Virtual Zika Transmission after the first US Case: Who said What and How It Spread on Twitter

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) or institutions on social media who can draw mass attention to the issue, trigger discussions about it and, in essence, shape the flow of information and nature of online chatter.[1] Given the reach of social media, the engagement of such amplifiers during public health emergencies like Zika, could affect people's perceptions of the disease as well as their resulting 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 between individual and institutional amplifiers through 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 to identify social media entities who amplify EIDO messages.

The purpose of this paper is to identify a) Zika-related themes that populated Twitter chatter, b) individuals and groups who attract most engagement on Twitter, and c) 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 US. Amplification[3] (pg. 152), is understood as the process by which the volume of Zika-related information on Twitter, the degree to which it is discussed and diffused, and the connotations attached to Zika are influenced by amplifiers (or user-groups). Our study contributes to ongoing efforts by public health agencies in the US[4] 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 US, 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, a message can spread through Twitter which, as of June 2016, reported approximately 328 million monthly active users[7], with 21% of its registered users residing in the US.[1] While 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 like H1N1 and Ebola.[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. While EIDOs give rise to a range of discussion themes on Twitter, the ebb and flow of chatter surrounding different aspects of EIDOs often follow critical events such as 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 manifest 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 like Twitter and other social media platforms. For instance, previous content analyses of Twitter data during H1N1[17] showed a temporal increase in tweets about resources and personal experiences during the 7-month study period. Ebola triggered a different set of sentiments as much public concern centered on knowing whether Ebola could be transmitted through air, self-protection strategies, and travel to Africa. During Ebola, 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 two 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, Zika is many conversations within one.[28]

Users who Transmit and Amplify Health Messages

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

sources with different agendas, appear on a user's 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. (2011)[32] found instead that planting content with a large number of potential influencers might return average effects, and ultimately prove to be a more efficient strategy for information diffusion. While 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 has provided new opportunities for individuals to engage in direct dialogue both with organizations and each other.[33,34] Given that social media platforms differ greatly in their conventions and characteristics, Guidry et al.[35] argued that it is crucial to understand what health-related information is available online and how to engage people on different social media platforms. For example, while Twitter provides two main ways for people to interact with tweets and with other Twitter users (i.e., "retweeting" and "favoriting"), Instagram provides two similar ways for people to interact with photos or videos, namely, "liking" and posting comments, and also facilitates the use of hashtags and mentions. By these modes, social media enable diffusion and amplification of messages.

The "behaviour" of endorsing or amplifying a message can be driven by both an assessment of the merit of the message as well as perceptions about the messenger.[36] Engagement through the above-mentioned modes has however becoming increasingly challenging, due to the co-existence of accurate health information and misinformation (including rumors), in the online space.[35] Misinformation refers to the presence of or belief in

objectively incorrect information or information counter to the scientific consensus, and can mislead audiences and pose a major problem.[37-39] Different from rumor, once solidified, misinformation is not easy to address, as it is often intertwined with strongly held attitudes and sentiments that can easily be spread to others.[40] Social media platforms and their interactive and engaging characteristics make the spread and effects of misinformation online even more frequently and of growing concern.[41]

The problems of rumor and misinformation propagation through online social networks like Twitter is particularly relevant during EIDOs. Computer scientists have discovered that high-influence users act as firewalls to circulating rumors or misinformation and thereby can help quell their diffusion.[42] This paper, however, is concerned with not only influential users but also influential content, and thus we use "amplification", a term that 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 user *and* content driven. For instance, Twitter 'mentions' – a function where a user's name is mentioned in a tweet using the '@' symbol – suggests that the user who was mentioned was perceived as being 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, being mentioned by many others who post Zika-related content. "Retweet", 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 enough 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 a 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 to

across the Twitter universe, as such messages will appear on the twitter feed of a large number of users. In these two 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 surrounding a previously little known infectious disease to engagement on social issues involving gender, stigma, and human rights.

RQ1: How did the volume of tweets vary across the key Zika themes?

RQ1a: How did the prevalence of themes fluctuate over the 3-month period?

As earlier discussed, 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.

RQ2: Did the modes of amplification vary by amplifier group?

RQ3: 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 (2012)[43] Steps in Web Mining of Tweets, we first selected the relevant key words prior to importing the data (Step 1 below). 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 then identified four types of amplification modes on Twitter (Step 3), and identified top users within each of these modes. Finally, each of these users were classified into 11 types of amplifier groups (Step 4). The main variables of interest in this section are: A) Themes: These refer to specific topics related to Zika, B) Amplification Modes: These refer to different mechanisms by which information spreads on Twitter and C) Amplifier Groups: These refer to users/user groups who are 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 three 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. While an earlier infection was reported in Puerto Rico in December, we used the first infection in the continental US as a point of reference, due to the low usage of Twitter in Puerto Rico[44] (6%) compared to the entire US adult population (21%) (Pew, 2016).[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 one of nine themes. Themes were determined by blending two approaches: 1) 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 & diagnosis and symptoms, and 2) a bottom-up approach – scrutinizing the data to identify themes beyond the ones suggested by the CDC– revealed additional themes such as social issues and conspiracy theories. Having identified the themes, an additional search for different sets of keywords using Boolean operators such as AND, OR, NOT was applied. For instance, to capture the conversation about symptoms, keywords 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. While 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, will be categorized with 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 three-month timespan and across a wide range of conversation themes, patterns of user engagement often change. The aim of this step was to capture different kinds of user engagement (which we refer to as Amplification Modes) vis-à-vis the users (whom we refer to as amplifiers) who drive it. We focused on four 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 one of **9** themes (T_1 to T_9). There were thus **14 * 9 = 126** weekly datasets (W_1T_1 to $W_{14}T_9$). Within each of these datasets, we then identified up to **10** users (amplifiers) with greatest values for each of the **4** amplification modes (mentions, popular, talkers, or Twitter-wide). Each week (let's say W_1T_1) thus had a maximum of 10*4 = 40 users (amplifiers). Across 126 weeks, that worked out to a maximum of: 126 * 40 = 5040 amplifiers. Because some weekly datasets were smaller or contained fewer large amplifiers, our final sample comprised N = 4,997 amplifiers (as opposed to 5,040) distributed across the four amplification modes as follows. 1) *Mentions (n=1241)*: Refers to 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 thus classified as Amplifiers; 2) *Retweets (n=1240)*: Refers to the number of retweets (shared tweets) for each post. Users of the most shared tweets were classified here as Amplifiers; 3) *Talkers (n=1258)*: Refers to the number of tweets each user posted. Users who posted the most tweets in each subset were classified here as Amplifiers; and 4) *Twitter-wide Amplifiers (n=1258)*: Refers to top users in terms of their "Klout score" (www.klout.com). 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 4,997 amplifiers. Each amplifier was classified into one of the following 12 user categories. 1. Health institution and its individual affiliates (@CDCgov; @Who); 2. Traditional news media and affiliated journalists (@Nytimes, @reuters); 3. Online-only news media and its affiliates (@HuffPostPol); 4. Advocacy organizations (@AmnestyWomenRts); 5. Non health-specific governmental organizations (@whitehouse); 6. Other organizations - organization not falling under previous categories (@Olympics); 7. Grassroots - Regular users (Bloggers; @OlderMommyStill); 8. Politicians/Elected Officials (@HillaryClinton); 9. Nonprofit (not advocacy); 10. Academic (@naturenews); 11. Other Media; and 12. User doesn't 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 α=0.89. The sample for reliability testing was selected through a stratified random sampling procedure, comprising 125 users from each of the four amplification modes.

Results

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

Theme	**# of tweets**	**% of total**	**% retweets**
Treatment	107,236	3.2	38
Travel	234,987	7	37
Transmission	1,943,813	52	42
Testing 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

Table 1: Distribution of tweets according to Zika-related 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%). Social issues (4.1%) and testing and diagnosis (4.0%) received an almost identical amount of attention on Twitter. Prevention of Zika (3.8%) was tweeted about slightly more as compared to treatment (3.2%). Conspiracy theories were tweeted (2.2%) more than Zika-related symptoms (1.9%) which was the least tweeted theme.

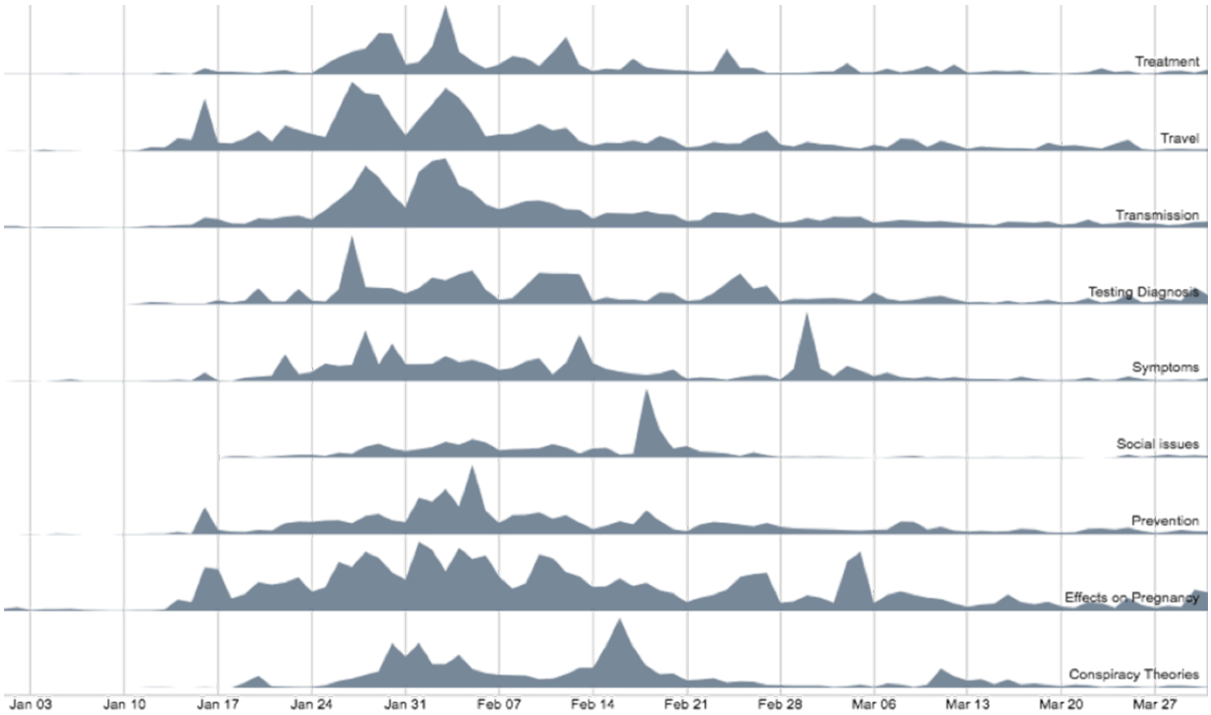

Figure 1: Temporal spread of Zika-related themes over the 14-week period

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 WHO declared Zika 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 towards the latter half of the timeline. Effects on pregnancy, however, witnessed constant attention throughout the timespan, except the last few weeks.

How did amplifier groups vary by the different modes of amplification?

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 four categories, traditional news media, grassroots users and health institutions were identified as the top three amplifier groups with most amplification ability.

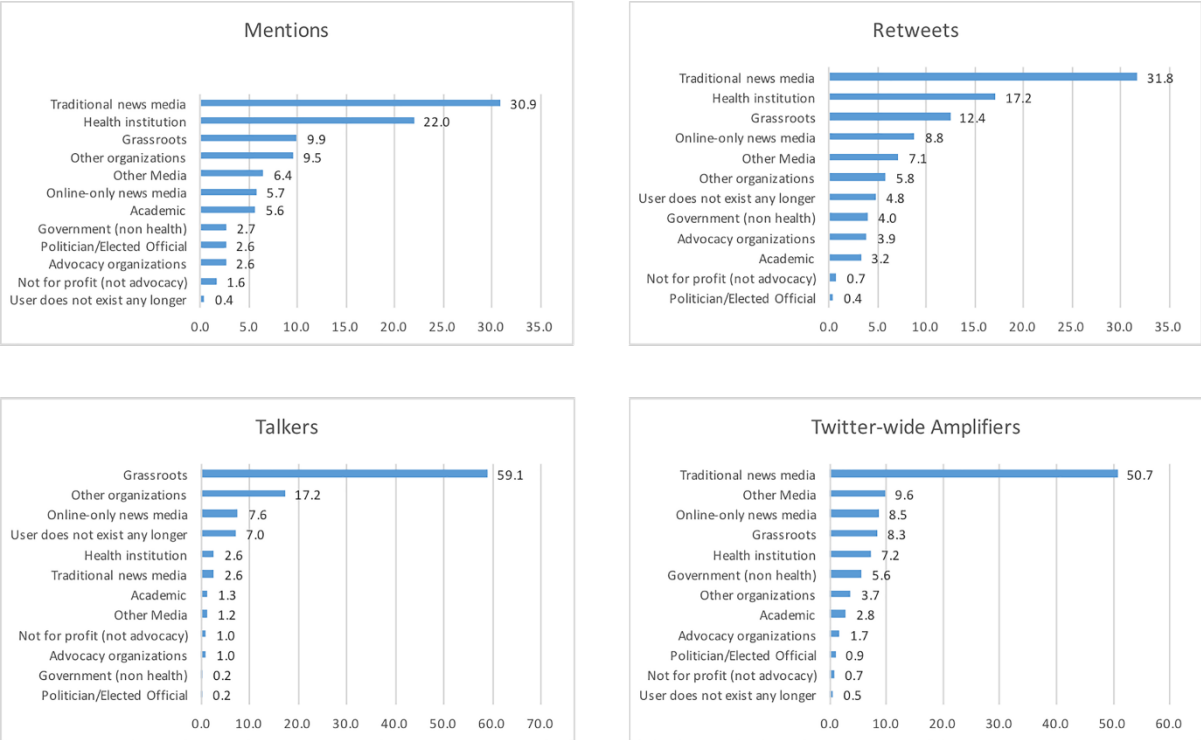

Figure 2: Distribution of different amplifier (user) groups by modes of amplification

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

We focused on the top-three groups with most amplification ability (identified in RQ2) 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 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) as compared to 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.

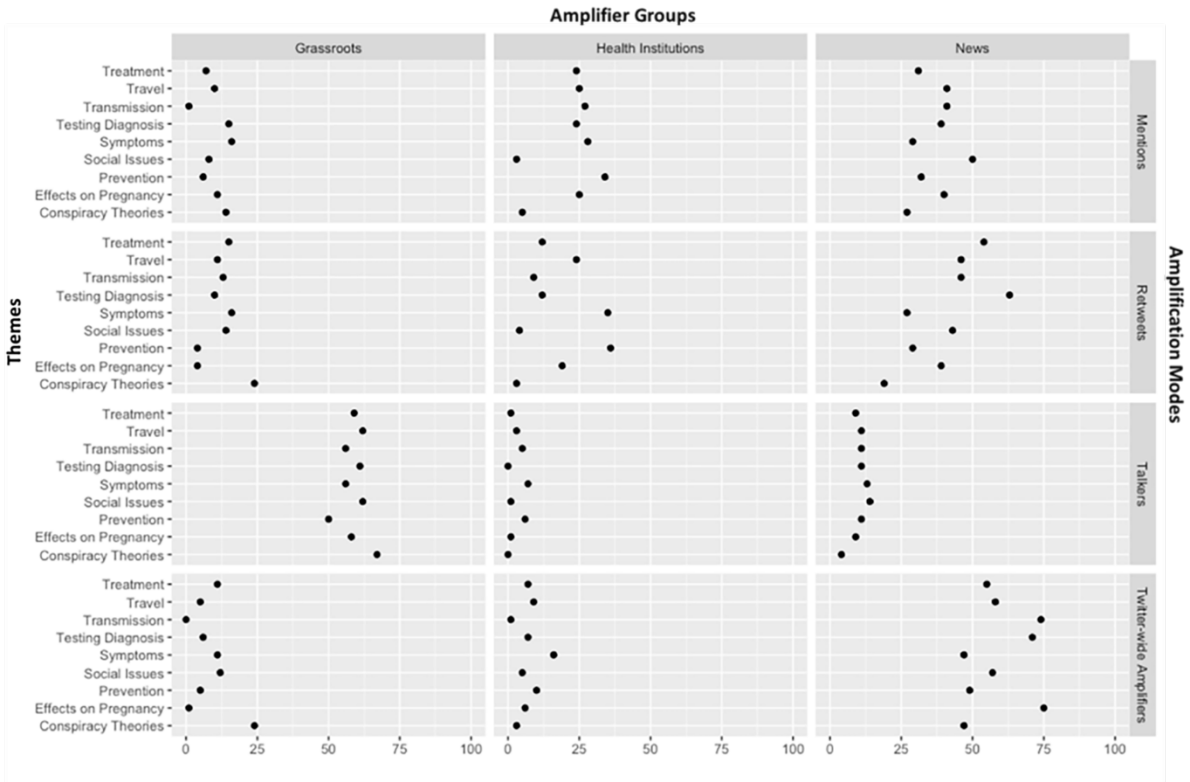

Figure 3: Distribution of different themes by top 3 amplifier groups and amplification modes

Discussion

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

RQ1 focused on the volume of tweets by specific themes and their temporal spread over the three-month period, and revealed inconsistencies between the two. Specifically, tweets with the highest volume were not necessarily ones that commanded sustained attention over time. A case in point is the theme related to the effects of Zika on pregnancy, whose volume was merely one fourth that of the most-tweeted theme (transmission), but one that witnessed sustained media chatter over the 3-month period (see Figure 2). In contrast, transmission and travel-related tweets ranked 1 and 3 by volume respectively, but chatter around these themes only increased occasionally, mainly during critical news events such as WHO's declaration of Zika as a PHEIC.[46] Other studies have found similar trends in volume of Twitter exchanges (i.e., initial increases followed by a decline after critical news events).[17,24] The results here suggest, however, that the temporal analysis performed here (as opposed to only looking at the 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 they attract media attention – can trigger tweeting. Consequently, it is recommended that they have partners ready to amplify tweets by re-tweeting and liking. Partners, for instance, include those they work with on infectious disease-related efforts in advocacy organizations, academic institutions, private sector healthcare organizations, and non-government organizations. Lastly, we recommend that public health institutions place priority on allaying the fears and anxieties triggered or potentially triggered by specific news events.

Findings related to RQ2 indicated that news media yielded the most influence on Zika-related tweets through three of the four modes of amplification, with public health institutions and grassroots users following second and third respectively. This means that despite their central role in preventing outbreaks and informing the public, 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 (2015)[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, in addition, news media also wielded Twitter-wide influence (through Klout scores) beyond chatter related to Zika. Possible explanations lie in news media informing and shaping public opinion across a range of social issues including health on a daily basis, as opposed to public health institutions' sporadic and topical engagement. However, public health institutions' role 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 RQ3 involved a thematic investigation of broad trends uncovered by RQ2, and found a profound influence of news media in amplifying all themes except prevention and symptoms, where public health institutions had greater visibility. Here, it is important to recognize that traditional news media have a more passive, as opposed to active, influence on amplification of themes. As demonstrated in Figure 5, frequency of tweets posted by news media (active influence) is relatively low. However, their passive influence is derived from the fact that a) they are mentioned to a greater volume in tweets related to those themes, b) their tweets are shared more, and c) their Twitter-wide influence is greater than other amplifier groups. Mentions and retweets indicate that users a) are either replying to tweets by the news media, b) consider news media as relevant to themes being referred to, and/or c) are actively sharing tweets posted by news media – implying greater responsiveness or engagement on part of their audience. In contrast,

while grassroots users were found to tweet with greatest frequency, they had limited impact, except on conspiracy theories, a theme likely of great concern to those managing the communication response to outbreaks.

Study findings have several implications from a policy standpoint in times when social media are integral components of public health institutions' outreach 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 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 that are 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 (like effects on pregnancy) surfaced by the 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 the potential of these to be swiftly amplified on social media.

Study Limitations: One limitation of this study is being able to analyze the entire population of tweets for RQ1, but only a sub-sample of tweets for RQs 2 and 3. This was done because categorizing tweets into themes is automated by Crimson Hexagon, while categorizing amplifier groups for RQs 2 and 3 required manual coding. Since around 3 million tweets were involved, human resource and time constraints necessitated coding a sizeable sub-sample of tweets

instead. It is possible not being able to code the entire population of tweets affected the results, but sampling is a tested approach and has been used by researchers previously for inquiries of a similar nature.[50] This study is also limited in providing an understanding of how communication between amplifier groups contributed to amplification effects. While such a network analysis was beyond the scope of this study, it would be helpful for future research to undertake such an effort.

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 the behaviors of individuals. Tweets are also identifiable and trackable, and as such, public health agencies and healthcare providers, particularly those responding to or affected by the emerging infectious disease threat, benefit from directly monitoring and analyzing Twitter data or having access to monitoring reports that do so. As shown here, social media analytics provide a sense of whether and how much attention an emerging infectious disease is garnering on Twitter; the information, beliefs, and sentiments being expressed; perceptions and potential acceptance of public health actions and advice; and misperceptions, including conspiracy theories, being shared.

As this study illustrated, 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 primary focus may be related to disease transmission and who is affected, and less so on prevention measures. Health institutions thus need to be mindful that while 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 the social conversations taking place. Grassroots users are quite visible on Twitter when it comes to EIDOs, but not so much to share transmission and prevention information. Rather, these entities and individuals are the sources and amplifiers of a range of "social"-related aspects, 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.

***Acknowledgments**: 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. 2016; 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. *Journal of Homeland Security and Emergency Management.* 2015;12(3):653-677.

3. Pidgeon N, Kasperson RE, Slovic P. *The social amplification of risk.* Cambridge University Press; 2003.

4. Centers for Disease Control and Prevention. Data and metrics. 2015; http://www.cdc.gov/socialmedia/data/index.html. Accessed September 25, 2016.

5. Organization WH. Zika strategic response framework and joint operations plan. 2016; 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.* 2015.

7. Statista. Number of monthly active Twitter users worldwide from 1st quarter 2010 to 2nd quarter 2017 (in millions). 2017; 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. *Journal of medical Internet research.* 2013;15(7):e147.

9. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection—harnessing the Web for public health surveillance. *New England Journal of Medicine.* 2009;360(21):2153-2157.

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):1.

11. Merchant RM, Elmer S, Lurie N. Integrating social media into emergency-preparedness efforts. *New England Journal of Medicine.* 2011;365(4):289-291.

12. Keim ME, Noji E. Emergent use of social media: a new age of opportunity for disaster resilience. *American journal of disaster medicine.* 2010;6(1):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. *Journal of medical Internet research.* 2009;11(4):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. *Journal of medical Internet research.* 2013;15(4):e85.

15. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. *Health promotion practice.* 2013;14(1):15-23.

16. Aramaki E, Maskawa S, Morita M. Twitter catches the flu: detecting influenza epidemics using Twitter. Paper presented at: Proceedings of the conference on empirical methods in natural language processing2011.

17. 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.

18. Fung IC-H, Tse ZTH, Cheung C-N, Miu AS, Fu K-W. Ebola and the social media. *The Lancet.* 2014;384(9961):2207.

19. Paul MJ, Dredze M. You are what you Tweet: Analyzing Twitter for public health. *ICWSM.* 2011;20:265-272.

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(5):e19467.

21. Odlum M, Yoon S. What can we learn about the Ebola outbreak from tweets? *American journal of infection control.* 2015;43(6):563-571.

22. Yoon S, Elhadad N, Bakken S. A practical approach for content mining of Tweets. *American journal of preventive medicine.* 2013;45(1):122-129.

23. Oyeyemi SO, Gabarron E, Wynn R. Ebola, Twitter, and misinformation: a dangerous combination? 2014.

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. *American journal of infection control.* 2016;44(12):1700-1702.

25. Khatua A, Khatua A. Immediate and long-term effects of 2016 Zika Outbreak: A Twitter-based study. Paper presented at: e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on2016.

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. *Emerging infectious diseases.* 2016;22(7):1320.

27. Stefanidis A, Vraga E, Lamprianidis G, et al. Zika in Twitter: Temporal Variations of Locations, Actors, and Concepts. *JMIR Public Health and Surveillance.* 2017;3(2).

28. The Guardian. Zika emergency pushes women to challenge Brazil's abortion law. 2016; 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. *Journal of health communication.* 2016;21(5):517-526.

30. Cha M, Haddadi H, Benevenuto F, Gummadi PK. Measuring User Influence in Twitter: The Million Follower Fallacy. *ICWSM.* 2010;10(10-17):30.

31. Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? Paper presented at: Proceedings of the 19th international conference on World wide web2010.

32. Bakshy E, Hofman JM, Mason WA, Watts DJ. Identifying influencers on twitter. Paper presented at: Fourth ACM International Conference on Web Seach and Data Mining (WSDM)2011.

33. McAllister-Spooner SM. Fulfilling the dialogic promise: A ten-year reflective survey on dialogic Internet principles. *Public Relations Review.* 2009;35(3):320-322.

34. Sommerfeldt EJ, Kent ML, Taylor M. Activist practitioner perspectives of website public relations: Why aren't activist websites fulfilling the dialogic promise? *Public Relations Review.* 2012;38(2):303-312.

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 Relations Review.* 2017.

36. Petrovic S, Osborne M, Lavrenko V. RT to Win! Predicting Message Propagation in Twitter. Paper presented at: ICWSM2011.

37. Fowler A, Margolis M. The political consequences of uninformed voters. *Electoral Studies.* 2014;34:100-110.

38. Reedy J, Wells C, Gastil J. How voters become misinformed: An investigation of the emergence and consequences of false factual beliefs. *Social Science Quarterly.* 2014;95(5):1399-1418.

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(38):4860-4865.

40. Weeks BE. Emotions, partisanship, and misperceptions: How anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *Journal of Communication.* 2015;65(4):699-719.

41. Bode L, Vraga EK. In related news, that was wrong: The correction of misinformation through related stories functionality in social media. *Journal of Communication.* 2015;65(4):619-638.

42. Borge-Holthoefer J, Moreno Y. Absence of influential spreaders in rumor dynamics. *Physical Review E.* 2012;85(2):026116.

43. Yoon S, Bakken S. Methods of knowledge discovery in tweets. Paper presented at: NI 2012: Proceedings of the 11th International Congress on Nursing Informatics2012.

44. Rosado L. Facebook continues to dominate social-network use in Puerto Rico. 2015; http://gaitherinternational.com/facebook-continues-to-dominate-social-network-use-in-puerto-rico/. Accessed August 1, 2017.

45. Greenwood S, Perrin A, Dugan M. Social media update 2016. 2016; 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. 2016; 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 journal.* 2016;22(3):470-478.

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. *Journal of the American Medical Informatics Association.* 2017:ocx028.

49. Allcott H, Gentzkow M. *Social media and fake news in the 2016 election.* National Bureau of Economic Research;2017.

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. *Journal of Medical Internet Research.* 2017;19(6):e193.

Appendix 1: 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 keyworks (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 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 keywrods associated with prevention (e.g., prevent, protect, avoid) or specific precautions ("long sleeve", repellant, 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.