Leveraging social networking sites for disease surveillance and public sensing: the case of the 2013 avian influenza A(H7N9) outbreak in China

Emma Xuxiao Zhang,[a*] Yinping Yang,[b*] Richard Di Shang,[c] Joseph John Pyne Simons,[b] Boon Kiat Quek,[b] Xiao Feng Yin,[b] Wanhan See,[a] Olivia Seen Huey Oh,[a] Khine Sein Tun Nandar,[a] Vivienne Ruo Yun Ling,[a] Pei Pei Chan,[a] Zhaoxia Wang,[b] Rick Siow Mong Goh,[b] Lyn James[a] and Jeannie Su Hui Tey[a]

Correspondence to Jeannie Su Hui Tey (e-mail: jeannie_tey@moh.gov.sg).

We conducted in-depth analysis on the use of a popular Chinese social networking and microblogging site, Sina Weibo, to monitor an avian influenza A(H7N9) outbreak in China and to assess the value of social networking sites in the surveillance of disease outbreaks that occur overseas. Two data sets were employed for our analysis: a line listing of confirmed cases obtained from conventional public health information channels and case information from Weibo posts. Our findings showed that the level of activity on Weibo corresponded with the number of new cases reported. In addition, the reporting of new cases on Weibo was significantly faster than those of conventional reporting sites and non-local news media. A qualitative review of the functions of Weibo also revealed that Weibo enabled timely monitoring of other outbreak-relevant information, provided access to additional crowd-sourced epidemiological information and was leveraged by the local government as an interactive platform for risk communication and monitoring public sentiment on the policy response. Our analysis demonstrated the potential for social networking sites to be used by public health agencies to enhance traditional communicable disease surveillance systems for the global surveillance of overseas public health threats. Social networking sites also can be used by governments for calibration of response policies and measures and for risk communication.

On 31 March 2013, China announced the world's first three human cases of avian influenza A(H7N9) in Shanghai and Anhui provinces.[1] This was followed by reports of further cases in over 16 provinces/municipalities of China and exportation of infection to China, Hong Kong Special Administrative Region, Malaysia and Taiwan, China. Most human cases of A(H7N9) infection were severe and were characterized by rapidly progressive pneumonia and acute respiratory distress syndrome.[2] There was significant international concern about the impact of this novel infection on global health and security.[3,4]

In Singapore, to follow the rapidly evolving A(H7N9) outbreak in China, we supplemented information obtained from conventional public health information channels with posts from Sina Weibo (www.weibo.com; Weibo). Weibo is a popular social networking site in China with more than 500 million registered users as of February 2013. It was one of the fastest social networking platforms to report breaking news on A(H7N9) and was leveraged by health authorities, media and the public to monitor outbreak-related information.[5]

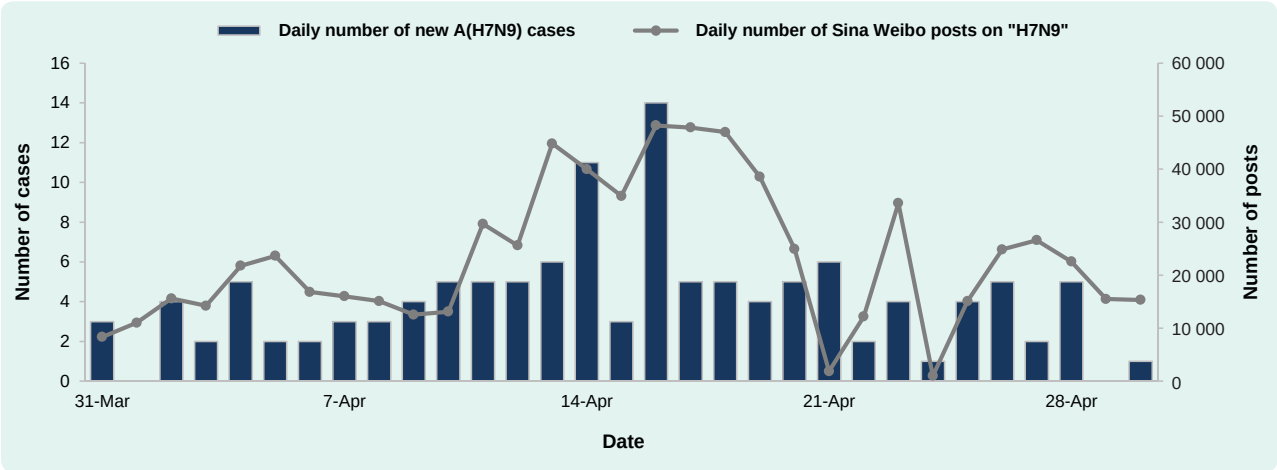
To evaluate the relevance of social networking sites as a new platform in the global surveillance of disease outbreaks external to Singapore, we carried out an in-depth analysis to review and verify the functions of Weibo in the monitoring of the A(H7N9) outbreak in China.

METHOD

For our analysis, we consolidated two data sets. The first included a line listing of confirmed cases obtained from conventional public health information channels, including the official website of Chinese

[a] Epidemiology and Disease Control Division, Public Health Group, Ministry of Health, Singapore.
[b] Social and Cognitive Computing Department and Computing Science Department, Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore.
[c] School of Business, Public Administration and Information Sciences, Long Island University Brooklyn, New York, United States of America.
* The first two authors contributed equally to this study.
Submitted: 28 January 2015; Published: 18 May 2015
doi: 10.5365/wpsar.2015.6.1.013

Figure 1. **The number of Weibo posts on "H7N9" and the number of of new A(H7N9) cases confirmed by authorities by day, 31 March to 30 April 2013**



National Health and Family Planning Commission (NHFPC); the Event Information Site of the World Health Organization (WHO) and email alerts from a leading international news agency, Agence France-Presse (AFP). The second data set included Weibo posts containing the search phrase "H7N9". We obtained the Weibo data set from an authorized provider. To ensure data quality and reduce data noise, we solicited posts from authenticated users whose identities had been verified by Weibo. We focused our analysis period from 31 March to 30 April 2013 for two reasons: (1) the outbreak started on 31 March and the majority of cases during the first wave of the outbreak were recorded in April (126 cases of the total of 133); (2) daily reporting of cases by the NHFPC website was only available during the first wave.

We performed two types of analysis: (1) quantitative analysis to compare the timeliness of reporting of new cases by the various information channels; (2) qualitative analysis on the Weibo users with the timeliest posts on new cases. Weibo posts that were the first to report the 126 cases were retrieved, and information including time of reporting, reporter account and epidemiological information of the cases was recorded. To identify the timeliest reporter of each case, we used the advanced search function of Weibo. Key phrase "h7n9" was used, and the search was restricted to authenticated users. Time duration was narrowed down to only one hour to allow retrieval of all posts as Weibo tends to automatically exclude posts if the volume is too large. The location of posts was left unspecified. The search results were compared against the line listing of cases confirmed by WHO to identify the earliest reporter.

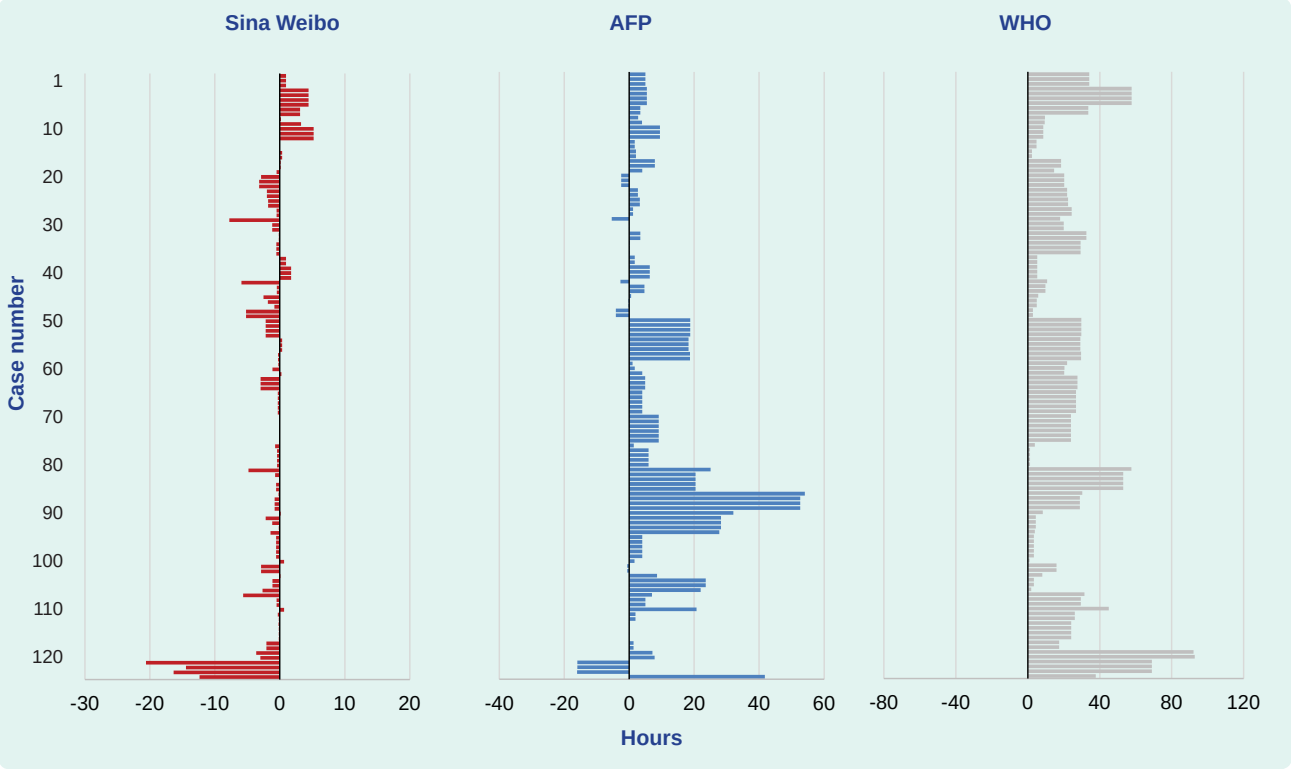
For comparison of timeliness of reporting by various channels, statistical analysis was performed to assess the significance of any temporal differences in the reporting. Non-parametric Wilcoxon signed-rank test was used as the temporal differences were not normally distributed. The analysis was performed in SPSS 16.0. A statistically significant result was defined as $p < 0.05$.

To explore the factors contributing to the timeliness of reporting by Weibo, we performed qualitative analysis to examine the characteristics of the users with the timeliest post on new cases. We also analysed the content of the posts to qualitatively assess other aspects in which Weibo was used. To ensure data reliability, we checked the information manually against the line listing of the corresponding cases confirmed by WHO.

RESULTS

Between 31 March and 30 April 2013, China reported 126 cases of A(H7N9). Correspondingly, 718 419 posts, or an average of about 23 175 per day, were posted on Weibo. The volume of social media discussion corresponded to the number of reported cases (**Figure 1**). The increasing number of cases from 10 April onward was accompanied by a surge in the number of posts, indicative of the public's heightened awareness of the disease while the outbreak appeared to be gaining momentum. A peak of 48 255 daily posts was recorded on 16 April when the highest daily number of 16 cases was reported. Weibo users' interest in A(H7N9) plummeted briefly on 21 April and 24 April, probably due to a shift in attention toward other major

Figure 2. **Timeliness of reports on A(H7N9) cases by Weibo, AFP and WHO compared with NHFPC, 31 March to 30 April 2013***



* The zero-hour baseline represents the time of reporting by the NHFPC.

AFP, Agence France-Presse; NHFPC, National Health and Family Planning Commission of China; WHO, World Health Organization.

events: an earthquake in China's Sichuan province on 21 April and social unrest in China's Xinjiang province on 24 April.

A comparison of reporting times revealed that Weibo was significantly faster in reporting new cases than the conventional public health channels including the NHFPC, AFP and WHO (*p*s < 0.001) (**Figure 2**). Reporting of new cases on Weibo was an average of 1 hour 2 minutes before the NHFPC website with a maximum lead time of 20 hours and 35 minutes. This lead was even greater when Weibo reporting was compared with that of AFP with the latter reporting an average of 8 hours and 14 minutes after the NHPFC. The average lead was more pronounced when compared to WHO that reported cases an average of 23 hours and 13 minutes after the NHPFC (**Figure 2**). The delay in the announcement by WHO was expected as WHO reports only cases that are notified to them by Member States (China in this case) in accordance with the International Health Regulations (2005).[6] In addition, further time lag would be incurred if clarification or confirmatory testing of the cases were required.

The Weibo users with the timeliest report on new cases comprised province/municipality-based news agencies, including Zhejiang Daily (the official newspaper of Zhejiang provincial government), Modern Express and China Exclusive (both belong to Xinhua News Agency, the official news agency of the Chinese central government) (**Table 1**). The veracity of the data reported from these users is likely to be high, and this was confirmed when manual verification of the information from these posts and those from official reports showed high level of concurrence. Upon release of information by the provincial/municipal health authorities, these news agencies posted the news on their Weibo account immediately, before reporting on their conventional websites. In contrast, announcement at the NHFPC website typically lagged behind and was probably due to the additional time taken to collate information from the various provincial/municipal health authorities for the Chinese central government's daily updates.

In addition to monitoring outbreak development, the use of Weibo enabled timely monitoring of other

outbreak-relevant information. On 24 April, WHO held a press conference on its investigation findings in China; the transcript was posted in real-time on the NHFPC's Weibo feed, allowing instant access to the information from anywhere in the world.

Weibo also provided access to additional crowd-sourced epidemiological information on infected cases, such as updates on patients' health conditions, exposure history and family contacts that were not readily available through official sources (**Table 2**). Such additional insights from Weibo usually came from informants in the community whose ready access to social media enabled them to actively participate in disease surveillance.

From the perspective of the Chinese health authorities, the rapid disclosure of information on social media appeared to have helped accelerate official response and reporting. For example, on 5 April, a Weibo user posted pictures of dead sparrows in a Nanjing residential area. The local authority promptly responded by cleaning the implicated premises and testing samples from the dead sparrows that were found to be negative for A(H7N9). In another case on 2 April, a medical document of a new case was disclosed by a Weibo user. This was soon followed by the official announcement of the case by the implicated hospital on Weibo along with official confirmation of four new cases by the NHFPC.

Social networking sites were leveraged by the Chinese health authorities as an interactive platform for risk communication with the general public. During the outbreak, the Chinese health authorities held many real-time question-and-answer sessions on Weibo. In these sessions, doctors and experts addressed queries from the public in a real-time and interactive manner.

Weibo was also used by the Chinese health authorities as a tool for assessing public sentiments to proposed outbreak response measures to guide policy decisions. In April 2013, a local news media conducted a survey on Weibo to seek citizens' views on permanent closure of live poultry markets (LPMs) in Shanghai.[7] A total of 28.6% of the respondents supported permanent closure, while 30.4% opposed the idea. Among all the

Table 1. **Timeliest reporting of new A(H7N9) cases on Weibo, 31 March to 30 April 2013**

Weibo user	Number of occasions in which the source was the first reporter of the event
Zhejiang Daily	8
Modern Express	6
Zhejiang Voice	6
China Exclusive	4
12320 China Health	4
CCTV News	2
Win in China News	2
Dajiang Net	2
People's Daily	2
Others (Qianjiang Evening, Yangzi Evening, etc)	24

Table 2. **Examples of crowd-sourced epidemiological information on Weibo**

Date of post	User	Content of post (translated)	Epidemiological information
8 April 2013	Youth Times (news media) http://weibo.com/qnsblh	Wife of the 67-year-old case from Hangzhou said during an interview, "Apart from shopping for groceries, my son and I avoid close contact with any person. After my husband fell sick, my son came to stay with me."	Status of family members of the case
9 April 2013	Gan Yuxiang (a celebrity) http://weibo.com/ganyuxiang	The 67-year-old male case resides near Wushan district with his wife. Both persons have hypertension. Their diet contains mainly fish and vegetables, and they did not eat chicken recently. Before the onset of illness, the case ate a quail he had bought from the Bin Sheng Market in Shang Chen district.	Exposure history; underlying co-morbidities of the case
11 April 2013	Zhejiang Mobile Newspaper (news media) http://weibo.com/zjsjb001	The female case from Huzhou reported yesterday is currently in a stable condition. The male case from Hangzhou has been placed under mechanical ventilation. The condition of another case, whose surname is Shen Tu, deteriorated rapidly.	Update on cases' situations

respondents, over 77% suggested enhanced animal surveillance and better management of poultry in farms and markets. The Shanghai health authority later ordered the temporary closure of LPMs during the peak of the outbreak. In addition, the authorities initiated various infection control measures, including enhanced poultry surveillance and restriction of live poultry trading to designated markets which were subjected to weekly closure for disinfection and cleaning. From 27 April to 4 May the China NHFPC conducted a poll on Weibo to survey public attitudes, concerns and expectations.[8] The results showed that 93.4% of the respondents were satisfied with the information released by NHFPC's Weibo page and expressed support for continued transparent information-sharing to be conducted by the NHFPC via Weibo.

DISCUSSION

Our study was conducted from the perspective of public health agencies involved in the global surveillance of overseas public health threats. The findings of our analysis demonstrate the potential for public health agencies to acquire time-sensitive information on rapidly evolving outbreaks occurring outside of their countries through social networking sites. In our analysis, Weibo served as a platform leveraged by central/provincial governments, local news agencies and the public for the timely release and retrieval of information. The Chinese social networking sites' timeliness of reporting are significantly better compared to international mainstream media in English, official websites of the central Chinese government and WHO. Information was released by the local news agencies on the social networking site in the local language before any reports in the conventional news websites or government websites. The information was subsequently amplified by the social networking sites through re-posting of the original report. This allowed the international community greater access to more detailed and timely information compared to that released at the central government level to the international media. The central Chinese government may have allowed the provincial/municipal health authorities to release information on new cases when available to ensure transparency and timeliness of public communications. The access to such timely, crowd-sourced information on infected cases greatly facilitated the understanding of the epidemiology of an unknown disease, which is key in developing effective prevention and control measures.

In view of the vast number of posts on various social networking sites, it would be important to employ the site(s) that have the most relevant user profiles, language medium and context to the country of interest. One limitation of the social media surveillance system is the initial difficulty in identifying reliable, consistent and timely information sources at the outset of surveillance since countless numbers of users would be posting on the topic of interest. Time is usually required to monitor the information put out by various users and to compare them against verified reports to insure their relevance for inclusion into the surveillance system. The other limitation is the veracity of the information. While the analyses of posts could be limited to those from verified users only, we observed that the study of posts from layman users provided an understanding of the actual situation and sentiments in the affected country. This additional viewpoint could have a significant impact on outbreak control and consequently influence the risk assessment of the outbreak. There is also a possibility of inaccurate or false information being purposely propagated through social media that could affect the quality of the intelligence acquired from this source. To overcome this, verifying information against credible sources, including WHO and the health authorities of the affected country, is necessary. For the social media platform to be effective, there is also a need for the disease to be sufficiently novel to warrant the interest and concern of the people in the affected country for significant re-posting of information to occur; the affected population must have a thriving social networking scene with high participation and connectivity. Despite this, re-posting can be still be insignificant due to apathy, low media coverage or diversion of public interest to other events as illustrated by the plunge in number of posts on 21 and 24 April in our study.

Our qualitative analysis showed the effective use of social media by the Chinese health authorities in risk communication as well as gathering public sentiments on response options is an innovative strategy in public education, social mobilization and garnering support for the outbreak response measures. This echoes a previous study that demonstrated that social media could be a useful tool for public health practitioners to understand public reaction to disease outbreak information released by health authorities.[9] An analysis of Internet data during the A(H7N9) outbreak suggested that the early stage of the outbreak was accompanied by rapidly increasing public attention and thus was considered the best time

frame for health authorities to engage the public, conduct education campaigns and control rumours.[10]

Although social media is considered a less formal platform, health authorities around the world are increasingly using it for information access and dissemination.[11–13] Social media has been used as a central platform for the retrieval of information from various official sources; such use was highlighted by epidemiologists from the United States Centers for Disease Control and Prevention who used social media to monitor the A(H7N9) outbreak.[14] There are, however, challenges to the use of social networking sites. Constant monitoring and real-time analysis of a large influx of data with a high level of background noise, including rumours and unrelated information, is labour-intensive. Identifying the most appropriate social media platform to use is also critical to ensure effectiveness. In this case, the use of Weibo for the A(H7N9) outbreak in China would be more appropriate than global social media platforms such as Twitter. Language barriers can pose additional challenges; our ability to access and accurately interpret information from Weibo was partly due to our being based in Singapore where Mandarin is an official language. Our experience showed that it may be beneficial for public health agencies to recruit and maintain a workforce of epidemiologists who are multilingual for international disease surveillance in a foreign language.

One limitation of our study is the selection of AFP to represent international mainstream news media, particularly since Chinese news media such as Xinhua News Agency would likely be faster in reporting new cases. AFP was included in our comparison because global surveillance is carried out primarily in English, and we noted that AFP was consistently one of the fastest global news agencies, along with Reuters, British Broadcasting Company, Cable News Network and Associated Press, to deliver accurate and comprehensive news on global disease outbreaks. While Xinhua News Agency may report news on outbreaks located in China in a timelier manner compared to AFP, we find it less relevant as a generic source of information for the surveillance of outbreaks outside of China.

We envisage significant potential for social media surveillance to be incorporated into mainstream disease surveillance and response systems. For international public health practitioners, social media surveillance could provide early warning for unusual public health events in a foreign country and serve as an additional source of epidemiological intelligence to complement conventional surveillance tools. For local public health authorities, social media surveillance could function as an effective platform for public education and social mobilization. The underlying value coupled with the challenges of using social media warrants future research and collaboration between public health agencies and computational scientists to enhance its use in disease outbreak surveillance and response.

Conflicts of interest

None declared.

Funding

This study was supported in part by the Social Technologies+ Programme funded by Agency for Science, Technology and Research (A*STAR) Joint Council Office, Singapore.

Acknowledgements

The authors are grateful to Dr Derrick Heng (Group Director of the Public Health Group, Singapore), Ms Ang Li Wei and Ms Cindy Lim of the Singapore Ministry of Health for helpful discussion.

References

1. *Questions and answers about human infection with A(H7N9) avian influenza virus*. Beijing, Chinese Center for Disease Control and Prevention (CDC), 2013 (http://www.chinacdc.cn/en/ne/201303/t20130331_79282.html, accessed 15 August 2014).
2. Li Q et al. Epidemiology of human infections with avian influenza A(H7N9) virus in China. *The New England Journal of Medicine*, 2014, 370:520–532. doi:10.1056/NEJMoa1304617 pmid:23614499
3. Uyeki TM, Cox NJ. Global concerns regarding novel influenza A(H7N9) virus infections. *The New England Journal of Medicine*, 2013, 368:1862–1864. doi:10.1056/NEJMp1304661 pmid:23577629
4. Meng Z et al. Possible pandemic threat from new reassortment of influenza A(H7N9) virus in China. *Euro Surveillance: European Communicable Disease Bulletin*, 2014,19(6):pii=20699. pmid:24556346
5. Salathé M et al. Influenza A (H7N9) and the importance of digital epidemiology. *The New England Journal of Medicine*, 2013, 369:401–404. doi:10.1056/NEJMp1307752 pmid:23822655

6. *International Health Regulations (2005), Second edition*. Geneva, World Health Organization, 2008 (http://www.who.int/ihr/9789241596664/en/, accessed 21 April 2015).
7. Survey on public opinion toward permanent closure of live poultry market in Shanghai. [In Chinese] *Jf Daily*, 2013 (http://newspaper.jfdaily.com/xwcb/html/2013-04/10/content_1004412.htm, accessed 18 August 2014).
8. *12320 public health hotline conducted survey on public opinions toward infection control measures against avian influenza A(H7N9).* [In Chinese] Beijing, China National Health and Family Planning Commission (NHFPC), 2013 (http://www.moh.gov.cn/zhuzhan/zsdwgzdt/201306/070c3ca66f604c1aa4ec5f65761ede76.shtml, accessed 18 August 2014).
9. Fung IC et al. Chinese social media reaction to the MERS-CoV and avian influenza A(H7N9) outbreaks. *Infectious Diseases of Poverty*, 2013, 2:31. doi:10.1186/2049-9957-2-31 pmid:24359669
10. Gu H et al. Importance of Internet surveillance in public health emergency control and prevention: evidence from a digital epidemiologic study during avian influenza A H7N9 outbreaks. *Journal of Medical Internet Research*, 2014, 16:e20. doi:10.2196/jmir.2911 pmid:24440770
11. *World Health Organization Twitter* (https://twitter.com/who, accessed 29 March 2015).
12. *European Center of Disease Prevention and Control Twitter* (https://twitter.com/ECDC_EU, accessed 29 March 2015).
13. *United States Centers for Disease Control and Prevention Twitter* (https://twitter.com/CDCgov, accessed 29 March 2015).
14. Fung ICH, Wong K. Efficient use of social media during the avian influenza A(H7N9) emergency response. *Western Pacific Surveillance and Response Journal*, 2013, 4:1–3. doi:10.5365/wpsar.2013.4.3.005 pmid:24478916