More diseases tracked by using Google Trends

PELAT, Camille, et al.


DOI: 10.3201/eid1508.090299
PMID: 19751610
Chung, J.-H., Song, K.S. Ko); and Asia-Pacific Research Foundation for Infectious Diseases, Seoul, South Korea (J.-H. Song, K.S. Ko)

DOI: 10.3201/eid1508.080772

References


Address for correspondence: Kwan Soo Ko, Department of Molecular Cell Biology, Sungkyunkwan University School of Medicine, 300 Chuncheon-dong, Jangan-gu, Suwon 440-746, South Korea; email: ksko@med.skku.ac.kr

More Diseases Tracked by Using Google Trends

To the Editor: The idea that populations provide data on their influenza status through information-seeking behavior on the Web has been explored in the United States in recent years (1,2). Two reports showed that queries to the Internet search engines Yahoo and Google could be informative for influenza surveillance (2,3).

Ginsberg et al. scanned the Google database and found that the sum of the results of 45 queries that most correlated with influenza incidences provided the best predictor of influenza trends (3). On the basis of trends of Google queries, these authors put their results into practice by creating a Web page dedicated to influenza surveillance. However, they did not develop the same approach for other diseases. To date, no studies have been published about the relationship of search engine query data with other diseases or in languages other than English.

We compared search trends based on a list of Google queries related to 3 infectious diseases (influenza-like illness, gastroenteritis, and chickenpox) with clinical surveillance data from the French Sentinel Network (4). Queries were constructed through team brainstorming. Each participant listed queries likely to be used for searching information about these diseases on the Web. The query time series from January 2004 through February 2009 for France were downloaded from Google Insights for Search, 1 of the 2 websites with Google Trends that enables downloading search trends from the Google database (5). Correlations with weekly incidence rates (no. cases/100,000 inhabitants) of the 3 diseases provided by the Sentinel Network were calculated for different lag periods (Pearson coefficient ρ).

The highest correlation with influenza-like illness was obtained with the query grippe –aviaire –vaccin, the French words for influenza, avian, and vaccine respectively (ρ = 0.82, p<0.001). The minus sign removed queries that contained the terms avian or vaccine. Use of the query word grippe alone resulted in a lower correlation (ρ = 0.34, p<0.001). The high double peak in 2005–2006 and the smaller peaks preceding annual epidemics observed with the query word grippe alone were decreased by this specification. However, the unusual double-peak shape of the 2005–2006 epidemic remained (online Appendix Figure, panel A, available from www.cdc.gov/EID/content/15/8/1327-appF.htm).

The highest correlation with acute diarrhea was obtained when we searched for the French word for gastroenteritis (ρ = 0.90, p<0.001). Various spellings were used to account for the presence/absence of an accent or a hyphen. The Google database was searched for gastro-enterite + gastro-entérite + gastroenterite + (gastro enterite) + (gastro entérite). The + sign coded for or, enabling searches for queries containing ≥1 of the terms. The second highest correlation was obtained when the keyword gastro (ρ = 0.88, p<0.001) (online Appendix Figure, panel B) was used. The highest correlation with chickenpox was obtained with the French word for chickenpox (varicelle) (ρ = 0.78, p<0.001) (online Appendix Figure, panel C).

A time lag of 0 weeks gave the highest correlations between the best
queries for influenza-like illness and acute diarrhea and the incidences of these diseases; the peak of the time series of Google queries occurred at the same time as that of the disease incidences. The best query for chickenpox had a 1-week lag, i.e., was 1 week behind the incidence time series.

In conclusion, for each of 3 infectious diseases, 1 well-chosen query was sufficient to provide time series of searches highly correlated with incidence. We have shown the utility of Internet search engine query data for surveillance of acute diarrhea and chickenpox in a non–English-speaking country. Thus, the ability of Internet search-engine query data to predict influenza in the United States presented by Ginsberg et al. (3) appears to have a broader application for surveillance of other infectious diseases in other countries.

This study was supported by the Institut National de la Santé et de la Recherche Médicale.

Camille Pelat, Clément Turbelin, Avner Bar-Hen, Antoine Flahault, and Alain-Jacques Valleron

Author affiliations: Institut National de la Santé et de la Recherche Médicale, Paris, France (C. Pelat, C. Turbelin, A.-J. Valleron); Université Pierre et Marie Curie–Paris 6, Paris (C. Pelat, C. Turbelin, A.-J. Valleron); Université Paris Descartes, Paris (A. Bar-Hen); and Ecole des Hautes Études en Santé Publique, Paris (A. Flahault)

DOI: 10.3201/eid1508.090299

References


Address for correspondence: Camille Pelat, Institut National de la Santé et de la Recherche Médicale, Unité Mixte de Recherche S 707, Faculté de Médecine Pierre et Marie Curie, Site Saint-Antoine, Porte 807, 27 rue Chaligny, 75571 Paris CEDEX 12, France; email: pelat@u707.jussieu.fr

To the Editor: In November 2007, a 76-year-old man with diabetes mellitus, chronic lymphocytic leukemia, and chronic obstructive pulmonary disease, who was being treated with prolonged corticosteroid therapy, received a diagnosis of invasive pulmonary aspergillosis. After 4 weeks of voriconazole therapy, cellulitis with substantial erythema, induration, and tenderness developed in his right bicep muscle. Bacterial cultures from a skin biopsy sample yielded methicillin-resistant Staphylococcus aureus (MRSA), resistant to trimethoprim/sulfamethoxazole, clindamycin, erythromycin, tetracycline, and ciprofloxacin. The patient received intravenous vancomycin for 3 weeks. After prolonged hospitalization, he was discharged but again hospitalized in February 2008 for cellulitis in the right ankle. Cultures of drainage around the ankle grew MRSA with a susceptibility pattern identical to that of the previous isolate. In April 2008, after the patient had received vancomycin for 1 week and the infection had resolved, a nasal swab showed carriage of MRSA with a susceptibility pattern identical to that of the previous isolates.

In late February 2008, the man’s 8-year-old spayed female Labrador retriever was examined for cellulitis and generalized abscessation of the neck area, which had not responded to empirical treatment with oral cephalaxin. In December 2007, she had undergone surgery for a ruptured cranial cruciate ligament (right tibial plateau-leveling osteotomy). She had chewed some sutures out after surgery, and cultures of a purulent discharge from the incision grew Pseudomonas aeruginosa;