American Journal of Infection Control ■■ (2017) ■■-■■

Contents lists available at ScienceDirect

American Journal of Infection Control

journal homepage: www.ajicjournal.org

Major Article

Virtual Zika transmission after the first U.S. case: who said what and how it spread on Twitter

Santosh Vijaykumar PhD, MA [a,*], Glen Nowak PhD, MA [b], Itai Himelboim PhD, MA [b], Yan Jin PhD, MA [b]

[a] Department of Psychology, Northumbria University, Newcastle upon Tyne, United Kingdom
[b] Grady College of Journalism and Mass Communication, University of Georgia, Athens, Georgia

Key Words:
Zika
Twitter
risk communication
social media
influence

Background: This paper goes beyond detecting specific themes within Zika-related chatter on Twitter, to identify the key actors who influence the diffusive process through which some themes become more amplified than others.

Methods: We collected all Zika-related tweets during the 3 months immediately after the first U.S. case of Zika. After the tweets were categorized into 12 themes, a cross-section were grouped into weekly datasets, to capture 12 amplifier/user groups, and analyzed by 4 amplification modes: mentions, retweets, talkers, and Twitter-wide amplifiers.

Results: We analyzed 3,057,130 tweets in the United States and categorized 4997 users. The most talked about theme was Zika transmission (~58%). News media, public health institutions, and grassroots users were the most visible and frequent sources and disseminators of Zika-related Twitter content. Grassroots users were the primary sources and disseminators of conspiracy theories.

Conclusions: Social media analytics enable public health institutions to quickly learn what information is being disseminated, and by whom, regarding infectious diseases. Such information can help public health institutions identify and engage with news media and other active information providers. It also provides insights into media and public concerns, accuracy of information on Twitter, and information gaps. The study identifies implications for pandemic preparedness and response in the digital era and presents the agenda for future research and practice.

BACKGROUND

Over the past decade, social media have assumed a greater role in the infectious disease landscape because of, among other things, their ability to *amplify* health issues through online information diffusion. Central to this process of amplification are some individuals (e.g., celebrities) and institutions on social media who can draw mass attention to an issue, trigger discussions about it, and in essence, shape the flow of information and the nature of online chatter.[1] Given the reach of social media, the engagement of such amplifiers during public health emergencies, such as the Zika virus outbreak, could affect people's perceptions of the disease, as well as their behavioral responses. The underlying processes of how health and risk information during an emerging infectious disease outbreak (EIDO) is produced and amplified through dissemination, diffusion, and exchange among individual and institutional amplifiers across multiple communication channels, interconnected by social media, are described by the Risk Amplification through Media Spread (RAMS) model.[2] In essence, the RAMS model helps those engaged in public communication efforts during EIDOs identify social media entities who amplify EIDO messages.

The purpose of this paper is to identify (1) Zika-related themes that populated Twitter chatter; (2) individuals and groups who attract the most engagement on Twitter; and (3) the modes by which they amplified specific themes within the larger Zika issue in the aftermath of the first reported case of Zika in the United States. Amplification[3] is understood as the process by which amplifiers (or user-groups) influence the volume of Zika-related information on Twitter, the degree to which it is discussed and diffused, and the connotations that become attached to Zika. Our study contributes to ongoing efforts by public health agencies in the United States[4]

* Address correspondence to Santosh Vijaykumar, PhD, MA, Department of Psychology, Northumbria University, Room 142, Northumberland Building, Newcastle upon Tyne NE1 8ST, United Kingdom.
E-mail address: santosh.vijaykumar@northumbria.ac.uk (S. Vijaykumar).
Conflicts of interest: None to report.

0196-6553/© 2017 Association for Professionals in Infection Control and Epidemiology, Inc. Published by Elsevier Inc. All rights reserved.
https://doi.org/10.1016/j.ajic.2017.10.015

and globally[5] to monitor the social media environment as a means to understand and respond to public sentiment, create greater awareness and better understanding of infectious disease threats and preventive measures, and influence risk perceptions.

Twitter's role during EIDOs

Public health agencies, including local health departments in the United States, now commonly use Twitter to communicate with the public during EIDOs.[6] Twitter is a popular micro-blogging platform where registered users can compose a message ("tweet") no longer than 140 characters and send it out to their followers. The automatically time-stamped tweets can be endorsed and shared, by those who "follow" the user, with their respective followers, by marking it as a favorite or "retweeting" it. In this way, messages can spread through Twitter, which as of June 2016, reported approximately 328 million monthly active users,[7] with 21% of registered users residing in the United States.[1]

Although social media platforms have assumed a growing presence in a range of public health functions,[8–15] the role of Twitter has been most extensively documented in the context of EIDOs such as the H1N1 and Ebola outbreaks.[16–22] Twitter data have been used to track the spread of EIDOs and to assess public knowledge and sentiment during such public health emergencies. Although EIDOs give rise to a range of discussion themes on Twitter, the ebb and flow of chatter surrounding various aspects of EIDOs often follows critical events such as World Health Organization (WHO) announcements, issuance of public health guidance, and news stories.[17] During an EIDO, public health agencies could, in real time, use Twitter data to understand public knowledge, beliefs, and sentiment as manifested on Twitter, and potentially disseminate messages through key influencers to extend the reach and impact of their communication efforts.

EIDO themes manifesting on Twitter

The communication environment in the initial stages of an EIDO is typically characterized by the need for more information and clarity about the nature and extent of the threat. Health agencies thus typically focus on communicating about modes of transmission, signs and symptoms, preventive methods, and diagnosis and treatment. With greater media coverage, EIDOs start generating public chatter on forums such as Twitter and other social media platforms. For instance, previous content analyses of Twitter data during the H1N1 outbreak[17] showed a temporal increase in tweets about resources and personal experiences during the 7-month study period. The Ebola outbreak triggered a different set of sentiments, as much public concern centered on whether Ebola could be transmitted through air, self-protection strategies, and travel to Africa. During the Ebola outbreak, Twitter also proved to be fertile ground for the proliferation of misinformation with sometimes-tragic consequences, including a rumor about the curative powers of saltwater consumption. This harmful, incorrect advice rapidly spread through social media, eventually claiming 2 lives in Nigeria.[23] Recent studies[24–27] demonstrate that the social media chatter surrounding Zika is bound to be varied and complex, given that its public narrative relates to sensitive issues, including pregnancy and women's rights, Zika's impact on child health, and its potential to be transmitted through sexual intercourse. As a result, concerns about Zika produce many conversations within one.[28]

Users who transmit and amplify health messages

Information sources such as doctors, scientists, health agencies, journalists, celebrities, and survivors have been historically critical in disseminating health information to the public. Twitter offers the health consumer the ability to receive messages from not only traditional sources, but also any Twitter user, including lay people who do not have specific expertise or an institutional affiliation. This information can prove problematic during a situation such as a Zika outbreak, in which new developments and information are rapidly unfolding and therefore inaccuracies and unconfirmed information can be prevalent. For instance, a public health agency might focus on ways to avoid being bitten by *Aedis Aegypti* mosquitoes, whereas a television news channel might focus on the dangers of Zika for pregnant women, showing graphic images and thereby amplifying the nature of the threat. When both these pieces of helpful information, from sources with differing agendas, appear on a Twitter feed, users might be confused about the magnitude and nature of the threat.

Inspired by the idea that a small number of highly influential users might possess the ability to diffuse content to a large audience, quantifying a Twitter user's social influence has attracted much interest in recent times.[29–31] Bakshy et al.[32] found instead that planting content with a large number of potential influencers can return average effects and ultimately prove to be a more efficient strategy for information diffusion. Although these studies bear important implications for public health agencies, they fail to provide a rubric for identifying potential groups or communities of Twitter users whose influence can be effectively utilized. Our study aims to fill this gap by showing how Twitter analyses can be used to identify specific amplifier groups and the extent of their amplification.

Ways in which content is amplified on Twitter

Social media have provided new opportunities for individuals to engage in direct dialogue, with both organizations and each other.[33,34] Given that social media platforms differ greatly in their conventions and characteristics, Guidry et al.[35] have argued that it is crucial to understand what health-related information is available online and how to engage people on various social media platforms. For example, Twitter provides 2 main ways for people to interact—via tweets and with other Twitter users (i.e., by retweeting, liking, replying or direct messaging). Instagram provides 2 similar ways for people to interact with photos or videos, namely, "liking" and posting comments, and facilitates use of hashtags and mentions. By these modes, social media enable diffusion and amplification of messages.

The "behavior" of endorsing or amplifying a message can be driven by both an assessment of the merit of the message and perceptions about the messenger.[36] Engagement through the previously mentioned modes has, however, becoming increasingly challenging, due to the coexistence of accurate health information and misinformation (including rumors) in the online space.[35] Misinformation is objectively incorrect information or information at odds with the scientific consensus, and it can mislead audiences and pose a major problem.[37–39] Different from rumor, misinformation, once solidified, is not easy to address, as it is often intertwined with strongly held attitudes and sentiments that can easily spread to others.[40] Social media platforms, with their interactive and engaging characteristics, make the spread and effects of misinformation online even more frequent and of growing concern.[41]

The problems of rumor and misinformation propagation through online social networks such as Twitter are particularly relevant during EIDOs. Computer scientists have discovered that high-influence users can act as firewalls against circulating rumors or misinformation, and thereby help quell their diffusion.[42] This paper, however, is concerned with not only influential users but also influential content, and thus we use the term "amplification," which denotes expansion of narratives (content) by communicators (users). Because public and other health information can be "amplified" in varying degrees through Twitter, via multiple modes, we conceptualize amplification as being user *and* content driven. For instance, Twitter "mentions"—a function that includes another user's name in a tweet, using the "@" symbol—suggests that the user mentioned is perceived as relevant or important to the conversation. A Zika-related tweet might thus be amplified, because the user who posted it is a key part of the conversation, that is, they are being mentioned by many others who post Zika-related content. Retweeting, a function that enables forwarding or sharing of a certain tweet, implies that a single piece of content—a tweet—was deemed relevant or worthy of being shared with one's social network (for further diffusion), thus amplifying the message through sharing. Whereas mentions and retweets reflect an actual amplification based on engagement with a message or user, other amplification modes capture the potential of information spread. Some users might aim to amplify Zika-related tweets, or certain themes within it, by actively increasing their frequency of tweets about that topic. In other instances, a tweet might gain amplification because it is associated with certain users whose influence extends beyond the limited boundaries of Zika conversations, across the Twitter universe, as such messages will appear on the Twitter feed of a large number of users. In these 2 cases, users may or may not actually engage with the content, but these amplification forms increase the potential of such engagement.

Research questions

The chatter surrounding Zika encompassed a range of themes, from concerns about the previously little known infectious disease to social issues involving gender, stigma, and human rights. We addressed 3 research questions. The first is: How did the volume of tweets vary across the key Zika themes? We also addressed a subcomponent of this question: How did the prevalence of themes fluctuate over the 3-month period? As discussed earlier, content posted by users can be amplified in many ways, and the types of users that play the role of content amplifiers may differ across the type of amplification. The second question is: Did the modes of amplification vary by amplifier group? The third question is: How did Zika-related themes vary in terms of key amplifier groups and modes of amplification?

METHODS

This study applied Twitter data mining and analysis to classify content and identify influential user groups within Zika-related chatter on Twitter. Adapting Yoon and Bakken's[43] Steps in Web Mining of Tweets, we first selected relevant key words prior to importing the data (Step 1). We then employed a mixture of top-down and bottom-up approaches to classify content into themes (Step 2). Based on Twitter interaction patterns, we identified 4 types of amplification modes that occur on Twitter (Step 3), and top users within each mode. Finally, each of these users was classified, using 11 types of amplifier groups (Step 4). The main variables of interest in this section are: (1) themes, which refer to specific topics related to Zika; (2) amplification modes, which refer to the various mechanisms by which information spreads on Twitter; and (3) amplifier groups, which refer to users/user groups responsible for spreading information through Twitter.

Step 1: data collection

All tweets that included the keywords "Zika," "#Zika," and "#Zikavirus" were collected for the first 3 months of 2016, using Crimson Hexagon, a social media analytics software and data library. The time period was selected to capture the Twitter conversations that preceded and followed the first reports of Zika infection in the continental United States on February 2nd, 2017. Although an earlier infection was reported in Puerto Rico in December, we used the first infection in the continental United States as a point of reference, due to the low usage of Twitter in Puerto Rico,[44] (6%) compared with the entire U.S. adult population (21%) (Pew Research Center).[45] For each tweet, the user's information was also collected. A total of 3,057,130 tweets were collected. The data captured a wide range of related, yet distinct, Zika themes.

Step 2: thematic classification

This collection of tweets was classified into 1 of 9 themes. Themes were determined by blending 2 approaches. The first was a top-down approach that used categories identified by the Centers for Disease Control and Prevention (CDC) and found on their website (http://www.cdc.gov/zika). These categories were transmission, pregnancy, travel, testing and diagnosis, and symptoms. The second was a bottom-up approach—scrutinizing the data to identify themes beyond the ones suggested by the CDC—which revealed additional themes, such as social issues and conspiracy theories. Once the themes were identified, an additional search for various sets of keywords was conducted, using Boolean operators such as AND, OR, and NOT. For instance, to capture the conversation about symptoms, keywords such as "rash," "joint pain," "conjunctivitis," and "red eyes" were used. For a conversation about travel, keywords such as "travel," "trip," "vacation," and "tourist" were applied. Each search syntax was tested against the resultant data and refined as needed. Although the themes were distinct, they were not mutually exclusive. For instance, a tweet about tips to avoid being infected with Zika while traveling to Brazil was categorized in both the "transmission" and "travel" themes. See Table 1 for the distribution of tweets by theme. An operational definition of individual themes, and a sample tweet for each, is included in Appendix 1.

Step 3: operationalizing amplification modes

In a 3-month timespan, and across a wide range of conversation themes, patterns of user engagement often change. The aim of this step was to capture various kinds of user engagement (which we refer to as "amplification modes") vis-à-vis the users who drive it (whom we refer to as "amplifiers"). We focused on 4 amplification modes—mentions, retweets, talkers, and Twitter-wide amplifiers—and the user-level metrics associated with each of them. Below, we describe the process for this step.

We collected tweets across 14 weeks (W_1 to W_{14}). Tweets in each week were further classified into 1 of 9 themes (T_1 to T_9), thereby creating $14 \times 9 = 126$ weekly datasets (W_1T_1 to $W_{14}T_9$). Within each of these datasets, we identified up to 10 users (amplifiers) who had the greatest values for each of the four amplification modes (mentions, popular, talkers, or Twitter-wide). Each week (for example, W_1T_1) thus had a maximum of $10 \times 4 = 40$ amplifiers. Across 126 weeks, that made a maximum of $126 \times 40 = 5040$ amplifiers.

Table 1
Distribution of tweets according to Zika-related themes

Theme	Tweets (n)	% of total	Retweets (%)
Treatment	107,236	3.2	38
Travel	234,987	7	37
Transmission	1,943,813	52	42
Testing and diagnosis	135,480	4	37
Symptoms	64,967	1.9	54
Social issues	136,482	4.1	43
Prevention	128,013	3.8	45
Effects on pregnancy	515,122	15.4	42
Conspiracy theories	72,349	2.2	47

Because some weekly datasets were smaller or contained fewer large amplifiers, our final sample comprised n = 4997 amplifiers (as opposed to 5040) distributed across the 4 amplification modes as follows. (1) The first mode was mentions (n = 1241)—the number of times each user was mentioned by others (i.e., using the @ symbol). Users with the highest "mentions" value in each subset were classified as amplifiers. (2) The second mode was retweets (n = 1240)—the number of retweets (shared tweets) for each post. Users of the most shared tweets were classified as amplifiers. (3) The third mode was talkers (n = 1258)—the number of tweets each user posted. Users who posted the most tweets in each subset were classified as amplifiers. (4) The fourth mode was twitter-wide amplifiers (n = 1258)—top users, based on "Klout score" (https://klout.com/corp/score). A Klout score calculates a user's Twitter-wide influence based on their ability to attract attention across social media platforms and topics of discussion. Content posted by users with high Klout scores reaches a large number of people who *may* choose to give it attention. Messages posted by these users therefore have greater potential to be amplified.

Step 4: classifying amplifier groups

We analyzed a total of 4997 amplifiers. Each amplifier was classified into 1 of the following 12 user categories:

- Health institution and its individual affiliates (@CDCgov; @Who);
- Traditional news media and affiliated journalists (@Nytimes, @reuters);
- Online-only news media and their affiliates (@HuffPostPol);
- Advocacy organizations (@AmnestyWomenRts);
- Non-health-specific governmental organizations (@whitehouse);
- Other organizations—organization not falling under previous categories (@Olympics);
- Grassroots—regular users (Bloggers; @OlderMommyStill);
- Politicians/elected officials (@HillaryClinton);
- Nonprofit (not advocacy);
- Academic (@naturenews);
- Other media; and
- User does not exist.

One of the authors (IH) and a graduate student coded 10% of the sample (n = 500), producing a high inter-rater reliability of Cohen's $\alpha = 0.89$. The sample for reliability testing was selected through a stratified random sampling procedure, comprising 125 users from each of the 4 amplification modes.

RESULTS

How did the volume of tweets vary across the key Zika themes?

Information or content related to the transmission of the Zika virus was the most tweeted theme (52%), followed by Zika's effects on pregnancy (15.4%) and travel (7%). The categories of social issues (4.1%) and testing and diagnosis (4.0%) received an almost identical amount of attention on Twitter. Prevention of Zika virus contraction (3.8%) was tweeted about slightly more, compared with treatment (3.2%). Conspiracy theories were tweeted about (2.2%) more than were Zika-related symptoms (1.9%), which was the theme receiving the fewest tweets.

How did the prevalence of themes fluctuate over the 3-month period?

Figure 1 shows the temporal changes in the prevalence of themes in the period January 1—March 31, 2016. All themes, except social issues and prevention, witnessed a spike around the last week of January, when the WHO declared Zika to be a public health emergency of international concern (PHEIC).[46] Bimodal spikes were witnessed for transmission and travel during and after the announcement, but both these themes declined in the weeks thereafter. Tweets related to conspiracy theories, social issues, and symptoms spiked toward the latter half of the timeline. Effects on pregnancy, however, witnessed constant attention throughout the timespan, except in the last few weeks.

How did amplifier groups vary by mode of amplification?

As seen in Figure 2, traditional news media were the most mentioned (30.9%) amplifier group in Zika-related tweets, followed by health institutions (22%) and grassroots users (9.9%). Similarly, tweets posted by traditional news media were most frequently retweeted (31.8%), followed by health institutions (17.2%) and grassroots users (12.4%). Grassroots users were found to post the largest number of tweets (talkers, 59.1%), followed by other organizations (17.2%) and online-only news media (7%). Traditional news media (50.7%) were found to have the greatest Twitter-wide amplifying potential, followed by other media (9.6%), online-only news media (8.5%), and grassroots users (8.3%). Across the 4 categories, traditional news media, grassroots users, and health institutions were identified as the top 3 amplifier groups with the most amplification ability.

How did Zika-related themes vary in terms of amplifier groups and modes of amplification?

We focused on the top 3 groups with the most amplification ability (identified in research question 2) and examined their influence across Zika-related themes. Figure 3 suggests that health institutions and news media were mentioned in almost equal proportion in tweets about Zika prevention and symptoms. However, news media were mentioned substantially more than were health institutions, in tweets related to transmission, effects on pregnancy, testing and diagnosis, travel, and treatment. News media were almost exclusively mentioned in tweets pertaining to Zika-related social issues. Similarly, tweets related to prevention and symptoms posted by health institutions were retweeted more frequently (or retweeted), compared with those by news media. However, tweets pertaining to transmission, testing and diagnosis, effects on pregnancy, social issues, travel, and treatment, posted by news media, were more widely retweeted, as opposed to similarly themed tweets posted by health institutions. Tweets pertaining to conspiracy theories posted by grassroots users were most widely retweeted. Lastly, grassroots-level users were found to post tweets most frequently across all Zika-related themes. News media were found to have the most Twitter-wide influence across all Zika-related themes.

DISCUSSION

This study examined Zika-related tweets in the period surrounding the first Zika case in the United States. We discovered variations in temporal spread of Zika-related themes and identified 3 main amplifier groups (traditional media, health institutions, and grassroots-level users) influencing amplification of specific themes through different modes.

Our first research question focused on the volume of tweets, by specific themes, and their temporal spread over the 3-month period; it revealed inconsistencies between the 2 factors. Specifically, tweets with the highest volume were not necessarily ones that commanded sustained attention over time. A case in point

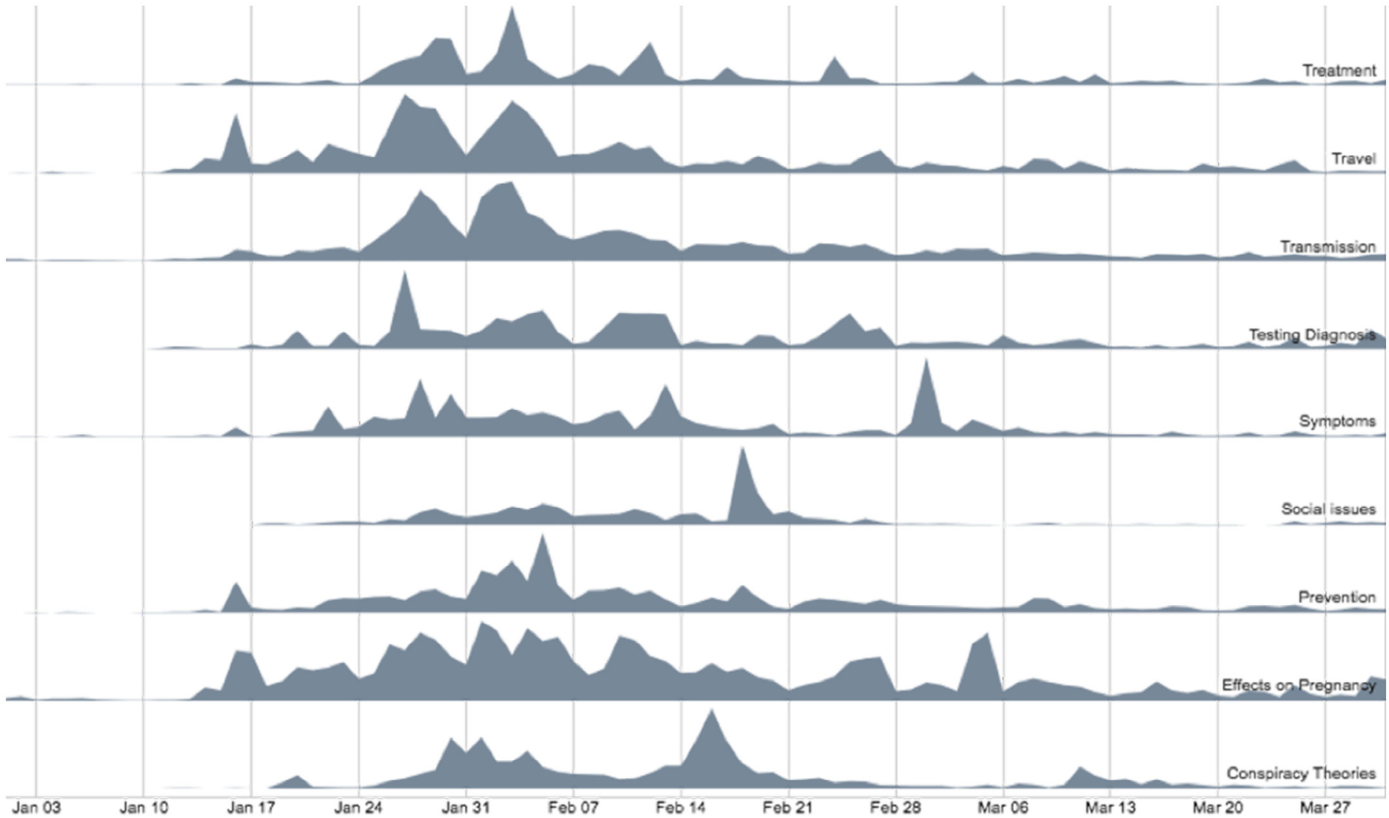

Fig 1. Temporal spread of Zika-related themes over the 14-week period.

is the theme related to the effects of Zika on pregnancy; the volume was one fourth that of the most-tweeted theme (transmission) but it received sustained media chatter over the 3-month period (Fig 1). In contrast, transmission and travel-related tweets ranked 1 and 3 by volume, respectively, but chatter around these themes increased only occasionally, mainly during critical news events such as the WHO's declaration of Zika as a PHEIC.[46] Other studies have found similar trends in the volume of Twitter exchanges (i.e., initial increases followed by a decline after critical news events).[17,24] The results suggest, however, that the temporal analysis performed here (as opposed to looking only at overall volume of tweets) can provide more granular insights about the ebb and flow of specific issues. Public health institutions might utilize similar analytic approaches to refine their social media strategies and recognize that their actions—to the extent that they attract media attention—can trigger tweeting. Consequently, a recommendation is that they have partners ready to amplify tweets by retweeting and "liking." Partners, for instance, include those they work with on infectious disease–related efforts in advocacy organizations, academic institutions, private sector health care organizations, and nongovernment organizations. Lastly, we recommend that public health institutions place priority on allaying the fears and anxieties triggered, or with potential to be triggered, by specific news events.

Findings related to our second research question indicated that news media yielded the most influence on Zika-related tweets followed by public health institutions and grassroots users . This finding means that, despite their central role in informing the public and preventing outbreaks, public health institutions trailed traditional news media in terms of their ability to *directly* influence social media chatter. These findings are consistent with those of Househ,[47] whose global study of Ebola-related tweets highlighted the influence of news media on Twitter feeds. Our study provides additional evidence by demonstrating that news media also wielded Twitter-wide influence (through Klout scores), beyond chatter related to Zika. Possible explanations lie in the fact that news media inform and shape public opinion across a range of social issues, including health, on a daily basis, whereas public health institutions have sporadic and topical engagement. However, the role of public health institutions as amplifiers remains critical, because much of what makes the news or attracts media attention during EIDOs is initiated by or related to their recommendations and actions, as well as their communications (e.g., announcements, media interviews, press conferences, release of reports and references, etc.).

Addressing our third research question involved a thematic investigation of broad trends uncovered by the second research question, and revealed a profound influence of news media in amplifying all themes, except prevention and symptoms, for which public health institutions had greater visibility. Here, the fact that traditional news media have a more passive, as opposed to active, influence on amplification of themes is important. The frequency of tweets posted by news media (active influence) is relatively low (Fig 2). However, their passive influence is derived from the following factors: (1) they are mentioned in greater volume in tweets related to those themes; (2) their tweets are shared more; and (3) their Twitter-wide influence is greater than that of other amplifier groups. Mentions and retweets indicate that users either are replying to tweets by the news media, consider news media to be relevant to themes being discussed, and/or are actively sharing tweets posted by news media—implying greater responsiveness or engagement on the part of their audience. In contrast, although grassroots users were found to tweet with greatest frequency, they

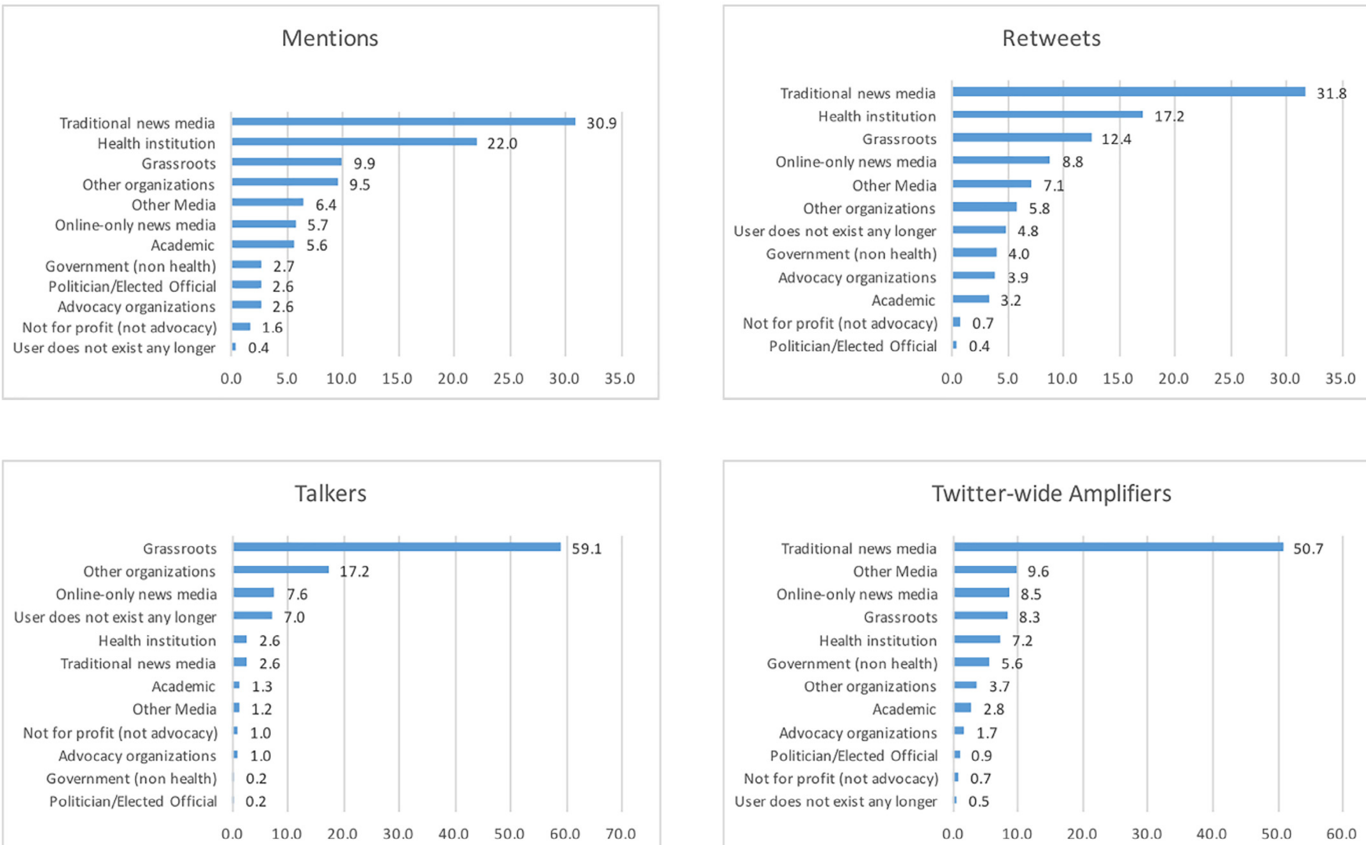

Fig 2. Distribution of different amplifier (user) groups by modes of amplification.

had limited impact, except on conspiracy theories, a theme likely of great concern to those managing the communication response to outbreaks.

The current study findings have several implications from a policy standpoint, at times when social media are integral components of outreach conducted by public health institutions during infectious disease outbreaks.[48] For instance, the findings suggest that public health institutions can utilize insights from social media monitoring to guide and refine their social media strategies. Social media monitoring is helpful for learning what topics and information are most visible and are being shared, including in the immediate aftermath of new developments or specific news events and coverage. Public health institutions can compare this to the messages and materials they are providing to identify gaps, or topics being inadequately addressed. They may also find they need to provide more information, including via social media channels, that puts risk into context and/or addresses fears and anxieties (such as effects on pregnancy) brought to the surface by social media analyses.

From the standpoint of public health agencies, the findings reaffirm the value of active engagement with the media during EIDOs. Whether and to what extent recent events, such as the "fake news" controversy,[49] that have brought the credibility of the news media into question might shape future engagements between health institutions and news media commands further research. Lastly, timely identification of conspiracy theories, rumors, and misinformation through real-time analytics is possible, should be done, and is critical, given their potential to be swiftly amplified on social media.

Limitations

One limitation of this study is that the entire population of tweets could be analyzed for the first research question, but only a subsample for the second and third. This approach was used because categorizing tweets into themes is automated by Crimson Hexagon, whereas categorizing amplifier groups for the second and third research questions required manual coding. Given that approximately 3 million tweets were involved, constraints on human resources and time necessitated coding a sizeable subsample of tweets instead. Using the subsample may have affected the results, but sampling is a tested approach and has been used by researchers for inquiries of a similar nature.[50] This study is also limited in that it did not provide an understanding of how communication among amplifier groups contributed to amplification effects. This type of network analysis was beyond the scope of this study but would be a helpful effort for future research.

CONCLUSIONS

During an infectious disease outbreak, particularly one that involves a new or previously unseen health threat, tweets can increase awareness, shape perceptions, and affect individual behavior. Tweets are also identifiable and trackable, so public health agencies and healthcare providers, particularly those responding to or affected by an emerging infectious disease threat, can benefit from directly monitoring and analyzing Twitter data or having access to reports that do so. As shown here, social media analytics provide a sense of the following factors: how much attention an emerging

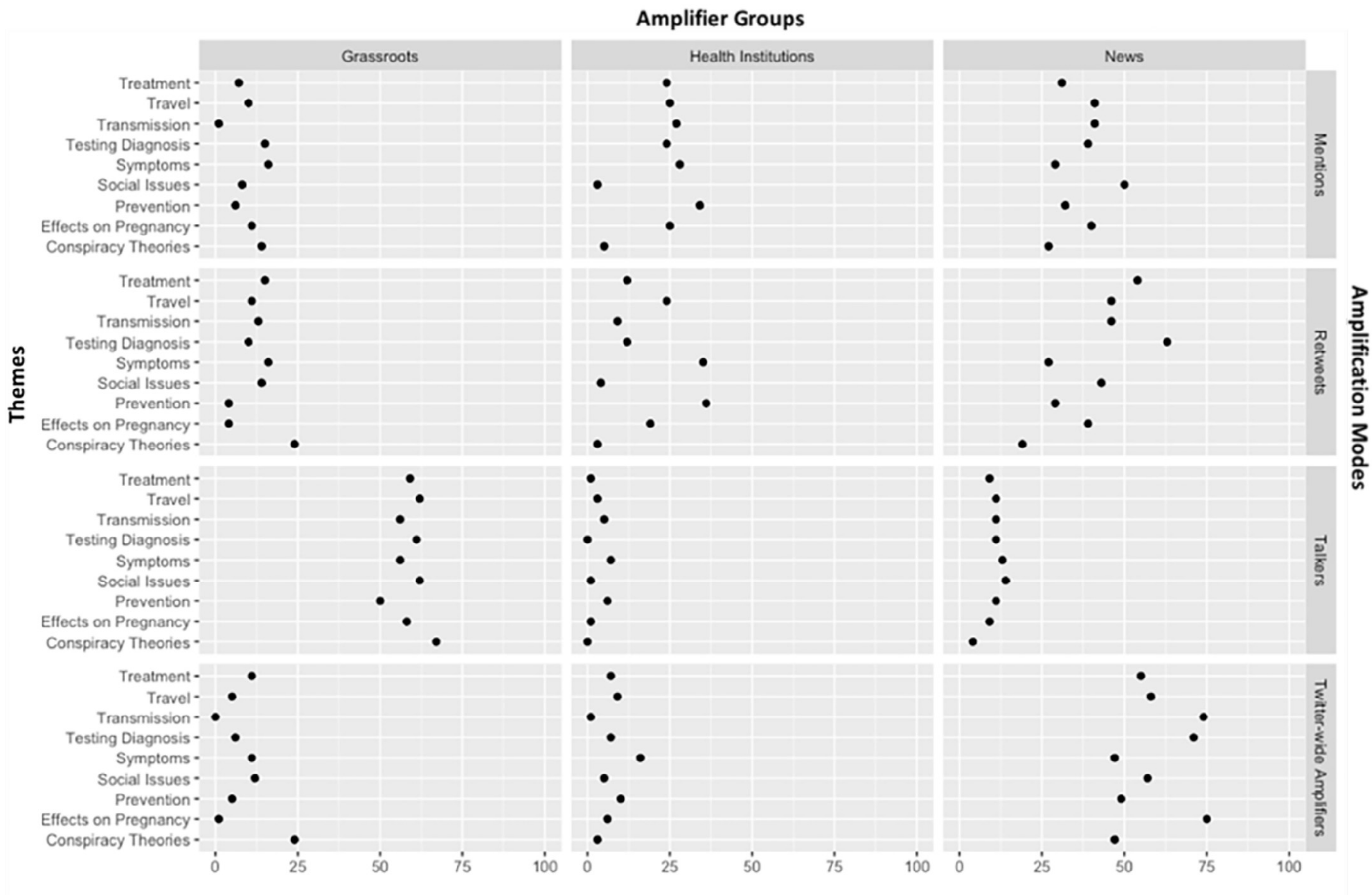

Fig 3. Distribution of different themes by top 3 amplifier groups and amplification modes.

infectious disease is garnering on Twitter; the information, beliefs, and sentiments being expressed; perceptions and potential acceptance of public health actions and advice; and the misperceptions, including conspiracy theories, that are being shared.

As this study illustrates, journalists and news media organizations play a significant role in disseminating and amplifying EIDO information, with much of the Twitter content related to public health updates, actions, and advice. As seen here, their focus may be primarily on disease transmission and who is affected, and less on prevention measures. Health institutions thus need to be mindful that although their major announcements and updates can fuel media tweeting on an EIDO, some important content—such as preventive measures—may not achieve much visibility or interest in the social conversation. This study also showed the value of knowing what social conversations are taking place. Grassroots users are quite visible on Twitter when it comes to EIDOs, but not with an emphasis on sharing transmission and prevention information. Rather, these entities and individuals are the sources and amplifiers of a range of "socially" related aspects of the situation, from social concerns to conspiracy theories. Knowing what is being disseminated and shared in this domain provides critical information for the health institutions engaged in the communication and education response.

Acknowledgements

We are grateful to Shivani Seshadri, Associate Professor, Impact School of Architecture, Bangalore and Dr. Venkat Chilukuri, Data Science Faculty, Information Arts and Information Design Practices, Shrishti Institute of Art, Design and Technology, Bangalore, India for designing the graphical abstract and Figure 3 respectively.

References

1. Wagner L. How "Zika-Proof" Hope Solo became the biggest villian of the Rio Olympics. Available from: http://www.slate.com/blogs/five_ring_circus/2016/08/12/how_zika_proof_hope_solo_became_the_biggest_villain_of_the_rio_olympics.html. Accessed September 25, 2016.
2. Vijaykumar S, Jin Y, Nowak G. Social media and the virality of risk: the risk amplification through media spread (RAMS) model. J Homel Secur Emerg Manag 2015;12:653-77.
3. Pidgeon N, Kasperson RE, Slovic P. The social amplification of risk. Cambridge, UK: Cambridge University Press; 2003.
4. Centers for Disease Control and Prevention. Data and metrics. Available from: http://www.cdc.gov/socialmedia/data/index.html. Accessed September 25, 2016.
5. World Health Organization. Zika strategic response framework and joint operations plan. Available from: http://apps.who.int/iris/bitstream/10665/204420/1/ZikaResponseFramework_JanJun16_eng.pdf. Accessed September 25, 2016.
6. Wong R, Harris JK, Staub M, Bernhardt JM. Local health departments tweeting about Ebola: characteristics and messaging. J Public Health Manag Pract 2017;23:e16-24.
7. Statista. Number of monthly active Twitter users worldwide from 1st quarter 2010 to 2nd quarter 2017 (in millions). Available from: https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/. Accessed August 1, 2017.
8. Bernardo TM, Rajic A, Young I, Robiadek K, Pham MT, Funk JA. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. J Med Internet Res 2013;15:e147.
9. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection—harnessing the Web for public health surveillance. N Engl J Med 2009;360:2153-7.

10. Thackeray R, Neiger BL, Smith AK, Van Wagenen SB. Adoption and use of social media among public health departments. BMC Public Health 2012; 12:1.
11. Merchant RM, Elmer S, Lurie N. Integrating social media into emergency-preparedness efforts. N Engl J Med 2011;365:289-91.
12. Keim ME, Noji E. Emergent use of social media: a new age of opportunity for disaster resilience. Am J Disaster Med 2010;6:47-54.
13. Chou W-YS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009;11:e48.
14. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013;15: e85.
15. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013;14:15-23.
16. Aramaki E, Maskawa S, Morita M. Twitter catches the flu: detecting influenza epidemics using Twitter. In Proceedings of the Conference on Empirical Methods in Natural Language Processing. Edinburgh, United Kingdom, July 27-31, 2011.
17. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of tweets during the 2009 H1N1 outbreak. PLoS ONE 2010;5:e14118.
18. Fung IC-H, Tse ZTH, Cheung C-N, Miu AS, Fu K-W. Ebola and the social media. Lancet 2014;384:2207.
19. Paul MJ, Dredze M. You are what you tweet: analyzing Twitter for public health. Proc Int AAAI Conf Weblogs Soc Media 2011;20:265-72.
20. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. PLoS ONE 2011;6:e19467.
21. Odlum M, Yoon S. What can we learn about the Ebola outbreak from tweets? Am J Infect Control 2015;43:563-71.
22. Yoon S, Elhadad N, Bakken S. A practical approach for content mining of tweets. Am J Prev Med 2013;45:122-9.
23. Oyeyemi SO, Gabarron E, Wynn R. Ebola, Twitter, and misinformation: a dangerous combination? BMJ 2014;349:g6178.
24. Fu K-W, Liang H, Saroha N, Tse ZTH, Ip P, Fung IC-H. How people react to Zika virus outbreaks on Twitter? A computational content analysis. Am J Infect Control 2016;44:1700-2.
25. Khatua A, Khatua A. Immediate and long-term effects of 2016 Zika Outbreak: a Twitter-based study. In Proceedings of the IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom). Munich, Germany, September 14-17, 2016.
26. Southwell BG, Dolina S, Jimenez-Magdaleno K, Squiers LB, Kelly BJ. Zika virus–related news coverage and online behavior, United States, Guatemala, and Brazil. Emerg Infect Dis 2016;22:1320.
27. Stefanidis A, Vraga E, Lamprianidis G, Radzikowski J, Delamater PL, Jacobsen KH, et al. Zika in Twitter: temporal variations of locations, actors, and concepts. JMIR Public Health Surveill 2017;3:e22.
28. The Guardian. Zika emergency pushes women to challenge Brazil's abortion law. Available from: https://www.theguardian.com/global-development/2016/jul/19/zika-emergency-pushes-women-to-challenge-brazil-abortion-law. Accessed September 26, 2016.
29. 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;21:517-26.
30. Cha M, Haddadi H, Benevenuto F, Gummadi PK. Measuring user influence in Twitter: the million follower fallacy. Proc Int AAAI Conf Weblogs Soc Media 2010;10:30.
31. Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? In: Proceedings of the 19th International Conference on World Wide Web. Raleigh, NC, April 26-30, 2010.
32. Bakshy E, Hofman JM, Mason WA, Watts DJ. Identifying influencers on Twitter. In: Proceedings of the Fourth ACM International Conference on Web Seach and Data Mining (WSDM). Hong Kong, China, February 9-12, 2011.
33. McAllister-Spooner SM. Fulfilling the dialogic promise: a ten-year reflective survey on dialogic Internet principles. Public Relat Rev 2009;35:320-2.
34. Sommerfeldt EJ, Kent ML, Taylor M. Activist practitioner perspectives of website public relations: why aren't activist websites fulfilling the dialogic promise? Public Relat Rev 2012;38:303-12.
35. Guidry JP, Jin Y, Orr CA, Messner M, Meganck S. Ebola on Instagram and Twitter: how health organizations address the health crisis in their social media engagement. Public Relat Rev 2017;43:477-86.
36. Petrovic S, Osborne M, Lavrenko V. RT to win! Predicting message propagation in Twitter. In: Proceedings of the 5th. International AAAI Conference on Weblogs and Social Media, Barcelona, Spain, July 17-21, 2011.
37. Fowler A, Margolis M. The political consequences of uninformed voters. Elect Stud 2014;34:100-10.
38. Reedy J, Wells C, Gastil J. How voters become misinformed: an investigation of the emergence and consequences of false factual beliefs. Soc Sci Q 2014;95:1399-418.
39. Tafuri S, Gallone M, Cappelli M, Martinelli D, Prato R, Germinario C. Addressing the anti-vaccination movement and the role of HCWs. Vaccine 2014;32: 4860-5.
40. Weeks BE. Emotions, partisanship, and misperceptions: how anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. J Commun 2015;65:699-719.
41. Bode L, Vraga EK. In related news, that was wrong: the correction of misinformation through related stories functionality in social media. J Commun 2015;65:619-38.
42. Borge-Holthoefer J, Moreno Y. Absence of influential spreaders in rumor dynamics. Phys Rev E 2012;85:026116.
43. Yoon S, Bakken S. Methods of knowledge discovery in tweets. In: NI 2012: Proceedings of the 11th International Congress on Nursing Informatics. Montreal, Canada, June 23-27, 2012.
44. Rosado L. Facebook continues to dominate social-network use in Puerto Rico. Available from: http://gaitherinternational.com/facebook-continues-to-dominate-social-network-use-in-puerto-rico/. Accessed August 1, 2017.
45. Pew Research Center. Social media update 2016. Available from: http://assets.pewresearch.org/wp-content/uploads/sites/14/2016/11/10132827/PI_2016.11.11_Social-Media-Update_FINAL.pdf. Accessed August 1, 2017.
46. World Health Organization. WHO statement on the first meeting of the International Health Regulations (2005) (IHR 2005) Emergency Committee on Zika virus and observed increase in neurological disorders and neonatal malformations. Available from: http://www.who.int/mediacentre/news/statements/2016/1st-emergency-committee-zika/en/. Accessed August 1, 2017.
47. Househ M. Communicating Ebola through social media and electronic news media outlets: a cross-sectional study. Health Informatics J 2016;22:470-8.
48. Vijaykumar S, Meurzec RW, Jayasundar K, Pagliari C, Fernandopulle Y. What's buzzing on your feed? Health authorities' use of Facebook to combat Zika in Singapore. J Am Med Inform Assoc 2017;24:1155-9.
49. Allcott H, Gentzkow M. Social media and fake news in the 2016 election. J Econ Perspect 2017;31:211-36.
50. van Lent LG, Sungur H, Kunneman FA, van de Velde B, Das E. Too far to care? Measuring public attention and fear for Ebola using Twitter. J Med Internet Res 2017;19:e193.

APPENDIX. OPERATIONAL DEFINITION OF THEMES AND SAMPLE TWEETS FOR EACH THEME

Theme	Description*	Sample Tweets†
Transmission and spread	Tweets containing words specific to the transmission of Zika (e.g., mosquito, sexual transmission, transmission, infecting, spread, bitten, catch)	RT @BBCBreaking First sexually transmitted case of Zika virus is confirmed by health authorities in Texas
Effects on pregnancy	Tweets containing words specific to pregnancy or its effects (e.g., microcephaly, birth defect, newborn, small head)	RT @blackstips Due to a Zika virus outbreak, an estimated 4,000 babies have been born with abnormally small heads in Brazil since October.
Travel	Tweets containing travel-related keywords (e.g., travel, trip, visit, airline, vacation, tourists)	RT @Adel__Almalki #News by #almalki : United to refund travel to regions hit with Zika virus
Social issues	Tweets containing keywords associated with social issues related to Zika, including abortion, climate change, contraceptive, and delaying pregnancy	RT @AP Pope suggests women who are threatened by Zika virus could use artificial contraception but not abort fetus
Testing and diagnosis	Tweets containing keywords associated with diagnosis (e.g., test, diagnosis, detect, tests)	Texas hospitals say they have developed rapid test for Zika
Prevention	Tweets containing keywords associated with prevention (e.g., prevent, protect, avoid) or specific precautions (long sleeve, repellent, condoms, etc.) or hazards (e.g., standing water)	RT @OlderMommyStill Take these 7 steps to avoid contracting the Zika Virus
Treatment	Tweets containing treatment-specific keywords (e.g. medication, cure, treatment, immunization)	RT @MoreScienceNews Scientists' path to usable Zika vaccine strewn with hurdles
Conspiracy theories	Tweets containing the word conspiracy or other conspiracy-related keywords (e.g., oxitec, monsanto, etc.)	RT @MassDeception1 Zika virus can be purchased over the internet; origins linked to Rockefeller Foundation
Symptoms	Tweets containing symptoms-like keywords (e.g., symptom, signs) or the actual symptons (e.g., rash, joint pain, red eyes, pinkeye, Guillain-Barre, fever)	RT @WHO Q: What are the symptoms of #Zika? A: Most ppl will get a slight fever, rash. Others may also get conjunctivitis, muscle pain, feel tired

*All searches were conducted within the larger Zika-specific Tweets dataset.
†Tweets include only text to illustrate the search process, and hyperlinks were removed to keep tweets short for this table.