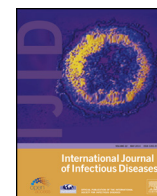


Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Short Communication

Correlation between reported human infection with avian influenza A H7N9 virus and cyber user awareness: what can we learn from digital epidemiology?

Tiansheng Xie^{a,b}, Zongxing Yang^{a,b}, Shigui Yang^{a,b}, Nanping Wu^{a,b}, Lanjuan Li^{a,b,*}^aState Key Laboratory for Diagnosis and Treatment of Infectious Diseases, the First Affiliated Hospital, School of Medicine, Zhejiang University, No. 79 Qingchun Road, Hangzhou 310031, China^bCollaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou, China

ARTICLE INFO

Article history:

Received 16 September 2013

Received in revised form 21 November 2013

Accepted 23 November 2013

Corresponding Editor: Eskild Petersen, Aarhus, Denmark

Keywords:

Avian influenza
Epidemiology
Cyber user awareness
Public health policy
Social networks

SUMMARY

Data on the topic of novel avian influenza A (H7N9) were collected based on the web analysis tool 'Baidu Index', a major Chinese search engine. We found a positive correlation between the volume of H7N9-related 'cyber user awareness' and the epidemic situation during the H7N9 outbreak in China ($r = 0.98$, $p < 0.01$, cumulative; $r = 0.43$, $p = 0.018$, daily) except in the early stage; the ranks of H7N9-related topics changed at different epidemic stages. This study may improve our understanding of the role of web-based media in infectious disease surveillance in China.

© 2014 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

On March 31, 2013, the Chinese government reported three human cases of avian influenza A H7N9 virus infection in Shanghai and Anhui. Further cases of human infection with H7N9 virus were then confirmed,^{1–3} along with the emergence of new clinical features.⁴ H7N9 has quickly become an important issue for Chinese public health.

By the end of June 2012, it was reported that there were 538 million internet users in China, and more than 95% of Chinese cyber users reported Baidu to be their first choice of online search engine.⁵ In this study, using the daily data available online from Baidu Index, we analyzed the correlation between the H7N9 epidemics and the volume of 'cyber user awareness'.

2. Methods

Baidu Index issues the search query volume for the entire country according to the daily values in the Baidu websites that are

open to the public. The data source for cyber user awareness was the search volume data of keywords issued by the Baidu search engine, and the value here has been weighted according to the search frequency. In this study, the keywords for searching were 'H7N9' or 'Qinliugan' (in Chinese), the value of cyber user awareness was shown on the daily and cumulative scale, and regional information was also recorded. All the H7N9-related topics were ranked according to the query frequency at the same time.

H7N9 infection and mortality information were acquired from the National Health and Family Planning Commission during April 1 and June 30 in 2013, together with the infection time and epidemic outbreak provinces. The Pearson correlation coefficient (r) was calculated between reported H7N9 epidemics and the volume of cyber user awareness on the daily or cumulative scale, with the statistical significance level set at 0.05.

3. Results

We found that during the first month of the H7N9 outbreak (April), the volume of cumulative awareness about H7N9 increased exponentially with the growth of the epidemics (Figure 1); the peak value of daily cyber user awareness (141 900) occurred on April 5. A significant correlation was seen between the cumulative reported cases and the volume of user awareness ($r = 0.98$,

* Corresponding author.

E-mail address: ljlj@zju.edu.cn (L. Li).

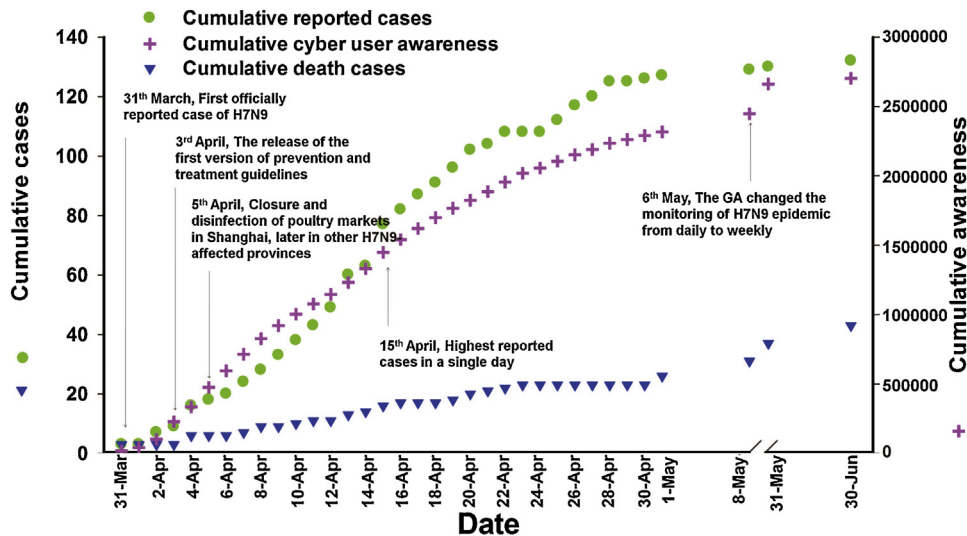


Figure 1. Cumulative numbers of avian influenza A H7N9 cases and cyber user awareness in China. The volume of cumulative awareness about H7N9 increased exponentially with the growth of the epidemic; there was a significant correlation between the cumulative reported cases and user awareness.

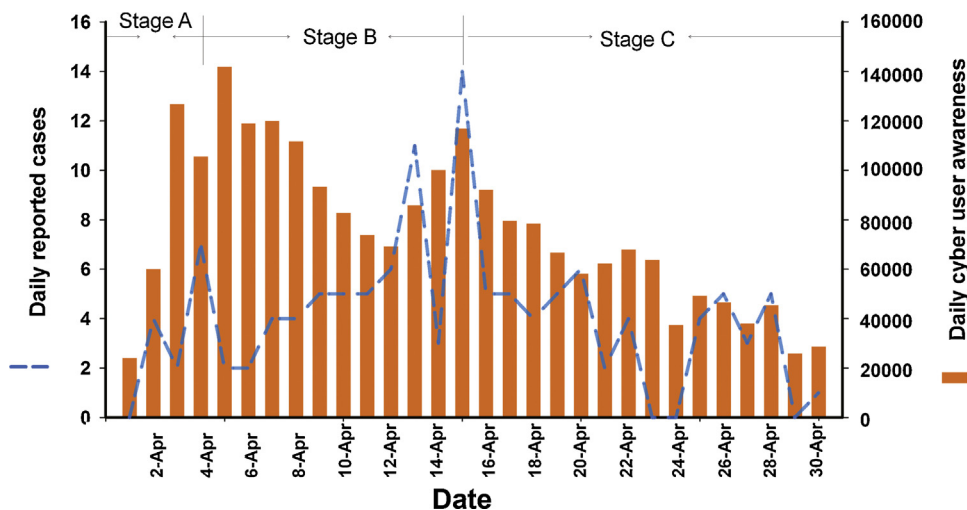


Figure 2. Distribution of daily reported cases and cyber user awareness during the months of the avian influenza A H7N9 outbreak. The H7N9 epidemic could be divided into three stages in April: stage A (April 1–4) represented a period of local spread; stage B (April 5–15) represented a period of increased but concentrated outbreaks in the four provinces of China's Yangtze River delta region; stage C (April 16–30) represented a period of widespread outbreak, as cases increased from 74 to 110 in the above region, and 15 new cases were reported in another five provinces.

$p < 0.01$), and between the daily reported cases and user awareness ($r = 0.43$, $p = 0.018$).

Our results suggested that the epidemics could be divided into three stages in April (Figure 2). Stage A (April 1–4) represented a period of local spread. The positive correlation was weak between the daily reported cases and user awareness ($r = 0.46$, $p = 0.43$). Stage B (April 5–15) represented a period of increased and concentrated outbreaks in the four provinces of China's Yangtze River delta region, with 74 of the 77 total cases occurring in this region. A negative correlation between the daily reported cases and user awareness ($r = -0.45$, $p = 0.17$) was seen. Stage C (April 16–30) represented a period of widespread outbreak, as cases increased from 74 to 110 in the abovementioned region, and 15 new cases were reported in another five provinces; a positive correlation between the daily reported cases and user awareness was seen ($r = 0.65$, $p = 0.012$).

We also found that the rank of H7N9-related topics changed at the different stages of the epidemics. In April, the first four topics that the public were most concerned with were disease

diagnosis, symptoms of H7N9 infection, mortality, and gene mutation. However, during May and June, the general public began to be more concerned about the prevention and treatment of the disease, such as 'common sense for avian influenza prevention'.

4. Discussion

Online search queries are a uniquely valuable source of historical and real-time information.^{6–8} Recent studies have shown that media coverage can produce positive changes, or prevent negative changes in health-related behaviors in large populations.^{9,10} In general, our data suggest that in April, trends in H7N9 epidemics were strongly correlated with the cumulative value of cyber user awareness ($p < 0.01$), and with the daily data ($p < 0.05$). Also in April, when H7N9 infection cases first occurred, the infection case number was low but awareness was high, showing no significant correlation between them; as it progressed, the case number increased but typically public awareness waned, with a

significant correlation (stage C in present study). Such results can also be seen in a previous report.¹⁰

This study may help to enhance the value of online information during emerging epidemics. We believe that digital epidemiology in the 'Network Era' will have a significant impact on the handling of emerging infectious diseases and will complement the present series of coping strategies, especially in the field of public health.

Acknowledgements

This study was supported by the Special Grant for National Key Technologies R&D Program for the 12th Five-Year Plan of China (No. 2013ZX10004-904) and the Independent Research Grant of State Key Laboratory for Diagnosis and Treatment of Infectious Disease (2012, Dr C.Z. Jin). We also appreciate Baidu Ltd who developed the Baidu index (http://www.baidu.com/search/index_help.html) as the data source of the present study.

Conflict of interest: No conflict of interest to declare.

References

1. Chen Y, Liang W, Yang S, Wu N, Gao H, Sheng J, et al. Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: clinical analysis and characterization of viral genome. *Lancet* 2013;**381**:1916–25.
2. Dai C, Jiang M. Understanding H7N9 avian flu. *BMJ* 2013;**346**:f2755.
3. Horby P. H7N9 is a virus worth worrying about. *Nature* 2013;**496**:399.
4. Gao HN, Lu HZ, Cao B, Du B, Shang H, Gan JH, et al. Clinical findings in 111 cases of influenza A (H7N9) virus infection. *N Engl J Med* 2013;**368**:2277–85.
5. China Internet Network Information Center. The statistical report on China's internet development status. People's Republic of China: Ministry of Information; 2012.
6. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection—harnessing the Web for public health surveillance. *N Engl J Med* 2009;**360**:2153–5.
7. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature* 2009;**457**:1012–4.
8. Reis BY, Brownstein JS. Measuring the impact of health policies using Internet search patterns: the case of abortion. *BMC Public Health* 2010;**10**:514.
9. Wakefield MA, Loken B, Hornik RC. Use of mass media campaigns to change health behaviour. *Lancet* 2010;**376**:1261–71.
10. Salathe M, Freifeld CC, Mekaru SR, Tomasulo AF, Brownstein JS. Influenza A (H7N9) and the importance of digital epidemiology. *N Engl J Med* 2013;**369**:401–4.