

Using Google Trends to Support Health Policy Analysis: A Case Study Examining the Correlation between Insurance-Related Search Patterns and Local-Area Characteristics

May 2017

AUTHORS

Caroline Au-Yeung
SHADAC

Christina Carberry
SHADAC

Sarah Gollust
University of Minnesota,
School of Public Health

ABOUT SHARE

State Health Access Reform Evaluation (SHARE) is a national program of the Robert Wood Johnson Foundation that supports rigorous research on issues surrounding state health reform. The SHARE program is managed by the State Health Access Data Assistance Center (SHADAC), an RWJF-funded research center in the Division of Health Policy and Management at the University of Minnesota, School of Public Health.

INTRODUCTION

As people increasingly use the internet to research and explore various topics related to health and health care, information about internet search patterns and trends is emerging as a valuable data source for health policy analyses involving a consumer behavior component. This brief describes a rich source of internet search data--Google Trends--and provides an example of how researchers are using these data to study health policy.

Google Trends

Google Trends is a particularly rich source of internet search data, since more than 65 percent of all internet searches are conducted using Google, resulting in 40,000 searches every second, on average (InternetLiveStats, 2017; Sullivan 2013). This wealth of information is steadily attracting the interest of researchers: In a 2014 systematic review of peer-reviewed studies using Google Trends data for health research, Sudhakar et al., found that Google Trends publications increased seven-fold from 2009 to 2013 (Nuti et al., 2014).



About Google Trends

Google Trends is a free, publicly available online portal that calculates the volume of searches for a user-specified term along the dimensions of time and geography. Geographic parameters in the U.S. include state, county, and various metropolitan areas. Google Trends software also gives users the ability to view search volume for a particular term filtered by search characteristics such as the use of a specific search portal (e.g., Google News, Google Images, YouTube, as opposed to a full Google search engine query) or designated search category (Arts & Entertainment, Health, Law & Government, etc.).

Advantages of Google Trends and Other Internet Search Data

Internet search data, like those available through Google Trends, offer some key advantages over traditional surveillance sources (e.g., surveys) as potential measures of health policies. First, internet search data are available in near-real-time, while traditional data can take months or years to collect. Second, search data can be automatically collected in an efficient way that requires no dedicated data collection infrastructure. Third, search data are subject to little oversight and control and therefore offer a high level of transparency. Finally, internet search data, at least from Google Trends, offer the advantages of logistically accessible, available to the public, and free of charge.

Limitations of Google Trends and Other Internet Search Data

There are also some limitations to relying on internet search data. Foremost among these is the variability of online access and internet search usage across different demographic, socioeconomic, and geographic subpopulations, such that certain groups of users might be over- or under-represented among internet search data.

Additionally, data collected over shorter periods of time may be subject to a variety of transient effects. These effects might not be a limitation in the case where the data aren't meant to generalize beyond the time period studied, but users with a longer time horizon might want to consider longer data collection periods in order to avoid these effects. There is also evidence that the search results are not stable over time, since Google constantly updates their algorithm; this instability challenges data replicability if searches are conducted at different time points (Lazer, Kennedy, King, & Vespignani, 2014).

Case Study: Geographic Characteristics and Searches for Health Insurance

A key feature of the ACA was the creation of marketplace health insurance exchanges that allow individuals to purchase health insurance plans directly from insurance companies while qualifying for premium subsidies and/or subsidized, reduced cost-sharing at certain household income levels. These exchanges were first opened for public enrollment on October 1, 2013, with coverage beginning January 1, 2014, and approximately 10 million Americans ultimately gained exchange-based health insurance coverage for calendar year 2014 (Sommers et al., 2014). Because both the marketplaces and the coverage options available through them were new, the public need for information about both was high during this first open enrollment period, data about internet search trends during the first ACA open enrollment period may provide insight about informational needs among uninsured populations and point to potential levers for enrollment outreach going forward.

Background

With funding from the Robert Wood Johnson Foundation's State Health Access Reform Evaluation (SHARE) grant program, a research team led by Dr. Sarah Gollust of the University of Minnesota and Erika Franklin Fowler of Wesleyan University has capitalized on the data available through Google Trends in order to evaluate geographic characteristics associated with Google searches during the ACA's first open enrollment period. Specifically, the team examined whether geographic regions with higher rates of uninsurance spent more time researching ACA marketplace plans and health insurance coverage generally. They also examined whether other local-level factors, including socio-demographic, political, and health insurance marketplace characteristics, were associated with higher search volumes. The team hypothesized that searches in general would be higher in areas with a higher number of uninsured people, but that the relationships may differ depending on the specific keyword searched.

HIGHLIGHTED RESEARCH

The research highlighted in this case study is detailed in the 2016 article "Search and You Shall Find: Geographic Characteristics Associated With Google Searches During the Affordable Care Act's First Enrollment Period," from Medical Care Research and Review.

Data and Methods

The researchers merged data from Google Trends with metro-area-level and state-level characteristics to examine factors associated with health insurance-related Google searches in the most populous U.S. metro areas during the for ACA open enrollment period. Each metro area was matched with its corresponding designated market area (DMA). Metro-area-level and state-level characteristics came from a variety of other data sources, including the Census Bureau's Small Area Health Insurance Estimates (SAHIE), the American Community Survey (ACS), the Bureau of Labor Statistics, county-level vote data, Kaiser Family Foundation data, and data from the Robert Wood Johnson Foundation. The team examined three key search terms: "Affordable Care Act," "Obamacare," and "health insurance."

Findings

The research team found that internet searches for health insurance terms vary greatly across the country, with areas of high pre-ACA rates of uninsurance more likely to search in higher volumes for the key terms "Obamacare" and "health insurance." This finding is adjusted for socio-demographic, political, and insurance market characteristics. Though the results show wide variations in search patterns, similar patterns are found across each of the three key search terms. In particular, geographically dispersed rural areas demonstrated a higher search volume than large urban areas, particularly urban areas on the coasts. Additionally, the probability of higher search volume for the terms "Obamacare" and "health insurance" rose with rising metro-area uninsurance rates.

Discussion

These results indicate that Google searches are an important way in which people gain information about health insurance coverage, especially in areas with greater need for coverage. Understanding how the uninsured are likely to research coverage options under the ACA can guide future outreach efforts by health advocates and policymakers, particularly as they move beyond initial broad outreach strategies to more targeted approaches tailored to specific audiences, including audiences that are more precisely geographically defined than in the past. The next step in research, according to the researchers, is to examine whether and how internet searching relates to actual insurance enrollment.

Using Google Trends Data for State Health Policy Research

As shown by the research highlighted here, Google Trends data can be used to study health policy at both the state and sub-state level. The geospatial perspective of Google Trends data is particularly useful for analyses of health insurance coverage, since health insurance markets vary markedly by geography and since states play a primary role in health systems policy and insurance market regulation. The research led by Drs. Gollust and Fowler offers a unique perspective on insurance-related information seeking patterns because it leverages Google Trends geospatial variables to look at local-level variation in ACA-related search patterns.

There is regional variation in other areas of health policy as well, including health care workforce capacity, demand for and access to care, costs of care, care quality, etc. Using Google Trends to examine regional differences on a range of topics can inform researchers and policymakers about the health and health care needs of the populations in particular geographies. Moreover, the capacity of Google Trends to explore search patterns on the dimension of time adds to its utility for analyses of policy impacts (i.e., through pre-post analyses).

Conclusion

Google Trends has the potential to be a valuable data source in examining temporal and regional variations in health-related searches. As policymakers work to improve the health and wellbeing of their constituents, further research on geographical differences in Google search terms could support and ultimately inform policy development and analysis. As the level of internet access and fluency increases across the nation, Google searches will be an increasing source of valuable information about the health and health care of the population.

REFERENCES

- Gollust, S.E., Qin, X., Wilcock, A., Baum, L.M., Barry, C.L., Niederdeppe, J., Fowler, E.F., & Karaca-Mandic, P. (2016). "Search and You Shall Find: Geographic Characteristics Associated With Google Searchers During the Affordable Care Act's First Enrollment Period." *Medical Care Research and Review*: 1-13. doi: 10.1177/1077558716660944
- Internet Live Stats. (2017). "Google Search Statistics." Internet Live Stats. Available at <http://www.internetlivestats.com/google-search-statistics/#ref-2>
- Lazer, D., Kennedy, R., King, G., & Vespignani, G.K. (2014). "The Parable of Google Flu: Traps in Big Data Analysis." *Science* 343: 1203-1205.
- Nuti, S.V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R.P., Serene, I.C., & Murugiah, K. (2014). "The Use of Google Trends in Health Care Research: A Systematic Review." *PLoS ONE*, 9 (10): e109583. <https://doi.org/10.1371/journal.pone.0109583>
- Sommers, B.D., Musco, T., Finegold, K., Gunja, M., Burke, A., & McDowell, A.M. (2014). "Health Reform and Changes in Health Insurance Coverage in 2014." *New England Journal of Medicine*, 371(9): 867-874.
- Sullivan, D. (2013). "Google Still World's Most Popular Search Engine by Far, but Share of Unique Searchers Dips Slightly." *Search Engine Land*. Available at <http://searchengineland.com/google-worlds-most-popular-search-engine-148089>