Spreading the (Fake) News: Exploring Health Messages on Social Media and the Implications for Health Professionals Using a Case Study

Silvia Sommariva*, Cheryl Vamos*, Alexios Mantzarlis*, Lillie Uyên-Loan Đào*, and Dinorah Martinez Tyson*

*University of South Florida; ‡Poynter Institute for Media Studies

ABSTRACT
Background: The importance of social networking sites (SNSs) as platforms to engage in the correction of “fake news” has been documented widely. More evidence is needed to understand the popularity of health-related rumors and how Health Educators can optimize their use of SNSs.

Purpose: The purpose of this study was to explore the spread of health rumors and verified information on SNSs using the Zika virus as a case study.

Methods: A content analysis of Zika-related news stories on SNSs between February 2016 and January 2017 was conducted to verify accuracy (phase 1). Phase 1 was followed by an analysis of volume of shares (phase 2) and a thematic analysis of headlines (phase 3).

Results: Rumors had three times more shares than verified stories. Popular rumors portray Zika as a conspiracy against the public and a low-risk issue and connect it to the use of pesticides.

Discussion: This study identifies the value of integrating in-depth analysis of popular health-related rumors into the development of communication strategies.

Translation to Health Education Practice: Misinformation on SNSs can hinder disease prevention efforts. This study shows how information circulating on SNSs can be analyzed from a quantitative and qualitative standpoint to help Health Educators maximize the use of online communication platforms.

CONTACT Silvia Sommariva sommarivas@health.usf.edu

ARTICLE HISTORY
Received 31 January 2018
Accepted 24 April 2018

Background
Over the past decade, social networking sites (SNSs), which are Web-based services that allow users to create a profile and connect with other individuals within the system,¹ have emerged as powerful health communication platforms.²³ With more people relying on SNSs for news,⁴⁵ these platforms act as a primary bridge between individuals and news sources, aggregating traditional and nontraditional media into one convenient feed.⁶ Due to their configuration, SNSs force Health Educators to move beyond the linearity of traditional communication frameworks and rethink the role of audience members as receiver-sources empowered by the platform to amplify, modify, and generate original messages.⁷⁸ The potential of SNSs to empower patients and communities has been increasingly recognized.⁹ Because empowerment is linked to informed decision making,¹⁰ Health Educators who engage in communication and promotion efforts on SNSs are motivated to capitalize on the uniqueness of these platforms, going beyond the use of social media pages as “bulletin boards” and promoting users as distributors of accurate message.¹¹

On SNSs, users can share content regardless of accuracy and, in some cases, reach as many readers as traditional media without filtering from third parties like editors or fact-checkers.¹² An increasingly documented component impacting information-sharing on SNSs is the so-called “fake news” phenomenon, although a unified definition of fake news is being actively debated and the term has been misused often.¹³¹⁴ The Harvard Kennedy School’s Shorenstein Center defines fake news as “misinformation that has the trappings of traditional news media, with the presumed associated editorial processes,” recognizing the need for the development of a nomenclature to help scholars study this phenomenon.¹³¹⁵ The effect of fake news on public discourse has been documented extensively, predominantly with respect to the practice of journalism, the political discourse and the election cycle.¹²¹⁶¹⁷¹⁸¹⁹²⁰ Though research has analyzed the importance of SNSs as platforms to engage in
correction of inaccurate health information, more evidence is needed regarding what drives people to share fake news and which strategies Health Educators can adopt to counter this behavior. Knowledge of the extent to which users are exposed to and decide to share messages of varying levels of accuracy on a given health topic is particularly crucial during outbreaks and public health emergencies, when educators are under pressure to address people’s concerns on the risks they face and promote disease-preventive behaviors.

A recent health crisis where risk communication has been challenging is the concurrent spread of the Zika virus in the Americas during 2015 and 2016. Traditionally, health authorities have not been concerned with Zika due to its habitually mild symptoms and low mortality that result from the infection. The lack of concern changed once evidence of a possible link between the infection and neonatal malformations emerged in late 2015. Studies later showed that the spread of Zika virus may be linked to an increase in incidence of Guillain-Barré syndrome in adults as well as neurological abnormalities in newborns. Though not a chronic disease per se, the Zika virus might cause long-lasting consequences on the neurological system.

In response to immediate recommendations of the World Health Organization’s (WHO) Emergency Committee, a dramatic rise of “risk communications” occurred in the countries affected to address the concerns of the public. As a result, public health authorities launched information campaigns promoting a shift in disease prevention, such as “Drain and Cover” in the United States and “Zika Zero” in Brazil. These campaigns were targeting specific behaviors; for instance, promoting the use of mosquito repellant, disposal of stagnant water, and use of protection during sexual intercourse. Alongside communication efforts by health professionals, a wide range of Zika-centered rumors spread across SNSs. This phenomenon even led the WHO to publish a web post titled “Dispelling Rumors Around Zika and Complications” to provide guidance over the state of available evidence. Therefore, we consider the Zika virus as an ideal case study to analyze the dynamics of health information and misinformation on social media.

Methods

We used a sequential mixed-methods approach to conduct this study, with each phase of the study building on the results of the previous phase. A qualitative content analysis of news stories was conducted to determine accuracy (journalistically referred to as “fact-checking”). We followed the analysis by a quantitative analysis of news stories’ engagements. Lastly, we conducted a thematic analysis of the headlines.

Data collection

This study analyzed data obtained using the social media content analysis tool Buzzsumo, which quantifies the reach of information across several SNSs. Data include volume of engagements or shares of the most popular content (Web links) on SNSs (Facebook, LinkedIn, Twitter, Pinterest, GooglePlus) that can be searched by keyword. Thus, this study used engagements (Facebook) and shares (LinkedIn, Twitter, Pinterest, GooglePlus) as “shares” in presenting our findings. Keywords used during the search in Buzzsumo were “Zika” and “Zika virus.” We retrieved monthly data for the top 10 Zika-related stories by popularity for the period from February 2016, when the WHO declared Zika a public health emergency of international concern, to January 2017. Overall, a sample of 120 stories was analyzed (top 10 for each month over a 12-month period). Popularity was measured by the number of times a specific story’s web link was shared across the SNSs monitored by Buzzsumo. Data were collected for all English-language sources.

Fake news classification system

The nomenclature and fake news classification system developed by Wardle guided the data analysis. Subsequently, this study distinguished Zika-related news stories as belonging to three main categories. The first category is verified news stories, which
represent content backed by the latest evidence and presented accurately. The second category is rumors. The WHO has used this term to generally refer to Zika-related information not backed by evidence.\textsuperscript{38} For this study, rumors include three of the following subcategories: (1) misleading content describes news stories that inaccurately use information to frame an issue or an individual (this includes Wardle’s categories of misleading content, false context, manipulated context); (2) false connection (similarly to Wardle’s category) is the mismatch between headline and content, where the content may be accurate but the headline is deceiving; and (3) fabricated content indicates material that is completely fake and not backed by evidence (similar to Wardle’s category). The third category Wardle developed is satire or parody, which are stories aimed at producing a comic effect. Wardle also defined the category “imposter content,” which refers to the attributes of the source of the news story rather than the content of the news story itself. This category was excluded because an in-depth analysis of the sources that produce fake news is beyond the scope of this study.

\textbf{Data analysis}

\textbf{Phase 1}

After the sample was collected, the content of Zika-related news stories were analyzed to determine accuracy. News stories were categorized as (1) “verified,” (2) “rumor” (which was subcategorized as “misleading content,” “false connection” or “fabricated content”), and (3) “satire.” To determine accuracy, the content of stories was compared to the evidence made available by official health authorities reporting. One such document was the WHO’s “Dispelling Rumors Around Zika and Complications” article.\textsuperscript{38} This fact-checking process was carried out by one author (SS), and a random sample (20\% of news stories) was checked by a second author (AM). Both co-authors are experienced fact-checkers, formerly employed by professional fact-checking projects. Intercoder reliability for the fact-checking process of news stories’ content was 100\%.

In this phase, the two authors also coded the sources of the news stories using the following nomenclature: “legacy media,” “digital media,” “alternative media” and “scientific/institutions.” Media outlets that were nondigital at launch (eg, broadcast, print) were categorized as “legacy media” and news sources that launched online and exclusively publish on that medium were categorized as “digital media” or “alternative media.” The difference between the two is that digital media maintain an editorial structure and verification process similar to that of print newspapers, whereas alternative media are mostly run as individual or collective blogs. Scientific institutions, governmental authorities, and research centers were classified as “scientific/institutions.”

\textbf{Phase 2}

A descriptive quantitative analysis of the volume of shares per news story was conducted. Volumes were analyzed by source and type of news story. Analysis of the volume of shares was conducted using Microsoft Excel. The analysis was conducted on annual data (February 2016–January 2017) and monthly volumes.

\textbf{Phase 3}

Finally, a thematic analysis of the headlines of all news stories was performed to explore the characteristics that drive stories’ popularity on social media. The analysis was performed on headlines only, rather than on the entire story content, because previous research has shown that users often share links they have not opened. For example, one study on social clicks on Twitter found that more than half (59\%) of the links are shared by users without being opened.\textsuperscript{40} All headlines of sampled news stories were coded using a codebook. The codebook was developed iteratively, composed of a priori and emerging codes.\textsuperscript{41,42} For each code, the codebook detailed the inclusion and exclusion criteria, a description of the code, and typical exemplars and atypical exemplars of news stories that would fall under the code “close but no” (indicating material that could be categorized mistakenly under that code) and subcodes.\textsuperscript{43} One author (SS) coded all of the headlines and two co-authors (DMT, CV) cumulatively coded a random sample of 20\% of the headlines. Analysis of shares was subsequently conducted by coding for verified stories and rumors to explore the characteristics driving stories’ popularity and neglected aspects of the Zika crisis. Intercoder reliability for the coding process was 91\%.

\textbf{Results}

\textbf{Popular Zika-related news stories}

A total of 120 news stories were analyzed. The peak of interest in Zika-related news stories during the study period occurred following the declaration of a public health emergency of international concern by the WHO in February 2016. During the year, the reach of Zika-related news stories declined, with a slight rise over the summer period in conjunction with the Brazil Olympics (which raised concerns around travel safety) and congressional approval of
US$1.1 billion in funding to fight the Zika virus (Figure 1).

Alternative media sources produced the largest proportion (66%) of the 120 most popular news stories on SNSs. Legacy media produced 25% of the top content, followed by digital media (6%) and scientific organizations or institutions (3%). On average, alternative media had the highest reach of news stories (44 673 shares per story), followed by digital media (36 340 shares per story), legacy media (12 482 shares per story), and scientific organizations or institutions (9656 shares per story). Table 1 reports the top 10 Zika-related news stories by number of shares over the course of the study period (February 2016–January 2017).

Prevalence of the fake news phenomenon in social media discourse around Zika

Of the 120 total news stories analyzed, 27 were categorized as rumors, 92 were verified news stories, and one story (“CDC Warns Man Buns Harbor Zika Virus”) by

Table 1. Top 10 most popular Zika-related news stories on social networking sites for the period February 2016–January 2017.

<table>
<thead>
<tr>
<th>Rank</th>
<th>News story title</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Larvicide Manufactured by Sumitomo, Not Zika Virus, True Cause of Brazil’s Microcephaly Outbreak</td>
<td>Tech Times</td>
</tr>
<tr>
<td>2</td>
<td>Millions of Bees Dead After State Sprays for Zika Mosquitoes</td>
<td>US Uncut</td>
</tr>
<tr>
<td>3</td>
<td>Argentine and Brazilian Doctors Name Larvicide As Potential Cause of Microcephaly</td>
<td>GMWatch</td>
</tr>
<tr>
<td>4</td>
<td>It’s Not the Zika Virus—Doctors Expose Monsanto Linked Pesticide As Cause of Birth Defects</td>
<td>The Free Thought Project</td>
</tr>
<tr>
<td>5</td>
<td>The Three Letter Word Missing From the Zika Virus Warnings—Men</td>
<td>The Guardian</td>
</tr>
<tr>
<td>6</td>
<td>India Has Developed The World’s First Vaccine for Zika, the Deadly Virus Terrifying the World</td>
<td>India Times</td>
</tr>
<tr>
<td>7</td>
<td>Pope Suggests Contraceptives OK to Slow Zika</td>
<td>CNN</td>
</tr>
<tr>
<td>8</td>
<td>Zika Virus Not to Blame? South American Doctors Groups Propose Man-Made Cause for Birth Defect Epidemic</td>
<td>Second Nexus</td>
</tr>
<tr>
<td>9</td>
<td>FSU Research Team Makes Zika Drug Breakthrough</td>
<td>Florida State University News</td>
</tr>
<tr>
<td>10</td>
<td>Zika Virus Now in Metro Manila</td>
<td>Trending news portal</td>
</tr>
</tbody>
</table>

Table 2. Number of news stories by code over total category (rumors and verified).

<table>
<thead>
<tr>
<th>Place/location</th>
<th>Rumor (%)</th>
<th>Verified (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevention</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Action</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>No Risk</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Fear</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Blame</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Scientific discoveries</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Current events</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Conspiracy</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>Epidemiology</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Spreading dynamics</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Politics</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Health</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Actor</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Organization</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Women</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Men</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Vague</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Policy</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Pesticide</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
the medical satirical website GomerBlog was categorized as satirical/parody. Given that only one story was categorized as satirical, the analysis focused on two categories: rumors and verified news stories. Among the rumors, fabricated content represented the largest share (81%), followed by misleading content (16%) and false connection (3%). Rumors obtained, on average, three times more shares than verified stories, as shown in Figure 2.

As illustrated in Figure 3, the proportion of rumors dropped substantially over the course of the period considered. The decline in reach of rumors was concurrent with the overall decline in shares of Zika-related news stories. As users’ interest in Zika faded, verified stories represented a higher proportion of the overall number of shares for Zika-related news stories.

Alternative media had the highest performance in terms of average reach of news stories, followed by digital media, legacy media, and scientific organizations or institutions. For alternative and legacy media, news stories classified as rumors were shared more often than verified stories (see Figure 4). All news stories from digital media and scientific institutions were verified.

### Popularity of news stories

Examining the number of stories per code (Table 2), most rumors portrayed Zika as a conspiracy against the public (approximately 20% of the total) or as a low-risk issue (approximately 15%) or discussed the role of pesticides in the epidemic (about 10%). Headlines of
verified stories tended to cover issues such as the spread of new cases of the virus (approximately 15%) and disease prevention or scientific discoveries (5%). Fifteen stories explicitly referred to women and two stories addressed concerns related to men’s health.

When considering the average number of shares by code (Table 3), the most popular rumors had headlines that covered issues such as blame (often associated with the actor or organization to be blamed) and pesticides. Headlines from the most popular verified stories also covered the role of pesticides in the crisis and, to a lesser extent, issues around disease prevention.

### Discussion

This study explored the spread of health information through SNSs, highlighting the role of the fake news phenomenon in Health Education practice. Fake news is not a concern unique to SNSs; however, our analysis showed that misinformation finds fertile soil in the fast-paced social media ecosystem, where the abundance of news sources and SNS platform structures can help misinformation reach a large audience. In the case study used in this article, the prevalence of health-related fake news in the social media landscape is substantial and cannot be disregarded. For instance, this study found that among the top 10 news stories about the Zika virus in 2016 (Table 1), half were classifiable as rumors. Data in this analysis also suggested a positive relationship between the popularity of a topic and the appearance of fake news related to that same topic. In fact, while Zika was a popular topic on SNSs such as Facebook and Twitter, rumors such as the connection between larvicides and microcephaly and the alleged contribution of the mass release of sterilized male mosquitoes to the spread of Zika in Brazil were circulated widely despite having been already disproven by fact-checkers. Among the top stories, shares of rumors outnumbered shares of verified stories by four to one.

These findings underscore the need for Health Educators to analyze misinformation and develop and implement information campaigns in the very early stages of social media strategy development. Early intervention can be beneficial in two ways. One, promptly exposing users to correct information can trigger social correction, which has been shown to be a viable strategy to counter misinformation. Monsanto, the multinational agrochemical corporation, adopted this strategy to counter a Zika-related rumor concerning its pesticides (see story 4 in Table 1). Second, an early understanding of rumors rising in popularity provides SNSs with information that can be used to flag misleading stories and reduce their visibility by adjusting algorithms. SNSs, most notably Facebook, recently have introduced strategies to reduce the visibility of fabricated information on their platforms by building on the work of fact-checkers and experts.

Furthermore, the content analysis of news stories conducted in this study suggests that moving beyond the mere analysis of volumes of shares and engagements is crucial to gaining an in-depth understanding of popular rumors and developing effective social media health communication strategies. For instance, several fabricated stories portrayed Zika as a low-risk issue; however, this type of framing did not appear in verified stories. Such framing represents a threat to the implementation of disease prevention efforts and prevention behaviors, because the virus may not be perceived by the public as a public health priority. This understanding could prompt Health Educators not only to disseminate content that refutes the “low-risk” view but also to try to meet the information needs of users by promoting accurate content on the risks of Zika. Moreover, information on which topics have been neglected in the social media debate would prove helpful to redirect communication efforts toward less-discussed issues. In our study, for instance, few verified news stories explicitly addressed the impact of Zika on men’s health, which suggests the need for Health Educators to focus on risk-related messages that target the male population.

Research on social media health information campaigns shows that it cannot be presumed that communication efforts on SNSs will necessarily reach a large audience just because these platforms have a wide user base. In some cases, SNS campaigns have proven to be...
ineffective.\textsuperscript{49-51} The ineffectiveness can be due to the overload of competing messages that circulate on social media,\textsuperscript{52} which can reduce visibility of public health information. Rumors are part of this highly competitive landscape. Misinformation is a by-product of SNSs, where health campaigns face a constant trade-off between reach and control.\textsuperscript{11} This trade-off is particularly true when health professionals share information on unfolding crises and events, where evidence is constantly updated and information circulates at a faster pace.\textsuperscript{53} Therefore, it can be hard to distinguish relevant information from background noise or fabricated information.\textsuperscript{54}

Inherent challenges exist for Health Educators and other public health professionals in addressing the fake news phenomenon. Pretesting messages is an important part of health communication practices.\textsuperscript{55,56} Pretesting may inject a time lag between the spread of rumors and the diffusion of messages that address health-related misinformation. Still, integrating early analysis of rumors as a formal step in the development of communication strategies on social media could help Health Educators understand competing messages on SNSs and increase the reach of verified information among users.

Despite providing insights on the spread of health messages through SNSs, this work presents several limitations. First, the data included news stories published by different media sources as separate news stories, even when the topic covered by the story was the same. For instance, if two media outlets published two articles reporting the same rumor on Zika and vaccines, they were counted twice. Though this approach did not affect our findings on the average volume of shares, it may have skewed our results on the popularity of thematic codes. Second, though informed by the current debate on fake news,\textsuperscript{14} the categorization used to describe this phenomenon needs to be further validated in the academic literature. Still, the categorization used in this article presents a systematic way to operationalize stories for analysis. Third, the decision to conduct the content analysis on news headlines instead of full articles was based on an assumption (that the content of the headline was the main driver of popularity) proven in one study\textsuperscript{40} but on which academic evidence is still limited. Fourth, the use of Buzzsumo is largely untested in the academic realm; therefore, potential limitations to its application may emerge. Moreover, among all SNSs considered, Facebook presents a significantly higher volume of shares; thus, our findings on fake news as a phenomenon were largely related to this platform.

Results from this analysis provide relevant insights on SNSs as platforms where Health Educators can disseminate accurate information. These findings can be of interest for public health professionals who want to further leverage SNSs to promote disease prevention. Though the centrality of users as drivers of content on SNSs is well understood,\textsuperscript{57,58} this study proposed and identified the value of integrating an in-depth analysis of popular rumors into the development of Health Education strategies.

Translation to Health Education Practice

Social networking sites represent an opportunity for Health Educators to promote user engagement and empowerment with respect to important public health issues and increase reach to target populations, particularly among vulnerable groups.\textsuperscript{59} However, the wider diffusion of health-related messages is not necessarily positive if the quality and accuracy of information are poor.

In this context, a first set of implications for health educators relates to the need for an understanding of popular fake news content circulating on core public health issues that extends beyond the quantitative study of volume of reach and engagements to provide in-depth qualitative understanding. Knowledge of rumors circulating on a given topic represents key information that can be incorporated into the development of messages. For instance, knowing that popular misinformation focuses on portraying Zika virus as a low-risk issue for men might lead institutions informing the public to consider tailoring their communication strategy in order to reach male audiences with messaging on risk perception.

A second implication relates to the potential for public Health Educators and promoters to engage with SNSs in improving the reach of accurate information. In particular, during public health crises, health professionals might consider working with professional fact-checkers in the context of their partnership with Facebook to flag and reduce the reach of false content.

Finally, the fast-paced spread of misinformation allowed by SNSs calls for fast-tracking testing processes for health messages and educational materials. SNSs can, in that sense, serve as the cause and the cure for this problem. For instance, public health professionals should leverage SNSs tools to test messages iteratively by using split testing, which allows to obtain immediate feedback from users on which messages work best.

Ultimately, fighting misinformation on social media is key to the promotion of media literacy, defined as the ability to critically evaluate health messages.\textsuperscript{60} Social media can be instrumental to promote health literacy.\textsuperscript{61} Promoting the integration of social media in health literacy strategies can, however, be ineffective if not
accompanied by an awareness of the magnitude and characteristics of the misinformation problem. This study shows how health-related misinformation can be analyzed from a quantitative and qualitative standpoint to help Health Educators maximize the use of SNSs as communication platforms.

**Conflict of interest**
The authors report no conflicts of interest.

**ORCID**
Dinorah Martinez Tyson @ http://orcid.org/0000-0002-9390-5364

**References**


