

BMJ Open Early detection of perceived risk among users of a UK travel health website compared with internet search activity and media coverage during the 2015–2016 Zika virus outbreak: an observational study

Jakob Petersen,¹ Hilary Simons,² Dipti Patel,¹ Joanne Freedman³

To cite: Petersen J, Simons H, Patel D, *et al.* Early detection of perceived risk among users of a UK travel health website compared with internet search activity and media coverage during the 2015–2016 Zika virus outbreak: an observational study. *BMJ Open* 2017;7:e015831. doi:10.1136/bmjopen-2017-015831

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2017-015831>).

Received 3 January 2017

Revised 17 June 2017

Accepted 13 July 2017



CrossMark

¹National Travel Health Network and Centre, University College London Hospital NHS Foundation Trust, London, UK

²National Travel Health Network and Centre, Liverpool School of Tropical Medicine, Liverpool, UK

³Travel and Migrant Health Section, Public Health England, London, UK

Correspondence to

Dr Jakob Petersen;
jakob.petersen1@nhs.net

ABSTRACT

Objectives The Zika virus (ZIKV) outbreak in the Americas in 2015–2016 posed a novel global threat due to the association with congenital malformations and its rapid spread. Timely information about the spread of the disease was paramount to public health bodies issuing travel advisories. This paper looks at the online interaction with a national travel health website during the outbreak and compares this to trends in internet searches and news media output.

Methods Time trends were created for weekly views of ZIKV-related pages on a UK travel health website, relative search volumes for ‘Zika’ on Google UK, ZIKV-related items aggregated by Google UK News and rank of ZIKV travel advisories among all other pages between 15 November 2015 and 20 August 2016.

Results Time trends in traffic to the travel health website corresponded with Google searches, but less so with media items due to intense coverage of the Rio Olympics. Travel advisories for pregnant women were issued from 7 December 2015 and began to increase in popularity (rank) from early January 2016, weeks before a surge in interest as measured by Google searches/news items at the end of January 2016.

Conclusions The study showed an amplification of perceived risk among users of a national travel health website weeks before the initial surge in public interest. This suggests a potential value for tools to detect changes in online information seeking behaviours for predicting periods of high demand where the routine capability of travel health services could be exceeded.

OBJECTIVES

The rapidly evolving Zika virus (ZIKV) outbreak associated with a cluster of congenital malformations in Brazil in 2015 led WHO to declare the outbreak a public health emergency of international concern on 1 February 2016.¹

Authoritative, accurate and timely information about the spread of ZIKV was paramount

Strengths and limitations of this study

- Infodemiological applications have to date primarily used online search engines or social media networks. This study is to our knowledge the first to use data on the online interactions with a professional resource.
- The study used a large and broad range of online news sources, which assisted in making the results generalisable.
- Formal validation of many Google tools is still pending on the disclosure of the underlying methodologies.
- Any system based on online information seeking behaviours is open to new biases such as the activity of robots designed to boost traffic to other websites for commercial gains.

to the work of public health bodies issuing travel advisories, especially to pregnant women, as the outbreak unfolded and spread across the Americas. This was important, first, to prevent harm in travellers at higher risk of complications from ZIKV, specifically pregnant women²; second, to support travel health services at a time of high demand; and third, to disseminate appropriate public health messages while media attention was high.³

Repositories of data on online behaviours carry rich potentials for surveillance and have been termed infodemiology, that is, the science of distribution and determinants of online information with the aim of informing public health and public policy.^{4–7} Google Flu Trends represents an early example of an infodemiological system that has been used for both early detection and estimation of outbreak magnitude.⁸ There have since been several applications in infectious

disease surveillance (dengue, Ebola virus, swine influenza, ZIKV), both in endemic and non-endemic settings, and exploiting a variety of data sources such as online search engines,^{9–14} social media networks,^{12 13 15 16} online encyclopaedias,¹⁷ media newsrooms and news aggregators^{3 11 14 18 19} and mediated surveillance networks.²⁰ Some of the lessons from these early works have been that information seeking behaviours wax and wane as new information becomes available and satisfies gaps in knowledge.^{11–13} In some situation, especially in endemic settings, there may be an equilibrium between information seeking and case numbers, which is useful for predicting the pressure on frontline services or in countries with limited surveillance capability.^{7 8 10 14 15} In this study, we look at information seeking behaviours among users of a UK national travel health website during the 2015–2016 ZIKV outbreak and compare this with time trends in concurrent internet searches and output from news media rooms. The motivation being that travel health services are staffed by healthcare professionals who routinely deal with traveller risk scenarios weeks and months ahead of departure. The paper in that way looks at the potential value of a professional network as a sentinel surveillance system for predicting periods of high demand on travel health services⁸ over commonly used sources such as internet search engines and news media.

METHODS

The National Travel Health Network and Centre (NaTHNaC) is a government-commissioned clinical advisory service supporting travel health services in England, Wales and Northern Ireland in consultation with Public Health England's Travel and Migrant Health unit. Its main information channel is a website (<https://travelhealthpro.org.uk>). Emerging threats such as ZIKV are covered in frequently updated news items and outbreak surveillance notices. Established threats are covered in factsheets with advice to travellers and health professionals. In the case of the ZIKV outbreak, the news items with travel advisories for pregnant women issued from 7 December 2015 onwards quickly became highly accessed. Various other news items on ZIKV and a fact-sheet (ZIKV risk assessment) were also available but were much less visited. Data on the weekly number of unique pageviews (defined as number of sessions in which the page was viewed at least once and from hereon referred to as 'views') for each page were extracted using Google Analytics, and the rank according to the weekly number of views was calculated. The access to the travelhealthpro.org.uk website was compared with Google searches (a measure of interest from the public) and news items aggregated by Google UK News (a measure of media output). The relative search volumes for 'Zika' were obtained from Google Trends UK (<https://www.google.co.uk/trends;country:UK>) and ZIKV-related news items on Google UK News (term: 'zika'; <https://news.google.co.uk/>; country:

Table 1 Google UK news items by category

Topics	Frequency	Per cent
General concerns	766	43.8
Rio Olympics 2016	663	37.9
Vector control	89	5.1
Vaccine development	53	3.0
Business news	42	2.4
First imported cases to a country	36	2.1
Sexual transmission news	23	1.3
Refund/cancellation practices of tourism industry	19	1.1
First deaths for a country	19	1.1
First congenital Zika syndrome cases for a country or region	18	1.0
First local transmission in a country	13	0.7
Other	9	0.5
Total	1750	100.0

UK, date of extraction: 1 September 2016). The media news items were included if the headline and excerpt were ZIKV related, and each item was categorised by topic (General concerns/Rio de Janeiro Olympics 2016/Vector control/Vaccine development/Business news/First imported cases to a country/Sexual transmission news/Refund/cancellation practices of tourism industry/First deaths for a country/First congenital Zika syndrome cases for a country or region/First local transmission in a country/Other). It was not uncommon for news items to include information on multiple aspects of the outbreak (ie, virus, disease, geographical spread, adverse outcomes, possible control measures, burden on society, etc). These items were grouped together under 'general concerns'. The study period began on 15 November 2015, the week WHO first published news about a microcephaly cluster in Northeast Brazil on 20 November 2015²¹ and ended on 20 August 2016, coinciding with the final week of the Rio Olympics 2016. Data for the entire period were extracted at weekly intervals (Sunday–Saturday weeks). Data analyses were carried out in Stata 14.²²

RESULTS

The analysis of ZIKV-related items from the Google UK news aggregator yielded a total of 1750 items (table 1). The majority of items were categorised under general concerns (43.8%) or the Rio Olympics 2016 (37.9%). Other popular topics, although on a smaller scale, concerned control measures such as vector control (5.1%) and vaccine development (3%).

The time series of the news items by category revealed that items about the Rio Olympics overtook all other topics from the beginning of June 2016 and culminating at the beginning of August, the week of the opening ceremony (figure 1).

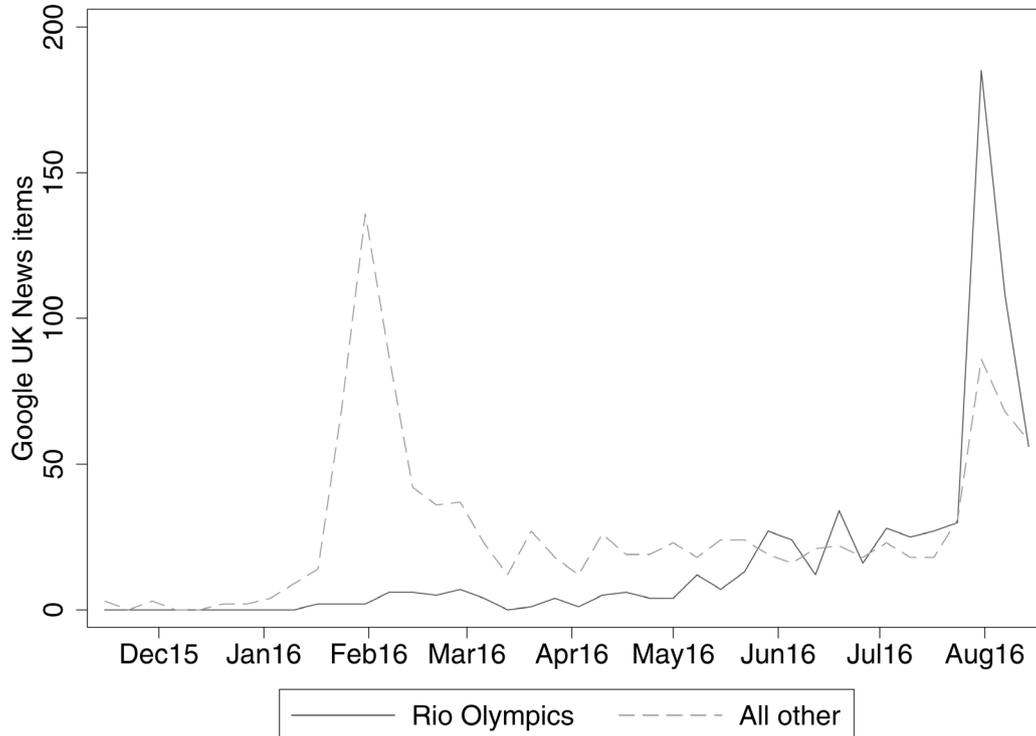


Figure 1 Zika-virus-related Google UK News items by whether the Rio Olympics was mentioned or not.

The time series of ZIKV-related browsing activities were characterised by an initial surge, at the end of January to the beginning of February 2016, punctuated by smaller spikes on travelhealthpro.org.uk and on Google searches in April and May, and again at the end of July (figure 2). The opening week of the Rio Olympics at the beginning

of August was associated with the highest number of news items of the study period. The rank of the first travel advisory to pregnant women published on 7 December 2015 quickly declined in popularity (rank) only to rise again in popularity from the beginning of January 2016 (figure 3). The third travel advisory, published on 26 January 2016,

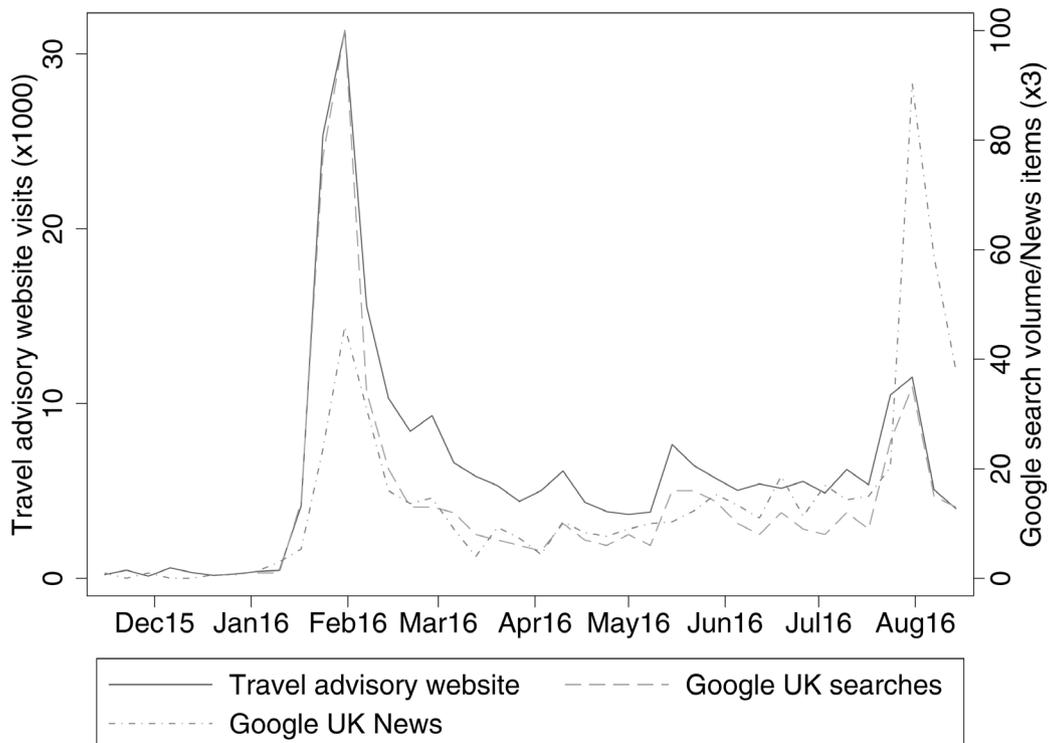


Figure 2 Professional browsing activity (travelhealthpro.org.uk), public searching activity (Google UK Trends) and number of Google UK News. Weekly data.

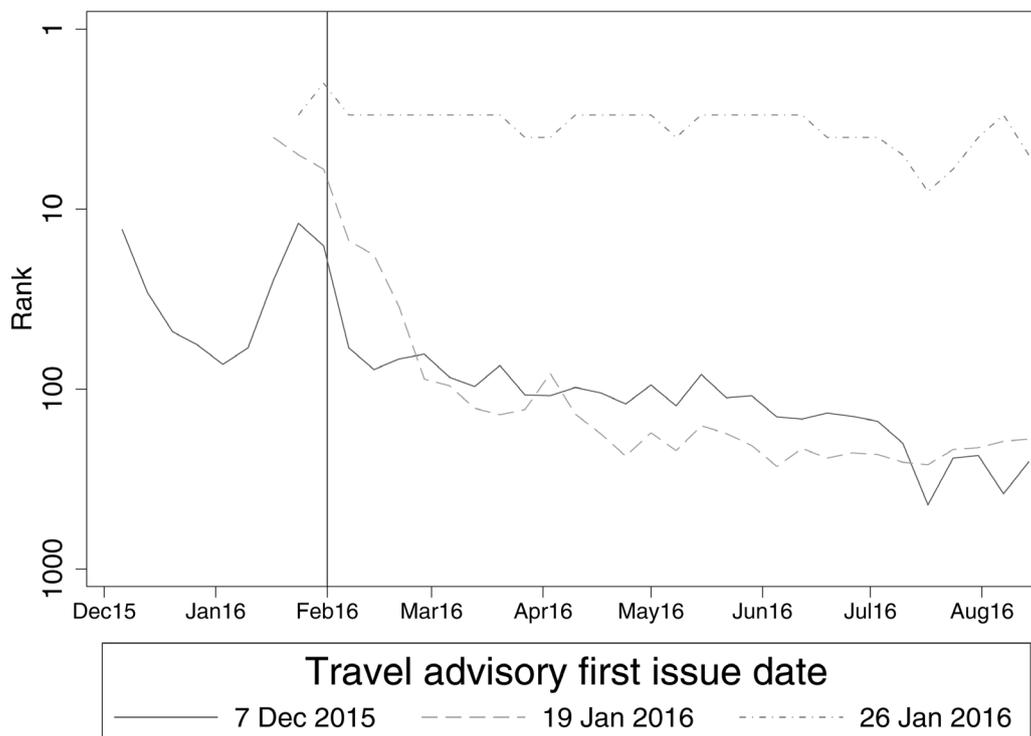


Figure 3 Rank among all pages on national travel health website, travelhealthpro.org.uk, for ZIKV travel advisories for pregnant women issued on 7 December 2015, 19 January and 26 January 2016, respectively (weekly data). Vertical line: WHO Public Health Emergency of International Concern on 1 February 2016.

remained among the 10 most popular pages on the website throughout the study period. The three pregnancy travel advisories had reached a total of 149 496 views by 20 August 2016. For comparison, the most popular item covering the developments of the Ebola virus outbreak in West Africa only reached 20 000 views between July and December 2014 (unpublished information, NaTHNaC, 1 September 2016).

DISCUSSION

Traditionally, public health bodies have been informed about changes in the burden of diseases mainly through indicator-based surveillance systems such as yearly incidence or prevalence rates of predefined diseases. In recent decades, however, increasing attention has been directed towards the development and use of complementary event-based surveillance systems tuned at detecting new and emerging threats of the month, week, day or even hour.^{23 24} Online information seeking behaviours (information demand), user-generated content and online media stories (information supply) have opened new and almost immediate opportunities for surveillance of emerging threats. This scientific discipline has been termed infodemiology or infoveillance.^{5 6} An early example of an infodemiological system was the Google Flu Trends, which could predict demand on healthcare systems based on mass searches for influenza symptoms and remedies on the Google search engine.⁸ Part of the early detection relies on the fact that many individuals will try to self-diagnose and self-treat before seeking help

from the healthcare services. Several studies have since explored the potential of various online data sources for early detection including online search engines,^{9–14} social media networks such as Twitter,^{12 13 15 16} the online encyclopaedia Wikipedia,¹⁷ media newsrooms and news aggregators^{3 11 18 19} and mediated surveillance networks such as Programme for Monitoring Emerging Diseases.²⁰ Particular diseases of interest have been swine/seasonal influenza,^{3 18 19} Ebola virus disease,^{9 11 13 16 20} dengue^{10 14} and, most recently, ZIKV.^{12 15}

Google Flu Trends has since been replaced by more sophisticated systems exploiting multiple data sources and calibrating the models to take regional variation into account (eg, internet connectivity, demographic factors and healthcare seeking behaviours).^{25 26} Infodemiological systems, including early versions of Google Flu Trends, have been subject to criticism on a number of other points too. First, some have been critical of the fact that many of the Google tools and underlying algorithms have not been published and have thus remained methodologically ‘black boxes’.²⁷ Second, and more fundamentally, less is known about internet and data-driven methodologies than traditional methods. New biases have to be defined and addressed, for example, that internet searches for a part are made by robots and not humans with genuine concerns about diseases.²⁷ Third, the issues around the fear and curiosity driven initial bursts of information seeking can mask de facto risk especially for new threats. An interesting example of this was the disproportional interest in the few Ebola cases

in the USA compared with the situation in West Africa during the 2014 Ebola epidemic.¹³ This and other studies have similarly found a very high initial interest in the ZIKV outbreak, especially as the first cases were imported to the USA and Europe.^{12,15} This phenomenon has been termed the social amplification of risk mechanism and known to be modified by volume, content and tone of coverage. Further attention can be triggered by key events such as first imported case, first death or, in the case of ZIKV, first pregnant case or first case of congenital Zika syndrome.³ In the present study, secondary spikes in online behaviours were observed in April, May and at the end of July in both the professional network and in Google searches. The spike in April coincided with reports about scientific consensus for a causal link between ZIKV infection and microcephaly and Guillian-Barré syndrome.²⁸ The spike at the end of July coincided with the first local cases of mosquito-borne ZIKV infection in Florida, USA. In the case of ZIKV and the UK media, the majority of ZIKV-related items from June to August were associated with an even bigger news generator, that is, the Rio Olympics with extensive coverage of, for example, top athletes dropping out due to the ZIKV outbreak. The epidemiological 'signal value' of media stories has in this case been weaker than for the online trends of Google searches or the information seeking behaviour on a travel health website. Chan *et al* similarly found that dengue-related Google searches were better correlated with case numbers in endemic countries, where outbreaks are newsworthy yet not media sensations.¹⁰ Fourth, many syndromic surveillance systems set up in the 2000s have been discontinued due to high false positive rates, which have led some observers to ask for systems incorporating higher quality sensors to inform early responders more effectively.²⁴ Finally, some have advocated for surveillance systems based on voluntary panels of members of the public that should offer timeliness and a tentative compromise between the safety and limitations of traditional study designs and the promises and perils of crowd-based systems.²⁹

Sections of the public planning to travel to potentially high risk disease areas may seek advice, vaccinations and other protection from travel health services, either voluntarily or due to mandatory requirements imposed by governments and the WHO.³⁰ Due to vaccination schedules, consultations may have to be arranged weeks or months before departure. Potentially vulnerable travellers such as pregnant women, young children, those with immunocompromise or other comorbidities are particularly likely to seek professional travel health advice.³¹ The online interaction with a national travel health information service could for that reason act as a sentinel system for early detection of an emerging outbreak—at least at the point where it starts to be considered a risk to travellers. The advantage of such a system would be to direct attention to emerging issues at an early stage before a surge in public/media interest threaten to overwhelm the routine capability of front-line services.

There are undeniably many examples of where the media, as a whole, delivers important health messages in ways that are not misleading or merely seeking to cause sensation.³² The present study, however, can show that trends in media reports on an infectious disease are not always related to risk as no exported cases of ZIKV were reported in connection with the Rio Olympics.³³

A possible limitation of a system based on online information seeking behaviours could be robots designed to boost traffic to other websites for commercial gains.²⁷ We have not been able to exclude this possibility, but the fact that the audience is relatively small and specialised makes this less likely than had it been big and broad in terms of interests.

The study used a large and broad range of news sources aggregated by Google News, which assists in making the results generalisable. The various Google tools have virtue in being widely used and the data easily accessible. Formal validation of many of the tools is, however, still pending on the disclosure of the underlying methodologies.²⁷

CONCLUSIONS

The study showed an amplification of perceived risk among users of a national travel health website weeks before the initial surge in public interest at the end of January 2016. This suggests a potential value for tools to detect changes in online information seeking behaviours for predicting periods of high demand where the routine capability of travel health services could be exceeded.

Contributors JP, HS, DP and JF were closely involved with the design, conduct, analysis, presentation and interpretation of the study findings.

Competing interests None declared.

Patient consent No patient data used.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Unpublished website usage data captured by Google Analytics, which can be made available upon request to the corresponding author.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

1. Plourde AR, Bloch EM. A literature review of Zika virus. *Emerg Infect Dis* 2016;22:1185–92.
2. Lucey DR, Gostin LO. The emerging Zika pandemic: enhancing preparedness. *JAMA* 2016;315:865–6.
3. Reintjes R, Das E, Klemm C, *et al*. "Pandemic Public Health Paradox": time series analysis of the 2009/10 Influenza A/H1N1 epidemiology, media attention, risk perception and public reactions in 5 European countries. *PLoS One* 2016;11:e0151258.
4. Charles-Smith LE, Reynolds TL, Cameron MA, *et al*. Using social media for actionable disease surveillance and outbreak management: a systematic literature review. *PLoS One* 2015;10:e0139701.
5. Eysenbach G. Infodemiology: The epidemiology of (mis)information. *Am J Med* 2002;113:763–5.

6. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med* 2011;40:S154–8.
7. Nuti SV, Wayda B, Ranasinghe I, et al. The use of Google trends in health care research: a systematic review. *PLoS One* 2014;9:e109583.
8. Ginsberg J, Mohebbi MH, Patel RS, et al. Detecting influenza epidemics using search engine query data. *Nature* 2009;457:1012–4.
9. Alicino C, Bragazzi NL, Faccio V, et al. Assessing Ebola-related web search behaviour: insights and implications from an analytical study of Google Trends-based query volumes. *Infect Dis Poverty* 2015;4:54.
10. Chan EH, Sahai V, Conrad C, et al. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. *PLoS Negl Trop Dis* 2011;5:e1206.
11. Hoffman SJ, Justicz V. Automatically quantifying the scientific quality and sensationalism of news records mentioning pandemics: validating a maximum entropy machine-learning model. *J Clin Epidemiol* 2016;75:47–55.
12. Southwell BG, Dolina S, Jimenez-Magdaleno K, et al. Zika virus-related news coverage and online behavior, United States, Guatemala, and Brazil. *Emerg Infect Dis* 2016;22:1320–1.
13. Towers S, Afzal S, Bernal G, et al. Mass media and the contagion of fear: the case of Ebola in America. *PLoS One* 2015;10:e0129179.
14. Wilder-Smith A, Cohn E, Lloyd DC, et al. Internet-based media coverage on dengue in Sri Lanka between 2007 and 2015. *Glob Health Action* 2016;9:31620.
15. Fu KW, Liang H, Saroha N, et al. How people react to Zika virus outbreaks on Twitter? A computational content analysis. *Am J Infect Control* 2016;44.
16. Tran T, Lee K. Understanding Citizen reactions and Ebola-related information propagation on social media, ArXiv160704318 Cs. <http://arxiv.org/abs/1607.04318> (accessed 17 Sep 2016).
17. Tausczik Y, Faasse K, Pennebaker JW, et al. Public anxiety and information seeking following the H1N1 outbreak: blogs, newspaper articles, and Wikipedia visits. *Health Commun* 2012;27:179–85.
18. Hilton S, Hunt K. UK newspapers' representations of the 2009–10 outbreak of swine flu: one health scare not over-hyped by the media? *J Epidemiol Community Health* 2011;65:941–6.
19. Kato M, Ishikawa H. Media Coverage of a global pandemic in Japan: content analysis of A/H1N1 influenza newspaper articles. *Journal of Mass Communication & Journalism* 2016;06.
20. Hossain L, Kam D, Kong F, et al. Social media in Ebola outbreak. *Epidemiol Infect* 2016;144:2136–43.
21. WHO. *Microcephaly – Brazil*: WHO, 2015. <http://www.who.int/csr/don/20-november-2015-microcephaly/en/> (accessed 4 Feb 2016).
22. StataCorp. *Stata Statistical Software*: Release 14. 2015.
23. Shmueli G, Burkom H. Statistical challenges facing early outbreak detection in biosurveillance. *Technometrics* 2010;52:39–51.
24. Velsko S, Bates T. A conceptual architecture for national biosurveillance: moving beyond situational awareness to enable digital detection of emerging threats. *Health Secur* 2016;14:189–201.
25. Klembczyk JJ, Jalalpour M, Levin S, et al. Google flu trends spatial variability validated against emergency department influenza-related visits. *J Med Internet Res* 2016;18:e175.
26. Santillana M, Nguyen AT, Louie T, et al. Cloud-based electronic health records for real-time, region-specific influenza surveillance. *Sci Rep* 2016;6:25732.
27. Müller O, Junglas I, Brocke J, et al. Utilizing big data analytics for information systems research: challenges, promises and guidelines. *European Journal of Information Systems* 2016;25:289–302.
28. WHO. *Zika situation report on 14 April 2016*: WHO, 2016. <http://www.who.int/emergencies/zika-virus/situation-report/14-april-2016/en/> (accessed 8 Sep 2016).
29. Wójcik OP, Brownstein JS, Chunara R, et al. Public health for the people: participatory infectious disease surveillance in the digital age. *Emerg Themes Epidemiol* 2014;11:7.
30. WHO. *International Health Regulations*: WHO, 2005. <http://www.who.int/ihr/publications/9789241580496/en/> (accessed 13 Sep 2016).
31. Zuckerman JN. *Principles and practice of travel medicine*: John Wiley & Sons, 2012.
32. Southwell BG, Reynolds BJ, Fowle K. Communication in infectious disease surveillance: PART 1: Communication, mass media relations, and infectious disease surveillance. In: M'ikanatha NM, Lynfield R, Beneden CAV, eds. *Et allInfectious disease surveillance*: John Wiley & Sons Ltd, 2013:607–17.
33. WHO. *Zika Situation Report on 25 August 2016*: WHO, 2016. <http://www.who.int/emergencies/zika-virus/situation-report/25-august-2016/en/> (accessed 16 May 2017).