Published in final edited form as:

Acad Radiol. 2014 September ; 21(9): 1172–1176. doi:10.1016/j.acra.2014.05.005.

Novel Data Sources for Women's Health Research: Mapping Breast Screening Online Information Seeking Through Google Trends

Soudabeh Fazeli Dehkordy, MD, MPH,

Department of Radiology, University of Michigan School of Medicine, Ann Arbor, Michigan

Ruth C. Carlos, MD, MS,

Department of Radiology, University of Michigan School of Medicine, Ann Arbor, Michigan.

University of Michigan Institute for Healthcare Policy and Innovation, Ann Arbor, Michigan

Kelli S. Hall, Ph.D., MS, and

Department of Obstetrics and Gynecology, University of Michigan School of Medicine, Ann Arbor, Michigan. Institute for Social Research, Population Studies Center, University of Michigan, Ann Arbor, Michigan

Vanessa K. Dalton, MD

Department of Obstetrics and Gynecology, University of Michigan School of Medicine, Ann Arbor, Michigan. University of Michigan Institute for Healthcare Policy and Innovation, Ann Arbor, Michigan

Abstract

Rationale and Objectives—Millions of people use online search engines every day to find health-related information and voluntarily share their personal health status and behaviors in various Web sites. Thus, data from tracking of online information seeker's behavior offer potential opportunities for use in public health surveillance and research. Google Trends is a feature of Google which allows internet users to graph the frequency of searches for a single term or phrase over time or by geographic region. We used Google Trends to describe patterns of information seeking behavior in the subject of dense breasts and to examine their correlation with the passage or introduction of dense breast notification legislation.

© 2014 Published by AUR.

Corresponding author: Soudabeh Fazeli Dehkordy, MD, MPH, Address: University Hospital Floor B2 Rm A209H, 1500 E Medical Center Dr SPC 5030, Ann Arbor MI 4810, soodabeh.fazeli@gmail.com, Phone: 734-936-4467, 704-550-8060, Fax: 734-764-241.

Reprint Address: Ruth C. Carlos, University Hospital Floor B2 Rm A209H, 1500 E Medical Center Dr SPC 5030, Ann Arbor MI 4810

The abstract has been accepted for poster presentation at the Women's Health 2014 annual meeting.

Author Disclosure Statement

No competing financial interests exist.

Disclosures:

Ruth C. Carlos is supported in part by the American College of Radiology as Deputy Editor of the JACR.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Materials and Methods—In order to capture the temporal variations of information seeking about dense breasts, the web search query “dense breast” was entered in the Google Trends tool. We then mapped the dates of legislative actions regarding dense breasts that received widespread coverage in the lay media to information seeking trends about dense breasts over time.

Results—Newsworthy events and legislative actions appear to correlate well with peaks in search volume of “dense breast”. Geographic regions with the highest search volumes have either passed, denied, or are currently considering the dense breast legislation.

Conclusions—Our study demonstrated that any legislative action and respective news coverage correlate with increase in information seeking for “dense breast” on Google, suggesting that Google Trends has the potential to serve as a data source for policy-relevant research.

Keywords

Dense breast; Legislation; information seeking; Google Trends

Introduction

Widespread access to the Internet in last few decades has made online social media and digital technologies one of the major sources of public information. Millions of people use online search engines (e.g. Google©) every day to find health-related information and voluntarily share their personal health status and behaviors in various Web sites (social networking sites, online disease support groups, etc.). Such data from tracking of online information seeker’s behavior offer potential opportunities for use in public health surveillance and research [1, 2]. A specific example of this is the examination of Google use patterns to understand timely public health issues including infectious disease hotspots, as has been shown in influenza [3, 4].

Google Trends (available at <http://google.com/trends/>) is a feature of Google which allows internet users to graph the frequency of searches for a single term or phrase. The fluctuations in the graph reflect changes in information seekers’ querying or use of the search term over time. Google Trends further provides options to compare graphs for different search terms or analyze regional differences for a specific term. Google Trends (or Google insights, the previous similar Google tool) has been primarily used as a real-time surveillance system for tracking infectious diseases such as Lyme disease [5], Tuberculosis [6], and Dengue [7]. Although Google Trends has potential implications for as a tool for surveillance and research on a variety of health topics, search query surveillance for non-infectious diseases [8] such as chronic disease or preventive health issues [9, 10], and especially for women’s health issues, has not been widely used. Therefore, we sought to demonstrate the use of Google Trends evaluating a current “hot topic” in women’s health - breast cancer screening.

The limitation of screening mammography in patients with dense breasts, along with the substantial increased risk for breast cancer, has made the issue of dense breasts a matter of great concern in recent years [11]. Dense breast notification legislation, which mandates patient’s notification of dense breast tissue in the mammography report, was first enacted in Connecticut in 2009. Since then 27 other states have passed, rejected or considered the

dense breast notification legislation as of Oct 22, 2013, as shown by Dehkordy and Carlos [12]. One of the potential consequences of legislation and attendant publicity is potential anxiety induced among women undergoing screening mammography leading to information seeking behavior. This type of behavior is difficult to quantify and describe using traditional data sources such as surveys. Google Trends has served as an alternative data source to inform practice and health policy, for example, pinpointing developing influenza hotspots ahead of traditional methods [3,4], depicting seasonal variation in mental health disorders [13], and correlating smoking cessation information seeking with smoking bans [14]. In this study, we review the use of Google Trends as a methodological innovation to broaden the women's health research toolkit by evaluating patterns for information seeking behavior in the subject of dense breasts in the United States, and its correlation to the passage or introduction of dense breast notification legislation as a case example.

Materials and Methods

Overview of Google Trends Methodology Using “Breast Cancer” As an Example

To highlight the methodological approach, we discuss the use of Google Trends using “breast cancer” as a search term in the United States from 2004-present and present this illustrative example in Figure 1. Time is represented on the x-axis and relative search frequency on the y-axis. Relative search frequency is derived in a two-step process. First, the search term volume (in this case “breast cancer”) is normalized relative to the total search volume on Google during the time period of interest. Then, the search term frequency at each time point is presented as a percentage of the highest volume of searches (for the term “breast cancer”) during the time period of interest, rather than absolute search volume. The peak volume within the time period of interest represents 100%, while the relative frequency at other time points is displayed as a proportion of this. For example, in Figure 1, the highest number of searches for “breast cancer” occurred in October 2012, represented as 100%. If the total volume of searches for the term does not reach a required threshold, estimated at a minimum of 1000 searches over the relevant time period and/or geographic region of interest, Google Trends will report the search volume index as zero. This allows one to capture the temporal variations of information seeking over time, as reflected in search volume.

Google Trends also identifies potential correlates or drivers of time-variant search patterns and can label the graph with a headline of a relevant but randomly selected news story, showing how news events potentially affect search frequency. For example, within each year, the peak search volume for “breast cancer” occurs in October, coinciding with Breast Cancer Awareness Month. Additionally, Google Trends allows display of search hot spots, the geographic regions with the highest search volumes, both as a heat map and a ranked list (e.g. Figure 1).

An Application of Google Trends to “Dense Breast”

Using the approach described above, we sought to apply this methodological tool to evaluate information seeking behavior regarding breast density using Google Trends.

To select proper terms for our search, we used the Google keyword tool (available at <https://adwords.google.com/o/KeywordTool>, accessed July 2013), which provided us with a list of keywords related to breast density with a count of how often each word had been searched. The web search query “dense breast” had the highest number of searches and was selected as our key term to evaluate information seeking behavior regarding breast density.

As mentioned previously, Google Trends temporally maps the highest volumes of Google news that contain search terms of interest, i.e. “dense breast”. To supplement Google Trends news information, we also mapped the dates of legislative actions regarding dense breasts that received coverage in the lay media [15–20] to inform trends regarding search behavior over time (Figure 2).

Results

Figure 2a shows the Google queries in the United States for “dense breast”. Generally, the mean search volume shows a rising trend over time. Newsworthy events and legislative actions appear to correlate well with peaks in search volume of “dense breast”, suggesting a corresponding trend between increased news about legislative action and increased information seeking about dense breasts. Figure 2b shows a heat map of Google queries for “dense breast”. The darker the blue, the higher the relative search volume in that state, compared to other states. In states where the total search volume for the term does not reach a required threshold, Google Trends will report the search volume index as zero, reflected as uncolored areas on the heat map. The top 10 states with the highest search volumes for “dense breast” and the status of dense breast legislation in those states are shown in Figure 2c. Regions with the highest search volumes have either passed, denied, or are currently considering the dense breast legislation. This suggests that any legislative activity, whether proposing a bill or even defeat of a bill, and associated news coverage in these states, appears to correspond with trends in increased information seeking about dense breasts, thus suggesting that legislation may be a potential driver of information seeking. Note that in states with small population and thus low search volumes, search volume index might have falsely been reported as zero, for example in Maine where the law was introduced in 2012.

Discussion

The issue of dense breasts has become a matter of great concern in recent years, leading to advocacy for policy change and legislation. Dense breast notification legislation has been introduced widely at the state-level and national-level.

Google Trends is a valuable and accessible tool which has been well used primarily for surveillance of communicable diseases and epidemics. We conducted a descriptive study to show a potential research use of Google Trends beyond describing infectious disease hotspots by exploring the relationship between information seeking behavior and passage or introduction of dense breast legislation. Our data suggest that any legislative action and respective news coverage appear to be correlated with an increase in information-seeking for “dense breast” on Google. Additionally, information-seeking regarding dense breasts varies geographically, most concentrated in the East and West Coasts where most of the legislative

action occurred [12] or in states with large populations. These data imply that Google Trends may be a useful tool for public health surveillance and research in preventive health, including women's health, which cannot be conducted using traditional data sources.

While information seeking behavior may be a surrogate measure for knowledge or awareness (or lack thereof), in this case for dense breasts and its consequences, this cannot be directly evaluated solely using Google Trends. Whether use of Google for health information results in increased knowledge or awareness of specific hot topics, like breast cancer, is unknown. Without access to individual search patterns, we cannot infer the quality of information that the search yielded or the types of information subsequently accessed by the searchers in order to evaluate potential knowledge gained. In addition, search reproducibility is dependent on the proprietary search algorithms remaining stable; Google continually optimizes their algorithms which may lead to variations in search results despite the use of the same search terms. Google Trends does not provide absolute numbers of searches, making comparative statistical analysis of correlations and relationships in the data difficult. While trends in data are useful, information seeking behavior cannot be directly attributed to legislative action or newsworthy events.

Despite these limitations, providing readily accessible and accurate on-line resources of information about breast density and educating women about efficient internet searching techniques can optimize the seeking, managing, and using the information. Further, understanding the information seeking behavior can inform health organizations, advocacy groups and health professionals regarding women's health information needs and the appropriate focus, content, and approach to education and counseling, especially in web-based venues. Additional research is needed to understand how researchers, practitioners and policy-makers can use Google Trends to measure women's health information seeking and its impact on clinically meaningful outcomes such as knowledge, anxiety, or changes in patient-provider communication practices.

Conclusion

We conducted a descriptive study to show research capabilities of Google Trends beyond describing infectious disease hotspots and specifically its potential for women's health research. Despite a number of limitations described above, Google Trends remains a potential tool to inform public health research and policy. Monitoring web queries is an efficient and timely source for identification of public health concerns and information needs, which can guide public health intervention policy. Public health organizations can also use this information to improve the accuracy and accessibility of their relevant health online sources.

References

1. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med.* 2011; 40(5 Suppl 2):S154–8. [PubMed: 21521589]
2. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PLoS One.* 2010; 5(11):e14118. [PubMed: 21124761]

3. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature*. 2009; 457(7232):1012–4. [PubMed: 19020500]
4. Valdivia A, Lopez-Alcalde J, Vicente M, Pichiule M, Ruiz M, Ordobas M. Monitoring influenza activity in Europe with Google Flu Trends: comparison with the findings of sentinel physician networks - results for 2009–10. *Euro Surveill*. 2010 Jul 22.15(29)
5. Seifter A, Schwarzwald A, Geis K, Aucott J. The utility of “Google Trends” for epidemiological research: Lyme disease as an example. *Geospat health*. 2010; 4(2):135–7. [PubMed: 20503183]
6. Zhou X, Ye J, Feng Y. Tuberculosis Surveillance by Analyzing Google Trends. *IEEE Trans Biomed Eng*. 2011 Aug.58(8)
7. Althouse BM, Ng YY, Cummings DA. Prediction of dengue incidence using search query surveillance. *PLoS Negl Trop Dis*. 2011; 5(8):e1258. [PubMed: 21829744]
8. Ayers JW, Althouse BM, Allem JP, Rosenquist JN, Ford DE. Seasonality in seeking mental health information on google. *Am J Prev Med*. 2013; 44(5):520–5. [PubMed: 23597817]
9. Ayers JW, Ribisl KM, Brownstein JS. Tracking the rise in popularity of electronic nicotine delivery systems (electronic cigarettes) using search query surveillance. *Am J Prev Med*. 2011; 40(4):448–53. [PubMed: 21406279]
10. Ayers JW, Althouse BM, Allem JP, Ford DE, Ribisl KM, Cohen JE. A novel evaluation of World No Tobacco day in Latin America. *J Med Internet Res*. 2012; 14(3):e77. [PubMed: 22634568]
11. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med*. 2007; 356(3):227–36. [PubMed: 17229950]
12. Dehkordy SF, Carlos R. Dense breast legislation in the United States: state of the states. *J Am Coll Radiol*. 2013; 10(12):899–902. [PubMed: 24295937]
13. Ayers JW, Althouse BM, Allem JP, Rosenquist JN, Ford DE. Seasonality in seeking mental health information on Google. *Am J Prev Med*. 2013; 44(5):520–5. [PubMed: 23597817]
14. Huang J, Zheng R, Emery S. Assessing the impact of the national smoking ban in indoor public places in china: evidence from quit smoking related online searches. *PLoS One*. 2013; 8(6):e65577. [PubMed: 23776504]
15. [Accessed July, 2013.] New Law May Help Women With Dense Breasts. NBC Connecticut. Available at: <http://goo.gl/11dcZ>
16. Vozzella, L. Breast-density bill to become law in Virginia. The Washington Post; Available at: <http://goo.gl/BgL0D> [Accessed July, 2013.]
17. NY bill would notify women of dense breast tissue. CBS; Available at: <http://goo.gl/0PDwH> [Accessed July, 2013.]
18. Grady, D. New Laws Add a Divisive Component to Breast Screening. The New York Times; Available at: <http://goo.gl/FHSrB> [Accessed July, 2013.]
19. Hutchison, C. Docs Weigh. ABC News; Should Women Be Warned About Breast Density?. Available at: <http://goo.gl/K3vda> [Accessed July, 2013.]
20. Colliver, V. [Accessed July, 2013.] New mammography law on breast density. San Francisco Chronicle. Available at: <http://goo.gl/dcSnY>
21. Hitti, M. Breast Density, Cancer Risk?. CBCNEWS; Available at: <http://www.cbcnews.com/stories/2007/01/17/health/webmd/main2369559.shtml> [Accessed July, 2013.]
22. Dense breast tissue hikes risk of cancer. NBCNews; Available at:<http://goo.gl/M2Up6> [Accessed July, 2013.]

Figure 1.

Google Trends output for Web search queries for the term “breast cancer” in the United States from January 2004 to January 2014. *top*: Search volume graph over time. Within each year, the peak search volume for “breast cancer” occurs in October, coinciding with Breast Cancer Awareness Month. *bottom*: Heat map and ranked list indicating regional interest during the entire time period considered.

Figure 2.

(a) Google queries for “dense breast” in the United States from January 2004 to June 2013. The X-axis represents time covering 2004–2013, the Y-axis, mean search volume (with 100 being the peak volume for any time point), Arrows represent noteworthy events that correspond to peaks in search volume for “dense breast”. These events by date of occurrence are as follows: January 2007: Boyd et al. [11] showed that women with higher breast density have more likelihood of breast cancer [21, 22]; October 2009: Connecticut law legislating dense breast communication went to effect [12]; Oct 2011: Dense breast notification legislation was introduced at federal level [12]; July 2012: New York dense breast communication bill was enacted [12], March 2013: California dense breast communication law went to effect [12]. (b) Google query volume for “dense breast” in different states of the United State from January 2004 to June 2013. 0–100 represents relative search volume. (c) States with the highest search volumes and their status of dense breast legislation