurate social network size [39]—the smallest circle indicated the smallest social network size (ie, one or two), and the largest size indicated the largest network size (ie, up to 100). Tables 2 and 3 provide an overview of the features specific to each of the sampling methodologies, how each feature is generally implemented, and considerations on how they can be modified for AKP.

Table 2. Time location sampling (TLS) methodology features, description of those features, and considerations for applying those features when conducting surveys of adolescent key populations (AKP).

Feature	Current implementation of features	Considerations of features when sampling AKP ^a
Mapping	TLS ^b requires a complete mapping of all venues in a sampling location.	Venues mapped for adults may not be the same as those mapped for AKP. Involve AKP and groups working with AKP to help map venues. Enumerating AKP at venues may result in some errors since it may be difficult to accurately know someone's age without speaking to them. Consider using AKP as enumerators during mapping as they may be best able to recognize who are adolescents.
Sampling times and periods	TLS requires that sampling times be randomly selected.	Make sure to include sampling times that conform to AKP's availability.
Recruitment	TLS involves researchers approaching participants. Sometimes this involves intercepting participants in the street or at a venue and then escort them to another location for an interview and/or testing.	Involve youth as interviewers to minimize intimidation of AKP interviewees. If taking AKP to another area for interviewing or testing, allow for AKP participants to be joined by a friend or trusted person. Make sure the person leading the participant to another location does not appear threatening to either the participant or the public.

^aAKP: adolescent key populations.

^bTLS: time location sampling.

Table 3. Respondent driven sampling (RDS) methodology features, description of those features, and considerations for applying those features when conducting surveys of adolescent key populations (AKP).

Feature	Current implementation of features	Considerations of features when sampling AKP ^a
Pre-survey research	Conduct pre-survey research to assess populations' social networks, RDS ^b acceptability, and logistics.	Organizations working with AKP can help identify sites where AKP spend time, meet friends, find new sex partners, or buy or sell drugs. Sites may or may not be the same as those frequented by adult key populations.
Seeds	Seeds are the initial participants that start the recruitment from within the network of interest.	In surveys of adolescents and adults together, select an ample number of AKP seeds who are more likely to recruit other AKP. In surveys of AKP only, special considerations may be needed to diversify seeds. Some adult diversifications, such as marital status or occupation, may not be applicable to AKP. Consider other diversity factors such as whether they live at home or not, are sexually active or not, in school or not, able to read and write or not, etc.
Incentives	A nominal incentive, usually in the form of local currency, is usually provided to participants who complete the survey process and for recruiting their peers.	Make sure a country's research ethical standards allow incentives to be provided to adolescents. Incentives used for adults may not be appropriate for adolescents. Involve organizations working with AKP to select appropriate incentives. Ensure that incentives do not induce participation (eg, if it were not for the incentive provided, the participant would not enroll in a survey or would not withdraw from the survey early, given his or her better judgment).
Measuring social network size	For RDS analysis, the number of eligible people that each subject "knows" and has seen during a specified period of time (eg, 2 weeks) is needed. This question is open-ended and gaining accurate responses can be difficult, especially from someone uncomfortable with counting.	Use special probing techniques to assist with and encourage accurate reporting of social network sizes. Some adolescents may need help with coming up with a number.
Interview sites	RDS usually requires recruits to be present at an interview site to undergo the survey.	Interview sites should be easy to access, safe, and comfortable for AKP. In surveys of adolescents and adults together, consider special hours for AKP to enroll when adult participants are not present. Avoid interview sites located close to schools, homes, police stations, prisons, sex work solicitation areas, high crime areas, etc. Consider what to do with family members who bring AKP to participate in a survey. Consider the use of WebRDS [45-47], whereby peers recruit their peers via messaging and email.
Staffing	RDS surveys usually have numerous staff members, including someone to screen for eligibility, interviewers, pre- and post-test HIV ^c counselors, coupon managers, and so on.	Ensure staff members are trained to interact with adolescents and can recognize signs of distress and able to respond appropriately. Retain a social worker familiar with AKP at the survey site to respond to difficult situations. Ensure staff members understand the consent form and are trained to acquire informed consent from adolescents and/or their parent or guardian without coercing the adolescent. Have a staff member to engage adolescents with activities while they are waiting for an interview. Support staff who may be emotionally impacted by their exposure to adolescents who sell sex, use drugs, have experienced violence, are living with HIV, or have other risks and vulnerabilities.
Coupons	Coupons are used by participants to recruit peers. Coupons convey important information about the survey location, operation hours, etc.	Use simple coupon designs and words that are easily understandable to adolescents. Use pictures in place of words or other creative means to convey important coupon information. If using WebRDS [45-47], coupons can be sent through Web-based applications.

^aAKP: adolescent key populations.^bRDS: respondent driven sampling.^cHIV: human immunodeficiency virus.

Both TLS and RDS require additional considerations when collecting biological and sensitive behavioral data from AKP (Table 4). Pre-survey assessments, data collection forms, interviewing, biological specimen procedures techniques, and eligibility descriptions may need to be modified when sampling youth populations. Additional staff may be needed to address the specific needs of AKP, especially those who have been in abusive situations. For instance, in some RDS surveys of adolescents, trained social workers were hired to respond to any participant needing assistance, feeling distress as a result of answering sensitive questions, or expressing that they were in a harmful or abusive situation [23,39]. Once data are analyzed, important findings should be shared with the population and validated. For instance, in the BBSS RDS surveys conducted

in the Philippines, the surprising finding that a large percentage of adolescent female sex workers and males who have sex with males were attending school resulted in a validation process [48]. Qualitative research conducted by the Department of Health in schools indicated that even though HIV education was part of the curriculum, few students received adequate HIV information in school. Identified reasons included that teachers were not receiving formal HIV training; that textbooks had outdated HIV information; and that some teachers were not comfortable discussing HIV and skipped lessons about HIV. This exercise resulted in plans to strengthen HIV education, thereby, reaching school-enrolled AKP and building upon HIV BBSS findings.

Table 4. Special considerations and suggestions when adapting biological behavioral surveillance surveys (BBSS) among adolescent key populations (AKP) in time location sampling (TLS) and respondent driven sampling (RDS).

Topic	Considerations and suggestions
Pre-survey assessment	An AKP ^a specific pre-survey assessment may include (1) meetings with and involvement from parents, guardians, or community gatekeepers to ensure AKP participate in surveys, (2) community meetings (without disclosing the full nature of the research that could result in further stigma or reprisals) to garner community acceptance of collecting information from AKP, (3) meetings with government officials, child advocates, and NGOs ^b working with adolescents to be fully aware of country laws and guidelines governing the involvement of children in research, and (4) research to determine if adolescent-friendly communication technologies (ie, mobile phone apps and Internet websites) or eliciting questions through cell phones and computers can enhance sampling (Note: Web-based surveys do not allow for biological data collection).
Questionnaires and other data collection forms	Materials for adolescents of higher age ranges (eg, 15-19 years) need to be adapted for those of lower age ranges (eg, 10-14 years). Keep questionnaires and materials short and simple, and include adolescent-appropriate language. Forms to be read by participants should be easy to read using language commonly used by adolescents. Some AKP are not able or do not want to read something—have the option that materials can be read to them. Consider the use of drawings to convey important information.
Interviews	Interviews among AKP should be short. Allow respondents to take breaks during an interview.
HIV and other testing	Minimize the invasiveness of HIV testing ^c and other testing procedures. Consider using oral swabs or finger pricks instead of venous blood.
Eligibility	Labels used by researchers to describe AKP (ie, people who inject drugs, male or female sex workers) may not be recognized by adolescents engaging in the same behaviors. Refer to the behavior, for example, selling sex in exchange for money, rather than the population group (sex workers).
Support during the survey	Have available a trained social worker or child advocate (a person or service provider with qualifications to provide information or support).
Dissemination and validation	Once data are analyzed, go back to AKP to disseminate and validate data.

^aAKP: adolescent key populations.

^bNGO: non-governmental organization.

^cHIV: human immunodeficiency virus.

Discussion

This paper is a product of an international meeting convened to share lessons learned, improvements and innovations, and outputs on HIV surveillance activities. This was the third such meeting (2004 in Addis Ababa, Ethiopia; 2009 and 2015 in Bangkok, Thailand) but the first to include a focus on AKP, reflecting recent concerns about their absence in HIV BBSS [49]. Data from surveys conducted among AKP thus far have provided essential information to respond to the specific needs

of AKP, including overlapping risks (ie, adolescents who inject drugs and sell sex), sexual behaviors, barriers to service uptake, violence, as well as HIV- and non-HIV-related (ie, reproductive and mental) health needs [14,48,50].

We note that there are several ethical barriers to sampling adolescents; however, using the strategies proposed here, AKP under the age of 18 years can be involved in research and still be protected from undue harm. Countries enacting laws allowing adolescents to access HIV testing without parental consent may serve as leverage to argue the benefits of conducting research

among AKP without parental consent, as long as there are sufficient ethical protections and access to needed care and treatment.

During the past decade or more, HIV prevalence and associated risk behaviors among adult key populations worldwide have been successfully measured using innovative sampling methods including TLS and RDS. However, it is unclear why the lessons learned from surveys of adult key populations are not being harnessed to capture strategic information about AKP in existing BBSS or in surveys targeting AKP. In some cases, more time and expense may be needed to capture AKP. Given that survey designs need to be tailored to reach AKP, they comprise a small part of the key populations and may be more hidden than their adult counterparts. Many surveys of adult key populations include adolescents as young as 15 years [11,12]; however, disaggregated analysis that might inform donors and governments about 15-19 year olds, including disaggregation by sex, is rarely presented [14]. Other important information including the age of initiation of drug use or sexual behavior from surveys of adult key populations have been useful for advocating new policies to benefit AKP.

As in most surveys of adult key populations, research among AKP should include the involvement of the research population in the survey planning, implementation, analysis, and dissemination. Researchers should endeavor to provide avenues to involve adolescents, especially AKP, in all phases of research as such involvement will increase their ownership of the data collected and a higher likelihood that the findings are used on their behalf to develop more useful national HIV testing,

treatment, care, and support strategies. When designing research, adolescents may have insight into current behavior trends such as which drugs are being used and which venues are being frequented by AKP, and which tools and language will be most effective. In addition, older adolescents can be included as part of the data collection and analysis team. Future directions for sampling AKP could include developing novel data collection strategies such as using game interfaces with colors, sounds, and levels of earning points when responding to questionnaires in an effort to keep adolescents engaged. As more BBSS are conducted among AKP using TLS and RDS, it is hoped that additional lessons learned will be shared to ensure the most optimal implementation and adaptation of these methods to sample these underserved high risk populations.

Conclusions

Despite the fact that deaths due to AIDS continue to increase for adolescents, while they have decreased for all other age groups since 2010 [13], little data are available to ensure that adolescents are provided adequate prevention, intervention, care, and treatment services. Although there are some ethical barriers to conducting surveillance among AKP, with effort and collective action, many of these can be overcome. There are currently effective methods for collecting and using data from surveys of adult key populations that could easily be used in AKP. At a minimum, existing data of adolescents in adult BBSS should be disaggregated into adolescent age groupings. It is necessary to take the extra time and resources to effectively collect meaningful data that will provide needed strategies to address the health needs of AKP.

Conflicts of Interest

None declared.

References

1. Idele P, Gillespie A, Porth T, Suzuki C, Mahy M, Kasedde S, et al. Epidemiology of HIV and AIDS among adolescents: current status, inequities, and data gaps. *J Acquir Immune Defic Syndr* 2014 Jul 01;66(Suppl 2):S144-S153. [doi: [10.1097/QAI.000000000000176](https://doi.org/10.1097/QAI.000000000000176)] [Medline: [24918590](https://pubmed.ncbi.nlm.nih.gov/24918590/)]
2. UNICEF. UNICEF. 2015. Adolescents under the radar in the Asia-Pacific AIDS response URL: http://www.unicef.org/eapro/Adolescents_Under_the_Radar_final.pdf [accessed 2017-06-04] [WebCite Cache ID [6qxSjoe0v](https://www.webcitation.org/6qxSjoe0v)]
3. United Nations General Assembly. Protocol to prevent, suppress and punish trafficking in persons, especially women and children, supplementing the United Nations convention against transnational organized crime. In: Tully S, editor. *International Documents On Corporate Responsibility*. New York: Edward Elgar Publishing; 2000.
4. McClure C, Chandler C, Bissell S. Responses to HIV in sexually exploited children or adolescents who sell sex. *Lancet* 2015 Jan 10;385(9963):97-99. [doi: [10.1016/S0140-6736\(14\)60979-X](https://doi.org/10.1016/S0140-6736(14)60979-X)] [Medline: [25059951](https://pubmed.ncbi.nlm.nih.gov/25059951/)]
5. Interagency Working Group on Key Populations. UNAIDS. 2014. HIV and young people who sell sex: a technical brief URL: http://www.unaids.org/sites/default/files/media_asset/2015_young_people_who_sell_sex_en.pdf [accessed 2017-06-04] [WebCite Cache ID [6qxTynUY4](https://www.webcitation.org/6qxTynUY4)]
6. Busza J, Mtetwa S, Mapfumo R, Hanisch D, Wong-Gruenwald R, Cowan F. Underage and underserved: reaching young women who sell sex in Zimbabwe. *AIDS Care* 2016 Mar;28(Suppl 2):14-20 [FREE Full text] [doi: [10.1080/09540121.2016.1176673](https://doi.org/10.1080/09540121.2016.1176673)] [Medline: [27391994](https://pubmed.ncbi.nlm.nih.gov/27391994/)]
7. Traore IT, Meda N, Hema NM, Ouedraogo D, Some F, Some R, et al. HIV prevention and care services for female sex workers: efficacy of a targeted community-based intervention in Burkina Faso. *J Int AIDS Soc* 2015;18:20088 [FREE Full text] [Medline: [26374604](https://pubmed.ncbi.nlm.nih.gov/26374604/)]
8. Curth NK, Hansson LN, Storm F, Lazarus JV. Select barriers to harm-reduction services for IDUs in eastern Europe. *Cent Eur J Public Health* 2009 Dec;17(4):191-197 [FREE Full text] [Medline: [20377047](https://pubmed.ncbi.nlm.nih.gov/20377047/)]

9. Family Health International. WHO. Guidelines for repeated behavioral surveys in populations at risk of HIV URL: <http://www.who.int/hiv/strategic/pubss/en/> [accessed 2017-06-03] [WebCite Cache ID 6qxU2FmeN]
10. UNAIDS. 2011. Guidelines on surveillance among populations most at risk for HIV URL: http://www.unaids.org/sites/default/files/en/media/unaids/contentassets/documents/epidemiology/2011/20110518_Surveillance_among_most_at_risk.pdf [accessed 2017-06-04] [WebCite Cache ID 6qxU4IsnR]
11. Malekinejad M, Johnston LG, Kendall C, Kerr LR, Rifkin MR, Rutherford GW. Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: a systematic review. *AIDS Behav* 2008 Jul;12(4 Suppl):S105-S130. [doi: [10.1007/s10461-008-9421-1](https://doi.org/10.1007/s10461-008-9421-1)] [Medline: [18561018](https://pubmed.ncbi.nlm.nih.gov/18561018/)]
12. Montealegre JR, Johnston LG, Murrill C, Monterroso E. Respondent driven sampling for HIV biological and behavioral surveillance in Latin America and the Caribbean. *AIDS Behav* 2013 Sep;17(7):2313-2340. [doi: [10.1007/s10461-013-0466-4](https://doi.org/10.1007/s10461-013-0466-4)] [Medline: [23568227](https://pubmed.ncbi.nlm.nih.gov/23568227/)]
13. UNICEF. UNICEF. 2016. For every child, end AIDS: seventh stocktaking report, 2016 URL: https://www.unicef.org/publications/files/Children_and_AIDS_Seventh_Stocktaking_Report_2016_EN.pdf.pdf [accessed 2017-06-04] [WebCite Cache ID 6qxUIpIx]
14. Dayton R, Nary P, Cunningham J, Plourde KF, Green K, Wambugu S, et al. How can we better serve adolescent key populations? strategies to encourage and inform future data collection, analysis, and use. *Glob Public Health* 2016 Apr 24:1-13 Epub ahead of print. [doi: [10.1080/17441692.2016.1170177](https://doi.org/10.1080/17441692.2016.1170177)] [Medline: [27109137](https://pubmed.ncbi.nlm.nih.gov/27109137/)]
15. UNESCO, UNFPA. UNICEF. 2013. Young key populations at higher risk of HIV in Asia and the Pacific: making the case with strategic information URL: http://www.unicef.org/eapro/Young_key_populations_at_high_risk_of_HIV_in_Asia_Pacific.pdf [accessed 2017-06-03] [WebCite Cache ID 6qxULbRN]
16. Barrett D, Hunt N, Stoicescu C. HRI. 2013. Injecting drug use among under 18s: a snapshot of available data URL: https://www.hri.global/files/2014/08/06/injecting_among_under_18s_snapshot_WEB.pdf [accessed 2017-06-03] [WebCite Cache ID 6qxUQPuAn]
17. UNICEF. 2007. Implementation handbook for the convention on the rights of the child: fully revised, third edition URL: http://www.unicef.org/french/publications/index_43110.html [accessed 2017-06-03] [WebCite Cache ID 6qxUSHr21]
18. Powell MA, Taylor N, Fitzgerald R. UNICEF. 2013. Ethical research involving children URL: <https://www.unicef-irc.org/publications/706/> [accessed 2017-06-03] [WebCite Cache ID 6qxUhInTj]
19. DiClemente RJ, Ruiz MS, Sales JM. Barriers to adolescents' participation in HIV biomedical prevention research. *J Acquir Immune Defic Syndr* 2010 Jul;54(Suppl 1):S12-S17 [FREE Full text] [doi: [10.1097/QAI.0b013e3181e1e2c0](https://doi.org/10.1097/QAI.0b013e3181e1e2c0)] [Medline: [20571418](https://pubmed.ncbi.nlm.nih.gov/20571418/)]
20. Nelson RM, Lewis LL, Struble K, Wood SF. Ethical and regulatory considerations for the inclusion of adolescents in HIV biomedical prevention research. *J Acquir Immune Defic Syndr* 2010 Jul;54(Suppl 1):S18-S24. [doi: [10.1097/QAI.0b013e3181e2012e](https://doi.org/10.1097/QAI.0b013e3181e2012e)] [Medline: [20571419](https://pubmed.ncbi.nlm.nih.gov/20571419/)]
21. Fisher CB, Arbeit MR, Dumont MS, Macapagal K, Mustanski B. Self-consent for HIV prevention research involving sexual and gender minority youth: reducing barriers through evidence-based ethics. *J Empir Res Hum Res Ethics* 2016 Feb;11(1):3-14. [doi: [10.1177/1556264616633963](https://doi.org/10.1177/1556264616633963)] [Medline: [26956988](https://pubmed.ncbi.nlm.nih.gov/26956988/)]
22. Gilbert AL, Knopf AS, Fortenberry JD, Hosek SG, Kapogiannis BG, Zimet GD. Adolescent self-consent for biomedical human immunodeficiency virus prevention research. *J Adolesc Health* 2015 Jul;57(1):113-119 [FREE Full text] [doi: [10.1016/j.jadohealth.2015.03.017](https://doi.org/10.1016/j.jadohealth.2015.03.017)] [Medline: [26095412](https://pubmed.ncbi.nlm.nih.gov/26095412/)]
23. Samonte G. Capturing younger cohorts through surveillance systems: Philippines. 2012 Presented at: Methodologies for Obtaining Strategic Information on Young People at Higher Risk of HIV Exposure; September 3-5, 2012; Bangkok, Thailand.
24. Worku EB, Davis AM, Morrow B. A critical review of health research ethical guidelines regarding caregiver consent for HIV research involving minors in South Africa: Ethical and legal issues. *S Afr J BL* 2016 Oct 18;9(2):78. [doi: [10.7196/SAJBL.2016.V9I2.458](https://doi.org/10.7196/SAJBL.2016.V9I2.458)]
25. Ho WW, Brandfield J, Retkin R, Laraque D. Complexities in HIV consent in adolescents. *Clin Pediatr (Phila)* 2005;44(6):473-478. [Medline: [16015393](https://pubmed.ncbi.nlm.nih.gov/16015393/)]
26. UNESCO. 2013. Young people and the law in Asia and the Pacific: a review of laws and policies affecting young people's access to sexual and reproductive health and HIV services URL: <http://unesdoc.unesco.org/images/0022/002247/224782E.pdf> [accessed 2017-06-03] [WebCite Cache ID 6qxUz5zIk]
27. World Health Organization. Youngpeopleandhiv. 2013. HIV and adolescents: guidance for HIV testing and counselling and care for adolescents living with HIV? recommendations for a public health approach and considerations for policy-makers and managers URL: http://www.youngpeopleandhiv.org/files/HIV_Testing_guideline.pdf [accessed 2017-06-03] [WebCite Cache ID 6qxVIpnWB]
28. Patton GC, Coffey C, Cappa C, Currie D, Riley L, Gore F, et al. Health of the world's adolescents: a synthesis of internationally comparable data. *Lancet* 2012 Apr 28;379(9826):1665-1675. [doi: [10.1016/S0140-6736\(12\)60203-7](https://doi.org/10.1016/S0140-6736(12)60203-7)] [Medline: [22538181](https://pubmed.ncbi.nlm.nih.gov/22538181/)]
29. Johnston LG, Sabin K, Prybylski D. Update for sampling most-at-risk and hidden populations for HIV biological and behavioral surveillance. *JHASE J HIV/AIDS Surveill Epidemiol* 2010;2(1):2 [FREE Full text]

30. Sabin KM, Johnston LG. Epidemiological challenges to the assessment of HIV burdens among key populations: respondent-driven sampling, time-location sampling and demographic and health surveys. *Curr Opin HIV AIDS* 2014 Mar;9(2):101-106. [doi: [10.1097/COH.0000000000000046](https://doi.org/10.1097/COH.0000000000000046)] [Medline: [24464090](https://pubmed.ncbi.nlm.nih.gov/24464090/)]
31. Karon JM, Wejnert C. Statistical methods for the analysis of time-location sampling data. *J Urban Health* 2012 Jun;89(3):565-586 [FREE Full text] [doi: [10.1007/s11524-012-9676-8](https://doi.org/10.1007/s11524-012-9676-8)] [Medline: [22421885](https://pubmed.ncbi.nlm.nih.gov/22421885/)]
32. Globalhealthsciences. 2010. Resource guide: time location sampling URL: <http://globalhealthsciences.ucsf.edu/sites/default/files/content/pphg/surveillance/modules/global-trainings/tls-res-guide-2nd-edition.pdf> [accessed 2017-06-04] [WebCite Cache ID [6qxW9unTC](https://www.webcitation.org/6qxW9unTC)]
33. MacKellar D, Valleroy L, Karon J, Lemp G, Janssen R. The Young Men's Survey: methods for estimating HIV seroprevalence and risk factors among young men who have sex with men. *Public Health Rep* 1996;111(Suppl 1):138-144 [FREE Full text] [Medline: [8862170](https://pubmed.ncbi.nlm.nih.gov/8862170/)]
34. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. *Social Problems* 1997 May;44(2):174-199. [doi: [10.1525/sp.1997.44.2.03x0221m](https://doi.org/10.1525/sp.1997.44.2.03x0221m)]
35. Heckathorn DD. 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)]
36. Johnston LG. WHO. Geneva, Switzerland: World Health Organization; 2013. Introduction to respondent-driven sampling URL: http://applications.emro.who.int/dsaf/EMRPUB_2013_EN_1539.pdf [accessed 2017-06-04] [WebCite Cache ID [6qxWVYM22](https://www.webcitation.org/6qxWVYM22)]
37. 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/)]
38. Decker MR, Marshall BD, Emerson M, Kalamar A, Covarrubias L, Astone N, et al. Respondent-driven sampling for an adolescent health study in vulnerable urban settings: a multi-country study. *J Adolesc Health* 2014 Dec;55(6 Suppl):S6-S12 [FREE Full text] [doi: [10.1016/j.jadohealth.2014.07.021](https://doi.org/10.1016/j.jadohealth.2014.07.021)] [Medline: [25454005](https://pubmed.ncbi.nlm.nih.gov/25454005/)]
39. Johnston LG, Thurman TR, Mock N, Nano L, Carcani V. Respondent-driven sampling: a new method for studying street children with findings from Albania. *Vulnerable Child Youth Stud* 2010 Apr 07;5(1):1-11. [doi: [10.1080/17450120903193923](https://doi.org/10.1080/17450120903193923)]
40. Chemnasiri T, Netwong T, Visarutratana S, Varangrat A, Li A, Phanuphak P, et al. Inconsistent condom use among young men who have sex with men, male sex workers, and transgenders in Thailand. *AIDS Educ Prev* 2010 Apr;22(2):100-109. [doi: [10.1521/aeap.2010.22.2.100](https://doi.org/10.1521/aeap.2010.22.2.100)] [Medline: [20387981](https://pubmed.ncbi.nlm.nih.gov/20387981/)]
41. Yi S, Tuot S, Yung K, Kim S, Chhea C, Saphonn V. Factors associated with risky sexual behavior among unmarried most-at-risk young people in Cambodia. *AJPHR* 2014 Jan 7;2(5):211-220. [doi: [10.12691/ajphr-2-5-5](https://doi.org/10.12691/ajphr-2-5-5)]
42. Kissin D. Slideserve. 2011. Using time-location sampling for HIV surveillance in street youth powerpoint presentation URL: <http://www.slideserve.com/Patman/using-time-location-sampling-for-hiv-surveillance-in-street-youth> [accessed 2017-06-03] [WebCite Cache ID [6qxWtIjky](https://www.webcitation.org/6qxWtIjky)]
43. UNICEF. UNICEF. 2014. Situational analysis of youth people at high risk for HIV exposure URL: http://www.unicef.org/thailand/UNICEF_study_on_HIV_infection_among_young_people_FINAL_ENGLISH.pdf [accessed 2017-06-03] [WebCite Cache ID [6qxX2COaB](https://www.webcitation.org/6qxX2COaB)]
44. Johnston LG, Mon MM, Steinhaus M, Sass J. Correlates of forced sex among young men who have sex with men in Yangon and Monywa, Myanmar. *Arch Sex Behav* 2016;46(4):1001-1010. [doi: [10.1007/s10508-016-0761-z](https://doi.org/10.1007/s10508-016-0761-z)] [Medline: [27278964](https://pubmed.ncbi.nlm.nih.gov/27278964/)]
45. Bauermeister JA, Zimmerman MA, Johns MM, Glowacki P, Stoddard S, Volz E. Innovative recruitment using online networks: lessons learned from an online study of alcohol and other drug use utilizing a web-based, respondent-driven sampling (webRDS) strategy. *J Stud Alcohol Drugs* 2012 Sep;73(5):834-838 [FREE Full text] [Medline: [22846248](https://pubmed.ncbi.nlm.nih.gov/22846248/)]
46. Bengtsson L, Lu X, Nguyen QC, Camitz M, Hoang NL, Nguyen TA, et al. Implementation of web-based respondent-driven sampling among men who have sex with men in Vietnam. *PLoS One* 2012 Nov;7(11):e49417 [FREE Full text] [doi: [10.1371/journal.pone.0049417](https://doi.org/10.1371/journal.pone.0049417)] [Medline: [23152902](https://pubmed.ncbi.nlm.nih.gov/23152902/)]
47. Wejnert C, Heckathorn D. Web-based network sampling. *Socio Meth Res* 2008 Aug;37(1):105-134. [doi: [10.1177/0049124108318333](https://doi.org/10.1177/0049124108318333)]
48. Mirano J. UNAIDS. 2016. Enhancing interventions for young key affected populations in selected cities in the Philippines - ALL in to #End Adolescent AIDS URL: http://www.unaids.org/sites/default/files/media_asset/20150217_ALL_IN_brochure.pdf [accessed 2017-06-09] [WebCite Cache ID [6r5cVNUWv](https://www.webcitation.org/6r5cVNUWv)]
49. Lall P, Lim SH, Khairuddin N, Kamarulzaman A. Review: an urgent need for research on factors impacting adherence to and retention in care among HIV-positive youth and adolescents from key populations. *J Int AIDS Soc* 2015;18(2 Suppl 1):19393 [FREE Full text] [Medline: [25724503](https://pubmed.ncbi.nlm.nih.gov/25724503/)]
50. Delany-Moretlwe S, Cowan FM, Busza J, Bolton-Moore C, Kelley K, Fairlie L. Providing comprehensive health services for young key populations: needs, barriers and gaps. *J Int AIDS Soc* 2015;18(2 Suppl 1):19833 [FREE Full text] [Medline: [25724511](https://pubmed.ncbi.nlm.nih.gov/25724511/)]

Abbreviations

AIDS: acquired immune deficiency syndrome

AKP: adolescent key populations
BBSS: biological behavioral surveillance surveys
HIV: human immunodeficiency virus
MSM: males who have sex with males
RDS: respondent driven sampling
TLS: time location sampling

Edited by P Mee; submitted 07.02.17; peer-reviewed by B Hensen, J Busza; comments to author 09.03.17; revised version received 30.03.17; accepted 27.04.17; published 20.06.17.

Please cite as:

Johnston LG, Sass J, Acaba J, Cheng WS, Mark Prabhu S

Ensuring Inclusion of Adolescent Key Populations at Higher Risk of HIV Exposure: Recommendations for Conducting Biological Behavioral Surveillance Surveys

JMIR Public Health Surveill 2017;3(2):e40

URL: <http://publichealth.jmir.org/2017/2/e40/>

doi: [10.2196/publichealth.7459](https://doi.org/10.2196/publichealth.7459)

PMID: [28634152](https://pubmed.ncbi.nlm.nih.gov/28634152/)

©Lisa Grazina Johnston, Justine Sass, Jeffrey Acaba, Wing-Sie Cheng, Shirley Mark Prabhu. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 20.06.2017. 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

HIV Strategic Information in Non-European Union Countries in the World Health Organization European Region: Capacity Development Needs

Ivana Bozicevic^{1*}, MSc, DrPH, MD; Senad Handanagic^{1*}, MSc, MD; Jurja Ivana Cakalo^{1*}, MD; Annemarie Rinder Stengaard^{2*}, MSc; George Rutherford^{3*}, MD

¹WHO Collaborating Centre for HIV Strategic Information, University of Zagreb School of Medicine, Zagreb, Croatia

²Joint Tuberculosis, HIV/AIDS and Hepatitis Programme, Division of Communicable Diseases, Health Security and Environment, WHO Regional Office for Europe, Copenhagen, Denmark

³Global Health Sciences, University of California, San Francisco, San Francisco, CA, United States

* all authors contributed equally

Corresponding Author:

Ivana Bozicevic, MSc, DrPH, MD
WHO Collaborating Centre for HIV Strategic Information
University of Zagreb School of Medicine
Rockefeller str. 4
Zagreb, 10000
Croatia
Phone: 385 992984138
Fax: 385 14590 100
Email: Ivana.Bozicevic@lshtm.ac.uk

Abstract

Background: Capacity building of the national HIV strategic information system is a core component of the response to the HIV epidemic as it enables understanding of the evolving nature of the epidemic, which is critical for program planning and identification of the gaps and deficiencies in HIV programs.

Objective: The study aims to describe the results of the assessment of the needs for further development of capacities in HIV strategic information systems in the non-European Union (EU) countries in the World Health Organization European Region (EUR).

Methods: Self-administered questionnaires were distributed to national AIDS programs. The first questionnaire was sent to all countries (N=18) to find out, among other issues, the priority level for strengthening a range of HIV surveillance areas and their key gaps and weaknesses. The second questionnaire was sent to 15 countries to more specifically determine capacities for the analysis of the HIV care cascade.

Results: Responses to the first questionnaire were received from 10 countries, whereas 13 countries responded to the second questionnaire. Areas that were most frequently marked as being of high to moderate priority for strengthening were national electronic patient monitoring systems, evaluation of HIV interventions and impact analysis, implementation science, and data analysis. Key weaknesses were lack of electronic reporting of HIV cases, problems with timeliness and completeness of reporting in HIV cases, under-estimates of the reported number of HIV-related deaths, and limited CD4 count testing at the time of HIV diagnosis. Migrant populations, internally displaced persons, and refugees were most commonly mentioned as groups not covered by surveillance, followed by clients of sex workers and men who have sex with men. The majority of countries reported that they were able to provide the number of people diagnosed with HIV who know their HIV status, which is important for the analysis of cross-sectional and longitudinal HIV care cascades. Ability to report on some of the key impact indicators of HIV programs—viral load suppression and mortality—should be considerably strengthened.

Conclusions: The assessment found a substantial need to invest in surveillance capacities, which is a cornerstone in the development of an evidence-informed response to HIV epidemics.

(*JMIR Public Health Surveill* 2017;3(2):e41) doi:[10.2196/publichealth.7357](https://doi.org/10.2196/publichealth.7357)

KEYWORDS

HIV; surveillance; evaluation; Europe; aptitude

Introduction**Recent Developments in HIV Surveillance**

Capacity development is a multidimensional concept with a variety of meanings. Goodman describes capacity as “the ability to carry out stated objectives” [1]. Similarly, Brown and colleagues define capacity development as a process that improves the ability of a person, group, organization, or system to meet its objectives and perform better [2]. The purpose of capacity building in HIV surveillance and monitoring and evaluation (M&E) is to improve the performance of the national HIV surveillance and M&E systems, which primarily means enhancing the ability to produce quality and timely data on the HIV epidemic and the HIV response [3]. Strengthening HIV information systems is the first strategic direction of the World Health Organization (WHO) Global Health Sector Strategy for HIV for 2016-2020 [4]. Using epidemiological and clinical data from these systems, countries should be able to construct cascades of HIV diagnosis, care, and treatment (HIV care cascades) and monitor the processes of reaching the 2020 Joint United Nations Programme on HIV/AIDS (UNAIDS) objectives of 90% of people living with HIV (PLHIV) knowing their HIV status, 90% of PLHIV on antiretroviral treatment (ART), and 90% of people on ART being virally suppressed [5]. The HIV care cascade is a framework for identifying and quantifying the magnitude of the gaps along the continuum of HIV diagnosis, care, and treatment [6,7]. The cascade analysis enables us to identify losses in the continuum of HIV services so that program implementers at facility, regional, or national levels can target resources and interventions more effectively, improve engagement in care for HIV positive individuals, and ultimately prevent new infections. Several data sources and indicators are needed to construct HIV care cascades, including estimates of the number of PLHIV, the number of PLHIV who are diagnosed (obtained from HIV case-based surveillance or surveys), the number of PLHIV who are receiving ART, and the number of

PLHIV who are virally suppressed (obtained from HIV patient monitoring systems). The HIV care cascade can be made at the national level but also at sub-national and facility levels and, whenever possible, it should be disaggregated by sex, age groups, and key populations. This enables us to monitor differences in access to services and treatment outcomes across geographical areas and population sub-groups and to improve the coverage and quality of services by providing targeted interventions [8]. The value of the cascade analysis lies in its use as a data and programmatic quality improvement tool. However, the ability to interpret the results of the cascade analysis depends heavily on the availability and the quality of data sources, both of which vary across European countries [9].

Overview of HIV Epidemics in Eastern Europe and Central Asia

There are several important reasons for the necessity of developing HIV surveillance and monitoring systems in Eastern Europe and Central Asia. The most recent report of the European

Centre for Disease Prevention and Control states that the HIV epidemic is still a major concern in Europe, particularly in the eastern part of the WHO European Region [10]. In 2015, 153,407 people were newly diagnosed with HIV in 50 of the 53 countries (no data were available from Bosnia and Herzegovina, Turkmenistan, and Uzbekistan) of the WHO European Region (EUR), which is the highest recorded number of newly diagnosed infections in one year since the start of reporting in the 1980s [10]. According to the UNAIDS estimates, between 2010 and 2015 the number of new HIV infections in Eastern Europe and Central Asia rose by 57%, totaling 190,000 (170,000-200,000) new infections, which is the highest increase globally [11]. In the same period, the number of AIDS-related deaths increased by 22%. HIV treatment coverage among the estimated 1.5 million (1.4 million-1.7 million) PLHIV in Eastern Europe and Central Asia in 2015 was only 21% (20-23%), the second lowest globally after the Middle East and North Africa [11].

National surveillance and M&E system capacity is a core component of the response to the HIV epidemic as it enables understanding of the evolving nature of the epidemic and the gaps and deficiencies in HIV programs, and it is critical for informing more strategic investments. Capacity development can broadly be classified as pre-service, which is training people before they enter the work force, and in-service, which is training people already in the workforce. In the field of HIV strategic information, pre-service capacity building can take the form of formal courses and tracks in established universities, such as have been developed in M&E at the Mzumbe University in Tanzania, Jimma University in Ethiopia, and the National School of Public Health in Brazil [12-14]. Post-doctoral full-time training programs have also provided substantial practical training opportunities in public health epidemiology through the Centers for Disease Control and Prevention’s Epidemic Intelligence Service and Field Epidemiology and Laboratory Training Programs and more academically focused fellowships, such as those provided through the Fogarty International Center and other institutes of the United States National Institutes of Health [15-18].

Since 2004, the WHO Collaborating Centre for HIV Strategic Information based at the Andrija Stampar School of Public Health at the University of Zagreb in Croatia has been providing in-service capacity-building activities in HIV surveillance and M&E of HIV programs, primarily via training workshops and technical assistance [19]. Twenty-eight training courses in HIV surveillance and M&E of national HIV programs and HIV interventions have been developed with partner institutions. The capacity development activities of the Collaborating Centre target diverse groups that work with and contribute to HIV surveillance, including public health professionals, health care providers, epidemiologists, clinicians, non-governmental organization (NGO) staff, and networks of PLHIV.

The aim of this paper is to describe results of an assessment of capacity development needs in HIV surveillance and M&E in the non-European Union (EU) countries in the WHO EUR.

Methods

Assessment of Capacities That Need to Be Developed in HIV Surveillance

To assess the needs for capacity development in HIV surveillance and strategic information systems, we distributed two self-administered questionnaires over email. The first questionnaire was sent to the directors of the national HIV programs of all (N=18) non-EU countries in the WHO EUR. This questionnaire was sent in May 2015 to assess the capacities that needed to be developed in HIV surveillance and addressed the following areas: (1) the types of training courses in HIV surveillance and M&E of HIV programs that national entities organized in 2012, 2013, or 2014 for staff working at the national level; (2) the priority level for strengthening of a range of HIV surveillance and M&E-related areas specified in the questionnaire; (3) the gaps and weaknesses in the areas that respondents marked as being of high or moderate priority for further development; (4) the population groups that might play an essential role in HIV transmission and that have not been covered with surveillance; (5) anticipated impact of declines in funding from the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM) on any of the HIV surveillance and M&E activities that have been developed through the support of GFATM.

Assessment of Availability of Data to Construct HIV Care Cascades

The second, separate questionnaire was sent after the training course in HIV care cascade analysis held in collaboration with WHO in June 2015. This questionnaire was sent to course participants from 15 countries of Eastern Europe and Central Asia (Armenia, Azerbaijan, Belarus, Croatia, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Moldova, Russian Federation, Serbia, Slovenia, Tajikistan, Turkey, and Ukraine) who were invited to the training course by WHO EURO and all worked at the national HIV programs as surveillance professionals. This questionnaire aimed to assess the availability of data to construct cross-sectional and longitudinal HIV care cascade analysis and identify the strengths and weaknesses of cascade indicators [20]. The questionnaire also assessed whether a cross-sectional HIV care cascade can be constructed for key populations, pregnant women, and geographical areas.

Results

Capacity Development Needs in HIV Surveillance

Responses to the first questionnaire were received from 10 of 18 countries (Albania, Azerbaijan, FYR Macedonia, Kyrgyzstan, Moldova, Montenegro, Serbia, Tajikistan, Turkey, and Ukraine). In 5 countries out of 10, a national entity had organized training courses in HIV surveillance in 2012, 2013, and 2014. Courses organized were on M&E of HIV programs (mentioned by 4 countries), followed by integrated bio-behavioral surveys (3

countries), key population size estimations (3 countries), using a national HIV database (2 countries), data analysis (2 countries), sentinel surveillance (1 country) and estimation and projection of the HIV epidemic (1 country).

HIV surveillance and M&E-related areas that were most frequently marked as being of high to moderate priority for strengthening were national electronic patient monitoring systems (8 countries), evaluation of HIV interventions and impact analysis (8 countries), HIV care cascade analysis (6 countries), implementation science (5 countries), and data analysis (5 countries). HIV case reporting (7 countries) and developing national HIV surveillance reports (7 countries) were of the lowest priority as these are already developed surveillance components. Among activities of low priority but not yet developed, one country mentioned HIV drug resistance surveillance and another estimation of the number of PLHIV and HIV incidence estimates. Respondents outlined a number of gaps and weaknesses that they felt should be addressed. In relation to HIV case reporting and mortality reporting these were under-estimates of the reported number of HIV-related deaths (7 countries), problems with timeliness and completeness of reporting of HIV cases (7 countries), limited CD4 count testing at the time of HIV diagnosis (4 countries), and lack of electronic reporting of HIV cases (3 countries). Other weaknesses respondents mentioned were lack of HIV drug resistance surveillance (4 countries); lack of capacity in data analysis, interpretation, and use (3 countries); lack of human resources to conduct HIV surveillance in key populations (3 countries); and lack of HIV incidence surveillance (2 countries). Single countries also responded that there was a lack of comprehensive evaluation and impact analysis, lack of capacity for fundraising, limited capacity to implement integrated bio-behavioral surveys and key population size estimates, challenges with estimating the number of PLHIV, and the heavy dependence of HIV surveillance activities on donor funding. Migrant populations, internally displaced persons and refugees and men who have sex with men (MSM) were most commonly mentioned as groups not covered by surveillance (5 countries), followed by clients of sex workers (4 countries). Seven of 8 countries responded that there will be an impact of declines in funding from GFATM on surveillance, and M&E activities and surveillance in key population reportedly have the lowest sustainability.

Progress in the HIV Continuum of Care Analysis

Responses to the second questionnaire were received from 13 out of 15 countries (all except Serbia and Slovenia). Tables 1 and 2 show the availability of HIV strategic information, which is needed to construct cross-sectional HIV care cascades and the weaknesses of the data, respectively. In terms of the availability of data to construct a cross-sectional HIV care cascade, 3 countries reported that they were unable to estimate the number of PLHIV, and 4 countries pointed out the need for this indicator but did not provide an explanation on whether they were able to obtain it. The most common weakness in obtaining the number of PLHIV was a lack of trust in modeling outputs and lack of high quality input data necessary for modeling tools to provide reliable results.

Table 1. Availability of HIV (human immunodeficiency virus) strategic information needed to construct cross-sectional HIV care cascades (n=13).

Type of indicator	Number of countries where available (n)	No response (n)
Estimated number of people living with HIV (human immunodeficiency virus) ^a	5	1
Number of people diagnosed with HIV who know their status	10	3
Number of PLHIV ^b who received HIV care in the past 12 months, including ART ^c	9	4
Number of PLHIV currently receiving ART	9	4
Number of PLHIV currently receiving ART who have a suppressed viral load (<1000 copies/mL)	9	4

^aFour countries reported that this indicator is needed but did not specify whether they were able to obtain it, and three reported that the indicator is not available.

^bPLHIV: people living with HIV.

^cART: antiretroviral treatment.

Table 2. Most frequently reported weaknesses of data and data sources used to construct cross-sectional HIV (human immunodeficiency virus) care cascades (n=13)

Indicator	Weakness	Number of countries
Estimated number of people living with HIV (human immunodeficiency virus)	A lack of good quality input data is necessary for modeling, which leads to low quality modeling outputs	3
	Used Spectrum ^c modeling tool but would like to use an additional tool (lack of trust in Spectrum estimates)	2
	Lack of trust in the modeling estimates	2
Number of people diagnosed with HIV who know their status	Poor mortality statistics: mortality statistics are not linked with HIV case reporting system	6
Number of PLHIV ^a who received HIV care in the past 12 months, including ART ^b	There is a lack of linkage of patient monitoring system with the HIV case reporting system	2
	It is challenging to ensure the quality of data at the local level	2
Number of PLHIV currently receiving ART	Lack of disaggregation by ART regimens and key populations	2
Number of PLHIV currently receiving ART who have a suppressed viral load (<1000 copies/mL)	Limited access to viral load diagnostics	4

^aPLHIV: people living with HIV.

^bART: antiretroviral treatment.

^cSpectrum: analytical tool used for mathematical modeling of HIV epidemics used by UNAIDS

The majority of the countries (n=10) have data on the number of people diagnosed with HIV who know their HIV status. The most common reported weakness of these data was poor quality mortality statistics, which makes it difficult to assess the cumulative number of people living with diagnosed HIV. In addition, multiple registrations of newly diagnosed HIV cases was mentioned by one country and under-reporting of newly diagnosed cases (due to not obtaining HIV positive test result when people tested for HIV) by another country.

Nine countries reported that they were able to provide the number of PLHIV who received HIV care in the past 12 months, including ART. Alongside the shortcomings shown in Table 2, an inability to record ART interruptions in the patient monitoring

system was reported by one country and a lack of clear criteria for reporting on patients enrolled in HIV care but not on ART by another country.

Nine countries reported that they were able to provide the number of PLHIV who are currently on ART, and the availability of a centralized HIV care and treatment database was mentioned by 3 countries as one of the major advantages in obtaining this indicator. The most commonly mentioned deficiency was limited disaggregation of data by ART regimens and key populations.

Nine countries reported that they were able to report on the number of PLHIV currently receiving ART who have a suppressed viral load (<1000 copies/mL). As a weakness, 3

countries indicated limited access to viral load diagnostics, while other challenges reported included a non-standardized definition of viral suppression at a national level (1 country) and a lack of an information system for collecting data on HIV viral load (1 country).

With regard to disaggregation of the HIV care cascade, only 3 countries out of 13 reported that they were able to construct a cascade for each of the key populations: people who inject drugs (PWID), MSM, and female sex workers (FSW). Eleven countries reported that the cascade can be created for PWID, 10 for MSM, and 3 for FSW. One country reported that the disaggregation was not possible, and three reported that geographical disaggregation into sub-national units was not possible. However, disaggregation for pregnant women was possible in 9 countries.

Twelve countries reported the availability of a longitudinal HIV patient monitoring system. Six countries reported that they could calculate the percentage of people diagnosed with HIV who were enrolled in HIV care within 12 months of HIV diagnosis. The same 6 countries reported that they were able to calculate the proportion of people diagnosed with HIV and on ART who were retained on ART for at least 12 months and have a suppressed viral load. They were also able to report the percentage of HIV-exposed infants who received a virological test within two months of birth and the proportion of HIV-exposed infants who were uninfected after cessation of breastfeeding.

Discussion

Principal Findings

We identified a diverse set of needs for capacity development in HIV strategic information among the countries that participated in the assessment. The greatest development needs were reported in the areas of national electronic patient monitoring systems, evaluation of HIV interventions and impact analysis, implementation science, HIV care cascade analysis, and data analysis. Of note is that, in the context of declines in funding for HIV surveillance, M&E activities and surveillance in key populations will have, as reported by the national HIV program staff, the lowest sustainability.

The countries reported a complex list of capacities, and six appeared to have very functional data systems able to report most aspects of the HIV treatment cascade. There was, however, substantially less ability to disaggregate data geographically or by key population, a capacity of great importance in a region where the most HIV epidemics are concentrated.

A common weakness was poor quality of mortality statistics, which limits the ability to assess how many PLHIV are

diagnosed. In addition, although the majority of surveyed countries have longitudinal HIV patient monitoring systems, only one-half could report on the proportion of patients who have a suppressed viral load 12 months after ART initiation. Understanding longer-term outcomes of the HIV continuum of care such as retention of treatment and viral suppression should be enabled through investment in infrastructure and human capacities with the aim to support national and global efforts to monitor progress towards the 90-90-90 targets [4].

This assessment points to the need to further invest in capacities in order to develop HIV strategic information systems that will provide necessary data for evidence-based decision making and for evaluating the access, coverage, and quality of interventions. A recent review identified several factors that drove the success in HIV treatment monitoring globally in the period from 2000-2015 [21]. These were commitment to invest in country data systems, aiming for 5-10% of program funds to be used to strengthen M&E, creating more demand for the data by conducting regular country program and epidemiological reviews, which were subsequently used as a basis for seeking funding from the donor agencies, and setting program targets [21].

Ability to report on some of the key impact indicators of HIV programs in Eastern Europe and Central Asia—viral load suppression and mortality—should be substantially strengthened, considering that the rates of viral load suppression among PLHIV are among the lowest globally [9,11]. Having in place uniform reporting practices and providing standardized guidelines and evaluation tools to surveillance programs were identified as policies that can increase effectiveness of surveillance, along with maintaining relationships with providers and laboratories to ensure complete and timely reporting [22]. Ongoing support from stakeholders to sites that provide HIV care and treatment and on-site training was also found to be important, particularly in settings where HIV treatment is being scaled-up and decentralized and patient monitoring systems established [23]. Timely provision of periodic reports from treatment sites also assisted in informing program management on the quality of patient care and enabled actions to be taken to improve treatment outcomes [23,24].

Limitations

Our data are subject to limitations. A few countries we attempted to survey did not participate, and we did not conduct site visits to verify reported capacity. Results of the assessment are applicable to the countries that responded to the questionnaires. Additionally, responses to some questions were missing. Nonetheless, we believe that our data will be helpful in planning the next round of strategic information capacity building in the region.

Acknowledgments

We acknowledge with gratitude all those who responded to the questionnaire from the National HIV Programs of the participating countries.

Conflicts of Interest

None declared.

References

1. Goodman RM, Speers MA, McLeroy K, Fawcett S, Kegler M, Parker E, et al. Identifying and defining the dimensions of community capacity to provide a basis for measurement. *Health Educ Behav* 1998 Jun;25(3):258-278. [doi: [10.1177/109019819802500303](https://doi.org/10.1177/109019819802500303)] [Medline: [9615238](https://pubmed.ncbi.nlm.nih.gov/9615238/)]
2. Brown L, LaFond A, Macintyre C. USAID. Chapel Hill, North Carolina: MEASURE Evaluation; 2001. Measuring capacity building URL: http://pdf.usaid.gov/pdf_docs/Pnacm119.pdf [accessed 2017-06-07] [WebCite Cache ID 6r3NuXjJJ]
3. UNAIDS. Geneva: United Nations Joint Programme on HIV/AIDS. 2010. Guidance on capacity building for HIV monitoring and evaluation URL: http://www.unaids.org/sites/default/files/sub_landing/files/5_4_MERG_Guidance_HIV_ME_Capacity_Buidling.pdf [WebCite Cache ID 6r3OZXAPV]
4. WHO. Geneva: World Health Organization; 2016. Global Health Sector Strategy for HIV for 2016-2020. Towards Ending AIDS URL: <http://apps.who.int/iris/bitstream/10665/246178/1/WHO-HIV-2016.05-eng.pdf?ua=1> [accessed 2017-06-07] [WebCite Cache ID 6r3OjAGt]
5. UNAIDS. Geneva: Joint United Nations Programme on HIV/AIDS; 2014. 90-90-90: an ambitious treatment target to help end the AIDS epidemic URL: http://www.unaids.org/sites/default/files/media_asset/90-90-90_en_0.pdf [accessed 2017-06-07] [WebCite Cache ID 6r3OteCqI]
6. WHO. Cairo: Regional Office for the Eastern Mediterranean; 2014. HIV test-treat-retain cascade analysis: guide and tools URL: http://applications.emro.who.int/dsaf/EMROPUB_2014_EN_1639.pdf [accessed 2017-06-07] [WebCite Cache ID 6r3PILAKa]
7. Haber N, Pillay D, Porter K, Bärnighausen T. Constructing the cascade of HIV care: methods for measurement. *Curr Opin HIV AIDS* 2016 Jan;11(1):102-108. [doi: [10.1097/COH.0000000000000212](https://doi.org/10.1097/COH.0000000000000212)] [Medline: [26545266](https://pubmed.ncbi.nlm.nih.gov/26545266/)]
8. Medland NA, McMahan JH, Chow EP, Elliott JH, Hoy JF, Fairley CK. The HIV care cascade: a systematic review of data sources, methodology and comparability. *J Int AIDS Soc* 2015;18:20634 [FREE Full text] [Medline: [26626715](https://pubmed.ncbi.nlm.nih.gov/26626715/)]
9. European Centre for Disease Prevention and Control. Stockholm: ECDC; 2015. Thematic report: HIV continuum of care. Monitoring implementation of the Dublin Declaration on Partnership to Fight HIV/AIDS in Europe and Central Asia: 2014 progress report URL: <http://ecdc.europa.eu/en/publications/Publications/dublin-declaration-continuum-of-care-2014.pdf> [accessed 2017-06-07] [WebCite Cache ID 6r3Q3C8sb]
10. European Centre for Disease Prevention and Control/WHO Regional Office for Europe. WHO. Stockholm: ECDC; 2016. HIV/AIDS surveillance in Europe 2015 URL: http://www.euro.who.int/_data/assets/pdf_file/0019/324370/HIV-AIDS-surveillance-Europe-2015.pdf [accessed 2017-06-07] [WebCite Cache ID 6r3QKlx5]
11. UNAIDS. Geneva: UNAIDS; 2016. Global HIV statistics. Fact sheet: latest statistics on the status of the AIDS epidemic URL: <http://www.unaids.org/en/resources/fact-sheet> [accessed 2017-06-07] [WebCite Cache ID 6r3QcwFb7]
12. Mzumbe University. Centre of excellence in health monitoring and evaluation URL: <http://coehme.mzumbe.ac.tz/msc-in-hme/> [accessed 2017-06-07] [WebCite Cache ID 6r3TOfl3L]
13. Jimma University. Health monitoring and evaluation department URL: <https://www.ju.edu.et/cphms/node/69?q=node/86> [accessed 2017-06-07] [WebCite Cache ID 6r3R6jtCT]
14. Fiocruz. Mestrado de Avaliação em Saúde, com novo cronograma, recebe inscrições URL: <http://portal.fiocruz.br/pt-br/content/mestrado-de-avaliacao-em-saude-com-novo-cronograma-recebe-inscricoes> [accessed 2017-06-07] [WebCite Cache ID 6r3RINMuV]
15. Thacker SB, Stroup DF, Sencer DJ. Epidemic assistance by the Centers for Disease Control and Prevention: role of the epidemic intelligence service, 1946-2005. *Am J Epidemiol* 2011 Dec 01;174(11 Suppl):S4-15. [doi: [10.1093/aje/kwr307](https://doi.org/10.1093/aje/kwr307)] [Medline: [22135393](https://pubmed.ncbi.nlm.nih.gov/22135393/)]
16. Rush T. Disease surveillance system evaluation as a model for improved integration and standardization of the laboratory component in the Field Epidemiology and Laboratory Training Program (FELTP) curriculum worldwide. *J Public Health Policy* 2012 Nov;33(4):390-400. [doi: [10.1057/jphp.2012.35](https://doi.org/10.1057/jphp.2012.35)] [Medline: [22971950](https://pubmed.ncbi.nlm.nih.gov/22971950/)]
17. Brickley DB, Mandel JS. International traineeships in AIDS prevention studies: 25 years of collaborative research. *AIDS Behav* 2014 Jan;18(Suppl 1):S1-S4 [FREE Full text] [doi: [10.1007/s10461-013-0658-y](https://doi.org/10.1007/s10461-013-0658-y)] [Medline: [24327184](https://pubmed.ncbi.nlm.nih.gov/24327184/)]
18. Heimbürger DC, Carothers CL, Blevins M, Warner TL, Vermund SH. Impact of global health research training on career trajectories: the Fogarty international clinical research scholars and fellows program. *Am J Trop Med Hyg* 2015 Sep;93(3):655-661 [FREE Full text] [doi: [10.4269/ajtmh.14-0705](https://doi.org/10.4269/ajtmh.14-0705)] [Medline: [26195466](https://pubmed.ncbi.nlm.nih.gov/26195466/)]
19. Whohub-zagreb. WHO collaborating centre for HIV strategic information, Zagreb, Croatia URL: <http://www.whohub-zagreb.org/> [accessed 2017-06-07] [WebCite Cache ID 6r3RsD8z7]
20. WHO. Geneva: World Health Organization; 2015. Consolidated strategic information guidelines for HIV in the health sector URL: http://apps.who.int/iris/bitstream/10665/164716/1/9789241508759_eng.pdf?ua=1&ua=1 [accessed 2017-06-07]

21. Low-Beer D, Beusenberg M, Hayashi C, Calleja T, Marsh K, Mamahit A, et al. Monitoring HIV treatment and the health sector cascade: from treatment numbers to impact. *AIDS Behav* 2017 Apr 11:- Epub ahead of print. [doi: [10.1007/s10461-017-1754-1](https://doi.org/10.1007/s10461-017-1754-1)] [Medline: [28401415](https://pubmed.ncbi.nlm.nih.gov/28401415/)]
22. Cohen SM, Gray KM, Ocfemia MC, Johnson AS, Hall HI. The status of the national HIV surveillance system, United States, 2013. *Public Health Rep* 2014;129(4):335-341 [[FREE Full text](#)] [doi: [10.1177/003335491412900408](https://doi.org/10.1177/003335491412900408)] [Medline: [24982536](https://pubmed.ncbi.nlm.nih.gov/24982536/)]
23. Halpern M, Lachmansingh B, Minior T, Hasbrouck LM, Persaud N, Foo A. Implementation of a standardized HIV patient monitoring system in Guyana. *Rev Panam Salud Publica* 2010 Aug;28(2):107-113 [[FREE Full text](#)] [Medline: [20963277](https://pubmed.ncbi.nlm.nih.gov/20963277/)]
24. Pham MD, Romero L, Parnell B, Anderson DA, Crowe SM, Luchters S. Feasibility of antiretroviral treatment monitoring in the era of decentralized HIV care: a systematic review. *AIDS Res Ther* 2017 Jan 19;14(1):3 [[FREE Full text](#)] [doi: [10.1186/s12981-017-0131-5](https://doi.org/10.1186/s12981-017-0131-5)] [Medline: [28103895](https://pubmed.ncbi.nlm.nih.gov/28103895/)]

Abbreviations

ART: antiretroviral treatment

HIV: human immunodeficiency virus

NGO: non-governmental organization

PLHIV: people living with HIV

Edited by B Rice; submitted 20.01.17; peer-reviewed by N Wagner, A Brown, R Simmons; comments to author 23.03.17; revised version received 28.04.17; accepted 23.05.17; published 23.06.17.

Please cite as:

Bozicevic I, Handanagic S, Cakalo JI, Rinder Stengaard A, Rutherford G

HIV Strategic Information in Non-European Union Countries in the World Health Organization European Region: Capacity Development Needs

JMIR Public Health Surveill 2017;3(2):e41

URL: <https://publichealth.jmir.org/2017/2/e41/>

doi: [10.2196/publichealth.7357](https://doi.org/10.2196/publichealth.7357)

PMID: [28645888](https://pubmed.ncbi.nlm.nih.gov/28645888/)

©Ivana Bozicevic, Senad Handanagic, Jurja Ivana Cakalo, Annemarie Rinder Stengaard, George Rutherford. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 23.06.2017. 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/>