ISDS 2018 Conference Abstracts

What value can Google search data add to existing syndromic surveillance systems?

Helen K. Green*[1], Obaghe Edeghere[1], Alex Elliot[1], Ingemar Cox[2, 3], Rachel McKendry[2] and Gillian Smith[1]

[1]Public Health England, Birmingham, United Kingdom; [2]University College London, London, United Kingdom; [3]University of Copenhagen, Copenhagen, Denmark

Objective

To carry out an observational study to explore what added value Google search data can provide to existing routine syndromic surveillance systems in England for a range of conditions of public health importance and summarise lessons learned for other countries.

Introduction

Globally, there have been various studies assessing trends in Google search terms in the context of public health surveillance[1]. However, there has been a predominant focus on individual health outcomes such as influenza, with limited evidence on the added value and practical impact on public health action for a range of diseases and conditions routinely monitored by existing surveillance programmes. A proposed advantage is improved timeliness relative to established surveillance systems. However, these studies did not compare performance against other syndromic data sources, which are often monitored daily and already offer early warning over traditional surveillance methods. Google search data could also potentially contribute to assessing the wider population health impact of public health events by supporting estimation of the proportion of the population who are symptomatic but may not present to healthcare services.

Methods

We sought to determine the added public health utility of Google search data alongside established syndromic surveillance systems in England[2] for a range of conditions of public health importance, including allergic rhinitis, scarlet fever, bronchitis, pertussis, measles, rotavirus and the health impact of heatwaves. Google search term selection was based on diagnostic and clinical codes underlying the syndromic indicators, with Google Trends[3] used to identify additional related internet search terms. Daily data was extracted from syndromic surveillance systems[2] and from the Google Health Trends Application Programming Interface (API) from 2012 to 2017 and a retrospective daily analysis undertaken during pre-identified public health events to identify a) whether signals were detected during these events and b) assess the correlation with analogous syndromic surveillance indicators through calculation of Spearman correlation coefficients and lag assessment to determine timeliness.

Results

We detected increases in Google search term frequency during public health events of interest. Good correlation was seen with comparable syndromic surveillance indicators on a daily timescale for several health outcomes, including the search terms *hayfever*, *scarlet fever*, *bronchiolitis* and *heatstroke*. Weaker correlation was seen for conditions which occur in small numbers and are vaccine preventable such as *measles* and *pertussis*. Lag analysis showed similar timeliness between daily syndromic and Google data, suggesting that, overall, Google data did not provide an earlier or delayed signal compared to syndromic surveillance indicators in England.

Conclusions

To the best of our knowledge this is the first time trends in Google search data have been compared against syndromic data for a range of public health conditions in England. These findings demonstrate the potential utility of internet search query data in conjunction with existing systems in England, with syndromic surveillance data found to be as timely as Google data. These findings also have important implications for countries where there are no such healthcare-based syndromic surveillance systems in place. Factors to consider with analyses of Google search trend data in the context of disease surveillance have been highlighted, including the choice of search terms and interpretation of the reasons behind searching the internet.

Keywords

Syndromic surveillance; Google; Public health

References

[1]Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, Murugiah K. The use of google trends in health care research: a systematic review. PLoS One. 2014 Oct 22;9(10):e109583.

[2]Public Health England. Syndromic surveillance: systems and analyses. 2017. Available online: https://www.gov.uk/government/collections/syndromic-surveillance-systems-and-analyses

[3]Google. 2017. Google Trends. Available online: https://trends.google.com/trends/

*Helen K. Green
E-mail: helen.k.green@phe.gov.uk