ORIGINAL RESEARCH

Quantifying Network Dynamics and Information Flow Across Chinese Social Media During the African Ebola Outbreak

Shihui Feng; Liaquat Hossain, PhD; John W. Crawford, PhD; Terry Bossomaier, PhD

ABSTRACT

Objective: Social media provides us with a new platform on which to explore how the public responds to disasters and, of particular importance, how they respond to the emergence of infectious diseases such as Ebola. Provided it is appropriately informed, social media offers a potentially powerful means of supporting both early detection and effective containment of communicable diseases, which is essential for improving disaster medicine and public health preparedness.

Methods: The 2014 West African Ebola outbreak is a particularly relevant contemporary case study on account of the large number of annual arrivals from Africa, including Chinese employees engaged in projects in Africa. Weibo (Weibo Corp, Beijing, China) is China's most popular social media platform, with more than 2 billion users and over 300 million daily posts, and offers great opportunity to monitor early detection and promotion of public health awareness.

Results: We present a proof-of-concept study of a subset of Weibo posts during the outbreak demonstrating potential and identifying priorities for improving the efficacy and accuracy of information dissemination. We quantify the evolution of the social network topology within Weibo relating to the efficacy of information sharing.

Conclusions: We show how relatively few nodes in the network can have a dominant influence over both the quality and quantity of the information shared. These findings make an important contribution to disaster medicine and public health preparedness from theoretical and methodological perspectives for dealing with epidemics. (*Disaster Med Public Health Preparedness.* 2018;12:26-37)

KeyWords: network dynamics, information flow social media, Ebola outbreak

Social media is a group of Internet-based applications that builds on the foundation of web 2.0 and that allows the creation and exchange of user-generated content.[1] In 2012, a survey by the Pew Research Center's Internet & American Life Project showed that 67% of Internet users use social networking sites and that women are more likely to use social media than men.[2] The user demographics suggest that most users of social media are under 50, and the majority are aged between 18 and 29 years. Those who live in urban areas are more likely than rural people to use social media.

The role of social media or web-based surveillance in disaster preparedness and response in large populations[3,4] is now regarded as an important area of research and practice, which could transform disaster response worldwide.[5-8] Crowd-sourcing through social media has potential for accessing locally situated information from the public in support of public health emergency preparedness and response.[8] A study of the adoption and use of social media among state public health departments suggested that the use of social media by state health departments is 60% (n = 30), whereas 87% had a Twitter (Twitter, San Francisco, CA) account, 56% had Facebook (Facebook, Menlo Park, CA), and 43% had been using YouTube (YouTube, San Bruno, CA) to promote health behaviors.[9] For example, the Twitter network was instrumental in providing information to the public about the availability in specific locations of vaccine against H1N1 during the 2009 influenza pandemic in Virginia.[8] Similarly, during the 1995 Chicago heat waves, a web-based buddy system allowed at-risk people to receive medical attention and social services in a timely manner.[10] As the world's most populous country, China faces particular challenges in dealing with emerging epidemic outbreaks in a sufficiently timely manner to safeguard the population from the deadliest local and global threats.

Ebola virus disease (EVD) is a life-threatening disease transmitted via blood or body fluids from infected patients or contaminated equipment. The outbreak of

EVD began in Guinea in March 2014 and developed strong foci in Guinea, Sierra Leone, and Liberia. The disease spread rapidly, affecting many countries in West Africa. According to the World Health Organization (WHO), as of December 7, 2014, 6388 people of the 17,942 confirmed infections had died in 7 countries around the world. This was by far the deadliest outbreak of EVD to date and aroused serious concern for imported infections in China as the number of annual arrivals from Africa and Chinese employees for cooperation projects in Africa is significant. On July 30, 2014, a Hong Kong woman returning from Kenya exhibited symptoms that were similar to EVD, and this suddenly focused the attention of the government and citizens of China who, up to that point, had regarded EVD as a distant threat.

In this study, we quantified the Chinese public's response to the African Ebola epidemic by analyzing relevant data from Weibo (Weibo Corp, Beijing, China), which is the most popular social media platform in China with over 100 million daily posts. The content of posts that relate to EVD is an important real-time barometer of public attitudes to an emerging crisis, and the capacity to follow the evolution of these attitudes over time is unique. Depending on the outcome of such analyses, the results can provide information on public attitudes, the quality of information being shared, and the span of public attention. Perhaps most importantly, the results may provide key insights into how best to ensure the public can keep themselves informed and be prepared to act appropriately in such circumstances.

To explore this potential, we first retrieved Weibo messages containing the key word "Ebola" and quantified the dynamics of the number of microblogs from March 2014 to March 2015. Second, we filtered those microblogs that had more than 100 comments and reposts and classified the posters of these microblogs as "influential propagators." By analyzing the content generated by influential propagators within the specified timeline, we can identify the major topics relating to Ebola that the public are concerned about and explore the reasons underlying any observed fluctuations in public attention to the issues. Third, to gain insight into the structure and evolution of communication channels during the outbreak, we analyzed the contact network for participants contributing to the comments section under the postings of one influential propagator. These analyses helped us to develop a better understanding about the usefulness and timeliness of social media networks in promoting public health awareness for preparedness and informed response to an emerging crisis.

METHODS
Social Media in Disease Surveillance
Online social media platforms have developed rapidly in the past decade. Contrary to traditional media, the interactivity between users of social media and the diverse content that is generated make these platforms an important potential source of information for disease surveillance relevant to early detection, improved preparedness, and a more effective response. More importantly, the huge quantity of time-sensitive data can serve as pool of information for conducting surveillance of rapidly emerging events. As a consequence of these potential advantages, several studies have evaluated different aspects of the use of social media in disease surveillance.

Digital disease surveillance is a recent advancement in early detection of infectious diseases or in research on other diseases (eg, mood disorders). This form of surveillance has branched out from Internet-based technologies in earlier years to mobile technologies.[11,12] Over the last 2 decades, current developments have gone beyond human health with applications extending to even animal health and monitoring of vector-borne diseases, transport modeling, as well as spatiotemporal mapping of unusual health events harnessing information on the Internet.[13-16] Digital disease surveillance is based on the extraction of data through formal (ie, government controlled and sanctioned information) and informal electronic information to identify abnormal or unusual events that may occur in the community, in a given population, or in a geographic locality during a specified period. Examples of Internet-based usage are mostly informal, such as the use of search query data,[17,18] online news,[19,20] and social media and networks,[17,20,21] whereas mobile technologies include formal data such as call detail records, which identify the location of the original call made, and even smart phones with global positioning systems (GPS).[12,15] The concept of online disease surveillance has recently experienced significant realizations and progress. Various online syndromic surveillance methods and models are being continuously developed, assessed, and improved with data from (1) search engine queries (such as Google queries), (2) social network blogs (eg, Twitter microblogs), and (3) health website accesses (eg, Wikipedia visits).

Twitter, as one of the most widely used and accessible microblogging services, is a major focus of studies on social networks. Most disease surveillance on social networks is based on Twitter data.[21] For example, Lampos et al[22] analyzed the content of Twitter feeds to track the prevalence of influenza-related illness in regions of the United Kingdom and presented an automated tool with a web interface to trace updated flu-score data from Twitter and official data. Signorini et al[21] tracked public reaction to H1N1 flu virus and showed that Twitter can be used to measure elements of public concern about health-related events. Aramaki et al[23] attempted to use Twitter to detect the incidence of flu. The high correlation between predicted and realized outcomes showed that Twitter texts have the potential to reflect the actual incidence of events and that natural language processing (or NLP) technology can be used for Twitter content. Using NLP to convert human language to machine-readable format has the potential to inform public health preparedness for targeted and timely intervention, which has been underutilized for improving public health preparedness.[24] Notwithstanding the promise of social media

approaches, relatively few studies have been conducted of the potential for Sina Weibo to support public health outcomes in China, despite the large number of users in China.

Sina Weibo was launched in September 2009 and is a microblog platform analogous to Twitter. Users can write posts of up to 140 characters attaching photos, videos, and emoticons and locating geographical information. Posts generated by users on Weibo are publicly accessible for viewing and interactions such as commenting, reposting, or liking. In a short time, Weibo has become a widely used and influential platform, largely because of easy accessibility and wide and rapid connectivity. The subsequent development of support for mobile applications also boosted the popularity of the platform. Weibo consists of 60% male users and 40% female users, with 40% of the users in the age group of 21 to 25 years and 32% in the age group of 26 to 30 years.[25] Various activities are conducted by Weibo users, including self-expression, contacting friends, following celebrities, and participating in interest groups or social activities. Information diffusion is a powerful function of the Weibo platform with plenty of strong and weak ties among users. A comparative study of user behavior and motivation of Weibo and Twitter suggest significant differences as well as new insight for multilingual and culture-awareness of user modeling.[26]

Current research using the Sina Weibo platform mainly focuses on the large number of user-generated content such as topic mining, sentiment analysis, and user classification. Topic mining deals with the use of text clustering algorithms or probability models to identify topics from short texts automatically, which can be used to detect spam information. Sentiment analysis is an important methodology for extracting user attitudes and feelings during an event, which is frequently used to monitor public opinion and analyze user preference in Weibo. User classification is used to understand users' behavior and preferences. Although these analysis methodologies do not specially focus on the application of disease surveillance, they provide insight into techniques that could work and some basic understanding of information diffusion and the potential for content analysis.

Relatively few studies have been done on the potential for analyses related to public health using Weibo. Fung et al[27] analyzed public awareness of H7N9 using Weibo data and demonstrated that Weibo can be a useful platform for exploring public reaction to disease outbreaks in China. The development of approaches to topic classification of microblogs provide us with the tools for classifying the social media data.[28] Cui et al[29] classified Weibo content to predict the outbreak of flu by improving on Chinese text processing methods. Wang et al[30] further identified health-related topics on Weibo by applying probabilistic topic models that can characterize each topic by a probability distribution over words by the Latent Dirichlet Allocation (LDA) model and further identify health topics by the Aliment Topic Aspect Model (ATAM).

Because the messages on Weibo are all in Chinese (only 1.6% of users are from other countries),[31] the developed applications of text analysis for English language are not applicable. Language is a major hindrance in transferring research results and methodologies from Twitter to Weibo.

Our research focuses on the development of new approaches specifically for use with Weibo data as a proof of concept for supporting public health awareness and response in China. We use public awareness of Ebola as a case study and identify the topics that concerned the public over time, identify the evolution of structures of the social networks, and quantify the role that influential propagators played in the spread of information. We believe that identification and exploration of the role of influential propagators (ie, who is able to influence others with their posting and comments about the subject) can be an important step in promoting public health awareness for improving our efforts in coordinating disaster medicine and public health preparedness.

Weibo Data Collection and Analysis

The choice of appropriate key words that relate to particular public health preparedness is a critical first step to ensure that useful information is extracted from posts for use in subsequent analyses. *Recall* is defined as the result of the number of documents that are relevant to the key words divided by the total number of documents that are relevant to the topic.[32] To explore the key words that can represent the response to EVD clearly and accurately, we downloaded the document "Frequently asked questions on Ebola virus diseases" from the website of the World Health Organization (WHO),[33] which describes 13 common questions about EVD, including the definition, infection reasons, symptoms, and preventive methods. We extracted key words from this document by inputting the document into the Natural Language Processing to Information Retrieval (NLPIR) system to extract key words automatically and calculated weighted values. In this analysis, 9 was the frequency and 14.71 (discussed below) was the weight value; the higher the weight value, the greater the importance of the key word.[34] Table 1 presents the top 5 key words.

As shown in Table 1, the highest-weighted key words contained "埃博拉" (Ebola). Thus, we can infer that the

TABLE 1

Top 5 Key Words in the Extraction Output		
Key Word (in Chinese)	**Key Word (in English)**	**Weight**
埃博拉出血热	Ebola hemorrhagic fever	14.71/9
埃博拉病毒	Ebola virus	14.31/16
埃博拉	Ebola	11.84/38
医务人员	Medical staff	9.21/13
公共卫生	Public health	7.92/7

search results using the key word *Ebola* can capture the content searched for by using the other 2 key words. We also tried searching content in Weibo by using only the key words "medical staff" or "public health," but the search results were mostly about common medical issues rather than the Ebola event. The public response to EVD differs from that to, for example, influenza H7N9, primarily because there was not any confirmed case of EVD in China. Therefore, people who talk about influenza may use various words to describe or express the symptoms instead of mentioning the disease name directly, whereas for EVD, people usually refer to the name directly in Weibo content. On the basis of the above analysis, we decided to choose "埃博拉" (Ebola) as the key word to retrieve related content in Weibo. We then searched the data by using this key word covering postings made by users between March 2014 and March 2015 and analyzed the posts that were retrieved. The total volume of Weibo posts retrieved in this way was 5,120,376 for the year, with an average weekly volume of 89,831. However, Weibo can only display 50 pages equaling 1000 records per search, which meant that we could collect only a subset of all Ebola-related messages for use in further NLP processing. To account for this, we used relative frequency, rather than absolute numbers of outputs returned, and assumed that the 1000 records returned per search were statistically representative of all relevant records. For each of the retrieval posts, we collected data on the information in each post and on the number of comments and reposts.

Observation of Trends

The weekly number of posts related to Ebola was collected throughout the whole period to quantify and visualize trend in the volume of messages. This reveals the change tendency in public attention and displays an overview of the public reaction to Ebola in the social network. On the basis of the trend analysis, we could also identify the peak points and analyze trends in the number of different "kinds" of users. For example, research published in the Chinese literature by Li and Liu classified Weibo users into 3 categories: observers, active participants, and propagators.[35] Observers mainly read information in Weibo and are passive information recipients. Active participants usually add comments or add links to their contacts. Weibo posts released by propagators have, by definition, a large number of comments and reposts. Propagators play an important role in the transmission of information that can be used to disseminate public health alerts to a wider population. In this research, we collected the number of propagators in each week to examine the correlation between the number of posts and the number of propagators. This was based on the collected data for examining the relevance between the number of microblog messages and the number of propagators.

Key Word Extraction

The aim of key word extraction was to discover the topics that the public was particularly concerned about at certain points in time and to discover the reasons for changes in public attention. Key words are succinct summary words or phrases that characterize widely shared content, which can provide a way to explore public awareness of specific epidemics so that public health preparedness can support the development and implementation of targeted and timely campaigns during outbreaks. The NLPIR system processed key word extraction and calculated weight values after conducting word segmentation. Word segmentation is the process of cutting Chinese character sequences into single words. Unknown words and word ambiguity are the 2 major problems that can affect the accuracy of segmentation.[36] The NLPIR system has strong performance on Chinese word segmentation,[34] and the key word *Ebola* was used for Weibo messages to extract relevant data, which was fundamental to our analysis of public awareness of Ebola outbreaks for improving public health preparedness. The posts generated by propagators are widespread and influential. According to the statistical distribution of the postings, we defined a propagator as someone whose posts have more than 100 comments and reposts. For "peak traffic events," we analyzed the hot topics based on the posts of propagators. By analyzing the messages in this way at specific time points, we were able to uncover trends in the topics that the public considered to be important about Ebola.

Social Network Analysis

Social network analysis has been found useful in describing, exploring, and understanding structural and relational aspects of health in a number of studies, in which the ecological questions in public health can be investigated by using the theoretical and methodological principles of social network analysis.[37] Besides the content, the link structure in the social network can also be analyzed to provide important additional information about how information is being shared about Ebola outbreaks. Social network analysis can be used to observe the dynamic relationships among active participants during the transmission of key events on Weibo, which we believe can be used to develop effective and timely dissemination plans for targeted populations who could be at risk. As a partial network to analyze the characteristics and evolution of network structure, we chose the network associated with the "People's Daily," because it is one of the most influential media bloggers in China. We analyzed properties of this network from July to November 2014. This was due to "peak traffic events," in which we selected 5 messages posted by the "People's Daily" from July to November 2014 as partial networks for analyzing dynamic changes on topics and participant interactions. In particular, we quantified network density and network centrality. Network density is important to disaster medicine and public health preparedness because it allows us to measure the connectivity of a network, which can be calculated by dividing the number of actual connections by the number of possible connections:

$$2L/n(n-1)$$

where L denotes the number of connections and n denotes the number of nodes for an undirected graph. For a directed graph (connections operate only in a specified direction of attachment from one node to another), the corresponding formula is

$$L/n(n-1)$$

A higher density implies a higher degree of interaction among individuals, which can serve as a scaffold on which to develop target interventions and create an information dissemination plan for improving public health preparedness. Network centrality identifies the number of connections a node has and can help to identify the most influential node in a social network,[38] which can help public health planners find the most effective and optimal pathways by

which to disseminate information to a wider population about specific outbreaks. Freeman[39] proposed 3 centrality measures: (1) the degree centrality of a node, which is calculated as the number of links that the node has; (2) the betweenness centrality, which quantifies the number of times a node is included in the shortest path between all other pairs of nodes in the network and which is a measure of the potential of an individual to play a brokerage role in the network; and (3) closeness, which is the reciprocal of farness, where *farness* is defined as the sum of the shortest distances of the node from all other nodes in the network (ie, the larger the closeness centrality of a node, the lower its distance to all other nodes). By analyzing the network density and network centrality, we can quantify important aspects of the nature of interaction of participants

FIGURE 1

(A) Weekly Number of Weibo Posts About Ebola From March 2014 to March 2015. (B) Daily Number of Weibo Posts About Ebola During the Week of July 26 to August 1.
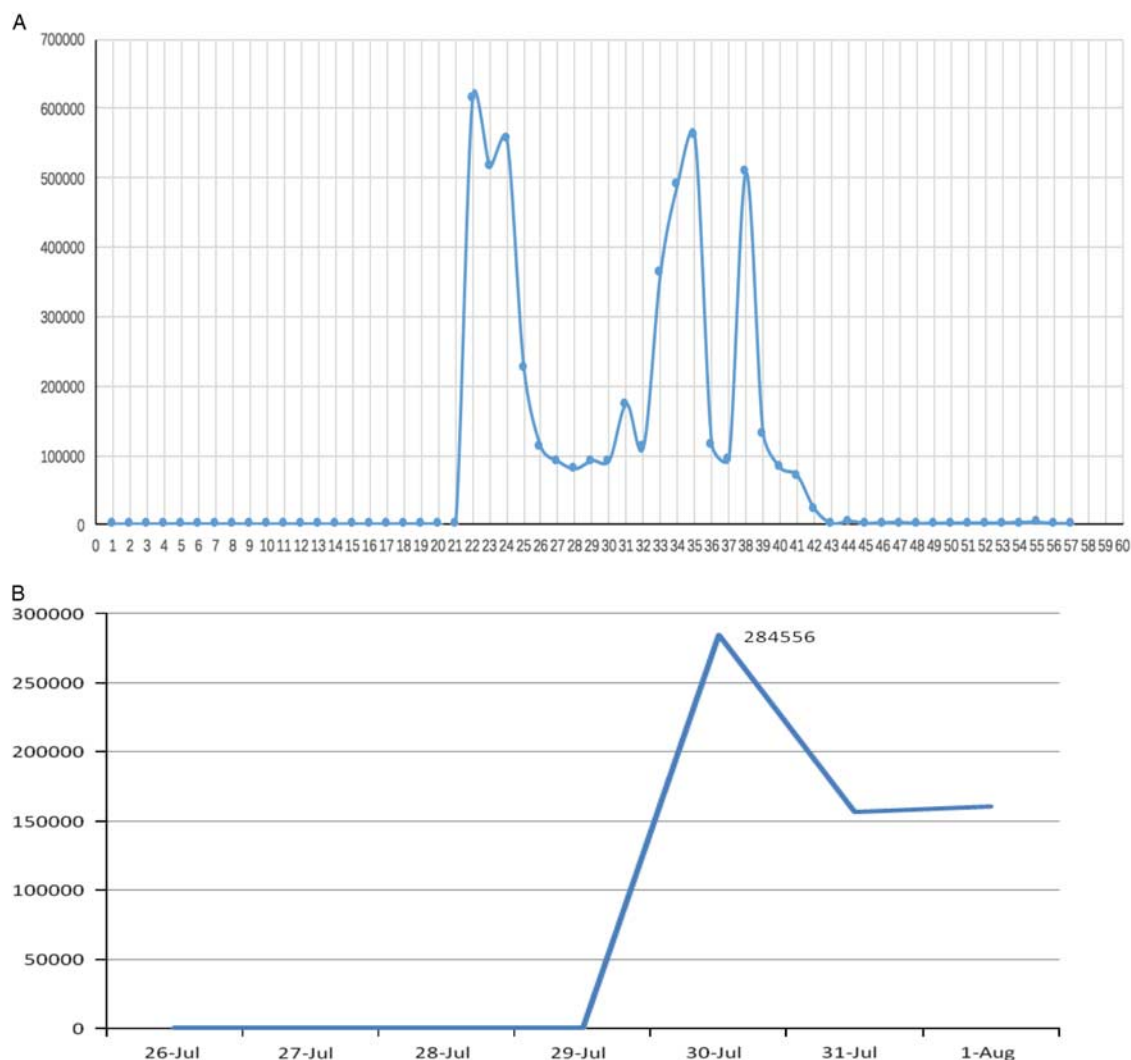
as well as the number of influential actors and the changes thereof over time.

RESULTS
Observation of Public Attention

Figure 1 lists the number of Weibo messages about Ebola for the weeks from March 2014 to March 2015 and the corresponding number of propagators in each period. The total volume of posts was 5,120,376 and the total number of propagators was 518 including duplicates. As previously discussed, because of the length limitations in downloading reports from Weibo, the number of propagators could only be calculated from a maximum of 1000 posts; therefore, these numbers are a measure of the *change in the proportion* of propagators in Weibo, rather than the absolute numbers.

The whole change tendency of public attention to Ebola on Weibo from March 2014 to March 2015 is shown in Figure 1A. On March 24, the WHO first reported an outbreak of EVD in Guinea, but we can see from the Weibo traffic that few people in China were sharing information about Ebola at that time. In fact, it was not until July 26, 21 weeks later, that traffic started to increase significantly. At that point, we have the first peak, which rose sharply during the week from July 25 to July 30. The number of posts that mention Ebola during that week was 614,476, the largest volume in the whole study period. It was during this week that a Hong Kong woman returning from Kenya was reported to be showing symptoms similar to EVD. The news was released on July 30. To see the detailed public reaction in this week, Figure 1B shows the daily number of posts from July 26 to August 1. We can see that the number of references to the disease peaked immediately on the day that the news was released. The data indicate that from July 30 onward, there was an increase in awareness among the Chinese public about the seriousness of Ebola. This period of raised awareness lasted from July 26 to December 19. There were 4 exploration weeks, respectively, July 26 to August 1, October 9 to October 15, October 25 to October 31, and November 15 to November 21. After December 20, the data indicated that public attention to the EVD event in the Weibo community gradually came to an end.

In addition to observing the change tendency in the number of posts, we also analyzed the correlation between the volume of posts and the number of propagators. In Figure 2, we plot the number of propagators against the volume of posts. The Spearman correlation coefficient was 0.899, which means that the number of propagators was significantly correlated with the volume of Weibo posts in the study period. After we removed duplicates in different weeks, the number of propagators was reduced from 518 to 252. The difference between the total number of propagators and the number of distinct actors showed that many propagators sent posts about Ebola repeatedly in different weeks. To some extent, the duration of their posts about one topic can influence the length of the topic transmission period. We observed that the number of posts decreased rapidly when influential posters ceased paying attention to the topic. We defined 3 types of propagators: ordinary users, journalists/media bloggers, and certified celebrities. Among the 252 total propagators, ordinary users made up 20% and the percentage of certified celebrities was 28%. At 52%, journalists or media bloggers

FIGURE 2

Distribution According to Volume of Posts and Number of Propagators.
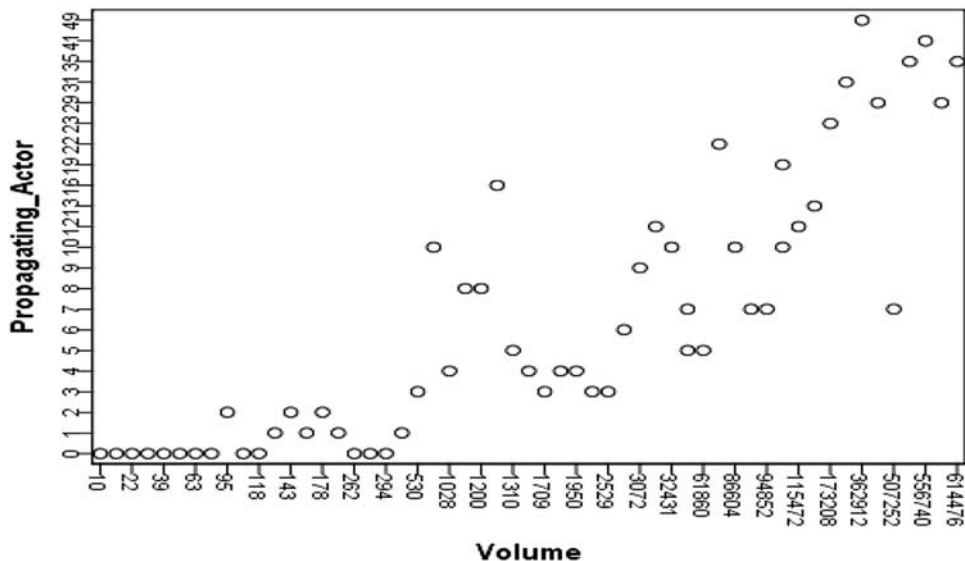
made up the majority of propagators. Thus, journalists or media bloggers were the major information posters and transmission nodes in the Ebola event. They tended to have a large number of followers and to post the newest information in a timely fashion. Many users commented on or reposted these propagators' microblogs. Compared to the certified celebrities and journalists or media bloggers, ordinary users may be less reliable sources of information, and therefore the widespread original content generated by common users could be monitored as a source of spreading rumors.

Analysis of Peak Traffic Events

We can see that there were 4 peaks, respectively, on July 26 to August 1, October 9 to October 15, October 25 to October 31, and November 15 to November 21. For each event, we filtered the messages above the threshold of 100 comments and reposts and extracted key words from the content. We postulated that the most frequently shared terms about Ebola may give insight into the reasons for the peak in traffic (see Table 2 for details). The posts generated by propagators were widespread and influential. The qualified dataset can minimize the spam information and improve the accuracy and confidence of the results. We extracted 45 threshold posts during the week of July 26 to August 1 and the key words therein. Combining the original messages, we could explain the key words. In this week, a Hong Kong woman was reported to be displaying symptoms similar to Ebola. The most widely shared blogs in this week mainly introduced basic knowledge about Ebola, such as "route of transmission," "fatality rate," and "therapy." The information from the WHO, reporting on the number of deaths in Africa, especially Sierra Leone, also attracted a lot attention.

The second peak appeared during the week of April 9 to April 15 and there were 45 posts above the threshold, with 2 popular topics. CCTV bloggers reported that China had the antibody gene of Ebola and would appoint 9 experts to aid Africa, that Canada would provide experimental vaccines, and the United States would treat Ebola with the vaccines in clinic trials in September. Additionally, the bloggers of the Foreign Office told Chinese citizens not to travel to the infected areas, such as Guinea and Liberia. After the second exploration point, the attention to Ebola declined but then peaked again during the week of October 25 to October 31. In October, the Annual Canton Fair, an international import and export fair, was held in Guangzhou and was attended by more than 10,000 delegates from Africa. The risk of introduction of Ebola into China was perceived to be high, and this was reflected in the disease-related traffic on Weibo. The news highlighting a large increase in Africans arriving in Guangzhou made up a fierce online discussion. Besides news about Guangzhou, there were popular news items about CCTV reporters wearing protective clothing to visit the observation center in Sierra Leone and an unusual behavior

TABLE 2

Key Words in the 4 Peak Periods[a]

Key Word (in Chinese)	Key Word (in English)	Weight	Frequency
July 26 to Aug 1			
埃博拉病毒	Ebola virus	7.55	56
致死率	Fatality rate	7.49	6
世卫组织	WHO	6.26	3
传播途径	Route of transmission	5.28	3
西医学	Western medicine	5.27	3
疾控中心	CDC	4.59	3
治疗方法	Therapy	4.17	3
塞拉利昂	Sierra Leone	3.77	11
死亡人数	Number of deaths	3.33	2
Oct 9 to Oct 15			
埃博拉病毒	Ebola virus	7.04	43
中国	China	5.53	29
医生	Doctor	5.42	19
疫情	Epidemic situation	4.88	21
抗体基因	Antibody gene	4.87	7
中国公民	Chinese citizen	4.52	6
疫苗	Vaccine	3.69	7
免疫反应	Immune reaction	3.33	2
大国博弈	Big-power politics	3.33	2
Oct 25 to Oct 31			
央视记者	CCTV reporter	10.01	4
防控	Prevent and control	7.73	5
埃博拉病毒	Ebola virus	6.61	29
医护人员	Medical staff	6.83	4
非洲	Africa	4.17	24
手语翻译	Sign language interpreter	4.16	3
广州	Guang Zhou	3.53	15
中国医疗队	Chinese medical team	3.33	2
疫情重灾区	Serious areas of epidemics	3.05	3
Nov 15 to Nov 21			
首都国际机场	Beijing Capital International Airport	5.93	10
原材料可可产量	Coco production raw material	5.27	8
埃博拉病毒	Ebola virus	5.14	25
医疗队伍	Medical team	4.54	10
巧克力	Chocolate	4.45	42
血液样本	Blood sample	4.16	3
印度	India	3.65	13
医疗队员	Medical staff	3.33	4
告别亲人	Farewell to relatives	3.23	5

[a]Abbreviations: CDC, Centers for Disease Control and Prevention; WHO, World Health Organization.

of a sign language interpreter in the press conference. These were all reflected in the key word search terms presented in Table 2. The last peak occurred during the week of November 15 to November 21, in which the following 4 topics were instrumental in stimulating this peak: Ebola influenced cocoa production in West Africa, and so the supply of chocolate may be at risk in 2020; China appointed 4 medical teams to provide aid in Africa; medical staff said farewell to their relatives in the Beijing Capital International Airport; and India found the first confirmed case of EVD, which was the first infection case in Asia.

Results of the Analysis of Network Evolution

The network of information transmission is dynamic, and changes in interaction among actors can reflect the evolution of topic development. Here, we selected 5 messages posted by "People's Daily" from July to November in 2014 as partial networks to analyze the dynamic changes in the participant interaction network. "People's Daily" has 394,300,000 followers and is regarded as an authoritative official media blogger. It posted 28 microblogs about Ebola, mainly during the period of July 30 to November 28 in 2014. We chose 5 messages out of the 28 posts in each month from July to November. Each message had the largest number of comments and reposts compared to other posts on Ebola in that month. The messages were chosen in each month for all influential posts, which were released as timely news about Ebola and are listed in Table 3.

Users can make a comment and reply to each other under the message. Weibo lists the hot comments that acquire many "likes" from other users and we identified the users who wrote the hot comments as the active participants. We constructed 5 undirected networks for the active participants of the 5 messages where each message had the largest number of comments and reposts compared to other posts on Ebola in that month. We constructed the network for the active participants under the 5 messages and developed a matrix of active participants and uploaded in Ucinet to get the network.

Active participants were regarded as nodes and links between nodes represented that they had interacted with each other including conversation or mention. For each of these networks, we calculated the number of nodes, the number of edges, the average density, the average distance, the distance based on cohesion, and the 3 measures of centrality as defined above: degree represented overall activity, betweenness showed the level of control, and closeness highlighted the level of independence. Table 4 shows the results of the analyses. The total number of active participants, indicated by the number of nodes, showed a declining trend during the period along with the number of interactions among them, as reflected by the reduction in the number of edges. On July 30, there were the highest number of participants and links for the whole period, but the network centralization was lowest. This suggests that there were many isolated actors who were not interacting with others. It can be seen that with fewer participants, the relationship among actors is closer. The biggest mean of node degree occurred in the network for the post on August 11 (see Table 4 for measures of network from July to November 2014).

Figure 3 shows the relative frequency of the degree of nodes in the network. The lines indicate a best-fit curve of the form $P(k) = ck^{-\alpha}$, where $P(k)$ is the probability of the nodes having k degree and c and α are constants. The correlation values are suggestive of the power-law characteristics of

TABLE 3

Five Messages Posted by People's Daily

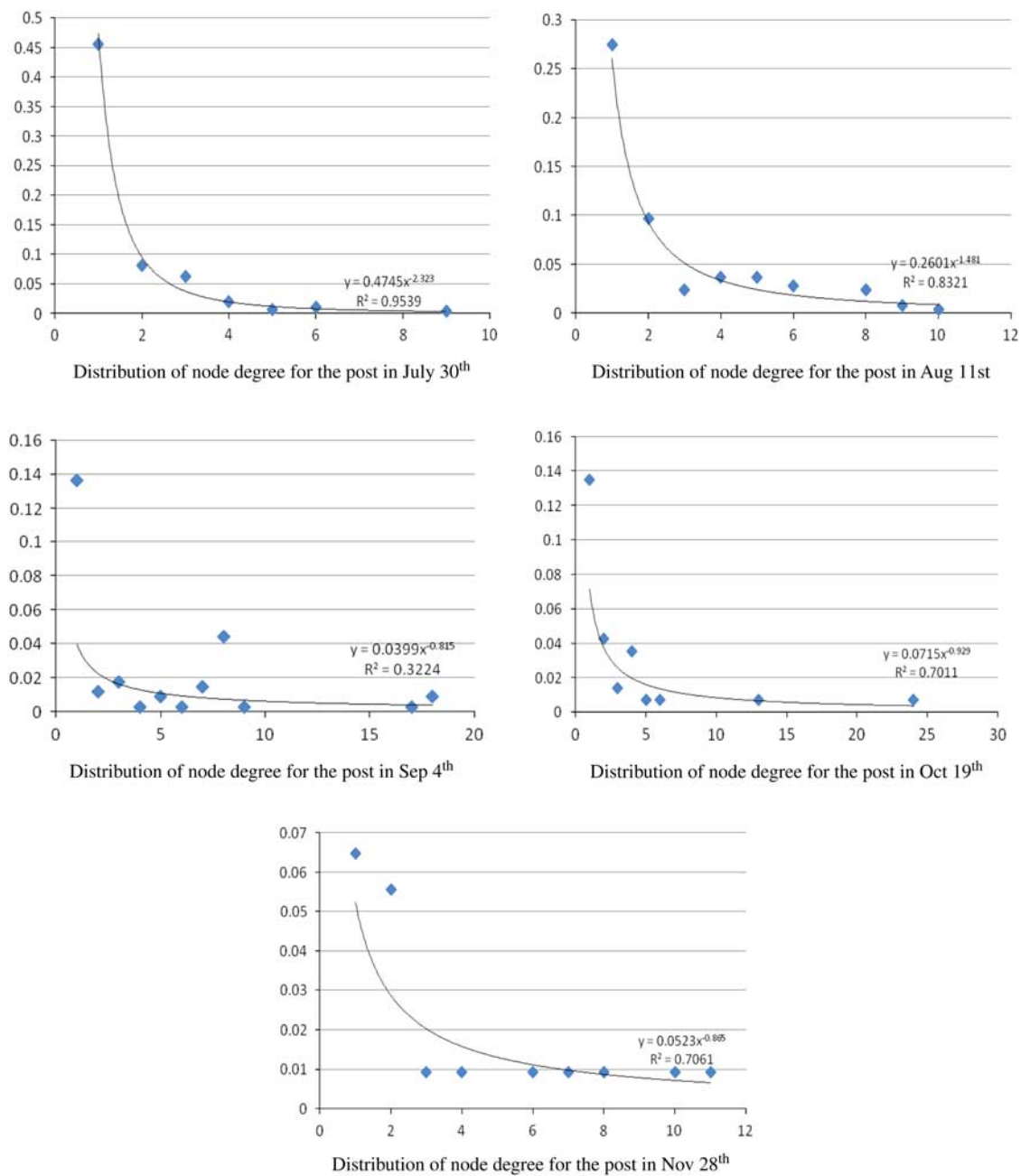| Post date	Content	Comments	Reposts
07/30/2014	The suspected infectious Hong Kong woman has been treated in isolation.	2579	13,529
08/11/2014	Common knowledge about Ebola, such as the definition, route of transmission, prevention, etc.	550	120
09/05/2014	In honor of great scientists and medical staff. By that time, 240 medical staff had been infected and half of them had died.	3193	1096
10/19/2014	The suspected infection of a Nigerian businessman in Guangzhou was excluded. There had been no infectious cases in Guangzhou.	208	185
11/29/2014	Chinese medical team in Sierra Leone had treated 167 observation patients; 80 confirmed cases out of them. The second medical team was ready to leave.	398	326

TABLE 4

Measures of the Network From July to November 2014

Measures	July	Aug	Sep	Oct	Nov
Total nodes	687	248	338	141	108
Total edges	668	312	310	52	36
Average density	0.0015	0.0054	0.0029	0.0054	0.0064
Average distance	1.107	1.106	1.547	3.3341	2.109
Distance-based cohesion	0.002	0.005	0.004	0.006	0.006
Network centralization	0.59	1.18	1.69	1.68	1.4
Mean node degree	1.000	1.339	0.982	0.752	0.685
Mean node betweeness	0.058	0.073	0.553	2.191	0.657

FIGURE 3

**Distribution of Node Degrees in the 5 Networks.**

Distribution of node degree for the post in July 30th

Distribution of node degree for the post in Aug 11st

Distribution of node degree for the post in Sep 4th

Distribution of node degree for the post in Oct 19th

Distribution of node degree for the post in Nov 28th

scale-free networks,[40] ie, most nodes have few connections, but a few nodes have a very large number of connections. Thus, a few actors with a large number of degrees have significant influence on information flow across the network, which can play an important role on diffusing information about Ebola safety for public health preparedness. Combining the relative distribution of node degree, it can be seen that there was a bigger probability of actors who own higher node degrees, which means that there were some close interactive groups in this network.

Figure 4 shows graphically the evolution of the network during the period. As indicated by Table 4, it can be seen that the number of nodes and links was decreasing through the whole period. In each map, there are many isolated nodes that represent the actors who had no interaction with others.
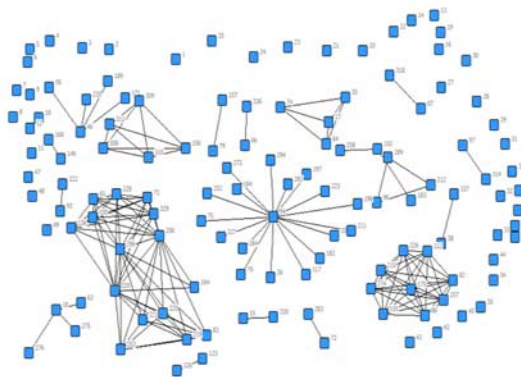
FIGURE 4

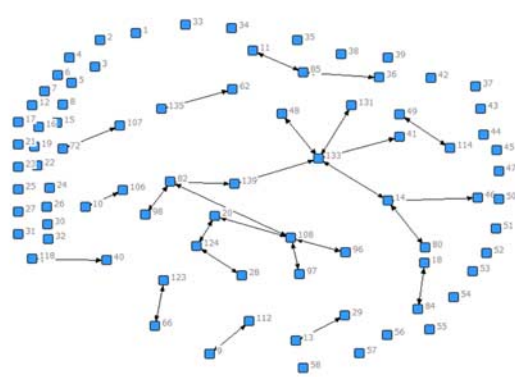**Visualization Maps of the 5 Networks From July to November.**
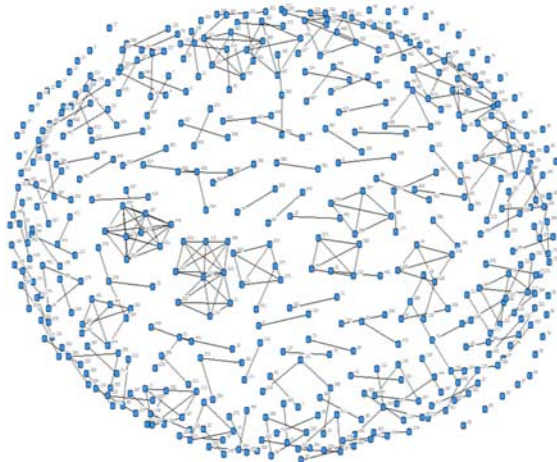
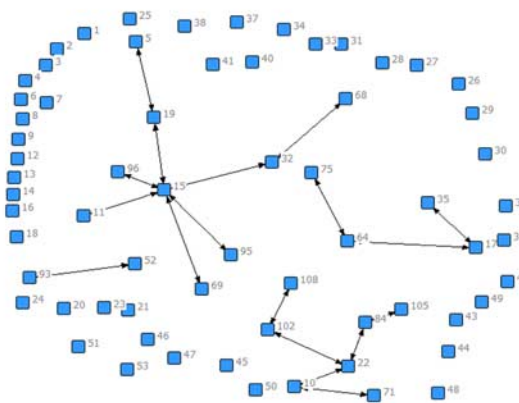Visualization map for the post in July 30th
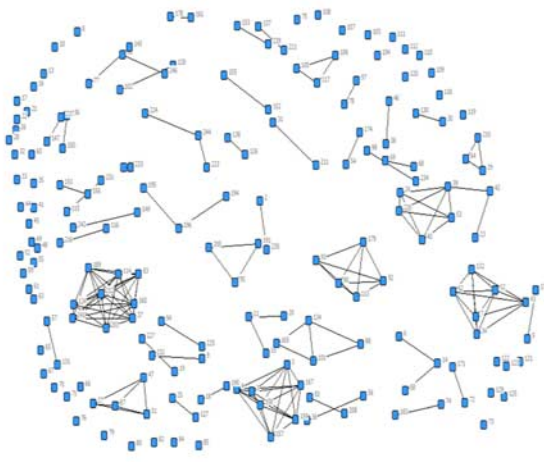
Visualization map for the post in Aug 11st

Visualization map for the post in Sep 4th

Visualization map for the post in Oct 19th

Visualization map for the post in Nov 28th

We can also see that the network for the post on September 4 is most centralized, consistent with the value of network centralization presented in Table 4. To conclude, the evolution of the network from July to November gradually showed a decreasing trend in the number of participants, but the increasing closeness showed that the level of relationship among the participants became closer with the reduction of network scale.

DISCUSSION

Social media as a real-time and widespread online platform provides a major information source and an information diffusion channel that can be used to observe social support networks for diseases.[41] Social media further promotes healthy behaviors and discourages unhealthy practices,[42] addresses public concern, and promotes early detection, allowing us to take corrective action.[43] Social media has been seen as playing a significant role in public health emergency in recent years and can allow regulators and citizens to promote and gain access to timely, accessible, and credible health information during an outbreak or public health emergency for preparedness and response.[44-46] From a theoretical and methodological perspective, researchers have been using a combination of theory and methods drawn from communication, sociology, and social anthropology as well as social psychology to understand the communication process and its impact in changing human behavior, which has a significant impact in a public health emergency.[47,48]

In this research, we analyzed the change tendency in public attention to Ebola, the hot topics which with the public was mainly concerned during peak time points, and the evolution of the participant network. We discovered that the Chinese public began to pay attention to Ebola on July 30, with quick and fierce reaction to the news about a potential suspected case in China, which was mainly concentrated during the period of July 30 to December 19, with 4 peaks, respectively, during the weeks of July 26 to August 1, April 9 to April 15, October 25 to October 31, and November 15 to November 21. After December 20, public attention to this event gradually ended. The hot topics that the Chinese public were mainly concerned with fell into 4 categories: epidemic situation ("Ebola outbreak," "affecting Ebola," "Ebola epidemic area," "West Africa region"); treatment against Ebola ("against Ebola," "Ebola medicine," "recover," "clinical trials"); China medical aid ("aid Africa," "medical staff," "Medical team aiding Liberia"); and Ebola prevention ("Ebola vaccine," "epidemic prevention").

The dynamic evolution of the participant network under the posts of the "People's Daily" reflected the decreasing active actors and links during the period of Ebola topic development. We observed that participants gradually decreased but the network centralization became closer with the distribution of the node degree of the 5 networks agreeing with the law power distribution of a scale-free network. We also observed that the number of propagators was positively correlated with the volume of posts in a period. More attention should be paid to the propagators who are ordinary users to prevent them from spreading rumors, especially for epidemics. Small-scale networks may have greater centralization than large-scale networks and there may be more isolated nodes in larger networks. The scale of the networks may influence the interaction level.

Compared with previous research in detecting the Chinese public's reaction to H7N9, we further analyzed hot topics, the performance of participants, and the relationship of propagators and the volume of posts instead of only discovering the change tendency of the public's reaction in presenting the first analysis of the Chinese public's reaction to Ebola by monitoring Sina Weibo. The methodology of this investigation could be used to judge and evaluate the public reaction to epidemics. Our results can provide some reference for governments and related authorities to take measures at the appropriate time. Data in social websites are a valuable resource that can reflect public opinions and moods quickly and truly and that can serve as effective references for the government to release announcements to the public by monitoring public reaction and taking appropriate measures essential for preventing and controlling epidemics.

About the Authors

Information Management, Division of Information and Technology Studies, The University of Hong Kong, Hong Kong (Ms Feng); Library and Information Management Program, Division of Information and Technology Studies, The University of Hong Kong, Hong Kong (Dr Hossain); Rothamsted Research, Harpenden, United Kingdom (Dr Crawford); Faculty of Business, Charles Sturt University, Albury, Australia (Dr Bossomaier).

Correspondence and reprint requests to Liaquat Hossain, PhD, Professor of Library and Information Management, Division of Information and Technology Studies, Room 113, Runme Shaw Building, The University of Hong Kong, Pokfulam Road, Hong Kong (e-mail: lhossain@hku.hk).

REFERENCES

1. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. *Bus Horiz.* 2010;53(1):59-68. https://doi.org/10.1016/j.bushor.2009.09.003.
2. Duggan M, Brenner J. *The Demographics of Social Media Users, 2012.* Vol 14. Washington, DC: Pew Research Center's Internet & American Life Project; 2013.
3. Arnold JL. Disaster medicine in the 21st century: future hazards, vulnerabilities, and risk. *Prehosp Disaster Med.* 2002;17(01):3-11. https://doi.org/10.1017/S1049023X00000042.
4. Bradt DA, Abraham K, Franks R. A strategic plan for disaster medicine in Australasia. *Emerg Med.* 2003;15(3):271-282. https://doi.org/10.1046/j.1442-2026.2003.00445.x.
5. Coiera E. Social networks, social media, and social diseases. *BMJ.* 2013;346:f3007. doi: https://doi.org/10.1136/bmj.f3007
6. Dufty N. Using social media to build community disaster resilience. *Aust J Emerg Manag.* 2012;27(1):40-45.
7. Keim ME, Noji E. Emergent use of social media: a new age of opportunity for disaster resilience. *Am J Disaster Med.* 2011;6(1):47-54.
8. Merchant RM, Elmer S, Lurie N. Integrating social media into emergency-preparedness efforts. *N Engl J Med.* 2011;365(4):289-291. https://doi.org/10.1056/NEJMp1103591.
9. 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):242. https://doi.org/10.1186/1471-2458-12-242.
10. Semenza JC, Rubin CH, Falter KH, et al. Heat-related deaths during the July 1995 heat wave in Chicago. *N Engl J Med.* 1996;335(2):84-90. https://doi.org/10.1056/NEJM199607113350203.

11. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection — harnessing the web for public health surveillance. *N Engl J Med.* 2009;360(21):2153-2157. https://doi.org/10.1056/NEJMp0900702.

12. Wesolowski A, Stresman G, Eagle N, et al. Quantifying travel behavior for infectious disease research: a comparison of data from surveys and mobile phones. *Sci Rep.* 2014;4:5678.

13. Khan K, McNabb SJN, Memish ZA, et al. Infectious disease surveillance and modelling across geographic frontiers and scientific specialties. *Lancet Infect Dis.* 2012;12(3):222-230. https://doi.org/10.1016/S1473-3099(11)70313-9.

14. Pfeiffer DU, Stevens KB. Spatial and temporal epidemiological analysis in the Big Data era. *Prev Vet Med.* 2015;122(1-2):213-220. https://doi.org/10.1016/j.prevetmed.2015.05.012.

15. Stevens KB, Pfeiffer DU. Sources of spatial animal and human health data: casting the net wide to deal more effectively with increasingly complex disease problems. *Spat Spatio-Temporal Epidemiol.* 2015;13:15-29. https://doi.org/10.1016/j.sste.2015.04.003.

16. Rényi A. On measures of entropy and information. In: *Fourth Berkeley Symposium on Mathematical Statistics and Probability.* Vol 1. Berkeley, CA: University of California Press; 1961:547-561.

17. St Louis C, Zorlu G. Can Twitter predict disease outbreaks? *BMJ.* 2012;344:e2353. doi: 10.1136/bmj.e2353.

18. Yom-Tov E, White RW, Horvitz E. Seeking insights about cycling mood disorders via anonymized search logs. *J Med Internet Res.* 2014;16(2):e65. https://doi.org/10.2196/jmir.2664.

19. Keller M, Blench M, Tolentino H, et al. Use of Unstructured event-based reports for global infectious disease surveillance. *Emerg Infect Dis.* 2009;15(5):689-695. https://doi.org/10.3201/eid1505.081114.

20. Velasco E, Agheneza T, Denecke K, et al. Social media and Internet-based data in global systems for public health surveillance: a systematic review. *Milbank Q.* 2014;92(1):7-33. https://doi.org/10.1111/1468-0009.12038.

21. 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. https://doi.org/10.1371/journal.pone.0019467.

22. Lampos V, De Bie T, Cristianini N. Flu detector - tracking epidemics on Twitter. In: *Machine Learning and Knowledge Discovery in Databases.* Berlin: Springer; 2010:599-602. DOI: 10.1007/978-3-642-15939-8_42

23. 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.* Stroudsburg, PA: Association for Computational Linguistics; 2011.

24. Mirhaji P. Public health surveillance meets translational informatics: a desiderata. *J Assoc Lab Autom.* 2009;14(3):157-170. https://doi.org/10.1016/j.jala.2009.02.007.

25. Zhang L, Pentina I. Motivations and usage patterns of Weibo. *Cyberpsychol Behav Soc Netw.* 2012;15(6):312-317. https://doi.org/10.1089/cyber.2011.0615.

26. Gao Q, Abel F, Houben GJ, Yu Y. A comparative study of users' microblogging behavior on Sina Weibo and Twitter. In: *International Conference on User Modeling, Adaptation, and Personalization.* Berlin: Springer; 2012:88-101. https://doi.org/10.1007/978-3-642-31454-4_8.

27. Fung IC-H, Fu K-W, Ying Y, et al. Chinese social media reaction to the MERS-CoV and avian influenza A (H7N9) outbreaks. *Infect Dis Poverty.* 2013;2(1):31. https://doi.org/10.1186/2049-9957-2-31.

28. Chen Y, Li Z, Nie L, et al. A semi-supervised Bayesian network model for microblog topic classification. In: *Proceedings of COLING 2012.* Powai, Mumbai, India: The COLING 2012 Organizing Committee Indian Institute of Technology Bombay; 2012:561-576.

29. Cui X, Yang N, Wang Z, et al. Chinese social media analysis for disease surveillance. *Pers Ubiquit Comput.* 2015;19:1125. doi: 10.1007/s00779-015-0877-5.

30. Wang S, Paul MJ, Dredze M. Exploring health topics in Chinese social media: an analysis of Sina Weibo. Presented at the Twenty-Eighth AAAI Conference on Artificial Intelligence; July 27-31, 2014; Québec City, Canada.

31. Fu K-W, Chau M. Reality check for the Chinese microblog space: a random sampling approach. *PloS One.* 2013;8(3):e58356. doi: 10.1371/journal.pone.0058356.

32. Ono T, Hishigaki H, Tanigami A, et al. Automated extraction of information on protein–protein interactions from the biological literature. *Bioinformatics.* 2001;17(2):155-161. https://doi.org/10.1093/bioinformatics/17.2.155.

33. World Health Organization website. Frequently asked questions on Ebola virus disease [in Chinese]. http://www.who.int/csr/disease/ebola/faq-ebola/zh/. Last updated May 2017. Accessed June 20, 2017.

34. Zhou L, Zhang D. NLPIR: A theoretical framework for applying natural language processing to information retrieval. *J Am Soc Inf Sci Technol.* 2003;54(2):115-123. https://doi.org/10.1002/asi.10193.

35. Lei L, Ji L. Empirical analysis of Micro-blogger behavior clustering on public opinion topics. *J Intell.* 2014;33(3):118-121.

36. Haizhou L, Baosheng Y. Chinese word segmentation. *Language.* 1998;212:217.

37. Luke DA, Harris JK. Network analysis in public health: history, methods, and applications. *Annu Rev Public Health.* 2007;28(1):69-93. https://doi.org/10.1146/annurev.publhealth.28.021406.144132.

38. Scott J. *Social Network Analysis.* Thousand Oaks, CA: Sage Publishing; 2012.

39. Freeman LC. Centrality in social networks conceptual clarification. *Soc Networks.* 1978-1979;1(3):215-239. https://doi.org/10.1016/0378-8733(78)90021-7.

40. Jordan IK, Mariño-Ramírez L, Wolf YI, et al. Conservation and coevolution in the scale-free human gene coexpression network. *Mol Biol Evol.* 2004;21(11):2058-2070. https://doi.org/10.1093/molbev/msh222.

41. Greene JA, Choudhry NK, Kilabuk E, et al. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. *J Gen Intern Med.* 2011;26(3):287-292. https://doi.org/10.1007/s11606-010-1526-3.

42. Randolph W, Viswanath K. Lessons learned from public health mass media campaigns: marketing health in a crowded media world. *Annu Rev Public Health.* 2004;25(1):419-437. https://doi.org/10.1146/annurev.publhealth.25.101802.123046.

43. Larson HJ, Cooper LZ, Eskola J, et al. Addressing the vaccine confidence gap. *Lancet.* 2011;378(9790):526-535. https://doi.org/10.1016/S0140-6736(11)60678-8.

44. Glik DC. Risk communication for public health emergencies. *Annu Rev Public Health.* 2007;28(1):33-54. https://doi.org/10.1146/annurev.publhealth.28.021406.144123.

45. Hawn C. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. *Health Aff (Millwood).* 2009;28(2):361-368. https://doi.org/10.1377/hlthaff.28.2.361.

46. McNab C. What social media offers to health professionals and citizens. *Bull World Health Organ.* 2009;87(8):566. https://doi.org/10.2471/BLT.09.066712.

47. Rimal RN, Lapinski MK. Why health communication is important in public health. *Bull World Health Organ.* 2009;87(4):247-247a. https://doi.org/10.2471/BLT.08.056713.

48. Wakefield MA, Loken B, Hornik RC. Use of mass media campaigns to change health behaviour. *Lancet.* 2010;376(9748):1261-1271. https://doi.org/10.1016/S0140-6736(10)60809-4.